Materials and Methods

Strains and culture conditions

C. thermocellum wild-type (WT) strains ATCC 27405 DSM 1313 were obtained from their respective culture collections. *C. thermocellum* ethanol adapted (EA) mutant derived from strain ATCC 27405 has been described previously (1). *C. thermocellum* strain DSM 1313 was cultured in modified M122 medium (2, 3) at 55°C and all *C. thermocellum* growth studies contained cellobiose as the sole carbon source. *Saccharomyces cerevisiae* InvSc1 was grown at 30°C in YPD medium prior to transformation and SD-ura medium upon transformation with URA3+ plasmids, and *E. coli* Top10 was cultured in LB medium at 37°C. To select for the presence of plasmids, thiamphenicol (10 mg/ml) and chloramphenicol (12 mg/ml) were used for *C. thermocellum* and *Escherichia coli*, respectively.

Resequencing

Resequencing was conducted essentially as described previously (4). Briefly, genomic DNA from *C. thermocellum* wild-type ATCC27405 or EA mutant was isolated using a Wizard Genomic DNA purification kit (Promega, Madison, WI). Genomic DNA of wild-type ATCC 27405 and EA mutant was sent to NimbleGen facility for CGS service following the company's procedure. Pyrosequencing using the Roche 454 GS FLX System **(**454 Life Sciences, Bradford, CT) was carried out using both shotgun and paired-end DNA library preparation methods and have been deposited in the National Center for Biotechnology Information (NCBI) short-read archive database (SRX030163.2 and SRX030164.1, respectively). The GSMapper application in the 454 GS FLX software package 1.1.03 (454 Life Sciences) was used to map the reads generated from GS FLX onto the *C. thermocellum* ATCC 27405 reference genome (GenBank accession: CP000568).

Analysis of distribution of mutations inferred from pyrosequencing data

Distribution of the mutations across the genome was analyzed by calculating quantities of each type of mutation (total mutations, non-synonymous, synonymous, indels, and number of mutations in the intergenic regions) within each 1000 and each 5000 nucleotides across the genome. For each position in the genome (Li) we also calculated the Local Mutation Frequency (LMP) as LMP=1/((Mup-Li)+(Li -Mdown))/2*100, where Lup is the start position of the closest upstream mutation and Ldown is the start position of the closest downstream mutation. Mutation hot sports were identified by manual curation of the distributions and by analysis of potential functional relationships between genes comprising each hot spot. In bacteria the neighboring gene or genes may comprise an operon and are likely involved in the same biological process or metabolic pathway. To find out if mutations target such functionally related genes we calculated the number of mutations in each pair of genes and in genes that belong to an operon with subsequent manual analysis of the effected protein products and their annotation with MetaCyc pathways (5) using *C. thermocellum* pathway genome database available in the BioEnergy Science Center (BESC) Knowledgebase (http://cricket.ornl.gov/cgi-bin/beocyc_home.cgi). At least one non-synonymous mutation in two adjacent genes or in a gene of the operon was required to consider mutations as functionally related. The intergenic region was considered as functionally related to a hot spot if the distance from a mutation in the region to a mutated gene/operon was less than 400 nucleotides.

Plasmid and strain construction

Yeast gap repair cloning (6) was used to insert the WT and EA *adhE* alleles into plasmid pAMG205 (2), deleting *pyrF* and creating an artificial operon with the antibiotic resistance gene, *cat,* for expression in *C. thermocellum*. Briefly, the *adhE* alleles were PCR amplified from *C. thermocellum* WT and EA using primers GGCCTAATGACTGGCTTTTATAATAAAGGAGGTCGACGTC**ATGACGAAAATAGCGAATAAATACG** and CAGTCTTCCGACTGAGCCTTTTGTTTCTCGAGGCCCGGG**TCATTTCTTCGCACCTCCGTAAT** (priming regions in **bold**). These PCR products were each transformed along with *Zra*I+*Sma*I digested pAMG205 into *S. cerevisiae* InvSc1 (Invitrogen, Carlsbad, CA) via a modified Lazy Bones protocol (7). Plasmid was isolated using a Zymoprep Yeast Plasmid Miniprep II kit (Zymo, Orange, CA, USA) and electroporated into *E. coli* Top10 for verification, resulting in pAMG242 (mutant *adhE*) and pAMG249 (WT *adhE*). Empty vector control plasmid pAMG226 (pAMG205□*pyrF*) was constructed via restriction digestion of pAMG205 with *Zra*I and *Sma*I, followed by self-ligation of the 6.9 kb fragment and transformation into *E. coli* Top10. Plasmids were then transformed into *C. thermocellum* DSM 1313 via electroporation as described (2)with modifications. To make *C. thermocellum* electrocompetent, 500 ml cultures were grown to an OD_{600} of 0.6 - 1.0 and placed on ice to cool. Cells were then centrifuged at 6000 × g for 15 minutes at room temperature in a Beckman Coulter Avanti J-25 centrifuge with a JA-10 rotor, the supernatant was removed, and the cell pellet was washed twice with electroporation buffer (250 mM sucrose, 10% glycerol, 100 µM MOPS pH 7.0, 0.5 mM $MgCl_2$, 0.5 mM $MgSO_4$) in a similar fashion. Cells were then brought into a Coy anaerobic chamber (Coy Laboratory Products, Grass Lake, MI) and resuspended in 500 ml electroporation buffer. Plasmid DNA (ca. 500 ng) was mixed with 20 ml washed cells in a pre-chilled 1 mm electroporation cuvette, and the cell/DNA mixture was subjected to a 1.2 kV, 1.5 msec square wave pulse using a BioRad GenePulser XCell. Electroporated cells were resuspended in 1 ml growth medium and mixed with ca. 25 ml growth medium supplemented with 0.8% agar and thiamphenicol. Plates were allowed to solidify and were placed in 2.5 L AnaeroPack Rectangular Jars (bioMerieux, Durham, NC, USA) and incubated at 51°C for up to one week. Colonies were picked into liquid growth medium containing thiamphenicol stored at -80°C for preservation.

C. thermocellum growth experiments

C. thermocellum strains carrying pAMG226 (empty vector), pAMG249 (WT *adhE*), and pAMG242 (mutant *adhE*) were grown to an OD_{600} of 1.0, and used as inocula (2% v/v) into 25 ml sealed Balch tubes containing 10 ml medium supplemented with 0, 1, 2, 3, and 5% (v/v) added ethanol. Cultures were incubated at 55°C, and growth was monitored via absorbance using a ThermoSpectric Genesys 10 vis spectrophotometer at 600 nm.

ADH enzyme assays

The ADH enzyme assays are based on previously described methods (8, 9). Briefly, half liter cultures of each strain were grown to OD_{600} = 0.6 and centrifuged at 4°C at 6000 x g in a Beckman Coulter Avanti J-25 centrifuge with a JA-10 rotor. The culture was brought into the anaerobic chamber, and the supernatant was removed. All further steps were carried out in the anaerobic chamber. The cell pellet was resuspended in 4 mL 100 mM Tris-HCl (pH 7.6) 0.1 mM dithiothreitol buffer, transferred to a 10 mL glass beaker, and sonicated for 8 minutes with 10 second pulses and 10 second pauses at 50% of the max intensity using a Misonix Sonicator 4000 with a microtip. Crude cell extract was centrifuged at 14000 x g for 25 min and stored on ice until assayed. The anaerobic reaction mixture contained 100 mM Tris-HCl (pH 7.6) 0.1 mM dithiothreitol buffer, 0.5 mM NAD(P)H, 55 mM acetaldehyde, and 2 – 50 μl cell extract in 1.2 ml total volume. Decrease in absorbance was monitored at 340 nm to follow NAD(P)H oxidation (extinction coefficient 6.22 mM^{-1} cm^{-1}) using an Agilent 8453 UV-Vis spectrophotometer with Peltier controlled heating set at 55°C. Protein concentration was determined using the Bradford method.

Homology modelling of *C. thermocellum*

AdhE. Using the protein sequence of the alcohol dehydrogenase (AdhE) domain of *C. thermocellum* ALDH/ADH, homology models of wild-type and double-mutant AdhE were constructed. The HHPRED webserver (10, 11), part of the Bioinformatics Toolkit webserver (12), was used to perform multiple sequence alignments of the *C. thermocellum* AdhE sequence to potential structural templates available in the Protein Data Bank (13). The 2.7 Å resolution X-ray structure of 1,3-propanediol dehydrogenase from *Klebsiella pneumoniae* (PDB ID 3BFJ) (14) and the 1.3 Å X-ray structure of Fe-containing alcohol dehydrogenase from *Thermotoga maritima* Tm0920 (PDB ID 1O2D)(15) were selected as templates based on their homology to AdhE, and inclusion of NAD(P), Fe, or both in the structures. The program MODELLER (16-19) was used to construct homology structures including the NAD cofactor and Fe ligand, and resulting models were assessed using the Discrete Optimized Protein Energy (DOPE) method (20). The fully automated I-TASSER webserver (21) was also used to construct a homology model, and it was found that MODELLER produced structures with a more favorable (i.e. lower) DOPE score if the I-TASSER model was included as an additional template structure. The reason this step improved the model is because the AdhE sequence has two regions that are not covered by either of the two PDB templates, and I-TASSER includes secondary structure prediction for sequences with no structural template. These regions are positioned on the same face of the protein as the N-terminus, which would be joined to the C-terminus of the ALDH domain in the full-length ALDH/ADH. Thus, it is likely that these two regions interact with ALDH, providing a possible explanation for why they are not present in other single-domain alcohol dehydrogenases. To provide an additional level of validation of the homology structures, the MolProbity server (22, 23) was used to perform a Ramachandran analysis on the final models. For wild-type AdhE, 92% of all residues were found to be in favored regions, and 98.3% were in allowed regions. Seven residues were found to be outliers, an acceptable result considering the somewhat low sequence identity between the AdhE sequence and structural templates.

Molecular dynamics simulations

Molecular dynamics simulations were performed on the homology models on wild-type and double mutant (P704L, H734R) AdhE using the program GROMACS (24) with the CHARMM 27 force field (25) and TIP3P water model (26). Previously Lennard-Jones parameters for Fe(II) were used (27). Energy minimization was performed on the homology structures using the steepest descent method for 1000 steps, and then each protein was solvated in a rectangular water box of with a minimum of 10 Å from the surface of the protein to the edge of the solvent box. Sodium cations were added to neutralize the charge of the system. Periodic boundary conditions were imposed, and the Particle Mesh Ewald (PME) method (28, 29) was used to describe long-range electrostatic interactions. MD simulations were carried out with an integration time step of 2 fs. To reach the target temperature (298 K) and pressure (1 bar), the Berendsen method was used with a relaxation time of 0.1 ps (30). After a 1 ns equilibration, production simulations were performed in the NPT ensemble using the Nose-Hoover thermostat (31, 32) and the Parrinello-Rahman barostat (33, 34) with relaxation times of 1.0 ps. The production run was carried out for 10 ns, and coordinates were saved every 1 ps for analysis.

References

1. Williams TI, Combs JC, Lynn BC, & Strobel HJ (2007) Proteomic profile changes in membranes of ethanol-tolerant *Clostridium thermocellum*. *Appl. Microbiol. Biotechnol.* 74:422-432.

2. Olson DG*, et al.* (2010) Deletion of the Cel48S cellulase from *Clostridium thermocellum*. *Proc. Natl. Acad. Sci. U. S. A.* 107:17727-17732.
3. Tyurin MV, Sullivan CR, & Lynd LR (2005) Role of spontaneous current oscillations during high-efficiency electrotransformation of thermophilic anaerobes. *Appl. Environ. Microbiol.* 71:8069-8076.
4. Burdette D & Zeikus JG (1994) Purification of acetaldehyde dehydrogenase and alcohol dehydrogenases from *Thermoanaerobacter ethanolicus* 39E and characterization of the secondary-alcohol dehydrogenase (2 degrees Adh) as a bifunctional alcohol dehydrogenase--acetyl-CoA reductive thioesterase. *Biochem. J.* 302:163-170.
5. Caspi R*, et al.* (2010) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* 38:D473-D479.
6. Burke D, Dawson D, & Stearns T (2000) *Methods in yeast genetics: a Cold Spring Harbor Laboratory course manual* (Cold Spring Harbor Press, Plainview, N.Y.).
7. Elble R (1992) A simple and efficient procedure for transformation of yeasts. *Biotechniques* 13:18-20.
8. Durre P, Kuhn A, Gottwald M, & Gottschalk G (1987) Enzymatic investigations on butanol dehydrogenase and butyraldehyde dehydrogenase in extracts of *Clostridium acetobutylicum*. *Appl. Microbiol. Biotechnol.* 26:268-272.
9. Shaw AJ, Jenney FE, Adams MWW, & Lynd LR (2008) End-product pathways in the xylose fermenting bacterium, *Thermoanaerobacterium saccharolyticum*. *Enz. Microbial Technol.* 42:453-458.
10. Söding J (2005) Protein homology detection by HMM-HMM comparison. *Bioinformatics* 21:2144-2144.
11. Söding J, Biegert A, & Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 33:W244-W248.
12. Biegert A, Mayer C, Remmert M, Söding J, & Lupas AN (2006) The MPI Bioinformatics Toolkit for protein sequence analysis. *Nucleic Acids Res.* 34:W335-W339.
13. Berman HM*, et al.* (2000) The Protein Data Bank. *Nucleic Acids Res.* 28:235-242.
14. Marcal D, Rego AT, Carrondo MA, & Enguita FJ (2009) 1,3-Propanediol dehydrogenase from Klebsiella pneumoniae: decameric quaternary structure and possible subunit cooperativity. *J. Bacteriol.* 191:1143-1151.
15. Schwarzenbacher R*, et al.* (2004) Crystal structure of an iron-containing 1,3-propanediol dehydrogenase (TM0920) from Thermotoga maritima at 1.3 angstrom resolution. *Proteins, Struct. Funct. Bioinf.* 54:174-177.
16. Eswar N*, et al.* (2006) Comparative Protein Structure Modeling Using Modeller. *Curr. Protoc. Bioinformatics*, (John Wiley & Sons, Inc.), Vol Supplement 15, pp 5.6.1-5.6.30.
17. Fiser A, Do RKG, & Sali A (2000) Modeling of loops in protein structures. *Protein Sci.* 9:1753-1773.
18. Marti-Renom MA*, et al.* (2000) Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* 29:291-325.
19. Sali A & Blundell TL (1993) Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.* 234:779-815.
20. Shen MY & Sali A (2006) Statistical potential for assessment and prediction of protein structures. *Protein Sci.* 15:2507-2524.
21. Roy A, Kucukural A, & Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protocols* 5:725-738.
22. Chen VB*, et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Cryst. D* 66:12-21.
23. Davis IW*, et al.* (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research* 35:W375-W383.
24. Hess B, Kutzner C, van der Spoel D, & Lindahl E (2008) GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* 4:435-447.
25. Brooks BR*, et al.* (1983) Charmm - a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4:187-217.
26. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, & Klein ML (1983) Comparison of simple potential functions for simulating Liquid water. *J. Chem. Phys.* 79:926-935.

27. Babu CS & Lim C (2006) Empirical force fields for biologically active divalent metal cations in water. *J. Phys. Chem. A* 110:691-699.
28. Darden T, York D, & Pedersen L (1993) Particle mesh Ewald - An N.log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089-10092.
29. Essmann U*, et al.* (1995) A smooth particle Mesh Ewald method. *J. Chem. Phys.* 103:8577-8593.
30. Berendsen HJC, Postma JPM, Vangunsteren WF, Dinola A, & Haak JR (1984) Molecular-dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684-3690.
31. Hoover WG (1985) Canonical dynamics - Equilibrium phase-space distributions. *Phys. Rev. A* 31:1695-1697.
32. Nose S & Klein ML (1983) Constant pressure molecular-Dynamics for molecular-systems. *Mol. Phys.* 50:1055-1076.
33. Nose S (1984) A molecular-dynamics method for simulations in the canonical ensemble. *Mol. Phys.* 52:255-268.
34. Parrinello M & Rahman A (1981) Polymorphic transitions in single-crystals - a new molecular-dynamics method. *J. Appl. Phys.* 52:7182-7190.

Table S1. High confidence differences identified in wild-type strain by pyrosequencing

															Pryosequencing results					
Replicon	**Start Pos**	**End Pos**	**Reference sequence**	**Observed sequence**	**Region**	**Change**	**Codon number**	**Orig_AA**	**New_AA**	**Locus_Tag**	**Gene start**	**Gene stop**	**Strand**	**Product**	**# reads fwd**	**# reads rev**	**#Var**	**#Tot**	**(%)**	**Found in EA**
CP000568	3073025	3073025	C	T	INTERGENIC	Mul									10	12	22	22	100	
CP000568	1463053	1463053	T	C	NON	One	391	T	A	Cthe1221	1455455	1464223	-	glycosyltransferase 36	9	17	26	26	100	1463053
CP000568	3833351	3833351	T	C	NON	Mul	274	W	R	Cthe3232	3832532	3837541	+	YD repeat protein	21	13	34	34	100	3833351
CP000568	2315283	2315283	C	-	NON	Del	334	G	-	Cthe1939	2314949	2316283	-	magnesium transporter	18	16	34	35	97	
CP000568	2088722	2088722	G	A	NON	One	149	A	V	Cthe1766	2088334	2089167	-	glutamate 5-kinase	15	20	35	40	87	
CP000568	3373938	3373938	A	G	SYN	One	226	G	G	Cthe2855	3373261	3374499	+	conserved hypothetical protein	12	14	26	51	51	3373938
CP000568	3373944	3373944	T	C	SYN	One	228	D	D	Cthe2855	3373261	3374499	+	conserved hypothetical protein	13	14	27	53	51	
CP000568	1198646	1198646	G	A	NON	One	116	S	L	Cthe1000	1198156	1198992	-	phosphatidate cytidylyltransferase	8	2	10	21	48	
CP000568	3342625	3342627	TGA	-	NON	Del	560	N	-	Cthe2825	3340946	3342793	+	conserved hypothetical protein precursor	11	9	20	47	43	3342625
CP000568	3342638	3342638	G	A	NON	Mul	565	E	K	Cthe2825	3340946	3342793	+	conserved hypothetical protein precursor	11	10	21	49	43	3342638
CP000568	2397516	2397516	A	C	NON	Mul	60	M	R	Cthe2018	2396459	2397694	-	conserved hypothetical protein	17	10	27	65	42	2397516
CP000568	2397410	2397410	T	C	SYN	One	95	E	E	Cthe2018	2396459	2397694	-	conserved hypothetical protein	12	11	23	56	41	2397410
CP000568	2397503	2397509	ATTCATG	GTCCATA	NON	One	64	N	-	Cthe2018	2396459	2397694	-	conserved hypothetical protein	16	10	26	64	41	2397503
CP000568	2371533	2371535	GCC	-	NON	Del	113	A	-	Cthe1990	2371523	2371870	-	hypothetical protein	1	3	4	10	40	2371533
CP000568	1198190	1198190	G	A	NON	One	268	P	L	Cthe1000	1198156	1198992	-	phosphatidate cytidylyltransferase	6	8	14	35	40	
CP000568	1776760	1776760	G	A	NON	One	189	P	S	Cthe1457	1776662	1777324	-	polar amino acid ABC transporter, inner membrane	7	7	14	35	40	
CP000568	2397524	2397526	TTG	GTC	NON	One	57	Q	-	Cthe2018	2396459	2397694	-	conserved hypothetical protein	17	10	27	67	40	2397524
CP000568	488360	488360	-	GATATGCAGGAATATATACTAC	NON	Ins	191	K	-	Cthe0392	487789	488838	+	inner-membrane translocator	7	9	16	41	39	
CP000568	3342604	3342621	GATAGGTGATAAAACAGG	AAAGGGTGTTAGATCAGGGGA	NON	One	553	L	-	Cthe2825	3340946	3342793	+	conserved hypothetical protein precursor	7	8	15	39	38	3342604
CP000568	2397551	2397554	TTGG	CTGA	NON	Mul	48	Q	-	Cthe2018	2396459	2397694	-	conserved hypothetical protein	16	8	24	64	37	2397551
CP000568	956162	956162	G	AGT	NON	One	234	S	-	Cthe0792	956159	956863	-	Protein of unknown function UPF0001	8	1	9	26	35	
CP000568	2219915	2219915	-	A	INTERGENIC	Ins									2	8	10	30	33	
CP000568	2397590	2397590	A	G	SYN	Mul	35	D	D	Cthe2018	2396459	2397694	-	conserved hypothetical protein	10	4	14	43	33	
CP000568	3342592	3342592	G	A	SYN	One	549	S	S	Cthe2825	3340946	3342793	+	conserved hypothetical protein precursor	7	7	14	42	33	3342592
CP000568	582135	582135	A	G	SYN	One	66	K	K	Cthe0467	581938	582729	+	Flagellar biosynthesis/type III secretory	2	6	8	25	32	
CP000568	2394647	2394647	C	T	INTERGENIC	Mul									2	7	9	28	32	
CP000568	2796652	2796652	-	A	NON	Ins	730	L	F	Cthe2346	2795812	2798841	-	O-antigen polymerase	7	3	10	32	31	
CP000568	2397467	2397467	A	G	SYN	One	76	P	P	Cthe2018	2396459	2397694	-	conserved hypothetical protein	12	8	20	64	31	
CP000568	3373687	3373689	TCC	GCT	NON	One	143	S	-	Cthe2855	3373261	3374499	+	conserved hypothetical protein	5	4	9	30	30	
CP000568	2397277	2397277	A	G	SYN	Mul	140	L	L	Cthe2018	2396459	2397694	-	conserved hypothetical protein	3	5	8	28	29	2397277
CP000568	3373707	3373707	T	C	SYN	Mul	149	N	N	Cthe2855	3373261	3374499	+	conserved hypothetical protein	5	3	8	28	29	3373707
CP000568	1563377	1563377	-	T	NON	Ins	379	A	T	Cthe1286	1562982	1564511	-	peptidase S1 and S6, chymotrypsin/Hap	7	2	9	31	29	
CP000568	3342959	3342959	A	G	SYN	One	48	Q	Q	Cthe2826	3342816	3344054	+	conserved hypothetical protein	7	4	11	41	27	3342959
CP000568	1276285	1276285	-	A	NON	Ins	88	D	Y	Cthe1069	1276058	1276546	-	protein of unknown function UPF0054	1	2	3	12	25	
CP000568	2290234	2290234	G	CGT	NON	One	13	G	-	Cthe1919	2289574	2290272	-	MgtC/SapB transporter	1	3	4	16	25	
CP000568	3374106	3374106	A	G	SYN	One	282	R	R	Cthe2855	3373261	3374499	+	conserved hypothetical protein	2	7	9	36	25	3374106
CP000568	311030	311030	C	T	SYN	One	99	G	G	Cthe0253	310734	311723	+	ATPase associated with various cellular	3	3	6	25	24	
CP000568	2110164	2110164	G	CGA	NON	One	42	S	-	Cthe1788	2110040	2111296	+	glycosyl transferase, family 2	1	5	6	25	24	
CP000568	3327479	3327482	TTAA	AAAT	INTERGENIC	One									4	2	6	25	24	3327479
CP000568	488356	488356	T	ATA	NON	One	190	S	-	Cthe0392	487789	488838	+	inner-membrane translocator	5	4	9	39	23	
CP000568	3342563	3342563	A	G	NON	One	540	T	A	Cthe2825	3340946	3342793	+	conserved hypothetical protein precursor	5	3	8	36	22	3342563
CP000568	829398	829398	-	TA	INTERGENIC	Ins									2	2	4	19	21	
CP000568	1597777	1597777	G	A	NON	One	150	P	S	Cthe1313	1597028	1598224	-	phosphopantothenoylcysteine	5	1	6	28	21	
CP000568	2565681	2565682	AT	TA	INTERGENIC	One									1	5	6	29	21	
CP000568	2197451	2197451	C	T	INTERGENIC	One									5	3	8	38	21	
CP000568	3342559	3342559	T	G	SYN	One	538	L	L	Cthe2825	3340946	3342793	+	conserved hypothetical protein precursor	5	3	8	38	21	3342559
CP000568	3342553	3342553	G	A	SYN	One	536	A	A	Cthe2825	3340946	3342793	+	conserved hypothetical protein precursor	4	3	7	36	19	3342553
CP000568	3374120	3374124	GATTC	AGTTT	NON	One	287	R	-	Cthe2855	3373261	3374499	+	conserved hypothetical protein	2	5	7	36	19	
CP000568	3112041	3112041	G	A	SYN	Mul	467	R	R	Cthe2632	3110641	3114177	+	transcription-repair coupling factor	1	3	4	22	18	
CP000568	2721434	2721435	TG	GTT	INTERGENIC	One									1	5	6	33	18	
CP000568	975105	975106	TC	CT	NON	One	145	E	-	Cthe0806	972830	975538	-	PAS/PAC sensor hybrid histidine kinase	3	1	4	23	17	
CP000568	3281497	3281498	GG	AA	INTERGENIC	One									2	2	4	23	17	
CP000568	1221337	1221337	G	A	NON	One	102	T	I	Cthe1020	1220262	1221641	-	extracellular solute-binding protein, family 1	1	4	5	29	17	
CP000568	1155791	1155791	C	ACT	INTERGENIC	One									2	1	3	19	16	
CP000568	835435	835435	T	C	NON	One	48	V	A	Cthe0679	835293	836507	+	Serine-type D-Ala-D-Ala carboxypeptidase	2	2	4	25	16	
CP000568	2451889	2451889	G	A	SYN	One	94	S	S	Cthe2060	2451337	2452170	-	RNA methyltransferase, TrmH family, group 3	3	3	6	37	16	
CP000568	2019375	2019375	G	A	NON	One	139	S	F	Cthe1695	2018315	2019790	-	Radical SAM	1	3	4	26	15	
CP000568	149324	149324	G	CGA	NON	One	77	S	-	Cthe0116	149095	150039	+	protein of unknown function DUF199	1	2	3	21	14	
CP000568	2199693	2199693	G	A	NON	One	60	A	V	Cthe1853	2199671	2199871	-	cold-shock DNA-binding domain protein	1	2	3	23	13	
CP000568	3373046	3373052	TTAGTGG	CTAAAGGGT	NON	One	553	L	-	Cthe2854	3371390	3373237	+	conserved hypothetical protein precursor	2	2	4	30	13	3373046
CP000568	3373064	3373067	AGAA	GGGG	NON	Mul	559	R	-	Cthe2854	3371390	3373237	+	conserved hypothetical protein precursor	2	2	4	31	13	
CP000568	1477522	1477522	G	T	NON	One	179	K	N	Cthe1233	1476986	1477525	+	hypothetical protein	2	1	3	24	12	
CP000568	1810682	1810682	-	A	SYN	Ins	14	C	C	Cthe1489	1810455	1810721	-	hypothetical protein	1	2	3	24	12	
CP000568	462676	462676	G	AGT	NON	One	284	M	-	Cthe0367	461825	462694	+	NADP oxidoreductase, coenzyme F420-dependent	1	3	4	33	12	
CP000568	3291905	3291905	G	A	SYN	One	143	K	K	Cthe2791	3291477	3292385	+	binding-protein-dependent transport systems	2	2	4	32	12	
CP000568	3373056	3373060	CTCGC	AGA	NON	Mul	556	S	-	Cthe2854	3371390	3373237	+	conserved hypothetical protein precursor	2	2	4	32	12	
CP000568	1112454	1112454	-	T	NON	Ins	122	A	T	Cthe0926	1111909	1112817	-	signal recognition particle-docking protein	2	1	3	28	11	
CP000568	3061145	3061146	AG	GA	NON	One	139	A	-	Cthe2589	3060203	3061561	-	hypothetical protein	3	1	4	38	11	
CP000568	1654031	1654031	G	A	SYN	One	299	L	L	Cthe1361	1653882	1654925	-	NAD-dependent epimerase/dehydratase	2	1	3	29	10	
CP000568	2421501	2421501	G	AGT	INTERGENIC	One									2	1	3	29	10	
CP000568	2905677	2905677	C	T	NON	One	237	H	Y	Cthe2431	2904969	2905934	+	electron transport complex, RnfABCDGE type, D	1	2	3	29	10	
CP000568	3373545	3373545	A	G	SYN	Mul	95	E	E	Cthe2855	3373261	3374499	+	conserved hypothetical protein	1	2	3	30	10	

Table S2. High confidence differences identified in strain EA by pyrosequencing

Replicon	Start Pos	End Pos	Reference sequence	Observed sequence	Region	Change	Codon number	Orig_AA	New_AA	Locus_Tag	Gene start	Gene stop	Strand	Product	Pryosequencing results					
															# reads fwd	# reads rev	#Var	#Tot	(%)	Found in WT
CP000568	2239468	2239468	T	C	NON	One	205	E	G	Cthe1886	2239245	2240081	-	Integrase, catalytic region	2	1	3	3	100	
CP000568	3373090	3373093	AAGT	G	SYN	Mul	567	V	V	Cthe2854	3371390	3373237	+	conserved hypothetical protein precursor	1	2	3	3	100	
CP000568	3373707	3373707	T	C	SYN	Mul	149	N	N	Cthe2855	3373261	3374499	+	conserved hypothetical protein	1	3	4	4	100	3373707
CP000568	3379264	3379264	G	A	NON	One	359	D	N	Cthe2860	3378190	3379900	+	hypothetical protein precursor	1	3	4	4	100	
CP000568	3786797	3786797	T	GCA	INTERGENIC	One	.				.	.			3	1	4	4	100	
CP000568	3800424	3800424	A	G	NON	One	277	V	A	Cthe3208	3800183	3801253	-	Integrase, catalytic region	2	2	4	4	100	
CP000568	3629527	3629527	A	G	SYN	One	1433	P	P	Cthe3078	3625229	3632170	+	cellulosome anchoring protein, cohesin region	4	1	5	5	100	
CP000568	3786785	3786793	GAAGCTATG	TTGAATT	INTERGENIC	One	.				.	.			3	2	5	5	100	
CP000568	157500	157500	A	G	SYN	One	350	R	R	Cthe0128	156451	157563	+	Integrase, catalytic region	4	2	6	6	100	
CP000568	3373046	3373067	TTAGTGGGTTCTCGCTCAAGAA	CTAAAGGGTGTTAGATCAGGGG	NON	One	553	L	-	Cthe2854	3371390	3373237	+	conserved hypothetical protein precursor	4	3	7	7	100	3373046
CP000568	3379279	3379279	A	G	NON	One	364	T	A	Cthe2860	3378190	3379900	+	hypothetical protein precursor	3	4	7	7	100	
CP000568	816559	816559	G	A	SYN	One	252	P	P	Cthe0661	815804	817519	+	Ricin B lectin	5	3	8	8	100	
CP000568	2245692	2245692	A	G	SYN	Mul	400	C	C	Cthe1890	2244759	2246891	-	cellulosome enzyme, dockerin type I	4	4	8	8	100	
CP000568	2245718	2245720	TTG	CTT	NON	Mul	392	N	-	Cthe1890	2244759	2246891	-	cellulosome enzyme, dockerin type I	4	4	8	8	100	
CP000568	3351245	3351245	A	G	SYN	One	358	E	E	Cthe2834	3350172	3351857	+	conserved hypothetical protein precursor	4	4	8	8	100	
CP000568	3790320	3790395	AAACTAGATTGCGAAAAGCCTGCCACATTGCAGGTCACAAGGTGTTGAAGTGGTACTTCCAGTAAAACAAGGATTG	-	INTERGENIC	Del	.				.	.			4	4	8	8	100	
CP000568	1675440	1675440	G	C	SYN	One	511	T	T	Cthe1373	1671219	1676972	-	YD repeat protein	4	5	9	9	100	
CP000568	1675446	1675446	C	T	SYN	One	509	A	A	Cthe1373	1671219	1676972	-	YD repeat protein	4	5	9	9	100	
CP000568	1855353	1855353	T	C	NON	One	214	Y	C	Cthe1529	1854740	1855992	-	transposase	4	5	9	9	100	
CP000568	3379321	3379321	A	G	NON	One	378	T	A	Cthe2860	3378190	3379900	+	hypothetical protein precursor	4	5	9	9	100	
CP000568	3379335	3379335	C	T	SYN	One	382	C	C	Cthe2860	3378190	3379900	+	hypothetical protein precursor	4	5	9	9	100	
CP000568	3623485	3623487	TCC	GAA	NON	One	1324	S	-	Cthe3077	3619516	3625077	+	cellulosome anchoring protein, cohesin region	5	4	9	9	100	
CP000568	3374106	3374106	A	G	SYN	One	282	R	R	Cthe2855	3373261	3374499	+	conserved hypothetical protein	5	5	10	10	100	3374106
CP000568	628597	628597	G	A	NON	One	326	A	T	Cthe0514	627622	628737	+	DNA methylase N-4/N-6	5	5	10	10	100	
CP000568	1231266	1231266	G	A	NON	One	272	A	V	Cthe1029	1231004	1232080	-	phosphate acetyltransferase	4	6	10	10	100	
CP000568	2137485	2137485	G	A	SYN	Mul	663	T	T	Cthe1806	2132940	2139473	-	cellulosome enzyme, dockerin type I	3	7	10	10	100	
CP000568	2137527	2137527	T	C	SYN	One	649	L	L	Cthe1806	2132940	2139473	-	cellulosome enzyme, dockerin type I	5	5	10	10	100	
CP000568	3379345	3379345	A	G	NON	One	386	R	G	Cthe2860	3378190	3379900	+	hypothetical protein precursor	4	6	10	10	100	
CP000568	56259	56259	G	A	SYN	One	629	P	P	Cthe0043	54373	56601	+	glycoside hydrolase, family 9	3	8	11	11	100	
CP000568	533551	533551	A	G	NON	One	734	H	R	Cthe0423	531351	533972	+	iron-containing alcohol dehydrogenase	6	5	11	11	100	
CP000568	1460496	1460496	T	C	NON	One	1243	E	G	Cthe1221	1455455	1464223	-	glycosyltransferase 36	6	5	11	11	100	
CP000568	1997171	1997171	G	A	NON	One	211	R	Q	Cthe1671	1996540	1998108	+	Recombinase	4	7	11	11	100	
CP000568	2012225	2012225	T	C	INTERGENIC	One	.				.	.			7	4	11	11	100	
CP000568	2137509	2137509	C	T	SYN	One	655	S	S	Cthe1806	2132940	2139473	-	cellulosome enzyme, dockerin type I	5	6	11	11	100	
CP000568	3526732	3526732	C	A	NON	One	77	C	F	Cthe3003	3525027	3526961	-	hydrogenases, Fe-only	6	5	11	11	100	
CP000568	3532242	3532242	A	G	NON	One	116	F	S	Cthe3008	3531917	3532588	-	Superoxide dismutase	7	4	11	11	100	
CP000568	401007	401007	C	T	NON	One	124	G	S	Cthe0317	400705	401376	-	hypothetical protein	4	8	12	12	100	
CP000568	410756	410756	A	G	SYN	One	53	T	T	Cthe0325	410598	412535	+	NAD+ synthetase	5	7	12	12	100	
CP000568	537435	537435	A	G	SYN	One	248	K	K	Cthe0427	536692	537861	+	serine phosphatase	4	8	12	12	100	
CP000568	763262	763262	G	A	NON	One	61	P	L	Cthe0622	762637	763443	-	methylthioadenosine phosphorylase	5	7	12	12	100	
CP000568	939962	939962	A	-	NON	Del	660	F	-	Cthe0777	939329	941941	-	DNA mismatch repair protein MutS	3	9	12	12	100	
CP000568	955054	955054	G	A	SYN	One	74	V	V	Cthe0789	954484	955275	-	RNA-binding S4	8	4	12	12	100	
CP000568	1311107	1311107	A	C	NON	One	324	N	K	Cthe1101	1310786	1312078	-	hypothetical protein	5	7	12	12	100	
CP000568	1606832	1606832	T	C	NON	One	24	E	G	Cthe1323	1606222	1606902	-	GrpE protein	6	6	12	12	100	
CP000568	1818170	1818170	C	T	SYN	One	461	A	A	Cthe1498	1817744	1819552	-	methyl-accepting chemotaxis sensory transducer	9	3	12	12	100	
CP000568	2504381	2504381	T	C	NON	Mul	807	H	R	Cthe2109	2504263	2506800	-	copper amine oxidase-like	4	8	12	12	100	
CP000568	37536	37536	G	T	NON	One	223	A	S	Cthe0030	36870	37580	+	phosphatidate cytidylyltransferase	6	7	13	13	100	
CP000568	462568	462568	G	A	SYN	One	248	R	R	Cthe0367	461825	462694	+	NADP oxidoreductase, coenzyme F420-dependent	5	8	13	13	100	
CP000568	535432	535432	C	T	NON	One	135	S	L	Cthe0426	535029	536699	+	putative PAS/PAC sensor protein	9	4	13	13	100	
CP000568	921216	921216	T	C	NON	One	174	N	D	Cthe0755	920548	921735	-	aminotransferase, class I and II	9	4	13	13	100	
CP000568	1194244	1194244	G	A	SYN	One	350	L	L	Cthe0997	1194233	1195291	-	1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate	5	8	13	13	100	
CP000568	1633288	1633288	G	A	NON	One	281	P	L	Cthe1341	1632612	1634129	-	Radical SAM	4	9	13	13	100	
CP000568	1713650	1713714	CTTGATAAAATAAACAATTTTTCTGTTGTAGATAGTTTTTATCGTACCTATGAGGAATTGAAACT	-	INTERGENIC	Del	.				.	.			8	5	13	13	100	
CP000568	1770389	1770389	T	C	NON	One	254	E	G	Cthe1450	1770364	1771149	-	conserved hypothetical protein	8	5	13	13	100	
CP000568	1872581	1872581	C	T	INTERGENIC	One	.				.	.			4	9	13	13	100	
CP000568	1953853	1953853	-	C	SYN	Ins	44	G	G	Cthe1625	1953574	1953984	-	phage protein, HK97 gp10 family	7	6	13	13	100	
CP000568	2491691	2491691	C	T	SYN	One	43	S	S	Cthe2095	2491052	2491819	-	hydrolase, TatD family	8	5	13	13	100	
CP000568	2595560	2595560	T	C	NON	One	186	D	G	Cthe2179	2593348	2596116	-	Pectate lyase/Amb allergen	5	8	13	13	100	
CP000568	164619	164619	A	G	INTERGENIC	One	.				.	.			10	4	14	14	100	
CP000568	356311	356311	C	T	SYN	One	389	V	V	Cthe0287	355145	358627	+	multi-sensor hybrid histidine kinase	9	5	14	14	100	
CP000568	432634	432634	A	G	SYN	One	137	A	A	Cthe0343	432224	432826	+	flavin reductase-like, FMN-binding	7	7	14	14	100	
CP000568	488363	488391	TATGCAGGAATATATACTACGTCGGAAGC	-	NON	Del	192	I	-	Cthe0392	487789	488838	+	inner-membrane translocator	8	6	14	14	100	
CP000568	600250	600250	-	T	SYN	Ins	270	V	V	Cthe0486	599442	600653	+	GTP-binding signal recognition particle SRP54,	5	9	14	14	100	
CP000568	933177	933177	A	G	NON	One	208	I	T	Cthe0770	932459	933799	-	signal recognition particle protein	3	11	14	14	100	
CP000568	1012867	1012867	G	A	NON	One	146	A	V	Cthe0831	1012416	1013303	-	Polyprenyl synthetase	6	8	14	14	100	
CP000568	1088316	1088316	T	C	NON	One	216	H	R	Cthe0911	1088123	1088962	-	protein of unknown function DUF152	7	7	14	14	100	
CP000568	1493070	1493070	C	T	NON	One	1989	E	K	Cthe1235	1478377	1499034	-	Cellulose 1,4-beta-cellobiosidase	5	9	14	14	100	
CP000568	1501585	1501585	G	A	NON	One	320	P	L	Cthe1237	1500066	1502543	-	leucyl-tRNA synthetase	6	8	14	14	100	
CP000568	1686740	1686740	T	C	INTERGENIC	One	.				.	.			8	6	14	14	100	
CP000568	1693277	1693277	T	C	NON	One	84	R	G	Cthe1385	1690794	1693526	-	preprotein translocase, SecA subunit	9	5	14	14	100	
CP000568	2863438	2863438	C	T	SYN	One	373	A	A	Cthe2397	2862320	2863540	+	conserved hypothetical protein	4	10	14	14	100	
CP000568	3494548	3494548	C	T	SYN	One	225	K	K	Cthe2975	3494458	3495222	-	DNA-directed RNA polymerase sigma factor	7	7	14	14	100	
CP000568	3679049	3679049	C	T	NON	One	186	T	M	Cthe3117	3678493	3679248	+	zinc/iron permease	10	4	14	14	100	
CP000568	3784408	3784408	T	C	SYN	One	780	F	F	Cthe3200	3782069	3784711	+	alanyl-tRNA synthetase	9	5	14	14	100	
CP000568	355031	355031	G	A	SYN	One	402	A	A	Cthe0286	353826	355100	+	response regulator receiver sensor signal	7	8	15	15	100	
CP000568	663719	663719	G	-	INTERGENIC	Del	.				.	.			6	9	15	15	100	
CP000568	717803	717803	A	G	INTERGENIC	One	.				.	.			5	10	15	15	100	
CP000568	766865	766865	A	G	NON	One	612	N	D	Cthe0624	765032	769837	+	glycoside hydrolase, family 9-like Ig-like	8	7	15	15	100	
CP000568	774279	774279	A	-	INTERGENIC	Del	.				.	.			8	7	15	15	100	
CP000568	1221346	1221346	G	A	NON	One	99	A	V	Cthe1020	1220262	1221641	-	extracellular solute-binding protein, family 1	9	6	15	15	100	
CP000568	1284343	1284343	T	C	INTERGENIC	One	.				.	.			9	6	15	15	100	
CP000568	1350542	1350542	-	C	SYN	Ins	171	G	G	Cthe1137	1350412	1351053	-	hypothetical protein	4	11	15	15	100	
CP000568	1357656	1357656	T	C	NON	One	368	Q	R	Cthe1141	1354697	1358758	-	hypothetical protein	6	9	15	15	100	
CP000568	1451061	1451061	T	C	NON	One	258	V	A	Cthe1217	1450289	1452619	+	ATP-dependent Clp protease, ATP-binding subunit	12	3	15	15	100	
CP000568	2008452	2008452	A	G	SYN	One	170	L	L	Cthe1684	2008356	2008961	-	transposase IS3/IS911	11	4	15	15	100	
CP000568	2366484	2366484	C	T	NON	One	697	S	N	Cthe1986	2366420	2368573	-	phage/plasmid primase, P4 family	10	5	15	15	100	
CP000568	3573194	3573194	C	T	SYN	One	278	G	G	Cthe3047	3572361	3573773	+	Peptidoglycan glycosyltransferase	9	6	15	15	100	
CP000568	3629428	3629428	G	A	SYN	Mul	1400	P	P	Cthe3078	3625229	3632170	+	cellulosome anchoring protein, cohesin region	10	5	15	15	100	
CP000568	3661256	3661256	C	T	SYN	One	99	D	D	Cthe3104	3660960	3661730	+	putative CoA-substrate-specific enzyme activase	7	8	15	15	100	
CP000568	3833351	3833351	T	C	NON	One	274	W	R	Cthe3232	3832532	3837541	+	YD repeat protein	7	9	16	16	100	3833351
CP000568	91329	91329	C	T	SYN	One	117	D	D	Cthe0059	90979	92439	+	type 3a, cellulose-binding	9	7	16	16	100	
CP000568	249888	249888	G	A	NON	One	115	V	M	Cthe0206	249546	251777	+	UvrD/REP helicase	7	9	16	16	100	
CP000568	1150774	1150774	G	A	INTERGENIC	One	.				.	.			12	4	16	16	100	
CP000568	1500998	1500998	C	T	NON	One	516	A	T	Cthe1237	1500066	1502543	-	leucyl-tRNA synthetase	5	11	16	16	100	
CP000568	1570314	1570314	T	C	NON	One	280	N	D	Cthe1293	1570144	1571151	-	metal dependent phosphohydrolase	10	6	16	16	100	
CP000568	1783212	1783212	T	C	NON	One	26	N	S	Cthe1464	1783007	1783288	-	Endoribonuclease L-PSP	10	6	16	16	100	
CP000568	1896686	1896686	A	T	NON	One	238	D	V	Cthe1568	1895974	1896849	+	Radical SAM	7	9	16	16	100	
CP000568	1900767	1900767	C	T	NON	One	34	R	W	Cthe1572	1900668	1901471	+	ABC transporter related	7	9	16	16	100	
CP000568	1900995	1900995	G	A	NON	One	110	G	R	Cthe1572	1900668	1901471	+	ABC transporter related	6	10	16	16	100	
CP000568	1957578	1957578	T	C	NON	One	119	R	G	Cthe1630	1956610	1957932	-	phage portal protein, HK97 family	8	8	16	16	100	
CP000568	2316023	2316023	G	A	SYN	One	87	D	D	Cthe1939	2314949	2316283	-	magnesium transporter	10	6	16	16	100	
CP000568	2322939	2322939	T	C	NON	One	115	D	G	Cthe1947	2322686	2323282	-	Redoxin	11	5	16	16	100	
CP000568	2895379	2895379	A	G	SYN	One	462	S	S	Cthe2423	2893768	2896764	-	hypothetical protein	8	8	16	16	100	
CP000568	3193128	3193128	G	A	SYN	One	111	Q	Q	Cthe2706	3192796	3193905	+	ABC transporter related	7	9	16	16	100	
CP000568	3639607	3639607	C	T	SYN	Mul	10	A	A	Cthe3084	3638500	3639636	-	putative RNA methylase	5	11	16	16	100	
CP000568	3711703	3711703	A	G	NON	One	711	K	E	Cthe3141	3709573	3712068	+	lipolytic enzyme, G-D-S-L	8	8	16	16	100	
CP000568	39120	39120	A	G	SYN	One	322	I	I	Cthe0032	38313	40085	-	glycoside hydrolase, family 26	9	8	17	17	100	
CP000568	58614	58614	A	G	INTERGENIC	One	.				.	.			10	7	17	17	100	
CP000568	137939	137939	T	C	NON	One	13	I	T	Cthe0106	137902	139143	+	GTP cyclohydrolase II	10	7	17	17	100	
CP000568	340690	340690	C	T	SYN	One	7	L	L	Cthe0274	339019	340710	-	glycoside hydrolase, family 9	8	9	17	17	100	
CP000568	565149	565149	C	T	SYN	One	118	N	N	Cthe0450	564796	565251	+	Ribonucleotide reductase regulator NrdR-like	4	13	17	17	100	
CP000568	651983	651983	A	-	NON	Del	165	L	-	Cthe0534	651489	653672	+	ABC-type bacteriocin transporter	10	7	17	17	100	
CP000568	691119	691119	C	T	NON	One	53	L	F	Cthe0564	690963	691919	+	Trans-hexaprenyltranstransferase	11	6	17	17	100	
CP000568	805314	805314	A	-	NON	Del	82	P	-	Cthe0651	805069	805407	+	hypothetical protein	7	10	17	17	100	
CP000568	950945	950945	A	T	NON	One	887	D	E	Cthe0787	950810	953605	-	isoleucyl-tRNA synthetase	4	13	17	17	100	
CP000568	1034043	1034043	A	G	NON	One	222	D	G	Cthe0858	1033379	1034368	+	protein of unknown function DUF1432	5	12	17	17	100	
CP000568	1220381	1220381	T	C	NON	One	421	S	G	Cthe1020	1220262	1221641	-	extracellular solute-binding protein, family 1	8	9	17	17	100	
CP000568	1411010	1411010	A	T	NON	One	356	I	N	Cthe1182	1410808	1412076	-	protein of unknown function DUF58	7	10	17	17	100	
CP000568	1480289	1480289	T	C	NON	One	6249	Y	C	Cthe1235	1478377	1499034	-	Cellulose 1,4-beta-cellobiosidase	7	10	17	17	100	
CP000568	1592620	1592620	A	G	NON	One	210	D	G	Cthe1309	1591992	1593065	+	Radical SAM	9	8	17	17	100	

CP000568	1678574	1678574	C	T	NON	One	322	G	S	Cthe1375	1678188	1679537	-	aspartate kinase	8	9	17	17	100	
CP000568	1778192	1778192	A	-	NON	Del	9	L	-	Cthe1458	1777408	1778217	-	extracellular solute-binding protein, family 3	11	6	17	17	100	
CP000568	1821198	1821198	T	C	NON	One	216	W	R	Cthe1499	1820553	1822307	+	ammonium transporter	10	7	17	17	100	
CP000568	2141861	2141861	G	A	NON	One	307	A	V	Cthe1807	2141116	2142780	-	conserved hypothetical protein	9	8	17	17	100	
CP000568	2460536	2460536	T	C	INTERGENIC	One	.				.	.			7	10	17	17	100	
CP000568	2657566	2657566	C	T	NON	One	84	G	D	Cthe2229	2656764	2657816	-	N-acetylneuraminate synthase	10	7	17	17	100	
CP000568	2661872	2661872	A	G	SYN	One	103	L	L	Cthe2234	2661795	2662178	-	hypothetical protein	10	7	17	17	100	
CP000568	2775727	2775727	T	C	NON	One	153	H	R	Cthe2329	2775348	2776184	-	hypothetical protein	7	10	17	17	100	
CP000568	2920132	2920132	T	C	NON	One	193	L	P	Cthe2448	2919555	2920538	+	inner-membrane translocator	7	10	17	17	100	
CP000568	3458321	3458321	G	A	NON	One	209	A	T	Cthe2943	3457697	3458914	+	ABC-2 type transporter	8	9	17	17	100	
CP000568	3496333	3496333	G	A	NON	One	254	E	K	Cthe2977	3495574	3496701	+	MAEBL, putative	7	10	17	17	100	
CP000568	3752468	3752468	A	G	NON	One	18	I	T	Cthe3174	3752131	3752520	-	conserved hypothetical protein	6	11	17	17	100	
CP000568	3373938	3373938	A	G	SYN	One	226	G	G	Cthe2855	3373261	3374499	+	conserved hypothetical protein	6	12	18	18	100	3373938
CP000568	1463053	1463053	T	C	NON	One	391	T	A	Cthe1221	1455455	1464223	-	glycosyltransferase 36	7	11	18	18	100	1463053
CP000568	232355	232355	A	G	NON	One	523	T	A	Cthe0191	230789	232588	+	proteinase inhibitor I4, serpin	10	8	18	18	100	
CP000568	580932	580932	G	A	NON	One	4	G	S	Cthe0466	580923	581945	+	flagellar motor switch protein FliG	9	9	18	18	100	
CP000568	681843	681843	G	A	INTERGENIC	One	.				.	.			7	11	18	18	100	
CP000568	815352	815352	C	T	NON	One	692	T	M	Cthe0660	813278	815599	+	glycoside hydrolase, family 81	11	7	18	18	100	
CP000568	915475	915475	A	G	NON	One	213	V	A	Cthe0749	915285	916112	-	binding-protein-dependent transport systems	12	6	18	18	100	
CP000568	1072740	1072740	C	T	NON	One	466	E	K	Cthe0896	1072333	1074135	-	DNA primase	6	12	18	18	100	
CP000568	1400010	1400010	G	A	NON	One	177	V	I	Cthe1172	1399482	1400630	+	protein of unknown function DUF362	9	9	18	18	100	
CP000568	1552089	1552089	C	T	NON	One	35	E	K	Cthe1275	1551754	1552191	-	H+-ATPase subunit H	8	10	18	18	100	
CP000568	1625686	1625686	C	T	NON	One	409	T	I	Cthe1334	1624461	1626023	+	FHA domain containing protein	10	8	18	18	100	
CP000568	1645595	1645595	T	C	SYN	One	21	V	V	Cthe1352	1644326	1645657	-	UDP-glucose 6-dehydrogenase	10	8	18	18	100	
CP000568	1734027	1734027	T	C	NON	One	33	Q	R	Cthe1413	1732208	1734124	-	hypothetical protein	12	6	18	18	100	
CP000568	1943386	1943386	T	C	NON	One	461	D	G	Cthe1613	1942344	1944767	-	glycosyl hydrolase-like	10	8	18	18	100	
CP000568	2032923	2032923	G	A	NON	One	412	L	F	Cthe1709	2031877	2034156	-	Phage-related protein-like	9	9	18	18	100	
CP000568	2139279	2139279	G	C	SYN	One	65	G	G	Cthe1806	2132940	2139473	-	cellulosome enzyme, dockerin type I	10	8	18	18	100	
CP000568	2419590	2419590	C	T	INTERGENIC	One	.				.	.			10	8	18	18	100	
CP000568	2739844	2739912	TTTCAATTCCTCATAGGTACGATAAAAACCCATTAGATAGAGAATTGAAAAATTGCCTGCAAATGTAGA	-	INTERGENIC	Del	.				.	.			8	10	18	18	100	
CP000568	3019185	3019185	C	T	NON	Mul	317	A	T	Cthe2550	3019150	3020133	-	glycosyltransferase sugar-binding region	15	3	18	18	100	
CP000568	3046517	3046517	-	A	SYN	Ins	22	R	R	Cthe2574	3046452	3047264	+	binding-protein-dependent transport systems	11	7	18	18	100	
CP000568	3309731	3309731	C	A	SYN	One	62	G	G	Cthe2808	3309546	3310583	+	transcriptional regulator, LacI family	9	9	18	18	100	
CP000568	3440083	3440083	C	T	NON	One	17	A	V	Cthe2917	3440034	3440432	+	ribosomal protein S8	8	10	18	18	100	
CP000568	3653236	3653236	T	C	NON	One	271	T	A	Cthe3096	3653111	3654046	-	hypothetical protein	3	15	18	18	100	
CP000568	92472	92472	A	-	INTERGENIC	Del	.				.	.			10	9	19	19	100	
CP000568	253589	253589	C	T	STOP	One	328	Q	*	Cthe0208	252608	254899	+	single-stranded-DNA-specific exonuclease RecJ	7	12	19	19	100	
CP000568	373018	373018	A	G	SYN	One	66	S	S	Cthe0298	372821	374623	+	methyl-accepting chemotaxis sensory transducer	9	10	19	19	100	
CP000568	507789	507789	G	A	NON	One	144	A	V	Cthe0406	507482	508219	-	conserved hypothetical protein	9	10	19	19	100	
CP000568	624149	624149	A	G	NON	Mul	119	N	D	Cthe0511	623795	625117	+	histidine kinase	11	8	19	19	100	
CP000568	689837	689837	A	G	NON	One	303	M	V	Cthe0561	688931	689977	+	ApbE-like lipoprotein	10	9	19	19	100	
CP000568	1031281	1031282	AA	-	INTERGENIC	Del	.				.	.			13	6	19	19	100	
CP000568	1284894	1284894	C	T	SYN	One	179	A	A	Cthe1077	1284358	1284987	+	hypothetical protein	7	12	19	19	100	
CP000568	2098330	2098330	T	C	NON	One	75	Y	C	Cthe1776	2098068	2098553	-	protein of unknown function UPF0079	10	9	19	19	100	
CP000568	2147594	2147594	-	T	SYN	Ins	62	I	I	Cthe1811	2147403	2147777	-	transcriptional repressor, CopY family	9	10	19	19	100	
CP000568	2238813	2238813	G	A	NON	One	125	D	N	Cthe1885	2238441	2239013	+	phage integrase-like SAM-like	9	10	19	19	100	
CP000568	2286170	2286170	C	T	SYN	One	553	T	T	Cthe1917	2285111	2287828	-	ATPase, P-type (transporting), HAD superfamily,	9	10	19	19	100	
CP000568	2478733	2478733	T	C	NON	One	745	N	D	Cthe2086	2478446	2480965	-	peptidase U32	12	7	19	19	100	
CP000568	2713030	2713030	G	A	NON	One	146	R	W	Cthe2282	2712653	2713465	-	MCP methyltransferase, CheR-type	8	11	19	19	100	
CP000568	3001848	3001848	G	A	NON	One	103	D	N	Cthe2532	3001542	3002390	+	sulfate ABC transporter, inner membrane subunit	11	8	19	19	100	
CP000568	3090411	3090411	G	A	NON	One	1103	D	N	Cthe2612	3087105	3090941	+	Fibronectin, type III	9	10	19	19	100	
CP000568	3132469	3132469	G	A	NON	One	368	P	L	Cthe2650	3131748	3133571	-	polysaccharide biosynthesis protein CapD	8	11	19	19	100	
CP000568	3227821	3227821	G	A	NON	One	258	V	I	Cthe2736	3227050	3228756	+	phosphoenolpyruvate-protein phosphotransferase	10	9	19	19	100	
CP000568	3236755	3236755	A	G	INTERGENIC	One	.				.	.			11	8	19	19	100	
CP000568	3811733	3811733	A	G	NON	One	37	D	G	Cthe3217	3811624	3812916	+	hypothetical protein	7	12	19	19	100	
CP000568	136133	136133	A	G	NON	One	35	D	G	Cthe0104	136030	137127	+	riboflavin biosynthesis protein RibD	8	12	20	20	100	
CP000568	351245	351245	G	A	SYN	One	143	P	P	Cthe0284	350817	352175	+	protein of unknown function DUF815	12	8	20	20	100	
CP000568	556152	556152	C	T	INTERGENIC	One	.				.	.			9	11	20	20	100	
CP000568	613076	613076	A	G	SYN	One	44	S	S	Cthe0502	612945	613811	+	hypothetical protein	11	9	20	20	100	
CP000568	812836	812836	A	-	NON	Del	1	L	-	Cthe0659	812448	812837	-	conserved hypothetical protein	11	9	20	20	100	
CP000568	1332697	1332697	T	C	SYN	One	127	S	S	Cthe1116	1330912	1333077	-	HMG-I and HMG-Y, DNA-binding	10	10	20	20	100	
CP000568	1580697	1580697	T	C	NON	One	143	I	M	Cthe1302	1579458	1581125	-	conserved hypothetical protein	9	11	20	20	100	
CP000568	1583731	1583731	T	C	NON	One	260	K	E	Cthe1304	1583135	1584508	-	PhoH-like protein	11	9	20	20	100	
CP000568	1796657	1796657	T	C	INTERGENIC	One	.				.	.			6	14	20	20	100	
CP000568	2168008	2168008	T	C	NON	One	485	D	G	Cthe1829	2166165	2169461	-	Chromosome segregation ATPases-like	9	11	20	20	100	
CP000568	2386244	2386244	G	A	NON	Mul	136	L	F	Cthe2006	2385984	2386649	-	hypothetical protein precursor	5	15	20	20	100	
CP000568	2517992	2517992	T	C	NON	One	275	I	V	Cthe2119	2516532	2518814	-	glycoside hydrolase, family 10	11	9	20	20	100	
CP000568	2692640	2692640	C	T	NON	One	58	P	S	Cthe2263	2692469	2692936	+	H+-transporting two-sector ATPase, C subunit	6	14	20	20	100	
CP000568	2971415	2971415	T	C	NON	One	25	D	G	Cthe2504	2970199	2971488	-	3D	9	11	20	20	100	
CP000568	3031718	3031718	G	A	STOP	One	156	W	*	Cthe2560	3031251	3031799	+	dTDP-4-dehydrorhamnose 3,5-epimerase	11	9	20	20	100	
CP000568	3064045	3064045	G	A	INTERGENIC	One	.				.	.			11	9	20	20	100	
CP000568	3082151	3082151	A	G	INTERGENIC	Mul	.				.	.			13	7	20	20	100	
CP000568	3304230	3304230	A	G	SYN	One	9	V	V	Cthe2803	3304204	3304983	+	binding-protein-dependent transport systems	5	15	20	20	100	
CP000568	3445873	3445873	G	A	NON	One	51	D	N	Cthe2927	3445723	3445941	+	translation initiation factor IF-1	11	9	20	20	100	
CP000568	3485280	3485280	T	G	NON	One	224	F	V	Cthe2966	3484611	3485357	+	hypothetical protein	10	10	20	20	100	
CP000568	3526997	3526997	A	G	NON	One	494	*	Q	Cthe3004	3526995	3528476	-	ferredoxin	13	7	20	20	100	
CP000568	3608159	3608159	G	A	NON	One	167	G	D	Cthe3066	3607660	3608385	+	ABC transporter related	9	11	20	20	100	
CP000568	3676841	3676841	C	A	SYN	One	233	V	V	Cthe3115	3676143	3677156	+	conserved hypothetical protein	8	12	20	20	100	
CP000568	3731348	3731348	G	A	SYN	One	178	V	V	Cthe3158	3730815	3732743	+	putative aconitate hydratase	10	10	20	20	100	
CP000568	104530	104530	G	A	NON	One	380	A	V	Cthe0071	102852	105668	-	Cellulose 1,4-beta-cellobiosidase	10	11	21	21	100	
CP000568	531270	531270	T	C	NON	One	195	M	T	Cthe0422	530687	531370	+	CoA-binding	13	8	21	21	100	
CP000568	632372	632372	C	G	NON	One	6	V	L	Cthe0517	632019	632387	-	hypothetical protein	6	15	21	21	100	
CP000568	1080738	1080738	G	A	NON	One	34	L	F	Cthe0903	1079881	1080837	-	protein-export membrane protein SecF	14	7	21	21	100	
CP000568	1113552	1113552	T	C	NON	One	959	T	A	Cthe0927	1112854	1116426	-	chromosome segregation protein SMC	10	11	21	21	100	
CP000568	1141582	1141582	A	G	SYN	One	304	H	H	Cthe0952	1141213	1142493	-	dihydroorotase, multifunctional complex type	11	10	21	21	100	
CP000568	1530181	1530181	C	T	NON	One	31	V	M	Cthe1259	1529729	1530271	-	CDP-alcohol phosphatidyltransferase	13	8	21	21	100	
CP000568	1593065	1593065	G	A	STOP	One	358	*	*	Cthe1309	1591992	1593065	+	Radical SAM	12	9	21	21	100	
CP000568	1726591	1726591	G	A	NON	One	1	M	I	Cthe1408	1726589	1727287	+	two component transcriptional regulator, winged	9	12	21	21	100	
CP000568	2248459	2248459	T	C	SYN	Mul	120	S	S	Cthe1893	2248150	2248818	-	conserved hypothetical protein	6	15	21	21	100	
CP000568	2604382	2604382	A	-	INTERGENIC	Del	.				.	.			16	5	21	21	100	
CP000568	2692670	2692670	A	G	NON	One	68	I	V	Cthe2263	2692469	2692936	+	H+-transporting two-sector ATPase, C subunit	8	13	21	21	100	
CP000568	2798536	2798536	T	C	NON	One	102	I	M	Cthe2346	2795812	2798841	-	O-antigen polymerase	8	13	21	21	100	
CP000568	2825521	2825521	T	C	NON	One	52	I	M	Cthe2361	2823193	2825676	-	DNA gyrase, A subunit	9	12	21	21	100	
CP000568	2825585	2825585	C	T	NON	One	31	S	N	Cthe2361	2823193	2825676	-	DNA gyrase, A subunit	9	12	21	21	100	
CP000568	2997878	2997878	C	T	SYN	One	2	L	L	Cthe2529	2997875	2998855	+	Porphobilinogen synthase	15	6	21	21	100	
CP000568	3004674	3004674	C	T	SYN	One	47	L	L	Cthe2535	3004536	3005255	+	adenylylsulfate reductase, thioredoxin	11	10	21	21	100	
CP000568	3467210	3467210	A	G	NON	One	430	V	A	Cthe2949	3466795	3468498	-	Pectinesterase	10	11	21	21	100	
CP000568	3491215	3491215	C	T	NON	One	310	A	T	Cthe2972	3490091	3492142	-	glycoside hydrolase, family 11	9	12	21	21	100	
CP000568	3501413	3501413	G	A	SYN	One	344	E	E	Cthe2983	3500382	3502037	+	methyl-accepting chemotaxis sensory transducer	12	9	21	21	100	
CP000568	3513247	3513247	C	T	STOP	One	89	Q	*	Cthe2992	3512983	3513507	+	RNA polymerase, sigma-24 subunit, ECF subfamily	13	8	21	21	100	
CP000568	3594982	3594982	A	G	NON	One	125	H	R	Cthe3059	3594609	3595157	+	phosphotransferase KptA/Tpt1	9	12	21	21	100	
CP000568	8023	8023	T	C	NON	One	1384	F	L	Cthe0004	3874	9702	+	YD repeat containing protein	11	11	22	22	100	
CP000568	314668	314668	A	-	NON	Del	265	K	-	Cthe0256	313876	316014	+	histidine kinase	13	9	22	22	100	
CP000568	493377	493377	G	A	SYN	Mul	474	Q	Q	Cthe0396	491956	493680	+	ABC transporter related	10	12	22	22	100	
CP000568	512193	512193	G	A	NON	One	783	D	N	Cthe0408	509847	512435	+	diguanylate cyclase/phosphodiesterase with	9	13	22	22	100	
CP000568	519958	519958	C	T	SYN	One	164	D	D	Cthe0413	519467	523141	+	glycoside hydrolase, family 9	11	11	22	22	100	
CP000568	588352	588352	G	A	NON	One	76	A	T	Cthe0473	588127	588516	+	flagellar operon protein	12	10	22	22	100	
CP000568	668438	668438	A	G	NON	One	64	M	V	Cthe0547	668249	669205	+	periplasmic solute binding protein	14	8	22	22	100	
CP000568	707138	707138	C	T	NON	One	385	G	S	Cthe0578	706080	708290	-	glycoside hydrolase, family 9	11	11	22	22	100	
CP000568	1028888	1028888	C	T	NON	One	233	A	T	Cthe0853	1027899	1029584	-	type II secretion system protein E	12	10	22	22	100	
CP000568	1181521	1181521	T	C	SYN	One	19	V	V	Cthe0987	1180639	1181577	-	riboflavin biosynthesis protein RibF	10	12	22	22	100	
CP000568	1683605	1683605	T	C	SYN	One	79	I	I	Cthe1379	1683369	1684457	+	oxidoreductase-like	9	13	22	22	100	
CP000568	2040360	2040360	T	C	NON	One	210	Q	R	Cthe1719	2039699	2040988	-	phage major capsid protein, HK97 family	11	11	22	22	100	
CP000568	2255649	2255650	TT	-	NON	Del	115	K	-	Cthe1902	2255114	2255992	-	Pyridoxal-5'-phosphate-dependent enzyme, beta	13	9	22	22	100	
CP000568	2393240	2393240	C	T	INTERGENIC	One	.				.	.			16	6	22	22	100	
CP000568	2514194	2514194	G	A	NON	One	129	V	I	Cthe2116	2513810	2514571	+	binding-protein-dependent transport systems	12	10	22	22	100	
CP000568	2917990	2917990	A	G	SYN	One	354	S	S	Cthe2446	2916929	2918047	+	ABC-type sugar transport system periplasmic	10	12	22	22	100	
CP000568	3324904	3324904	G	A	NON	One	250	P	L	Cthe2815	3324051	3325652	-	lysyl-tRNA synthetase	11	11	22	22	100	
CP000568	3383311	3383311	A	G	NON	One	79	K	R	Cthe2865	3383076	3385229	+	phage/plasmid primase, P4 family	13	9	22	22	100	
CP000568	3682096	3682096	A	G	NON	One	498	N	D	Cthe3120	3680605	3684141	+	pyruvate flavodoxin/ferredoxin	10	12	22	22	100	
CP000568	3710176	3710176	G	A	NON	One	202	A	T	Cthe3141	3709573	3712068	+	lipolytic enzyme, G-D-S-L	12	10	22	22	100	
CP000568	281873	281873	C	T	INTERGENIC	One	.				.	.			10	13	23	23	100	

CP000568	385361	385361	A	G	SYN	One		75	G	G	Cthe0307	385137	385514	+	conserved hypothetical protein	12	11	23	23	100
CP000568	556255	556255	A	G	INTERGENIC	One	.					.	.			12	11	23	23	100
CP000568	1015063	1015063	A	G	NON	One		76	S	P	Cthe0834	1014851	1015288	-	NusB antitermination factor	7	16	23	23	100
CP000568	1771468	1771468	C	T	NON	One		120	R	K	Cthe1451	1771353	1771826	-	GCN5-related N-acetyltransferase	14	9	23	23	100
CP000568	1917309	1917309	T	C	SYN	One		159	P	P	Cthe1586	1916130	1917785	-	binding-protein-dependent transport systems	12	11	23	23	100
CP000568	2178645	2178645	T	C	SYN	One		2	S	S	Cthe1835	2175342	2178650	-	Viral A-type inclusion protein repeat	10	13	23	23	100
CP000568	2210573	2210573	C	T	SYN	One		240	N	N	Cthe1864	2209854	2210768	+	acetylglutamate kinase	12	11	23	23	100
CP000568	2795536	2795536	G	A	NON	One		73	T	I	Cthe2345	2794605	2795753	-	DegT/DnrJ/EryC1/StrS aminotransferase	14	9	23	23	100
CP000568	2831640	2831640	T	-	NON	Del		95	I	-	Cthe2366	2831302	2831922	-	single-stranded nucleic acid binding R3H	15	8	23	23	100
CP000568	2908769	2908769	C	T	NON	One		55	R	K	Cthe2436	2908696	2908932	-	hypothetical protein	10	13	23	23	100
CP000568	3116739	3116739	C	T	SYN	One		80	L	L	Cthe2635	3116500	3116934	+	pyridoxamine 5'-phosphate oxidase-related,	10	13	23	23	100
CP000568	3131416	3131416	C	T	SYN	One		39	Q	Q	Cthe2649	3130705	3131532	-	HpcH/HpaI aldolase	15	8	23	23	100
CP000568	3147607	3147607	A	G	SYN	Mul		188	G	G	Cthe2666	3145576	3148170	-	ATPase, P-type (transporting), HAD superfamily,	13	10	23	23	100
CP000568	3537881	3537881	G	A	NON	One		97	A	V	Cthe3013	3537193	3538170	-	hydrogenase expression/formation protein HypE	10	13	23	23	100
CP000568	96306	96306	A	G	SYN	One		62	A	A	Cthe0063	96121	96456	+	Rubredoxin-type Fe(Cys)4 protein	17	7	24	24	100
CP000568	908630	908630	G	A	SYN	One		392	N	N	Cthe0744	906953	909805	-	copper amine oxidase-like	9	15	24	24	100
CP000568	1006265	1006265	T	C	INTERGENIC	One	.					.	.			14	10	24	24	100
CP000568	1244897	1244897	G	A	NON	One		427	R	C	Cthe1041	1244775	1246175	-	UDP-N-acetylmuramoylalanine--D-glutamate ligase	9	15	24	24	100
CP000568	1636484	1636484	T	C	SYN	One		334	V	V	Cthe1344	1635311	1637485	-	(p)ppGpp synthetase I, SpoT/RelA	12	12	24	24	100
CP000568	1833211	1833211	C	T	STOP	One		363	W	*	Cthe1509	1833063	1834298	-	protein of unknown function DUF438	8	16	24	24	100
CP000568	2126854	2126854	A	G	NON	One		146	N	S	Cthe1800	2126418	2127929	+	Peptidoglycan-binding LysM	11	13	24	24	100
CP000568	2318080	2318080	G	A	NON	One		151	S	L	Cthe1942	2317632	2318531	-	hypothetical protein	12	12	24	24	100
CP000568	2362406	2362406	T	C	INTERGENIC	One	.					.	.			10	14	24	24	100
CP000568	2880313	2880313	C	T	SYN	One		390	D	D	Cthe2412	2879144	2881795	+	SMC protein-like	12	12	24	24	100
CP000568	3098565	3098565	G	A	INTERGENIC	One	.					.	.			14	10	24	24	100
CP000568	83602	83602	G	A	NON	One		2737	D	N	Cthe0056	75394	89286	+	Ig-like, group 2	18	7	25	25	100
CP000568	91153	91153	G	A	NON	One		59	A	T	Cthe0059	90979	92439	+	type 3a, cellulose-binding	15	10	25	25	100
CP000568	211988	211988	C	-	NON	Del		31	P	-	Cthe0175	211898	213034	+	polysaccharide deacetylase	17	8	25	25	100
CP000568	488241	488241	A	G	SYN	One		151	S	S	Cthe0392	487789	488838	+	inner-membrane translocator	14	11	25	25	100
CP000568	1157787	1157787	C	T	NON	One		682	E	K	Cthe0968	1157710	1159830	-	UvrD/REP helicase	13	12	25	25	100
CP000568	1455798	1455798	G	A	NON	One		2809	A	V	Cthe1221	1455455	1464223	-	glycosyltransferase 36	11	14	25	25	100
CP000568	1612350	1612350	T	C	SYN	One		244	P	P	Cthe1328	1611804	1613081	-	Stage II sporulation P	14	11	25	25	100
CP000568	1894413	1894413	T	C	NON	One		49	S	P	Cthe1566	1894269	1895612	+	Nitrogenase	12	13	25	25	100
CP000568	2087622	2087622	T	C	SYN	One		215	L	L	Cthe1765	2087115	2088266	-	hypothetical protein	11	14	25	25	100
CP000568	2421931	2421931	T	C	SYN	One		668	P	P	Cthe2038	2421511	2423934	-	cellulosome enzyme, dockerin type I	10	15	25	25	100
CP000568	2449572	2449572	G	A	SYN	One		53	N	N	Cthe2058	2449203	2449730	-	Appr-1-p processing	13	12	25	25	100
CP000568	2662099	2662099	T	C	NON	One		27	Q	R	Cthe2234	2661795	2662178	-	hypothetical protein	11	14	25	25	100
CP000568	2766062	2766062	T	-	NON	Del		195	N	-	Cthe2319	2765686	2766645	-	restriction endonuclease (HaeIII)	16	9	25	25	100
CP000568	2837314	2837314	A	G	NON	One		34	Q	R	Cthe2374	2837214	2838323	+	DNA replication and repair protein RecF	13	12	25	25	100
CP000568	2995953	2995953	G	A	NON	One		236	S	N	Cthe2527	2995247	2996125	+	porphobilinogen deaminase	9	16	25	25	100
CP000568	3304557	3304557	T	C	SYN	One		118	F	F	Cthe2803	3304204	3304983	+	binding-protein-dependent transport systems	14	11	25	25	100
CP000568	3470913	3470913	A	-	NON	Del		65	Q	-	Cthe2951	3470720	3471445	+	hypothetical protein	9	16	25	25	100
CP000568	1212630	1212630	T	C	INTERGENIC	One	.					.	.			21	5	26	26	100
CP000568	2160407	2160407	C	T	NON	One		768	G	D	Cthe1825	2159911	2162709	-	multi-sensor hybrid histidine kinase	16	10	26	26	100
CP000568	2216332	2216332	C	T	NON	One		599	P	L	Cthe1868	2214537	2217740	+	carbamoyl-phosphate synthase, large subunit	18	8	26	26	100
CP000568	2466234	2466234	G	A	INTERGENIC	One	.					.	.			14	12	26	26	100
CP000568	2843031	2843031	C	T	NON	One		165	P	L	Cthe2378	2842538	2843407	+	parB-like partition proteins	13	13	26	26	100
CP000568	2884831	2884831	A	G	NON	One		201	N	D	Cthe2415	2884231	2884932	+	cell wall hydrolase, SleB	12	14	26	26	100
CP000568	2888660	2888660	C	T	NON	One		20	A	T	Cthe2418	2887674	2888717	-	ATPase	15	11	26	26	100
CP000568	2900823	2900823	G	A	STOP	One		243	W	*	Cthe2427	2900096	2901052	+	protein of unknown function DUF1385	15	11	26	26	100
CP000568	3029890	3029890	A	G	NON	One		49	D	G	Cthe2559	3029745	3031118	+	DegT/DnrJ/EryC1/StrS aminotransferase	11	15	26	26	100
CP000568	3087899	3087899	G	A	SYN	One		265	L	L	Cthe2612	3087105	3090941	+	Fibronectin, type III	14	12	26	26	100
CP000568	3223060	3223060	C	T	NON	One		177	A	V	Cthe2730	3222531	3223733	+	translation elongation factor Tu	13	13	26	26	100
CP000568	3400528	3400528	A	G	NON	One		502	Q	R	Cthe2876	3399024	3401249	+	ATP-dependent DNA helicase PcrA	12	14	26	26	100
CP000568	3420674	3420674	A	G	NON	One		91	I	V	Cthe2894	3420404	3420934	+	conserved hypothetical protein	10	16	26	26	100
CP000568	3707682	3707682	A	G	INTERGENIC	One	.					.	.			16	10	26	26	100
CP000568	1390700	1390700	C	T	NON	One		19	A	T	Cthe1165	1389543	1390754	-	YbbR-like	9	18	27	27	100
CP000568	1715927	1715927	A	G	INTERGENIC	One	.					.	.			16	11	27	27	100
CP000568	1828037	1828037	T	C	NON	One		253	M	T	Cthe1504	1827280	1828092	+	Linocin_M18 bacteriocin protein	12	15	27	27	100
CP000568	2903728	2903728	G	-	NON	Del		27	M	-	Cthe2430	2903648	2904967	+	electron transport complex, RnfABCDGE type, C	14	13	27	27	100
CP000568	1506895	1506895	-	T	SYN	Ins		26	K	K	Cthe1243	1506510	1506971	-	GCN5-related N-acetyltransferase	15	13	28	28	100
CP000568	3086217	3086217	A	G	NON	Mul		1337	N	D	Cthe2611	3082209	3087095	+	Fibronectin, type III	16	12	28	28	100
CP000568	893075	893075	C	A	STOP	One		245	G	*	Cthe0735	893001	893807	-	cellulosome anchoring protein, cohesin region	19	10	29	29	100
CP000568	15883	15883	T	C	NON	One		1110	V	A	Cthe0009	12555	18077	+	YD repeat containing protein	17	13	30	30	100
CP000568	182829	182829	G	A	NON	One		175	A	T	Cthe0143	182307	183608	+	Phosphopyruvate hydratase	19	12	31	31	100
CP000568	2432297	2432297	T	C	INTERGENIC	One	.					.	.			16	15	31	31	100
CP000568	3325323	3325323	G	A	SYN	One		110	Y	Y	Cthe2815	3324051	3325652	-	lysyl-tRNA synthetase	17	14	31	31	100
CP000568	2975548	2975548	A	G	NON	One		834	I	M	Cthe2506	2973047	2976088	+	S-layer-like domain containing protein	13	19	32	32	100
CP000568	3031364	3031364	T	C	SYN	One		38	A	A	Cthe2560	3031251	3031799	+	dTDP-4-dehydrorhamnose 3,5-epimerase	13	19	32	32	100
CP000568	2225182	2225182	C	T	NON	One		69	A	T	Cthe1873	2224043	2225386	-	HMG-I and HMG-Y, DNA-binding	18	16	34	34	100
CP000568	2351019	2351019	C	T	SYN	One		83	E	E	Cthe1968	2348067	2351267	-	cell divisionFtsK/SpoIIIE	16	18	34	34	100
CP000568	3693510	3693510	C	T	INTERGENIC	One	.					.	.			14	20	34	34	100
CP000568	2947995	2947995	A	G	NON	One		361	S	G	Cthe2475	2946915	2948348	+	phage portal protein, SPP1 family	18	26	44	44	100
CP000568	2943643	2943643	G	A	NON	One		333	G	D	Cthe2471	2942646	2943926	+	conserved hypothetical protein	20	31	51	51	100
CP000568	2900141	2900141	-	G	SYN	Ins		16	G	G	Cthe2427	2900096	2901052	+	protein of unknown function DUF1385	16	15	31	32	97
CP000568	3240362	3240362	G	A	NON	One		42	R	Q	Cthe2746	3240238	3240708	+	protein of unknown function DUF402	11	11	22	23	96
CP000568	3742942	3742942	C	T	SYN	One		37	V	V	Cthe3167	3742832	3743953	+	glucose-1-phosphate adenylyltransferase, GlgD	10	12	22	23	96
CP000568	1561620	1561620	-	T	SYN	Ins		393	K	K	Cthe1285	1561401	1562798	-	metal dependent phosphohydrolase	13	10	23	24	96
CP000568	2461889	2461889	-	G	SYN	Ins		68	R	R	Cthe2069	2461687	2462463	+	hypothetical protein	14	9	23	24	96
CP000568	536206	536206	-	A	SYN	Ins		393	E	E	Cthe0426	535029	536699	+	putative PAS/PAC sensor protein	14	10	24	25	96
CP000568	2116605	2116605	T	-	NON	Del		123	I	-	Cthe1793	2116267	2116971	-	protein of unknown function DUF1624	14	10	24	25	96
CP000568	2434214	2434214	-	T	SYN	Ins		23	K	K	Cthe2044	2433680	2434282	-	hypothetical protein	8	16	24	25	96
CP000568	187474	187474	A	G	NON	One		37	T	A	Cthe0147	187366	187797	+	protein of unknown function DUF523	10	15	25	26	96
CP000568	8811	8811	G	A	SYN	One		1646	R	R	Cthe0004	3874	9702	+	YD repeat containing protein	12	14	26	27	96
CP000568	2476065	2476065	-	T	SYN	Ins		320	K	K	Cthe2083	2475825	2477024	-	DNA-directed DNA polymerase	15	11	26	27	96
CP000568	2833067	2833067	A	G	SYN	One		110	F	F	Cthe2369	2833016	2833396	-	ribonuclease P protein component	14	12	26	27	96
CP000568	3823553	3823553	A	-	INTERGENIC	Del	.					.	.			17	9	26	27	96
CP000568	190910	190910	G	A	SYN	One		348	P	P	Cthe0150	189867	191171	+	RNA modification enzyme, MiaB family	9	9	18	19	95
CP000568	668476	668476	A	G	SYN	One		76	S	S	Cthe0547	668249	669205	+	periplasmic solute binding protein	12	6	18	19	95
CP000568	1249068	1249068	T	C	NON	One		396	K	E	Cthe1046	1248913	1250253	-	extracellular solute-binding protein, family 1	6	12	18	19	95
CP000568	2138801	2138801	-	T	SYN	Ins		225	R	R	Cthe1806	2132940	2139473	-	cellulosome enzyme, dockerin type I	11	7	18	19	95
CP000568	3505725	3505725	C	-	INTERGENIC	Del	.					.	.			7	11	18	19	95
CP000568	189608	189608	C	T	NON	One		44	R	H	Cthe0149	189481	189738	-	HPrNtr domain containing protein	8	11	19	20	95
CP000568	421600	421600	-	A	INTERGENIC	Ins	.					.	.			11	8	19	20	95
CP000568	997314	997315	AT	-	INTERGENIC	Del	.					.	.			5	14	19	20	95
CP000568	1382074	1382074	C	T	INTERGENIC	One	.					.	.			9	10	19	20	95
CP000568	2039632	2039632	C	T	NON	One		19	E	K	Cthe1718	2039246	2039686	-	hypothetical protein	7	12	19	20	95
CP000568	7341	7341	C	T	SYN	One		1156	D	D	Cthe0004	3874	9702	+	YD repeat containing protein	10	10	20	21	95
CP000568	94420	94420	C	T	NON	One		479	G	R	Cthe0062	93965	95854	-	sulfatase	13	7	20	21	95
CP000568	225962	225962	C	T	SYN	One		137	C	C	Cthe0187	225552	226724	+	hypothetical protein	10	10	20	21	95
CP000568	435560	435560	T	A	INTERGENIC	One	.					.	.			10	10	20	21	95
CP000568	2068104	2068104	C	T	NON	One		372	E	K	Cthe1748	2067664	2069217	-	hypothetical protein	8	12	20	21	95
CP000568	3672053	3672053	C	T	NON	One		50	S	L	Cthe3111	3671905	3672384	+	conserved hypothetical protein	13	7	20	21	95
CP000568	1368045	1368045	G	A	INTERGENIC	One	.					.	.			6	15	21	22	95
CP000568	2340346	2340346	G	A	INTERGENIC	One	.					.	.			11	10	21	22	95
CP000568	225230	225230	C	T	SYN	One		282	S	S	Cthe0186	224385	225449	+	UDP-glucose 4-epimerase	7	9	16	17	94
CP000568	1799818	1799818	G	A	NON	One		569	S	N	Cthe1480	1798113	1800218	+	conserved hypothetical protein	8	8	16	17	94
CP000568	3391700	3391700	C	T	SYN	One		38	F	F	Cthe2871	3391587	3392099	+	TipAS antibiotic-recognition	12	4	16	17	94
CP000568	395168	395168	A	G	INTERGENIC	One	.					.	.			10	7	17	18	94
CP000568	610845	610845	A	G	SYN	One		437	K	K	Cthe0498	609535	611163	+	protein of unknown function DUF342	8	9	17	18	94
CP000568	1763704	1763704	G	A	NON	One		32	R	C	Cthe1445	1763552	1763797	-	uncharacterized Zn-finger protein	10	7	17	18	94
CP000568	3450332	3450332	C	-	NON	Del		243	N	-	Cthe2934	3449604	3450503	+	ABC transporter related	13	4	17	18	94
CP000568	442148	442148	C	T	NON	One		458	S	F	Cthe0351	440776	442167	+	PpiC-type peptidyl-prolyl cis-trans isomerase	6	7	13	14	93
CP000568	770018	770018	C	T	INTERGENIC	One	.					.	.			7	6	13	14	93
CP000568	2062134	2062134	-	T	SYN	Ins		191	K	K	Cthe1740	2061942	2062706	-	BRO-like	9	4	13	14	93
CP000568	702366	702366	A	G	NON	One		305	T	A	Cthe0574	701454	703565	+	serine/threonine protein kinase	6	8	14	15	93
CP000568	839332	839332	-	A	INTERGENIC	Ins	.					.	.			3	11	14	15	93
CP000568	1615176	1615176	C	T	SYN	One		1000	K	K	Cthe1329	1613874	1618175	-	putative CoA-substrate-specific enzyme activase	3	11	14	15	93
CP000568	3227762	3227762	A	-	NON	Del		238	E	-	Cthe2736	3227050	3228756	+	phosphoenolpyruvate-protein phosphotransferase	6	8	14	15	93
CP000568	2920701	2920701	G	A	SYN	One		54	P	P	Cthe2449	2920540	2921154	+	Phosphoglycerate mutase	16	10	26	28	93
CP000568	3102781	3102781	T	-	INTERGENIC	Del	.					.	.			14	13	27	29	93
CP000568	659458	659458	T	C	NON	One		291	I	T	Cthe0540	658587	661193	+	protein of unknown function DUF214	2	9	11	12	92

CP000568	1273130	1273130	-	T	INTERGENIC	Ins	.					.	.			4	7	11	12	92	
CP000568	1495972	1495972	-	T	SYN	Ins		1021	K	K	Cthe1235	1478377	1499034	-	Cellulose 1,4-beta-cellobiosidase	7	4	11	12	92	
CP000568	1946531	1946531	T	C	NON	One		531	T	A	Cthe1616	1945599	1948121	-	phage minor structural protein	3	8	11	12	92	
CP000568	780355	780355	C	T	SYN	One		20	P	P	Cthe0633	780296	780781	+	hypothetical protein	6	6	12	13	92	
CP000568	870718	870718	T	C	NON	One		560	Y	H	Cthe0714	869041	871125	+	hydroxymethylbutenyl pyrophosphate reductase	6	6	12	13	92	
CP000568	1913840	1913840	G	A	NON	One		155	R	C	Cthe1584	1913523	1914302	-	two component transcriptional regulator, AraC	8	4	12	13	92	
CP000568	2702003	2702003	T	-	NON	Del		21	N	-	Cthe2272	2701941	2702855	+	conserved hypothetical protein	16	6	22	24	92	
CP000568	2782553	2782553	G	A	NON	One		264	L	F	Cthe2334	2781498	2783342	-	polysaccharide biosynthesis protein CapD	13	9	22	24	92	
CP000568	321569	321569	G	A	NON	One		671	E	K	Cthe0260	319559	321682	+	peptidase S1 and S6, chymotrypsin/Hap	15	9	24	26	92	
CP000568	1205955	1205955	G	A	NON	One		187	E	K	Cthe1008	1205397	1206518	+	aminodeoxychorismate lyase	12	8	20	22	91	
CP000568	196095	196095	-	A	SYN	Ins		335	K	K	Cthe0156	195093	196934	+	Radical SAM	8	13	21	23	91	
CP000568	1627581	1627581	G	A	INTERGENIC	One	.					.	.			13	8	21	23	91	
CP000568	2017515	2017515	-	C	INTERGENIC	Ins	.					.	.			13	8	21	23	91	
CP000568	3665258	3665258	G	A	INTERGENIC	One	.					.	.			12	9	21	23	91	
CP000568	2706499	2706499	-	T	INTERGENIC	Ins	.					.	.			12	18	30	33	91	
CP000568	2900249	2900249	-	A	SYN	Ins		52	K	K	Cthe2427	2900096	2901052	+	protein of unknown function DUF1385	16	14	30	33	91	
CP000568	1403036	1403036	A	G	NON	One		170	T	A	Cthe1175	1402529	1403338	+	hypothetical protein	3	6	9	10	90	
CP000568	533461	533461	C	T	NON	One		704	P	L	Cthe0423	531351	533972	+	iron-containing alcohol dehydrogenase	4	4	8	9	89	
CP000568	3379326	3379326	T	C	SYN	One		379	S	S	Cthe2860	3378190	3379900	+	hypothetical protein precursor	4	4	8	9	89	
CP000568	244074	244074	-	A	SYN	Ins		15	Q	Q	Cthe0201	244031	244753	+	glutamate synthase, alpha subunit-like	7	10	17	19	89	
CP000568	1194190	1194190	C	T	NON	One		9	R	K	Cthe0996	1189887	1194215	-	DNA polymerase III, alpha subunit	8	9	17	19	89	
CP000568	3749144	3749144	G	A	SYN	One		777	V	V	Cthe3171	3746814	3749261	+	S-layer domain-like	5	12	17	19	89	
CP000568	97575	97575	T	C	INTERGENIC	One	.					.	.			8	7	15	17	88	
CP000568	1330160	1330160	C	T	NON	Mul		254	G	R	Cthe1115	1329264	1330919	-	Tn7-like transposition protein C	4	11	15	17	88	
CP000568	3159478	3159478	-	T	SYN	Ins		57	K	K	Cthe2678	3159114	3159647	-	hypothetical protein	8	7	15	17	88	
CP000568	1248961	1248961	-	T	SYN	Ins		431	K	K	Cthe1046	1248913	1250253	-	extracellular solute-binding protein, family 1	9	13	22	25	88	
CP000568	2636341	2636341	-	C	SYN	Ins		187	G	G	Cthe2207	2635725	2636900	-	isoaspartyl dipeptidase	10	13	23	26	88	
CP000568	2791484	2791484	-	A	SYN	Ins		285	F	F	Cthe2341	2791172	2792338	-	glycosyl transferase, family 2	13	10	23	26	88	
CP000568	1103031	1103031	-	T	SYN	Ins		741	K	K	Cthe0918	1101624	1105253	-	cellulosome enzyme, dockerin type I	4	3	7	8	87	
CP000568	949261	949261	C	T	NON	One		42	H	Y	Cthe0785	949138	949440	+	hypothetical protein	7	7	14	16	87	
CP000568	3432750	3432750	A	-	INTERGENIC	Del	.					.	.			14	6	20	23	87	
CP000568	863219	863219	-	A	SYN	Ins		66	K	K	Cthe0705	863024	863260	+	hypothetical protein	10	11	21	24	87	
CP000568	2725547	2725547	-	A	SYN	Ins		63	F	F	Cthe2291	2725445	2725735	-	hypothetical protein	8	13	21	24	87	
CP000568	2320021	2320021	G	A	NON	One		43	S	F	Cthe1944	2319600	2320148	-	protein of unknown function DUF458	12	14	26	30	87	
CP000568	2382985	2382985	-	T	INTERGENIC	Ins	.					.	.			14	14	28	32	87	
CP000568	3379300	3379300	T	G	NON	One		371	C	G	Cthe2860	3378190	3379900	+	hypothetical protein precursor	3	3	6	7	86	
CP000568	2238760	2238760	-	A	SYN	Ins		107	E	E	Cthe1885	2238441	2239013	+	phage integrase-like SAM-like	7	11	18	21	86	
CP000568	2873484	2873484	-	A	SYN	Ins		245	K	K	Cthe2407	2872752	2874374	+	AAA ATPase, central region	11	7	18	21	86	
CP000568	2872117	2872117	-	A	INTERGENIC	Ins	.					.	.			14	5	19	22	86	
CP000568	138885	138885	C	T	SYN	One		328	P	P	Cthe0106	137902	139143	+	GTP cyclohydrolase II	8	16	24	28	86	
CP000568	269914	269914	-	T	INTERGENIC	Ins	.					.	.			6	11	17	20	85	
CP000568	1020112	1020112	-	T	SYN	Ins		3	N	N	Cthe0841	1018890	1020119	-	mutants block sporulation after engulfment	11	6	17	20	85	
CP000568	2892757	2892757	-	T	SYN	Ins		278	K	K	Cthe2422	2892517	2893590	-	hypothetical protein	5	11	16	19	84	
CP000568	530864	530864	G	A	NON	One		60	G	S	Cthe0422	530687	531370	+	CoA-binding	5	5	10	12	83	
CP000568	3758418	3758418	-	A	INTERGENIC	Ins	.					.	.			7	11	18	22	82	
CP000568	82364	82364	-	A	SYN	Ins		2324	E	E	Cthe0056	75394	89286	+	Ig-like, group 2	7	6	13	16	81	
CP000568	3507977	3507977	G	A	SYN	One		590	K	K	Cthe2989	3506208	3509162	+	glycosyltransferase 36	5	12	17	21	81	
CP000568	2509967	2509967	G	A	INTERGENIC	One	.					.	.			9	13	22	27	81	
CP000568	2152897	2152897	-	T	SYN	Ins		10	N	N	Cthe1817	2152557	2152925	-	urease, beta subunit	7	7	14	18	78	
CP000568	3613913	3613913	-	A	SYN	Ins		2	K	K	Cthe3073	3613910	3614581	+	HAD-superfamily hydrolase, subfamily IA, variant	11	3	14	18	78	
CP000568	3630169	3630172	ATCT	GTCA	NON	One		1647	P	-	Cthe3078	3625229	3632170	+	cellulosome anchoring protein, cohesin region	2	1	3	4	75	
CP000568	2245709	2245712	TACC	CTCT	NON	One		395	S	-	Cthe1890	2244759	2246891	-	cellulosome enzyme, dockerin type I	2	4	6	8	75	
CP000568	2263439	2263439	T	-	NON	Del		86	K	-	Cthe1905	2262463	2263695	-	glycosyl transferase, family 28	6	6	12	16	75	
CP000568	1329105	1329118	TTGAAGGCAAGTCC	-	NON	Del		52	N	-	Cthe1114	1327372	1329258	-	Tn7-like transposition protein D	3	5	8	11	73	
CP000568	2245662	2245662	T	C	SYN	One		410	L	L	Cthe1890	2244759	2246891	-	cellulosome enzyme, dockerin type I	4	1	5	7	71	
CP000568	2245669	2245671	TCT	CCG	NON	One		408	E	-	Cthe1890	2244759	2246891	-	cellulosome enzyme, dockerin type I	4	1	5	7	71	
CP000568	1132986	1132986	T	C	SYN	One		49	G	G	Cthe0945	1132140	1133132	-	metallophosphoesterase	7	5	12	17	71	
CP000568	3065054	3065055	GA	TT	INTERGENIC	Mul	.					.	.			9	7	16	30	53	
CP000568	3065049	3065051	GGT	TGC	NON	One		155	W	-	Cthe2592	3064586	3065053	+	hypothetical protein	9	7	16	31	52	
CP000568	542230	542230	T	C	INTERGENIC	One	.					.	.			2	4	6	12	50	
CP000568	1476664	1476664	A	G	INTERGENIC	One	.					.	.			6	3	9	18	50	
CP000568	3643837	3643837	C	T	NON	One		69	A	T	Cthe3089	3643649	3644041	-	UspA	4	9	13	26	50	
CP000568	2125777	2125777	C	T	NON	One		490	A	V	Cthe1799	2124309	2126225	+	ABC transporter related	4	7	11	23	48	
CP000568	290766	290766	-	TATA	INTERGENIC	Ins	.					.	.			2	3	5	11	45	
CP000568	1083495	1083495	T	C	NON	Mul		295	E	G	Cthe0906	1083026	1084378	-	Radical SAM	3	3	6	14	43	
CP000568	2495816	2495818	CGA	-	NON	Del		158	G	-	Cthe2099	2495655	2496287	-	nucleoside recognition	1	4	5	12	42	
CP000568	2125349	2125349	C	T	SYN	One		347	N	N	Cthe1799	2124309	2126225	+	ABC transporter related	5	3	8	19	42	
CP000568	2874591	2874591	A	G	SYN	One		41	A	A	Cthe2408	2874469	2875197	+	phage shock protein A, PspA	5	4	9	22	41	
CP000568	1090364	1090364	C	T	NON	One		719	E	K	Cthe0912	1089285	1092518	-	glycoside hydrolase, family 10	4	2	6	15	40	
CP000568	3342638	3342638	G	A	NON	Mul		565	E	K	Cthe2825	3340946	3342793	+	conserved hypothetical protein precursor	8	7	15	39	38	3342638
CP000568	1851137	1851137	A	G	NON	One		306	Y	H	Cthe1525	1850184	1852052	-	protein of unknown function DUF214	4	5	9	24	37	
CP000568	3342959	3342959	A	G	SYN	One		48	Q	Q	Cthe2826	3342816	3344054	+	conserved hypothetical protein	5	6	11	30	37	3342959
CP000568	3342625	3342627	TGA	-	NON	Del		560	N	-	Cthe2825	3340946	3342793	+	conserved hypothetical protein precursor	8	7	15	41	37	3342625
CP000568	1855366	1855366	A	G	NON	One		209	S	P	Cthe1529	1854740	1855992	-	transposase	1	3	4	11	36	
CP000568	509539	509539	-	T	SYN	Ins		6	K	K	Cthe0407	508249	509556	-	Radical SAM	1	5	6	17	35	
CP000568	3342604	3342621	GATAGGTGATAAAACAGG	AAAGGGTGTTAGATCAGGGGA	NON	One		553	L	-	Cthe2825	3340946	3342793	+	conserved hypothetical protein precursor	7	6	13	37	35	3342604
CP000568	3342592	3342592	G	A	SYN	One		549	S	S	Cthe2825	3340946	3342793	+	conserved hypothetical protein precursor	8	5	13	38	34	3342592
CP000568	2371533	2371535	GCC	-	NON	Del		113	A	-	Cthe1990	2371523	2371870	-	hypothetical protein	1	2	3	9	33	2371533
CP000568	790216	790216	G	A	SYN	One		233	S	S	Cthe0639	789712	790914	-	glycoside hydrolase, family 13-like	1	3	4	12	33	
CP000568	3327878	3327881	ATTT	TTAA	INTERGENIC	One	.					.	.			4	3	7	21	33	
CP000568	3336374	3336374	A	G	NON	One		66	T	A	Cthe2821	3336179	3337273	+	response regulator receiver modulated CheB	4	2	6	19	32	
CP000568	980941	980941	T	C	INTERGENIC	One	.					.	.			4	5	9	28	32	
CP000568	619235	619242	CTTTTTTC	TCTTTTTTT	NON	One		214	S	-	Cthe0507	619228	619875	-	Exonuclease, RNase T and DNA polymerase III	2	2	4	13	31	
CP000568	1627485	1627529	TTTGTCGAATAATTATTTGTCGAATAATTATTTGTCGAATAATTA	-	INTERGENIC	Del	.					.	.			3	2	5	16	31	
CP000568	2282632	2282632	G	A	NON	One		113	A	V	Cthe1915	2281578	2282969	-	periplasmic sensor signal transduction histidine	1	9	10	32	31	
CP000568	2100481	2100481	C	T	NON	One		173	A	T	Cthe1778	2100038	2100997	-	copper amine oxidase-like	4	3	7	23	30	
CP000568	3135064	3135064	C	T	NON	One		2	S	L	Cthe2653	3135060	3135809	+	conserved hypothetical protein	2	5	7	23	30	
CP000568	1802051	1802051	C	T	NON	One		615	A	V	Cthe1481	1800208	1802520	+	membrane protein-like	1	4	5	17	29	
CP000568	1787290	1787290	G	A	NON	One		61	D	N	Cthe1470	1787110	1787727	+	RNA polymerase, sigma-24 subunit, ECF subfamily	5	2	7	25	28	
CP000568	1068550	1068550	-	T	NON	Ins		267	G	D	Cthe0892	1068243	1069349	-	protein of unknown function DUF34	1	2	3	11	27	
CP000568	1190435	1190435	-	TT	NON	Ins		1261	C	-	Cthe0996	1189887	1194215	-	DNA polymerase III, alpha subunit	2	2	4	15	27	
CP000568	395731	395731	G	A	NON	One		109	E	K	Cthe0313	395407	396438	+	conserved hypothetical protein	3	3	6	22	27	
CP000568	3680727	3680727	T	C	SYN	One		41	T	T	Cthe3120	3680605	3684141	+	pyruvate flavodoxin/ferredoxin	2	4	6	22	27	
CP000568	129793	129793	C	G	NON	One		213	A	G	Cthe0097	129156	130031	+	peptidase M23B	4	1	5	19	26	
CP000568	3342972	3342977	GGTACA	ATTACG	NON	One		53	G	-	Cthe2826	3342816	3344054	+	conserved hypothetical protein	5	2	7	27	26	
CP000568	3342981	3342986	AAATTA	GCTGAC	NON	Mul		56	K	-	Cthe2826	3342816	3344054	+	conserved hypothetical protein	5	2	7	27	26	
CP000568	2788094	2788094	T	C	NON	One		193	Y	C	Cthe2338	2787439	2788671	-	glycosyl transferase, group 1	4	4	8	31	26	
CP000568	3342563	3342563	A	G	NON	One		540	T	A	Cthe2825	3340946	3342793	+	conserved hypothetical protein precursor	5	4	9	34	26	3342563
CP000568	3342559	3342559	T	G	SYN	One		538	L	L	Cthe2825	3340946	3342793	+	conserved hypothetical protein precursor	5	4	9	34	26	3342559
CP000568	1897646	1897646	G	A	NON	One		176	A	T	Cthe1569	1897121	1898437	+	O-acetylhomoserine	2	3	5	20	25	
CP000568	3426133	3426133	C	T	NON	One		140	R	W	Cthe2900	3425716	3426474	+	RNA polymerase, sigma 28 subunit	1	4	5	20	25	
CP000568	2397503	2397509	ATTCATG	GTCCATA	NON	One		64	N	-	Cthe2018	2396459	2397694	-	conserved hypothetical protein	6	3	9	36	25	2397503
CP000568	3342994	3342994	T	G	NON	Mul		60	M	R	Cthe2826	3342816	3344054	+	conserved hypothetical protein	5	2	7	29	24	
CP000568	3342553	3342553	G	A	SYN	One		536	A	A	Cthe2825	3340946	3342793	+	conserved hypothetical protein precursor	4	4	8	34	24	3342553
CP000568	2397524	2397526	TTG	GTC	NON	One		57	Q	-	Cthe2018	2396459	2397694	-	conserved hypothetical protein	6	3	9	38	24	2397524
CP000568	2397516	2397516	A	C	NON	Mul		60	M	R	Cthe2018	2396459	2397694	-	conserved hypothetical protein	6	3	9	38	24	2397516
CP000568	2397410	2397410	T	C	SYN	One		95	E	E	Cthe2018	2396459	2397694	-	conserved hypothetical protein	5	3	8	36	22	2397410
CP000568	1560959	1560959	T	C	NON	One		101	H	R	Cthe1284	1559794	1561260	-	glycogen/starch synthases, ADP-glucose type	2	1	3	14	21	
CP000568	2334233	2334233	-	T	INTERGENIC	Ins	.					.	.			2	1	3	14	21	
CP000568	797819	797819	C	T	NON	One		636	A	V	Cthe0645	795913	798327	+	homocysteine S-methyltransferase	2	2	4	19	21	
CP000568	2782783	2782783	-	T	NON	Ins		187	I	N	Cthe2334	2781498	2783342	-	polysaccharide biosynthesis protein CapD	3	1	4	19	21	
CP000568	346813	346813	-	G	NON	Ins		148	I	S	Cthe0278	346371	347066	+	pseudouridine synthase	1	2	3	15	20	
CP000568	1083477	1083480	AGCC	G	NON	One		301	L	P	Cthe0906	1083026	1084378	-	Radical SAM	1	2	3	15	20	
CP000568	3327479	3327482	TTAA	AAAT	INTERGENIC	One	.					.	.			3	2	5	25	20	3327479
CP000568	3247112	3247112	-	A	INTERGENIC	Ins	.					.	.			1	4	5	25	20	
CP000568	3693449	3693449	G	A	INTERGENIC	One	.					.	.			2	5	7	35	20	
CP000568	219891	219891	-	T	INTERGENIC	Ins	.					.	.			1	2	3	16	19	
CP000568	3186158	3186158	-	T	INTERGENIC	Ins	.					.	.			1	2	3	16	19	
CP000568	492822	492823	CG	-	NON	Del		289	F	-	Cthe0396	491956	493680	+	ABC transporter related	3	1	4	21	19	
CP000568	2397551	2397554	TTGG	CTGA	NON	Mul		48	Q	-	Cthe2018	2396459	2397694	-	conserved hypothetical protein	5	2	7	36	19	2397551
CP000568	2454212	2454212	T	C	NON	One		428	V	A	Cthe2062	2452930	2454465	+	GerA spore germination protein	1	2	3	17	18	
CP000568	3662398	3662398	G	A	NON	One		126	G	D	Cthe3105	3662022	3662690	+	exsB protein	1	2	3	17	18	

CP000568	371791	371791	A	G	NON	One	322	Y	C	Cthe0296	370827	372302	+	hypothetical protein	1	2	3	18	17	
CP000568	2412393	2412393	G	A	NON	Mul	984	P	L	Cthe2032	2409329	2415343	-	hypothetical protein	1	2	3	18	17	
CP000568	445662	445662	C	T	NON	One	110	T	I	Cthe0355	445334	446512	+	conserved hypothetical protein	1	3	4	24	17	
CP000568	2600988	2600988	G	AGT	NON	One	515	A	-	Cthe2182	2599446	2602235	+	Ig-like, group 2	1	2	3	19	16	
CP000568	3144465	3144467	TAA	GGAACCG	NON	One	361	L	-	Cthe2664	3143384	3145105	+	ABC-1	1	2	3	19	16	
CP000568	614571	614571	C	-	INTERGENIC	Del	.				.	.			1	2	3	21	14	
CP000568	2127018	2127020	GCA	C	NON	One	201	A	P	Cthe1800	2126418	2127929	+	Peptidoglycan-binding LysM	1	2	3	22	14	
CP000568	3310989	3310989	T	C	SYN	One	99	R	R	Cthe2809	3310693	3314658	+	glycoside hydrolase, family 16	1	2	3	23	13	
CP000568	124029	124029	C	T	SYN	One	11	S	S	Cthe0091	123997	126126	+	Peptidoglycan glycosyltransferase	1	2	3	24	12	
CP000568	1958491	1958491	G	A	STOP	One	435	R	*	Cthe1631	1958009	1959793	-	conserved hypothetical protein	2	1	3	25	12	
CP000568	3007869	3007869	A	G	NON	One	482	N	D	Cthe2537	3006426	3008225	+	sulfate adenylyltransferase, large subunit	1	2	3	24	12	
CP000568	3016361	3016362	TG	-	NON	Del	110	F	-	Cthe2548	3016032	3017543	+	Alpha-N-arabinofuranosidase	1	2	3	25	12	
CP000568	1260236	1260236	-	T	SYN	Ins	12	Y	Y	Cthe1055	1260203	1261411	+	protein of unknown function DUF58	1	2	3	27	11	
CP000568	2397277	2397277	A	G	SYN	One	140	L	L	Cthe2018	2396459	2397694	-	conserved hypothetical protein	1	3	4	40	10	2397277

Table S3. Genetic differences identified in EA by comparative genome sequencing (CGS) via microarray analysis

SNPs in coding sequences

GENOME_POSITION	REF	SNP	CONFIDENCE	PROBABILITY	TYPE	NMER_SCORE	UNIQ	HIGHEST_RANK	FORWARD_PBC_SCORE	REVERSE_PBC_SCORE	FORWARD_INTENSITY	REVERSE_INTENSITY	AA_CHANGE	ORIG_AA	SNP_AA	GENE_POSITION	LOCUS_TAG	START	END	STRAND	PRODUCT	NOTE
8023	T	C	MEDIUM	0.975	coding	0	0	1.56	0.942	0.794	3370	2608	NON	F	L	4150	Cthe_0004	3874	9702	1	YD repeat containing protein	TIGRFAM: YD repeat protein PFAM: YD repeat KEGG: mba:Mbar_A3399 hypothetical protein
8027	T	C	MEDIUM	0.915	coding	1.54	0	1.56	0.946	0.953	435	418	NON	V	A	4154	Cthe_0004	3874	9702	1	YD repeat containing protein	TIGRFAM: YD repeat protein PFAM: YD repeat KEGG: mba:Mbar_A3399 hypothetical protein
8802	T	C	MEDIUM	1	coding	3.29	1	0.24	0.77	0.964	570	3041	SYN	D	D	4929	Cthe_0004	3874	9702	1	YD repeat containing protein	TIGRFAM: YD repeat protein PFAM: YD repeat KEGG: mba:Mbar_A3399 hypothetical protein
8811	G	A	MEDIUM	0.985	coding	0	0	0.24	0.735	0.915	1266	1997	SYN	R	R	4938	Cthe_0004	3874	9702	1	YD repeat containing protein	TIGRFAM: YD repeat protein PFAM: YD repeat KEGG: mba:Mbar_A3399 hypothetical protein
39120	A	G	HIGH	0.98	coding	0	1	-0.54	0.977	0.985	6953	7748	SYN	I	I	966	Cthe_0032	38313	40085	-1	glycoside hydrolase, family 26	PFAM: glycoside hydrolase, family 26; cellulosome enzyme, dockerin type I KEGG: sde:Sde_3691 mannan endo-1,4-beta-mannosidase
56259	G	A	HIGH	0.97	coding	0	1	10000.78	0.798	0.807	2571	1677	SYN	P	P	1887	Cthe_0043	54373	56601	1	glycoside hydrolase, family 9	PFAM: glycoside hydrolase, family 9; type 3a, cellulose-binding; cellulosome enzyme, dockerin type I KEGG: bld:BLi01880 hypothetical protein
83602	G	A	HIGH	0.995	coding	0	1	10002.42	0.446	0.678	1032	1981	NON	D	N	8209	Cthe_0056	75394	89286	1	Ig-like, group 2	PFAM: S-layer-like region; Ig-like, group 2 KEGG: pfl:PFL_0133 calcium-binding outer membrane-like protein
91153	G	A	HIGH	0.965	coding	0	1	0.56	0.906	0.971	3378	16934	NON	A	T	175	Cthe_0059	90979	92439	1	type 3a, cellulose-binding	PFAM: type 3a, cellulose-binding KEGG: bld:BLi01500 Ykrl
91329	C	T	HIGH	0.99	coding	0	1	0.41	0.821	0.924	6473	2994	SYN	D	D	351	Cthe_0059	90979	92439	1	type 3a, cellulose-binding	PFAM: type 3a, cellulose-binding KEGG: bld:BLi01500 Ykrl
94420	C	T	HIGH	0.945	coding	0	1	-0.52	0.953	0.92	2581	2325	NON	G	R	1435	Cthe_0062	93965	95854	-1	sulfatase	PFAM: sulfatase KEGG: ctc:CTC01943 alkaline phosphatase superfamily enzyme
104530	G	A	HIGH	0.995	coding	0	1	10001.39	0.738	0.986	1655	4976	NON	A	V	1139	Cthe_0071	102852	105668	-1	Cellulose 1,4-beta-cellobiosidase	PFAM: glycoside hydrolase, family 48; type 3a, cellulose-binding KEGG: hch:HCH_02465 RTX toxins and related Ca2+-binding protein
136133	A	G	HIGH	0.99	coding	0	1	0.28	0.892	0.988	1783	6178	NON	D	G	104	Cthe_0104	136030	137127	1	5-amino-6-(5-phosphoribosylamino)uracil reductase / diaminohydroxyphosphoribosylaminopyrimidine deaminase	KEGG: ctc:CTC00671 5-amino-6-(5-phosphoribosylamino)uracil reductase / diaminohydroxyphosphoribosylaminopyrimidine deaminase TIGRFAM: riboflavin biosynthesis protein RibD; Riboflavin-specific deaminase-like PFAM: CMP/dCMP deaminase, zinc-binding; bifunctio
137939	T	C	HIGH	0.98	coding	0	1	-0.9	0.957	0.566	11357	2528	NON	I	T	38	Cthe_0106	137902	139143	1	GTP cyclohydrolase II	PFAM: 3,4-dihydroxy-2-butanone 4-phosphate synthase; GTP cyclohydrolase II KEGG: mta:Moth_0917 3,4-dihydroxy-2-butanone 4-phosphate synthase
182829	G	A	HIGH	0.995	coding	0	1	1.5	0.869	0.971	1149	3569	NON	A	T	523	Cthe_0143	182307	183608	1	enolase	PFAM: enolase KEGG: tte:TTE1759 enolase
187474	A	G	HIGH	0.99	coding	0	1	0.72	0.815	0.985	1801	12636	NON	T	A	109	Cthe_0147	187366	187797	1	protein of unknown function DUF523	PFAM: protein of unknown function DUF523 KEGG: cac:CAC2339 uncharacterized protein, YBBK B.subtilis ortholog
189608	C	T	HIGH	0.985	coding	0	1	10002.3	0.917	0.73	1791	631	NON	R	H	131	Cthe_0149	189481	189738	-1	HPrNtr domain containing protein	KEGG: tte:TTE0115 hypothetical protein
190910	G	A	HIGH	0.96	coding	0	1	-0.99	0.991	0.981	16517	3707	SYN	P	P	1044	Cthe_0150	189867	191171	1	RNA modification enzyme, MiaB family	KEGG: tte:TTE0962 2-methylthioadenine synthetase TIGRFAM: RNA modification enzyme, MiaB family; MiaB-like tRNA modifying enzyme PFAM: deoxyribonuclease/rho motif-related TRAM; protein of unknown function UPF0004; Radical SAM SMART: Elongator protein 3/MiaB
225962	C	T	HIGH	0.95	coding	0	1	-0.61	0.628	0.913	2814	8890	SYN	C	C	411	Cthe_0187	225552	226724	1	hypothetical protein	KEGG: mta:Moth_2506 spore germination protein-like
253589	C	T	HIGH	0.99	coding	0	1	1.06	0.799	0.883	3233	6133	NON	Q	*	982	Cthe_0208	252608	254899	1	single-stranded-DNA-specific exonuclease RecJ	TIGRFAM: single-stranded-DNA-specific exonuclease RecJ PFAM: phosphoesterase, RecJ-like; phosphoesterase, DHHA1 KEGG: sat:SYN_02529 single-stranded DNA-specific exonuclease
321569	G	A	HIGH	1	coding	0	1	-0.68	0.614	0.522	1459	2193	NON	E	K	2011	Cthe_0260	319559	321682	1	peptidase S1 and S6, chymotrypsin/Hap	PFAM: peptidase S1 and S6, chymotrypsin/Hap KEGG: cme:CMT075C DegP protease precursor Pfam: Trypsin Peptidase_S29 PDZ PROSITE: PDZ
340690	C	T	HIGH	0.96	coding	0	1	0.44	0.984	0.991	7250	5581	SYN	L	L	21	Cthe_0274	339019	340710	-1	glycoside hydrolase, family 9	PFAM: glycoside hydrolase, family 9; cellulosome enzyme, dockerin type I KEGG: cac:CAC0917 and cellulose-binding endoglucanase family 9; CelL ortholog; dockerin domain
351245	G	A	HIGH	0.97	coding	0	1	-1	0.814	0.811	2599	2053	SYN	P	P	429	Cthe_0284	350817	352175	1	protein of unknown function DUF815	PFAM: protein of unknown function DUF815 SMART: ATPase KEGG: mta:Moth_2317 protein of unknown function DUF815
355031	G	A	HIGH	0.965	coding	0	1	0.52	0.908	0.815	5299	2261	SYN	A	A	1206	Cthe_0286	353826	355100	1	response regulator receiver sensor signal transduction histidine kinase	PFAM: response regulator receiver; ATP-binding region, ATPase-like; histidine kinase A-like KEGG: cac:CAC0903 sensory transduction histidine kinase
356311	C	T	HIGH	0.99	coding	0	1	10002.02	0.618	0.979	3259	5684	SYN	V	V	1167	Cthe_0287	355145	358627	1	GAF sensor hybrid histidine kinase	PFAM: response regulator receiver; GAF; ATP-binding region, ATPase-like; histidine kinase A-like KEGG: cvi:CV1579 probable two-component hybrid sensor and regulator
373018	A	G	HIGH	0.975	coding	0	1	-0.31	0.801	0.988	5765	6538	SYN	S	S	198	Cthe_0298	372821	374623	1	methyl-accepting chemotaxis sensory transducer	PFAM: histidine kinase, HAMP region; chemotaxis sensory transducer KEGG: gsu:GSU1140 methyl-accepting chemotaxis protein
385361	A	G	HIGH	0.99	coding	0	1	-0.95	0.923	0.988	4285	2778	SYN	G	G	225	Cthe_0307	385137	385514	1	conserved hypothetical protein	KEGG: fra:Francci3_3054 hypothetical protein
401007	C	T	HIGH	0.985	coding	0	1	0.83	0.986	0.986	7728	24154	NON	G	S	370	Cthe_0317	400705	401376	-1	hypothetical protein	
410756	A	G	HIGH	0.995	coding	0	1	1.5	0.431	0.672	619	2747	SYN	T	T	159	Cthe_0325	410598	412535	1	NH(3)-dependent NAD(+) synthetase	TIGRFAM: NAD+ synthetase PFAM: NAD+ synthase KEGG: bfs:BF3022 putative glutamine-dependent NAD+ synthetase
432634	A	G	HIGH	0.995	coding	0	1	-0.94	0.989	0.968	10070	6014	SYN	A	A	411	Cthe_0343	432224	432826	1	flavin reductase-like, FMN-binding	PFAM: flavin reductase-like, FMN-binding KEGG: fnu:FN1468 flavoredoxin
462568	G	A	HIGH	0.96	coding	0	1	0.1	0.888	0.965	3204	4742	SYN	R	R	744	Cthe_0367	461825	462694	1	NADP oxidoreductase, coenzyme F420-dependent	PFAM: NADP oxidoreductase, coenzyme F420-dependent; NAD-dependent glycerol-3-phosphate dehydrogenase-like KEGG: chy:CHY_2378 hypothetical protein
488241	A	G	HIGH	0.98	coding	0	1	-0.2	0.801	0.983	4030	5930	SYN	S	S	453	Cthe_0392	487789	488838	1	inner-membrane translocator	PFAM: inner-membrane translocator KEGG: sme:SMb20502 putative sugar ABC transporter permease protein
493377	G	A	HIGH	1	coding	0	1	10001.9	0.667	0.976	513	5599	SYN	Q	Q	1422	Cthe_0396	491956	493680	1	ABC transporter related protein	PFAM: ABC transporter, transmembrane region; ABC transporter related SMART: ATPase KEGG: cac:CAC2393 uncharacterized ABC transporter, ATPase component
507789	G	A	HIGH	0.97	coding	0	1	-0.5	0.641	0.501	1008	2375	NON	A	V	431	Cthe_0406	507482	508219	-1	conserved hypothetical protein	KEGG: bth:BT4483 hypothetical protein
512193	G	A	HIGH	0.98	coding	0	1	10001.28	0.817	0.68	5852	1976	NON	D	N	2347	Cthe_0408	509847	512435	1	diguanylate cyclase/phosphodiesterase with PAS/PAC sensor(s)	KEGG: dar:Daro_2048 PAS:GGDEF TIGRFAM: PAS sensor protein; diguanylate cyclase PFAM: conserved hypothetical protein; EAL; extracellular solute-binding protein, family 3; PAS fold-4; PAS fold SMART: PAS
519958	C	T	HIGH	0.98	coding	0	1	10001.39	0.897	0.979	4508	11743	SYN	D	D	492	Cthe_0413	519467	523141	1	glycoside hydrolase, family 9	PFAM: glycoside hydrolase, family 9; type 3a, cellulose-binding; cellulosome enzyme, dockerin type I; Carbohydrate-binding, CenC-like; glycoside hydrolase, family 9-like Ig-like KEGG: cac:CAC0561 cellulase CelE ortholog; dockerin domain
531270	T	C	HIGH	0.98	coding	0	1	0.63	0.98	0.987	36228	18904	NON	M	T	584	Cthe_0422	530687	531370	1	CoA-binding protein	PFAM: CoA-binding; Rex DNA-binding-like KEGG: tde:TDE1912 DNA-binding protein p25
533461	C	T	HIGH	0.945	coding	0	1	-0.64	0.986	0.991	5278	10221	NON	P	L	2111	Cthe_0423	531351	533972	1	iron-containing alcohol dehydrogenase	PFAM: iron-containing alcohol dehydrogenase; aldehyde dehydrogenase KEGG: cpe:CPE2531 alcohol dehydrogenase / acetaldehyde dehydrogenase
533551	A	G	HIGH	0.99	coding	0	1	-0.15	0.97	0.968	6670	7170	NON	H	R	2201	Cthe_0423	531351	533972	1	iron-containing alcohol dehydrogenase	PFAM: iron-containing alcohol dehydrogenase; aldehyde dehydrogenase KEGG: cpe:CPE2531 alcohol dehydrogenase / acetaldehyde dehydrogenase
535432	C	T	HIGH	0.96	coding	0	1	0.42	0.956	0.99	4853	11043	NON	S	L	404	Cthe_0426	535029	536699	1	putative PAS/PAC sensor protein	PFAM: 4Fe-4S ferredoxin, iron-sulfur binding; putative Fe-S cluster SMART: PAS KEGG: tte:TTE0697 ferredoxin 2
565149	C	T	HIGH	0.98	coding	0	1	0.01	0.984	0.977	20777	11320	SYN	N	N	354	Cthe_0450	564796	565251	1	Ribonucleotide reductase regulator NrdR-like protein	TIGRFAM: Ribonucleotide reductase regulator NrdR-like PFAM: ATP-cone KEGG: tte:TTE1634 predicted transcriptional regulator, consists of a Zn-ribbon and ATP-cone domains
580932	G	A	HIGH	0.975	coding	0	1	10001.22	0.896	0.971	2479	3433	NON	G	S	10	Cthe_0466	580923	581945	1	flagellar motor switch protein FliG	PFAM: flagellar motor switch protein FliG KEGG: tte:TTE1441 flagellar motor switch protein
588352	G	A	HIGH	0.955	coding	0	1	-0.7	0.987	0.79	7953	3145	NON	A	T	226	Cthe_0473	588127	588516	1	flagellar operon protein	TIGRFAM: flagellar operon protein KEGG: tte:TTE1434 hypothetical protein
600252	A	T	HIGH	0.94	coding	0	1	-0.98	0.986	0.984	545	647	NON	I	L	811	Cthe_0486	599442	600653	1	GTP-binding signal recognition particle SRP54, G-domain	PFAM: GTP-binding signal recognition particle SRP54, G-domain SMART: ATPase KEGG: cac:CAC2146 flagellar GTP-binding protein, FlhF
610845	A	G	HIGH	1	coding	0	1	10001.53	0.486	0.878	1576	5315	SYN	K	K	1311	Cthe_0498	609535	611163	1	protein of unknown function DUF342	PFAM: protein of unknown function DUF342 KEGG: tte:TTE1412 hypothetical protein
613076	A	G	HIGH	0.985	coding	0	1	-0.21	0.583	0.862	2875	3896	SYN	S	S	132	Cthe_0502	612945	613811	1	hypothetical protein	KEGG: bba:Bd3636 hypothetical protein
628597	G	A	HIGH	0.995	coding	0	1	-0.52	0.994	0.885	9133	7870	NON	A	T	976	Cthe_0514	627622	628737	1	DNA methylase N-4/N-6	PFAM: DNA methylase N-4/N-6 KEGG: adenine-specific methyltransferase
632372	C	G	HIGH	1	coding	0	1	-0.89	0.425	0.582	678	1044	NON	V	L	16	Cthe_0517	632019	632387	-1	hypothetical protein	KEGG: dsy:DSY3291 hypothetical protein
659458	T	C	HIGH	0.995	coding	0	1	0.57	0.89	0.916	2295	1641	NON	I	T	872	Cthe_0540	658587	661193	1	protein of unknown function DUF214	PFAM: protein of unknown function DUF214 KEGG: ctc:CTC01915 ABC transporter ATP-binding protein
689837	A	G	HIGH	0.995	coding	0	1	-0.03	0.936	0.976	4522	8863	NON	M	V	907	Cthe_0561	688931	689977	1	ApbE-like lipoprotein	PFAM: ApbE-like lipoprotein KEGG: dsy:DSY3906 hypothetical protein
691119	C	T	HIGH	0.985	coding	0	1	-0.01	0.45	0.572	1390	2764	NON	L	F	157	Cthe_0564	690963	691919	1	Trans-hexaprenyltranstransferase	PFAM: Polyprenyl synthetase KEGG: tte:TTE2483 geranylgeranyl pyrophosphate synthase
702366	A	G	HIGH	0.995	coding	0	1	0.72	0.847	0.988	6696	4586	NON	T	A	913	Cthe_0574	701454	703565	1	serine/threonine protein kinase	PFAM: protein kinase; tyrosine protein kinase; PASTA domain containing protein SMART: Serine/threonine protein kinase KEGG: tte:TTE1500 serine/threonine protein kinase
707138	C	T	HIGH	0.975	coding	0	1	-0.96	0.811	0.705	2047	2134	NON	G	S	1153	Cthe_0578	706080	708290	-1	glycoside hydrolase, family 9	PFAM: glycoside hydrolase, family 9; type 3a, cellulose-binding; cellulosome enzyme, dockerin type I KEGG: cac:CAC0913 possible non-processive endoglucanase family 9, secreted; CelG ortholog; dockerin and cellulose-binding domain
763262	G	A	HIGH	0.96	coding	0	1	0.46	0.618	0.846	1095	2642	NON	P	L	182	Cthe_0622	762637	763443	-1	methylthioadenosine phosphorylase	TIGRFAM: methylthioadenosine phosphorylase PFAM: purine phosphorylase, family 2 KEGG: sai:Saci_0997 5'-methylthioadenosine phosphorylase
766865	A	G	HIGH	0.98	coding	0	1	-0.26	0.641	0.566	1251	623	NON	N	D	1834	Cthe_0624	765032	769837	1	glycoside hydrolase, family 9-like Ig-like	PFAM: PKD domain containing protein; glycoside hydrolase, family 9; cellulosome enzyme, dockerin type I; glycoside hydrolase, family 9-like Ig-like KEGG: cac:CAC0915 endoglucanase A precursor (endo-1,4-beta-glucanase) (cellulase A), secreted; dockerin doma
780355	C	T	HIGH	0.985	coding	0	1	-0.72	0.431	0.425	396	615	SYN	P	P	60	Cthe_0633	780296	780781	1	hypothetical protein	
816559	G	A	HIGH	0.99	coding	0	1	10003.19	0.854	0.449	3307	2209	SYN	P	P	756	Cthe_0661	815804	817519	1	Ricin B lectin	PFAM: Ricin B lectin; cellulosome enzyme, dockerin type I; glycoside hydrolase, family 43 KEGG: sma:SAV2109 beta-xylosidase
893075	C	A	HIGH	0.95	coding	0	1	-0.98	0.801	0.991	4193	3312	NON	G	*	733	Cthe_0735	893001	893807	-1	cellulosome anchoring protein, cohesin region	PFAM: cellulosome anchoring protein, cohesin region
908630	G	A	HIGH	0.975	coding	0	1	2.37	0.911	0.779	1713	2205	SYN	N	N	1176	Cthe_0744	906953	909805	-1	copper amine oxidase-like protein	PFAM: peptidase S41; copper amine oxidase-like KEGG: cac:CAC3039 hypothetical secreted protein
915475	A	G	HIGH	0.99	coding	0	1	10001.39	0.984	0.832	7863	5904	NON	V	A	638	Cthe_0749	915285	916112	-1	binding-protein-dependent transport systems inner membrane component	PFAM: binding-protein-dependent transport systems inner membrane component KEGG: dsy:DSY1462 hypothetical protein
933177	A	G	HIGH	0.975	coding	0	1	-0.61	0.353	0.974	1205	5781	NON	I	T	623	Cthe_0770	932459	933799	-1	signal recognition particle subunit FFH/SRP54 (srp54)	KEGG: tte:TTE1462 signal recognition particle GTPase TIGRFAM: signal recognition particle protein PFAM: GTP-binding signal recognition particle SRP54, G-domain; Signal peptide binding (SRP54) M-domain SMART: ATPase
949261	C	T	HIGH	0.98	coding	0	1	-0.78	0.481	0.484	849	624	NON	H	Y	124	Cthe_0785	949138	949440	1	hypothetical protein	
950945	A	T	HIGH	0.99	coding	0	1	0.29	0.485	0.566	432	496	NON	D	E	2661	Cthe_0787	950810	953605	-1	Isoleucyl-tRNA synthetase	TIGRFAM: isoleucyl-tRNA synthetase KEGG: tte:TTE1594 isoleucyl-tRNA synthetase
955054	G	A	HIGH	0.965	coding	0	1	10001.65	0.971	0.758	6787	3615	SYN	V	V	222	Cthe_0789	954484	955275	-1	RNA-binding S4	PFAM: RNA-binding S4 KEGG: tte:TTE1597 hypothetical protein
1012867	G	A	HIGH	0.985	coding	0	1	0.3	0.951	0.949	3056	6502	NON	A	V	437	Cthe_0831	1012416	1013303	-1	farnesyl-diphosphate synthase	PFAM: Polyprenyl synthetase KEGG: tte:TTE1296 Geranylgeranyl pyrophosphate synthase
1015063	A	G	HIGH	0.995	coding	0	1	1.24	0.531	0.852	1111	2874	NON	S	P	226	Cthe_0834	1014851	1015288	-1	NusB antitermination factor	TIGRFAM: transcription antitermination factor NusB PFAM: NusB/RsmB/TIM44 KEGG: cac:CAC2084 transcription termination factor NusB
1028888	C	T	HIGH	0.98	coding	0	1	0.28	0.986	0.666	4207	2046	NON	A	T	697	Cthe_0853	1027899	1029584	-1	type II secretion system protein E	PFAM: type II secretion system protein E; General secretory system II, protein E-like KEGG: tte:TTE1262 predicted ATPases involved in pili biogenesis, PilB homologs
1072740	C	T	HIGH	0.95	coding	0	1	0.53	0.928	0.775	4105	3407	NON	E	K	1396	Cthe_0896	1072333	1074135	-1	DNA primase	KEGG: tte:TTE1756 DNA primase (bacterial type) TIGRFAM: DNA primase PFAM: zinc finger, CHC2-type; TOPRIM domain containing protein; DNA primase catalytic core-like SMART: Toprim subdomain
1088316	T	C	HIGH	0.97	coding	0	1	0.08	0.933	0.942	2232	5966	NON	H	R	647	Cthe_0911	1088123	1088962	-1	protein of unknown function DUF152	PFAM: protein of unknown function DUF152 KEGG: tte:TTE0179 conserved hypothetical protein
1113552	T	C	HIGH	0.99	coding	0	1	0.95	0.886	0.833	2416	2853	NON	T	A	2875	Cthe_0927	1112854	1116426	-1	condensin subunit Smc	TIGRFAM: GTP-binding; chromosome segregation protein SMC PFAM: SMC protein-like; SMCs flexible hinge KEGG: cac:CAC1751 chromosome segregation SMC protein, ATPase
1141582	A	G	HIGH	0.99	coding	0	1	-0.6	0.659	0.982	4662	12044	SYN	H	H	912	Cthe_0952	1141213	1142493	-1	dihydroorotase	KEGG: gsu:GSU1272 dihydroorotase, multifunctional complex type TIGRFAM: dihydroorotase, multifunctional complex type PFAM: amidohydrolase
1157787	C	T	HIGH	0.99	coding	0	1	2.05	0.711	0.73	1229	1420	NON	E	K	2044	Cthe_0968	1157710	1159830	-1	UvrD/REP helicase	PFAM: UvrD/REP helicase KEGG: ctc:CTC00692 DNA helicase II
1181521	T	C	HIGH	0.995	coding	0	1	0.12	0.842	0.743	3061	1931	SYN	V	V	57	Cthe_0987	1180639	1181577	-1	riboflavin kinase / FMN adenylyltransferase	TIGRFAM: riboflavin biosynthesis protein RibF PFAM: Riboflavin kinase / FAD synthetase KEGG: cpe:CPE1682 riboflavin kinase / FMN adenylyltransferase
1194190	C	T	HIGH	0.98	coding	0	1	10002.29	0.336	0.538	454	2236	NON	R	K	26	Cthe_0996	1189887	1194215	-1	DNA polymerase III catalytic subunit, PolC type	KEGG: cpe:CPE1691 DNA polymerase III alpha subunit TIGRFAM: DNA polymerase III, epsilon subunit; DNA polymerase III, alpha subunit PFAM: PHP-like; nucleic acid binding, OB-fold, tRNA/helicase-type; Exonuclease, RNase T and DNA polymerase III SMART: Phospho
1194244	G	A	HIGH	0.975	coding	0	1	0.27	0.53	0.525	1732	3392	SYN	L	L	1048	Cthe_0997	1194233	1195291	-1	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	KEGG: cpe:CPE1692 peptidoglycan acetylation protein TIGRFAM: 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase PFAM: IspG protein
1231266	G	A	HIGH	0.99	coding	0	1	-0.42	0.989	0.98	12642	4198	NON	A	V	815	Cthe_1029	1231004	1232080	-1	phosphotransacetylase	TIGRFAM: phosphate acetyltransferase PFAM: phosphate acetyl/butaryl transferase KEGG: mma:MM0496 phosphate acetyltransferase
1244897	G	A	HIGH	0.98	coding	0	1	-0.58	0.982	0.874	11213	1858	NON	R	C	1279	Cthe_1041	1244775	1246175	-1	UDP-N-acetylmuramoylalanine--D-glutamate ligase	TIGRFAM: UDP-N-acetylmuramoylalanine--D-glutamate ligase PFAM: cytoplasmic peptidoglycan synthetases-like; Mur ligase, middle region KEGG: ctc:CTC00213 UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase
1284894	C	T	HIGH	0.97	coding	0	1	-1.04	0.807	0.677	3364	1781	SYN	A	A	537	Cthe_1077	1284358	1284987	1	hypothetical protein	
1330160	C	T	HIGH	0.96	coding	0	1	0.16	0.914	0.797	3848	1803	NON	G	R	760	Cthe_1115	1329264	1330919	-1	Tn7-like transposition protein C	KEGG: bca:BCE0176 Tn7-like transposition protein C
1357656	T	C	HIGH	0.985	coding	0	1	-0.17	0.98	0.756	5819	2692	NON	Q	R	1103	Cthe_1141	1354697	1358758	-1	hypothetical protein	KEGG: pfa:PFE1095w hypothetical protein
1390700	C	T	HIGH	0.965	coding	0	1	-0.13	0.982	0.99	2854	1537	NON	A	T	55	Cthe_1165	1389543	1390754	-1	YbbR-like protein	PFAM: YbbR-like KEGG: tte:TTE2207 hypothetical protein
1400010	G	A	HIGH	0.99	coding	0	1	-0.66	0.861	0.85	2512	2747	NON	V	I	529	Cthe_1172	1399482	1400630	1	protein of unknown function DUF362	PFAM: 4Fe-4S ferredoxin, iron-sulfur binding; protein of unknown function DUF362 KEGG: dsy:DSY4491 hypothetical protein
1411010	A	T	HIGH	0.98	coding	0	1	10001.06	0.991	0.984	10281	8931	NON	I	N	1067	Cthe_1182	1410808	1412076	-1	protein of unknown function DUF58	PFAM: protein of unknown function DUF58 KEGG: sth:STH2554 hypothetical protein
1451061	T	C	HIGH	0.995	coding	0	1	10000.91	0.941	0.988	7814	15580	NON	V	A	773	Cthe_1217	1450289	1452619	1	ATP-dependent Clp protease ATP-binding subunit ClpA	KEGG: cac:CAC1824 ATP-dependent Clp proteinase TIGRFAM: ATP-dependent Clp protease, ATP-binding subunit clpA PFAM: AAA ATPase, central region; Clp, N terminal; ATPase associated with various cellular activities, AAA_5; ATPase AAA-2 SMART: ATPase
1455798	G	A	HIGH	0.95	coding	0	1	10000.22	0.956	0.859	4451	2090	NON	A	V	8426	Cthe_1221	1455455	1464223	-1	glycosyltransferase 36	PFAM: putative carbohydrate binding; glycosyltransferase 36; glycosyltransferase 36 associated KEGG: tte:TTE2146 cyclic beta 1-2 glucan synthetase
1460496	T	C	HIGH	0.995	coding	0	1	0.1	0.99	0.989	7635	3763	NON	E	G	3728	Cthe_1221	1455455	1464223	-1	glycosyltransferase 36	PFAM: putative carbohydrate binding; glycosyltransferase 36; glycosyltransferase 36 associated KEGG: tte:TTE2146 cyclic beta 1-2 glucan synthetase
1480289	T	C	HIGH	0.98	coding	0	1	-0.84	0.939	0.687	6253	7473	NON	Y	C	18746	Cthe_1235	1478377	1499034	-1	Cellulose 1,4-beta-cellobiosidase	PFAM: PKD domain containing protein; S-layer-like region; Hyalin; Fibronectin, type III; Cna B-type; APHP KEGG: psp:PSPPH_4807 glycosyl hydrolase, family 18
1493070	C	T	HIGH	0.99	coding	0	1	0.25	0.335	0.687	790	895	NON	E	K	5965	Cthe_1235	1478377	1499034	-1	Cellulose 1,4-beta-cellobiosidase	PFAM: PKD domain containing protein; S-layer-like region; Hyalin; Fibronectin, type III; Cna B-type; APHP KEGG: psp:PSPPH_4807 glycosyl hydrolase, family 18
1500998	C	T	HIGH	0.995	coding	0	1	-0.89	0.336	0.447	880	850	NON	A	T	1546	Cthe_1237	1500066	1502543	-1	leucyl-tRNA synthetase	TIGRFAM: leucyl-tRNA synthetase KEGG: mta:Moth_0568 leucyl-tRNA synthetase
1501585	G	A	HIGH	0.95	coding	0	1	0	0.868	0.735	7615	6167	NON	P	L	959	Cthe_1237	1500066	1502543	-1	leucyl-tRNA synthetase	TIGRFAM: leucyl-tRNA synthetase KEGG: mta:Moth_0568 leucyl-tRNA synthetase
1530181	C	T	HIGH	0.975	coding	0	1	10000.65	0.958	0.99	11211	6753	NON	V	M	91	Cthe_1259	1529729	1530271	-1	CDP-alcohol phosphatidyltransferase	PFAM: CDP-alcohol phosphatidyltransferase KEGG: tte:TTE0757 Phosphatidylglycerophosphate synthase
1552089	C	T	HIGH	0.955	coding	0	1	-0.63	0.694	0.929	4289	5226	NON	E	K	103	Cthe_1275	1551754	1552191	-1	H+-ATPase subunit H	KEGG: tte:TTE1485 archaeal/vacuolar-type H+-ATPase subunit H
1580697	T	C	HIGH	0.995	coding	0	1	10001.38	0.95	0.993	7201	18769	NON	I	M	429	Cthe_1302	1579458	1581125	-1	conserved hypothetical protein	TIGRFAM: conserved hypothetical protein PFAM: beta-lactamase-like; RNA-metabolising metallo-beta-lactamase KEGG: tte:TTE1253 predicted hydrolase of the metallo-beta-lactamase superfamily
1583731	T	C	HIGH	0.995	coding	0	1	-0.67	0.988	0.737	3638	794	NON	K	E	778	Cthe_1304	1583135	1584508	-1	PhoH-like protein	PFAM: PhoH-like protein SMART: Nucleotide binding protein, PINc KEGG: cpe:CPE1665 probable phosphate starvation inducible protein
1593065	G	A	HIGH	0.96	coding	0	1	10001.05	0.829	0.989	2090	5338	SYN	*	*	1074	Cthe_1309	1591992	1593065	1	Radical SAM	PFAM: Radical SAM SMART: Elongator protein 3/MiaB/NifB KEGG: mbu:Mbur_1233 Fe-S protein, radical SAM family
1606832	T	C	HIGH	0.995	coding	0	1	-0.03	0.967	0.946	5237	3534	NON	E	G	71	Cthe_1323	1606222	1606902	-1	GrpE protein	PFAM: GrpE protein KEGG: cac:CAC1281 molecular chaperone GrpE
1612350	T	C	HIGH	0.995	coding	0	1	10000.32	0.967	0.985	8312	12279	SYN	P	P	732	Cthe_1328	1611804	1613081	-1	Stage II sporulation P	PFAM: Stage II sporulation P KEGG: tte:TTE0947 stage II sporulation protein P
1615176	C	T	HIGH	0.99	coding	0	1	10001.56	0.826	0.816	5620	8825	SYN	K	K	3000	Cthe_1329	1613874	1618175	-1	putative CoA-substrate-specific enzyme activase	TIGRFAM: putative CoA-substrate-specific enzyme activase PFAM: ATPase, BadF/BadG/BcrA/BcrD type KEGG: cac:CAC2401 activator of 2-hydroxyglutaryl-CoA dehydratase (duplicated HSP70 class ATPase domain) fused to uncharacterized conserved protein
1633288	G	A	HIGH	1	coding	0	1	-0.93	0.391	0.389	1790	4129	NON	P	L	842	Cthe_1341	1632612	1634129	-1	coproporphyrinogen III oxidase, anaerobic	PFAM: Radical SAM SMART: Elongator protein 3/MiaB/NifB KEGG: cpe:CPE1935 probable coproporphyrinogen III oxidase
1636484	T	C	HIGH	0.995	coding	0	1	10003.4	0.993	0.983	9930	9597	SYN	V	V	1002	Cthe_1344	1635311	1637485	-1	(p)ppGpp synthetase I, SpoT/RelA	KEGG: tte:TTE1195 Guanosine polyphosphate pyrophosphohydrolases/synthetases TIGRFAM: RelA/SpoT family protein PFAM: amino acid-binding ACT; TGS; metal-dependent phosphohydrolase, HD subdomain; RelA/SpoT SMART: Metal-dependent phosphohydrolase, HD region
1645595	T	C	HIGH	0.985	coding	0	1	10000.64	0.991	0.618	8477	1848	SYN	V	V	63	Cthe_1352	1644326	1645657	-1	UDP-glucose 6-dehydrogenase	PFAM: UDP-glucose/GDP-mannose dehydrogenase KEGG: bja:bll8129 UDP-glucose 6-dehydrogenase
1675113	C	T	MEDIUM	0.97	coding	0	0	10001.32	0.733	0.823	1486	3645	NON	A	S	1860	Cthe_1373	1671219	1676972	-1	YD repeat protein	TIGRFAM: YD repeat protein PFAM: YD repeat KEGG: mba:Mbar_A3399 hypothetical protein
1675115	C	A	MEDIUM	0.985	coding	0	0	10001.32	0.602	0.777	484	560	NON	A	S	1858	Cthe_1373	1671219	1676972	-1	YD repeat protein	TIGRFAM: YD repeat protein PFAM: YD repeat KEGG: mba:Mbar_A3399 hypothetical protein
1675135	G	C	MEDIUM	0.94	coding	6.58	0	10001.32	0.989	0.974	8387	2827	NON	A	G	1838	Cthe_1373	1671219	1676972	-1	YD repeat protein	TIGRFAM: YD repeat protein PFAM: YD repeat KEGG: mba:Mbar_A3399 hypothetical protein
1675440	G	C	LOW	0.985	coding	0	0	-0.06	0.456	0.599	864	669	SYN	T	T	1533	Cthe_1373	1671219	1676972	-1	YD repeat protein	TIGRFAM: YD repeat protein PFAM: YD repeat KEGG: mba:Mbar_A3399 hypothetical protein
1675446	C	T	LOW	0.985	coding	0	0	-0.06	0.43	0.595	688	545	NON	A	S	1527	Cthe_1373	1671219	1676972	-1	YD repeat protein	TIGRFAM: YD repeat protein PFAM: YD repeat KEGG: mba:Mbar_A3399 hypothetical protein
1675448	C	A	LOW	0.945	coding	3.29	0	-0.44	0.544	0.693	1352	390	NON	A	S	1525	Cthe_1373	1671219	1676972	-1	YD repeat protein	TIGRFAM: YD repeat protein PFAM: YD repeat KEGG: mba:Mbar_A3399 hypothetical protein
1677616	T	C	LOW	0.96	coding	2.93	0	-0.83	0.995	0.998	9625	12198	SYN	Q	Q	288	Cthe_1374	1677115	1677903	-1	copper amine oxidase-like protein	PFAM: copper amine oxidase-like KEGG: mta:Moth_0517 N-acetylmuramoyl-L-alanine amidase
1678574	C	T	HIGH	0.98	coding	0	1	-0.36	0.466	0.579	1114	2098	NON	G	S	964	Cthe_1375	1678188	1679537	-1	aspartate kinase	KEGG: bha:BH1500 homoserine dehydrogenase TIGRFAM: aspartate kinase PFAM: aspartate/glutamate/uridylate kinase; amino acid-binding ACT
1683605	T	C	HIGH	0.995	coding	0	1	-0.88	0.606	0.579	1911	1504	SYN	I	I	237	Cthe_1379	1683369	1684457	1	oxidoreductase-like protein	PFAM: oxidoreductase-like; Oxidoreductase-like KEGG: tma:TM0312 hypothetical protein
1693277	T	C	HIGH	0.985	coding	0	1	10001.71	0.985	0.827	7579	3868	NON	R	G	250	Cthe_1385	1690794	1693526	-1	protein translocase subunit secA	TIGRFAM: preprotein translocase, SecA subunit PFAM: SEC-C motif; SecA DEAD-like; SecA Wing and Scaffold; SecA preprotein cross-linking region KEGG: tte:TTE0522 Preprotein translocase subunit SecA (ATPase, RNA helicase)
1710654	C	T	HIGH	0.95	coding	0	1	0.08	0.702	0.783	3453	3314	NON	G	R	280	Cthe_1399	1710415	1710933	-1	C_GCAxxG_C_C family protein	TIGRFAM: C_GCAxxG_C_C family protein KEGG: bfs:BF2583 hypothetical protein
1723086	T	C	MEDIUM	0.99	coding	0	0	0.72	0.99	0.986	6654	11103	NON	I	M	1887	Cthe_1405	1722843	1724972	-1	Spore coat assembly protein-like protein	KEGG: oih:OB2511 hypothetical protein
1723100	C	T	MEDIUM	0.94	coding	0	0	10000.9	0.982	0.987	5856	4720	NON	D	N	1873	Cthe_1405	1722843	1724972	-1	Spore coat assembly protein-like protein	KEGG: oih:OB2511 hypothetical protein
1723108	C	T	MEDIUM	0.935	coding	0	0	10000.9	0.986	0.981	817	686	NON	R	Q	1865	Cthe_1405	1722843	1724972	-1	Spore coat assembly protein-like protein	KEGG: oih:OB2511 hypothetical protein
1723114	C	T	MEDIUM	0.96	coding	3.29	0	10000.9	0.973	0.971	3800	6730	NON	G	D	1859	Cthe_1405	1722843	1724972	-1	Spore coat assembly protein-like protein	KEGG: oih:OB2511 hypothetical protein
1723122	T	C	MEDIUM	0.995	coding	0	0	10000.78	0.979	0.987	2353	5532	NON	I	M	1851	Cthe_1405	1722843	1724972	-1	Spore coat assembly protein-like protein	KEGG: oih:OB2511 hypothetical protein
1723136	C	T	MEDIUM	0.945	coding	0	0	10001.8	0.99	0.986	4855	7284	NON	D	N	1837	Cthe_1405	1722843	1724972	-1	Spore coat assembly protein-like protein	KEGG: oih:OB2511 hypothetical protein
1723150	T	C	MEDIUM	0.995	coding	0	0	10001.8	0.987	0.981	6833	15798	NON	D	G	1823	Cthe_1405	1722843	1724972	-1	Spore coat assembly protein-like protein	KEGG: oih:OB2511 hypothetical protein
1723164	T	C	LOW	0.965	coding	0	0	-0.32	0.992	0.982	8411	20453	NON	I	M	1809	Cthe_1405	1722843	1724972	-1	Spore coat assembly protein-like protein	KEGG: oih:OB2511 hypothetical protein
1726591	G	A	LOW	0.97	coding	0	0	-0.82	0.997	0.984	5407	4972	NON	M	I	3	Cthe_1408	1726589	1727287	1	two component transcriptional regulator, winged helix family	PFAM: response regulator receiver; transcriptional regulatory protein-like KEGG: oih:OB0584 two-component response regulator
1763704	G	A	HIGH	0.98	coding	0	1	-0.14	0.698	0.629	2272	1907	NON	R	C	94	Cthe_1445	1763552	1763797	-1	uncharacterized Zn-finger protein	KEGG: cac:CAC2470 uncharacterized Zn-finger protein
1771468	C	T	HIGH	0.965	coding	0	1	10000.81	0.947	0.943	5625	8250	NON	R	K	359	Cthe_1451	1771353	1771826	-1	GCN5-related N-acetyltransferase	PFAM: GCN5-related N-acetyltransferase KEGG: sth:STH1999 putative acetyltransferase
1799818	G	A	HIGH	0.995	coding	0	1	10002.48	0.987	0.984	3702	8509	NON	S	N	1706	Cthe_1480	1798113	1800218	1	conserved hypothetical protein	KEGG: oih:OB0347 hypothetical protein
1818170	C	T	HIGH	0.955	coding	0	1	-0.38	0.857	0.986	2450	10397	SYN	A	A	1383	Cthe_1498	1817744	1819552	-1	methyl-accepting chemotaxis sensory transducer	PFAM: histidine kinase, HAMP region; chemotaxis sensory transducer KEGG: cac:CAC0120 membrane-associated methyl-accepting chemotaxis protein with HAMP domain
1821198	T	C	HIGH	1	coding	0	1	10000.92	0.967	0.993	4006	29525	NON	W	R	646	Cthe_1499	1820553	1822307	1	ammonium transporter	TIGRFAM: ammonium transporter PFAM: nitrogen regulatory protein P-II; Rh-like protein/ammonium transporter KEGG: ade:Adeh_2218 ammonium transporter
1828037	T	C	HIGH	0.99	coding	0	1	-0.94	0.506	0.456	422	388	NON	M	T	758	Cthe_1504	1827280	1828092	1	Linocin_M18 bacteriocin protein	PFAM: Linocin_M18 bacteriocin protein KEGG: gka:GK0916 hypothetical protein
1833211	C	T	HIGH	0.975	coding	0	1	-0.27	0.409	0.546	1464	1087	NON	W	*	1088	Cthe_1509	1833063	1834298	-1	protein of unknown function DUF438	PFAM: protein of unknown function DUF438; Hemerythrin HHE cation binding region KEGG: dsy:DSY1536 hypothetical protein
1894413	T	C	HIGH	1	coding	0	1	10001.4	0.972	0.971	3008	6101	NON	S	P	145	Cthe_1566	1894269	1895612	1	Nitrogenase	PFAM: oxidoreductase/nitrogenase, component 1 KEGG: rru:Rru_A0793 nitrogenase
1896686	A	T	HIGH	0.99	coding	0	1	-0.5	0.987	0.852	4832	1658	NON	D	V	713	Cthe_1568	1895974	1896849	1	Radical SAM	PFAM: Radical SAM SMART: Elongator protein 3/MiaB/NifB KEGG: rpa:RPA2634 putative nitrogenase iron-molybdenum cofactor biosynthesis protein NifB
1900767	C	T	HIGH	0.965	coding	0	1	0.4	0.742	0.932	2369	4283	NON	R	W	100	Cthe_1572	1900668	1901471	1	ABC transporter related protein	PFAM: ABC transporter related SMART: ATPase KEGG: dsy:DSY3499 hypothetical protein
1900995	G	A	HIGH	0.945	coding	0	1	10001.74	0.989	0.982	8518	5346	NON	G	R	328	Cthe_1572	1900668	1901471	1	ABC transporter related protein	PFAM: ABC transporter related SMART: ATPase KEGG: dsy:DSY3499 hypothetical protein
1917309	T	C	HIGH	0.995	coding	0	1	-0.84	0.976	0.875	14538	3813	SYN	P	P	477	Cthe_1586	1916130	1917785	-1	binding-protein-dependent transport systems inner membrane component	PFAM: binding-protein-dependent transport systems inner membrane component KEGG: bcl:ABC3648 Fe3+ ABC transporter permease
1943386	T	C	HIGH	0.975	coding	0	1	0.14	0.958	0.894	6415	5438	NON	D	G	1382	Cthe_1613	1942344	1944767	-1	glycosyl hydrolase-like protein	KEGG: bha:BH3513 hypothetical protein
1946531	T	C	HIGH	0.995	coding	0	1	0.18	0.805	0.79	1371	2312	NON	T	A	1591	Cthe_1616	1945599	1948121	-1	phage minor structural protein	TIGRFAM: phage minor structural protein KEGG: bha:BH3515 hypothetical protein
1957578	T	C	HIGH	0.995	coding	0	1	0.34	0.962	0.585	4924	915	NON	R	G	355	Cthe_1630	1956610	1957932	-1	phage portal protein, HK97 family	TIGRFAM: phage portal protein, HK97 family PFAM: phage portal protein KEGG: bca:BCE0398 phage portal protein, HK97 family
1997171	G	A	HIGH	1	coding	0	1	0.22	0.414	0.321	1203	1161	NON	R	Q	632	Cthe_1671	1996540	1998108	1	Recombinase	PFAM: Resolvase-like; Recombinase KEGG: bca:BCE0419 DNA recombinase, putative
2039632	C	T	HIGH	0.995	coding	0	1	-0.84	0.649	0.992	947	3949	NON	E	K	55	Cthe_1718	2039246	2039686	-1	hypothetical protein	
2040360	T	C	HIGH	1	coding	0	1	-0.6	0.99	0.749	6856	3457	NON	Q	R	629	Cthe_1719	2039699	2040988	-1	phage major capsid protein, HK97 family	TIGRFAM: phage major capsid protein, HK97 family PFAM: phage major capsid protein, HK97 KEGG: rsp:RSP_2358 possible phage phi-C31 GP36-like protein / major capsid protein, HK97 family
2068104	C	T	HIGH	0.97	coding	0	1	10002.16	0.726	0.978	4111	6267	NON	E	K	1114	Cthe_1748	2067664	2069217	-1	hypothetical protein	
2098330	T	C	HIGH	0.98	coding	0	1	-0.86	0.995	0.798	8715	3393	NON	Y	C	224	Cthe_1776	2098068	2098553	-1	protein of unknown function UPF0079	PFAM: protein of unknown function UPF0079 KEGG: cpe:CPE2164 conserved hypothetical protein
2126854	A	G	HIGH	0.985	coding	0	1	0.1	0.529	0.66	3630	5775	NON	N	S	437	Cthe_1800	2126418	2127929	1	Peptidoglycan-binding LysM	PFAM: glycoside hydrolase, family 18; Peptidoglycan-binding LysM; protein of unknown function DUF1058; SH3, type 3 SMART: SH3-like region; Chitinase II KEGG: cac:CAC1556 chitinase family protein
2137476	C	T	LOW	0.975	coding	0.98	0	-0.58	0.989	0.977	792	2253	SYN	E	E	1998	Cthe_1806	2132940	2139473	-1	cellulosome enzyme, dockerin type I	PFAM: cellulosome enzyme, dockerin type I SMART: Ig-like, group 2 KEGG: dme:CG31439-PA CG31439 gene product from transcript CG31439-RA
2137485	G	A	LOW	1	coding	3.29	0	-0.58	0.745	0.946	2001	3788	SYN	T	T	1989	Cthe_1806	2132940	2139473	-1	cellulosome enzyme, dockerin type I	PFAM: cellulosome enzyme, dockerin type I SMART: Ig-like, group 2 KEGG: dme:CG31439-PA CG31439 gene product from transcript CG31439-RA
2137509	C	T	LOW	0.975	coding	3.29	0	0.02	0.617	0.852	1542	2107	SYN	S	S	1965	Cthe_1806	2132940	2139473	-1	cellulosome enzyme, dockerin type I	PFAM: cellulosome enzyme, dockerin type I SMART: Ig-like, group 2 KEGG: dme:CG31439-PA CG31439 gene product from transcript CG31439-RA
2137527	T	C	MEDIUM	0.995	coding	3.29	1	0.32	0.992	0.888	12017	2133	SYN	L	L	1947	Cthe_1806	2132940	2139473	-1	cellulosome enzyme, dockerin type I	PFAM: cellulosome enzyme, dockerin type I SMART: Ig-like, group 2 KEGG: dme:CG31439-PA CG31439 gene product from transcript CG31439-RA
2139279	G	C	HIGH	0.955	coding	0	1	0.42	0.986	0.962	3915	4520	SYN	G	G	195	Cthe_1806	2132940	2139473	-1	cellulosome enzyme, dockerin type I	PFAM: cellulosome enzyme, dockerin type I SMART: Ig-like, group 2 KEGG: dme:CG31439-PA CG31439 gene product from transcript CG31439-RA
2141861	G	A	HIGH	1	coding	0	1	-0.01	0.989	0.982	9352	13409	NON	A	V	920	Cthe_1807	2141116	2142780	-1	conserved hypothetical protein	KEGG: gka:GK1403 hypothetical protein
2160407	C	T	HIGH	0.985	coding	0	1	-0.03	0.695	0.936	544	1441	NON	G	D	2303	Cthe_1825	2159911	2162709	-1	Hpt sensor hybrid histidine kinase	PFAM: response regulator receiver; ATP-binding region, ATPase-like; histidine kinase, HAMP region; histidine kinase A-like; Hpt KEGG: gsu:GSU2314 sensory box histidine kinase/response regulator
2168008	T	C	HIGH	0.985	coding	0	1	-1.05	0.986	0.835	6139	2581	NON	D	G	1454	Cthe_1829	2166165	2169461	-1	Chromosome segregation ATPases-like protein	KEGG: pfa:PF11_0207 hypothetical protein
2169035	G	A	MEDIUM	0.975	coding	0	0	0.05	0.408	0.566	354	559	SYN	L	L	427	Cthe_1829	2166165	2169461	-1	Chromosome segregation ATPases-like protein	KEGG: pfa:PF11_0207 hypothetical protein
2171586	T	G	LOW	0.98	coding	0.55	0	10013.19	0.377	0.295	351	482	SYN	S	S	537	Cthe_1831	2171091	2172122	-1	hypothetical protein	
2171925	C	T	LOW	0.98	coding	0	0	10023.28	0.391	0.402	522	415	SYN	A	A	198	Cthe_1831	2171091	2172122	-1	hypothetical protein	
2173127	G	C	MEDIUM	0.965	coding	0	0	4.3	0.392	0.517	893	461	SYN	L	L	681	Cthe_1832	2172278	2173807	-1	hypothetical protein	
2173457	A	T	LOW	1	coding	0	0	10008.27	0.39	0.38	1061	631	SYN	S	S	351	Cthe_1832	2172278	2173807	-1	hypothetical protein	
2173466	C	T	LOW	0.985	coding	0	0	10008.99	0.389	0.471	671	424	SYN	V	V	342	Cthe_1832	2172278	2173807	-1	hypothetical protein	
2174271	A	G	MEDIUM	0.975	coding	0	0	10002.11	0.402	0.355	373	536	SYN	V	V	390	Cthe_1833	2173851	2174660	-1	copper amine oxidase-like protein	PFAM: copper amine oxidase-like KEGG: mta:Moth_0517 N-acetylmuramoyl-L-alanine amidase
2174280	A	G	MEDIUM	0.945	coding	0	0	10002.11	0.671	0.728	345	654	SYN	G	G	381	Cthe_1833	2173851	2174660	-1	copper amine oxidase-like protein	PFAM: copper amine oxidase-like KEGG: mta:Moth_0517 N-acetylmuramoyl-L-alanine amidase
2174493	T	C	LOW	0.985	coding	3.29	0	10006.06	0.981	0.988	5826	9257	SYN	P	P	168	Cthe_1833	2173851	2174660	-1	copper amine oxidase-like protein	PFAM: copper amine oxidase-like KEGG: mta:Moth_0517 N-acetylmuramoyl-L-alanine amidase
2174514	C	T	LOW	0.98	coding	3.29	1	10006.06	0.968	0.939	10676	10915	SYN	A	A	147	Cthe_1833	2173851	2174660	-1	copper amine oxidase-like protein	PFAM: copper amine oxidase-like KEGG: mta:Moth_0517 N-acetylmuramoyl-L-alanine amidase
2174532	C	T	MEDIUM	0.94	coding	0	0	10001.61	0.409	0.33	377	351	SYN	E	E	129	Cthe_1833	2173851	2174660	-1	copper amine oxidase-like protein	PFAM: copper amine oxidase-like KEGG: mta:Moth_0517 N-acetylmuramoyl-L-alanine amidase
2176500	G	A	MEDIUM	1	coding	3.29	1	2.58	0.863	0.927	14651	5887	SYN	G	G	2151	Cthe_1835	2175342	2178650	-1	Viral A-type inclusion protein repeat containing protein	PFAM: Viral A-type inclusion protein repeat KEGG: ehi:3.t00034 Viral A-type inclusion protein repeat, putative Pfam: CH GBP_C Myosin_tail_1 DUF593 Filament E-MAP-115 DUF827 CHASE3 KE2 PROSITE: GLU_RICH LYS_RICH CH
2210573	C	T	HIGH	0.98	coding	0	1	0.4	0.987	0.889	10928	2008	SYN	N	N	720	Cthe_1864	2209854	2210768	1	N-acetylglutamate kinase	KEGG: tte:TTE2496 acetylglutamate kinase TIGRFAM: acetylglutamate kinase PFAM: aspartate/glutamate/uridylate kinase
2216332	C	T	HIGH	0.97	coding	0	1	-0.02	0.85	0.628	1797	1656	NON	P	L	1796	Cthe_1868	2214537	2217740	1	carbamoyl-phosphate synthase large subunit	TIGRFAM: carbamoyl-phosphate synthase, large subunit PFAM: phosphoribosylglycinamide synthetase; ATP-dependent carboxylate-amine ligase-like, ATP-grasp; protein of unknown function DUF201; Carbamoyl-phosphate synthase L chain, ATP-binding; Carbamoyl-phosph
2225182	C	T	HIGH	0.995	coding	0	1	10002.65	0.929	0.99	6564	8636	NON	A	T	205	Cthe_1873	2224043	2225386	-1	HMG-I and HMG-Y, DNA-binding	KEGG: ava:Ava_3487 HMG-I and HMG-Y, DNA-binding
2238813	G	A	HIGH	0.975	coding	0	1	0.92	0.884	0.908	1986	2935	NON	D	N	373	Cthe_1885	2238441	2239013	1	phage integrase-like SAM-like	PFAM: phage integrase-like SAM-like KEGG: cac:CAC1170 tyrosine recombinase
2245658	C	T	LOW	0.915	coding	3.29	0	-0.92	0.885	0.902	4055	4042	NON	E	K	1234	Cthe_1890	2244759	2246891	-1	cellulosome enzyme, dockerin type I	PFAM: cellulosome enzyme, dockerin type I KEGG: ehi:421.t00002 BspA-like leucine rich repeat protein, putative Pfam: FNIP
2245662	T	C	LOW	0.99	coding	0	0	-0.92	0.788	0.924	2008	932	SYN	L	L	1230	Cthe_1890	2244759	2246891	-1	cellulosome enzyme, dockerin type I	PFAM: cellulosome enzyme, dockerin type I KEGG: ehi:421.t00002 BspA-like leucine rich repeat protein, putative Pfam: FNIP
2245692	A	G	LOW	0.99	coding	11.39	0	-0.43	0.988	0.981	10552	9423	SYN	C	C	1200	Cthe_1890	2244759	2246891	-1	cellulosome enzyme, dockerin type I	PFAM: cellulosome enzyme, dockerin type I KEGG: ehi:421.t00002 BspA-like leucine rich repeat protein, putative Pfam: FNIP
2248459	T	C	HIGH	0.985	coding	0	1	-0.77	0.981	0.979	6500	2659	SYN	S	S	360	Cthe_1893	2248150	2248818	-1	conserved hypothetical protein	KEGG: dsy:DSY4370 hypothetical protein
2286170	C	T	HIGH	0.99	coding	0	1	10000.69	0.982	0.993	17716	36372	SYN	T	T	1659	Cthe_1917	2285111	2287828	-1	ATPase, P-type (transporting), HAD superfamily, subfamily IC	TIGRFAM: ATPase, P-type (transporting), HAD superfamily, subfamily IC PFAM: cation transporting ATPase-like; Haloacid dehalogenase-like hydrolase; E1-E2 ATPase-associated region KEGG: tte:TTE0746 Cation transport ATPases
2316023	G	A	HIGH	0.995	coding	0	1	-1	0.94	0.431	4929	1474	SYN	D	D	261	Cthe_1939	2314949	2316283	-1	magnesium transporter	TIGRFAM: magnesium transporter PFAM: CBS domain containing protein; MgtE integral membrane region; MgtE intracellular region KEGG: ctc:CTC01350 Mg2+ transporter mgtE
2318080	G	A	HIGH	0.985	coding	0	1	10001.66	0.775	0.776	3220	6997	NON	S	L	452	Cthe_1942	2317632	2318531	-1	hypothetical protein	KEGG: mta:Moth_0516 spore germination protein-like
2322939	T	C	HIGH	0.985	coding	0	1	0.01	0.929	0.927	1904	3285	NON	D	G	344	Cthe_1947	2322686	2323282	-1	Redoxin	PFAM: Redoxin; Thioredoxin domain KEGG: dsy:DSY1195 hypothetical protein
2351019	C	T	HIGH	0.995	coding	0	1	-0.82	0.561	0.53	986	1510	SYN	E	E	249	Cthe_1968	2348067	2351267	-1	cell divisionFtsK/SpoIIIE	PFAM: cell divisionFtsK/SpoIIIE KEGG: cac:CAC0039 DNA segregation ATPase FtsK/SpoIIIE family protein, contains FHA domain
2368028	C	T	LOW	0.96	coding	3.29	0	-0.53	0.995	0.997	4464	2135	SYN	L	L	546	Cthe_1986	2366420	2368573	-1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
2421931	T	C	HIGH	0.995	coding	0	1	-0.6	0.921	0.797	3703	5318	SYN	P	P	2004	Cthe_2038	2421511	2423934	-1	cellulosome enzyme, dockerin type I	PFAM: cellulosome enzyme, dockerin type I
2449572	G	A	HIGH	0.97	coding	0	1	10001.16	0.987	0.893	15452	12390	SYN	N	N	159	Cthe_2058	2449203	2449730	-1	Appr-1-p processing protein	PFAM: Appr-1-p processing KEGG: noc:Noc_1830 Appr-1-p processing
2478733	T	C	HIGH	0.995	coding	0	1	10001.93	0.916	0.918	3745	6492	NON	N	D	2233	Cthe_2086	2478446	2480965	-1	peptidase U32	PFAM: peptidase U32 KEGG: chy:CHY_1565 peptidase, U32 family
2491691	C	T	HIGH	0.965	coding	0	1	-0.09	0.742	0.854	1896	2142	SYN	S	S	129	Cthe_2095	2491052	2491819	-1	hydrolase, TatD family	TIGRFAM: hydrolase, TatD family PFAM: TatD-related deoxyribonuclease; amidohydrolase 2 KEGG: bld:BLi00052 similar to proteins; RBL03161
2504381	T	C	HIGH	0.98	coding	0	1	10000.66	0.67	0.732	3358	2759	NON	H	R	2420	Cthe_2109	2504263	2506800	-1	copper amine oxidase-like protein	PFAM: copper amine oxidase-like KEGG: dsy:DSY2326 hypothetical protein
2514194	G	A	HIGH	0.995	coding	0	1	10001.42	0.991	0.989	4384	14625	NON	V	I	385	Cthe_2116	2513810	2514571	1	binding-protein-dependent transport systems inner membrane component	PFAM: binding-protein-dependent transport systems inner membrane component KEGG: cac:CAC1399 ABC transporter, permease component
2517992	T	C	HIGH	0.985	coding	0	1	10002.82	0.914	0.842	3186	7723	NON	I	V	823	Cthe_2119	2516532	2518814	-1	glycoside hydrolase, family 10	PFAM: glycoside hydrolase, family 10 KEGG: putative endo-1,4-beta-xylanase X-1
2595560	T	C	HIGH	1	coding	0	1	1.64	0.977	0.926	9733	10784	NON	D	G	557	Cthe_2179	2593348	2596116	-1	Pectate lyase/Amb allergen	PFAM: Pectate lyase/Amb allergen; cellulosome enzyme, dockerin type I; Carbohydrate binding family 6 KEGG: sco:SCO7228 polysaccharide lyase
2657566	C	T	HIGH	0.985	coding	0	1	10001.94	0.748	0.97	3646	5559	NON	G	D	251	Cthe_2229	2656764	2657816	-1	N-acetylneuraminate synthase	PFAM: N-acetylneuraminic acid synthase-like KEGG: cac:CAC2187 sialic acid synthase
2662099	T	C	HIGH	0.995	coding	0	1	0.27	0.912	0.827	4847	4813	NON	Q	R	80	Cthe_2234	2661795	2662178	-1	hypothetical protein	KEGG: ctc:CTC01713 hypothetical protein
2692640	C	T	HIGH	0.98	coding	0	1	-0.33	0.602	0.443	2545	890	NON	P	S	172	Cthe_2263	2692469	2692936	1	H+-transporting two-sector ATPase, C subunit	PFAM: H+-transporting two-sector ATPase, C subunit KEGG: efa:EF1494 V-type ATPase, subunit K
2692670	A	G	HIGH	0.99	coding	0	1	10001.92	0.893	0.995	6759	29205	NON	I	V	202	Cthe_2263	2692469	2692936	1	H+-transporting two-sector ATPase, C subunit	PFAM: H+-transporting two-sector ATPase, C subunit KEGG: efa:EF1494 V-type ATPase, subunit K
2775727	T	C	HIGH	0.995	coding	0	1	10003.31	0.989	0.987	4750	4002	NON	H	R	458	Cthe_2329	2775348	2776184	-1	hypothetical protein	
2782553	G	A	HIGH	0.985	coding	0	1	-0.35	0.645	0.502	2180	494	NON	L	F	790	Cthe_2334	2781498	2783342	-1	polysaccharide biosynthesis protein CapD	PFAM: NAD-dependent epimerase/dehydratase; short-chain dehydrogenase/reductase SDR; 3-beta hydroxysteroid dehydrogenase/isomerase; CoA-binding; polysaccharide biosynthesis protein CapD; dTDP-4-dehydrorhamnose reductase; Male sterility-like KEGG: tte:TTE066
2795536	G	A	HIGH	0.995	coding	0	1	0.3	0.666	0.599	814	1239	NON	T	I	218	Cthe_2345	2794605	2795753	-1	DegT/DnrJ/EryC1/StrS aminotransferase	PFAM: DegT/DnrJ/EryC1/StrS aminotransferase; aromatic amino acid beta-eliminating lyase/threonine aldolase KEGG: bca:BCE5394 aminotransferase family protein
2798536	T	C	HIGH	1	coding	0	1	10001.18	0.986	0.988	18777	5020	NON	I	M	306	Cthe_2346	2795812	2798841	-1	O-antigen polymerase	PFAM: TPR repeat; O-antigen polymerase SMART: Tetratricopeptide region KEGG: tte:TTE0644 TPR-repeat-containing proteins
2825521	T	C	HIGH	1	coding	0	1	10000.8	0.725	0.597	1165	2304	NON	I	M	156	Cthe_2361	2823193	2825676	-1	DNA gyrase subunit A	TIGRFAM: DNA gyrase, A subunit PFAM: DNA gyrase/topoisomerase IV, subunit A; DNA gyrase C-terminal repeat, beta-propeller KEGG: chy:CHY_2704 DNA gyrase, A subunit
2825585	C	T	HIGH	0.985	coding	0	1	10000.82	0.689	0.594	3510	3415	NON	S	N	92	Cthe_2361	2823193	2825676	-1	DNA gyrase subunit A	TIGRFAM: DNA gyrase, A subunit PFAM: DNA gyrase/topoisomerase IV, subunit A; DNA gyrase C-terminal repeat, beta-propeller KEGG: chy:CHY_2704 DNA gyrase, A subunit
2833067	A	G	HIGH	0.995	coding	0	1	-1	0.983	0.991	3976	6001	SYN	F	F	330	Cthe_2369	2833016	2833396	-1	ribonuclease P protein component	TIGRFAM: ribonuclease P protein component PFAM: ribonuclease P protein KEGG: tte:TTE2801 RNase P protein component
2837314	A	G	HIGH	1	coding	0	1	-0.38	0.905	0.702	5148	1070	NON	Q	R	101	Cthe_2374	2837214	2838323	1	DNA replication and repair protein RecF	TIGRFAM: DNA replication and repair protein RecF PFAM: SMC protein-like KEGG: cpe:CPE0004 DNA repair and genetic recombination protein
2843031	C	T	HIGH	0.96	coding	0	1	-0.28	0.78	0.869	2751	2792	NON	P	L	494	Cthe_2378	2842538	2843407	1	chromosome segregation DNA-binding protein	TIGRFAM: parB-like partition protein PFAM: ParB-like nuclease KEGG: tte:TTE2790 predicted Transcriptional regulator
2863438	C	T	HIGH	0.98	coding	0	1	10002.37	0.651	0.885	2445	9750	SYN	A	A	1119	Cthe_2397	2862320	2863540	1	conserved hypothetical protein	KEGG: bce:BC5428 hypothetical protein
2880313	C	T	HIGH	1	coding	0	1	10001.63	0.782	0.988	5688	8126	SYN	D	D	1170	Cthe_2412	2879144	2881795	1	SMC protein-like protein	PFAM: SMC protein-like KEGG: cpe:CPE2159 hypothetical protein
2884831	A	G	HIGH	0.98	coding	0	1	10000.8	0.823	0.91	4225	4195	NON	N	D	601	Cthe_2415	2884231	2884932	1	cell wall hydrolase, SleB	PFAM: Peptidoglycan-binding domain 1; cell wall hydrolase, SleB KEGG: ctc:CTC00350 spore-cortex-lytic enzyme precursor
2888660	C	T	HIGH	1	coding	0	1	10001.76	0.805	0.986	5410	8541	NON	A	T	58	Cthe_2418	2887674	2888717	-1	ATPase	SMART: ATPase KEGG: tte:TTE2693 DNA replication protein DnaC
2895379	A	G	HIGH	0.985	coding	0	1	10001.79	0.551	0.615	1739	1965	SYN	S	S	1386	Cthe_2423	2893768	2896764	-1	hypothetical protein	
2900823	G	A	HIGH	0.96	coding	0	1	10002.12	0.982	0.962	6912	5273	NON	W	*	728	Cthe_2427	2900096	2901052	1	protein of unknown function DUF1385	PFAM: protein of unknown function DUF1385 KEGG: tte:TTE0141 hypothetical protein
2908769	C	T	HIGH	0.975	coding	0	1	-0.99	0.84	0.687	1979	856	NON	R	K	164	Cthe_2436	2908696	2908932	-1	hypothetical protein	
2917990	A	G	HIGH	0.975	coding	0	1	10004.43	0.972	0.99	5555	6496	SYN	S	S	1062	Cthe_2446	2916929	2918047	1	ABC-type sugar transport system periplasmic component-like protein	KEGG: sma:SAV5318 putative D-ribose ABC transporter permease protein
2920132	T	C	HIGH	0.995	coding	0	1	10003.47	0.735	0.659	1175	1328	NON	L	P	578	Cthe_2448	2919555	2920538	1	inner-membrane translocator	PFAM: inner-membrane translocator KEGG: oih:OB2573 ribose ABC transporter permease
2943643	G	A	HIGH	0.985	coding	0	1	10000.65	0.527	0.497	787	3824	NON	G	D	998	Cthe_2471	2942646	2943926	1	conserved hypothetical protein	KEGG: ctc:CTC00388 hypothetical protein
2947995	A	G	HIGH	0.995	coding	0	1	0.87	0.667	0.903	1298	4187	NON	S	G	1081	Cthe_2475	2946915	2948348	1	phage portal protein, SPP1 family	TIGRFAM: phage portal protein, SPP1 family PFAM: phage portal protein, SPP1 KEGG: sha:SH2357 hypothetical protein
2971415	T	C	HIGH	0.99	coding	0	1	0.27	0.891	0.988	1852	10657	NON	D	G	74	Cthe_2504	2970199	2971488	-1	conserved hypothetical protein	PFAM: 3D KEGG: ctc:CTC00245 hypothetical membrane associated protein
2975548	A	G	HIGH	0.98	coding	0	1	0.1	0.591	0.995	1390	40231	NON	I	M	2502	Cthe_2506	2973047	2976088	1	S-layer-like domain containing protein	PFAM: S-layer-like region; type 3a, cellulose-binding SMART: LamG-like jellyroll fold KEGG: bps:BPSS0105 hypothetical protein
2995953	G	A	HIGH	0.995	coding	0	1	10001.29	0.985	0.992	21492	32815	NON	S	N	707	Cthe_2527	2995247	2996125	1	hydroxymethylbilane synthase	KEGG: mca:MCA3060 porphobilinogen deaminase TIGRFAM: porphobilinogen deaminase PFAM: Porphobilinogen deaminase

2997878	C	T	HIGH	1	coding	0	1	10008.31	0.863	0.99	3616	10219	SYN	L	L	4	Cthe_2529	2997875	2998855	1	porphobilinogen synthase	PFAM: delta-aminolevulinic acid dehydratase KEGG: fnu:FN0460 delta-aminolevulinic acid dehydratase
3001848	G	A	HIGH	0.99	coding	0	1	10002.28	0.953	0.899	6661	7516	NON	D	N	307	Cthe_2532	3001542	3002390	1	sulfate ABC transporter, inner membrane subunit CysT	TIGRFAM: sulfate ABC transporter, inner membrane subunit; sulfate ABC transporter, inner membrane subunit CysT PFAM: binding-protein-dependent transport systems inner membrane component KEGG: dsy:DSY2957 hypothetical protein
3004674	C	T	HIGH	0.99	coding	0	1	10001.58	0.836	0.931	1996	2836	SYN	L	L	139	Cthe_2535	3004536	3005255	1	phosphoadenylylsulfate reductase (thioredoxin)	KEGG: dsy:DSY2953 phosphoadenosine phosphosulfate reductase TIGRFAM: adenylylsulfate reductase, thioredoxin dependent PFAM: phosphoadenosine phosphosulfate reductase
3029890	A	G	HIGH	0.985	coding	0	1	10000.48	0.963	0.981	4563	8589	NON	D	G	146	Cthe_2559	3029745	3031118	1	DegT/DnrJ/EryC1/StrS aminotransferase	PFAM: DegT/DnrJ/EryC1/StrS aminotransferase KEGG: pca:Pcar_1130 pyridoxal phosphate-dependent enzyme apparently involved in regulation of cell wall biogenesis-like
3031364	T	C	HIGH	0.995	coding	0	1	10001.21	0.906	0.983	2723	8227	SYN	A	A	114	Cthe_2560	3031251	3031799	1	dTDP-4-dehydrorhamnose 3,5-epimerase	KEGG: pca:Pcar_1129 dTDP-4-dehydrorhamnose 3,5-epimerase TIGRFAM: dTDP-4-dehydrorhamnose 3,5-epimerase PFAM: dTDP-4-dehydrorhamnose 3,5-epimerase related
3031718	G	A	HIGH	0.985	coding	0	1	1.15	0.789	0.805	2371	3339	NON	W	*	468	Cthe_2560	3031251	3031799	1	dTDP-4-dehydrorhamnose 3,5-epimerase	KEGG: pca:Pcar_1129 dTDP-4-dehydrorhamnose 3,5-epimerase TIGRFAM: dTDP-4-dehydrorhamnose 3,5-epimerase PFAM: dTDP-4-dehydrorhamnose 3,5-epimerase related
3086217	A	G	HIGH	1	coding	0	1	10002.19	0.989	0.967	2643	1263	NON	N	D	4009	Cthe_2611	3082209	3087095	1	Fibronectin, type III	SMART: Fibronectin, type III KEGG: tte:TTE2615 subtilisin-like serine protease
3087899	G	A	HIGH	0.96	coding	0	1	10003.4	0.991	0.986	4054	5279	SYN	L	L	795	Cthe_2612	3087105	3090941	1	Fibronectin, type III	SMART: Fibronectin, type III
3090411	G	A	HIGH	0.97	coding	0	1	1.28	0.939	0.64	4580	2728	NON	D	N	3307	Cthe_2612	3087105	3090941	1	Fibronectin, type III	SMART: Fibronectin, type III
3116739	C	T	HIGH	0.99	coding	0	1	0.4	0.528	0.481	889	2121	SYN	L	L	240	Cthe_2635	3116500	3116934	1	pyridoxamine 5'-phosphate oxidase-related, FMN-binding	PFAM: pyridoxamine 5'-phosphate oxidase-related, FMN-binding KEGG: mba:Mbar_A3327 hypothetical protein
3131416	C	T	HIGH	0.985	coding	0	1	0.24	0.596	0.719	1410	2675	SYN	Q	Q	117	Cthe_2649	3130705	3131532	-1	HpcH/HpaI aldolase	PFAM: HpcH/HpaI aldolase
3132469	G	A	HIGH	0.955	coding	0	1	10003.86	0.986	0.977	2774	2594	NON	P	L	1103	Cthe_2650	3131748	3133571	-1	polysaccharide biosynthesis protein CapD	PFAM: NAD-dependent epimerase/dehydratase; short-chain dehydrogenase/reductase SDR; 3-beta hydroxysteroid dehydrogenase/isomerase; polysaccharide biosynthesis protein CapD; dTDP-4-dehydrorhamnose reductase; Male sterility-like KEGG: ctc:CTC02268 CapD prote
3147607	A	G	HIGH	1	coding	0	1	10002.48	0.862	0.977	2590	3554	SYN	G	G	564	Cthe_2666	3145576	3148170	-1	ATPase, P-type (transporting), HAD superfamily, subfamily IC	TIGRFAM: ATPase, P-type (transporting), HAD superfamily, subfamily IC PFAM: cation transporting ATPase-like; Haloacid dehalogenase-like hydrolase; E1-E2 ATPase-associated region KEGG: efa:EF1938 cation-transporting ATPase, E1-E2 family
3193128	G	A	HIGH	0.965	coding	0	1	10002.1	0.865	0.82	4688	3715	SYN	Q	Q	333	Cthe_2706	3192796	3193905	1	ABC transporter related protein	PFAM: ABC transporter related SMART: ATPase KEGG: tte:TTE2145 ABC-type multidrug transport system, ATPase component
3223060	C	T	HIGH	0.975	coding	0	1	10001.83	0.712	0.92	4882	8130	NON	A	V	530	Cthe_2730	3222531	3223733	1	translation elongation factor 1A (EF-1A/EF-Tu)	TIGRFAM: translation elongation factor Tu; small GTP-binding protein PFAM: protein synthesis factor, GTP-binding; elongation factor Tu-like; elongation factor Tu, domain 2 KEGG: gka:GK0104 translation elongation factor Tu (EF-Tu)
3227821	G	A	HIGH	0.99	coding	0	1	10003.16	0.99	0.923	8771	6377	NON	V	I	772	Cthe_2736	3227050	3228756	1	phosphoenolpyruvate--protein phosphotransferase	TIGRFAM: phosphoenolpyruvate-protein phosphotransferase PFAM: PEP-utilizing enzyme; PEP-utilising enzyme, mobile region; PEP-utilising enzyme-like KEGG: tte:TTE2334 phosphoenolpyruvate-protein kinase (PTS system EI component in bacteria)
3240362	G	A	HIGH	0.95	coding	0	1	-0.24	0.843	0.691	4778	4366	NON	R	Q	125	Cthe_2746	3240238	3240708	1	protein of unknown function DUF402	PFAM: protein of unknown function DUF402 KEGG: pfu:PF1626 hypothetical protein
3304230	A	G	HIGH	0.98	coding	0	1	0.2	0.983	0.984	4560	8823	SYN	V	V	27	Cthe_2803	3304204	3304983	1	binding-protein-dependent transport systems inner membrane component	PFAM: binding-protein-dependent transport systems inner membrane component KEGG: mta:Moth_1974 binding-protein-dependent transport systems inner membrane component
3304557	T	C	HIGH	1	coding	0	1	10001.4	0.862	0.992	3381	14562	SYN	F	F	354	Cthe_2803	3304204	3304983	1	binding-protein-dependent transport systems inner membrane component	PFAM: binding-protein-dependent transport systems inner membrane component KEGG: mta:Moth_1974 binding-protein-dependent transport systems inner membrane component
3309731	C	A	HIGH	0.955	coding	0	1	10000.86	0.774	0.988	2814	15413	SYN	G	G	186	Cthe_2808	3309546	3310583	1	transcriptional regulator, LacI family	PFAM: regulatory protein, LacI; periplasmic binding protein/LacI transcriptional regulator KEGG: tte:TTE0349 transcriptional regulator
3324904	G	A	HIGH	0.97	coding	0	1	10002.81	0.931	0.969	5576	5147	NON	P	L	749	Cthe_2815	3324051	3325652	-1	lysyl-tRNA synthetase	KEGG: bce:BC2535 class 1 lysyl-tRNA synthetase TIGRFAM: lysyl-tRNA synthetase PFAM: Lysyl-tRNA synthetase, class 1c
3351913	T	G	HIGH	0.98	coding	0	1	10.84	0.436	0.969	1654	7911	NON	F	C	44	Cthe_2835	3351870	3352448	1	hypothetical protein	
3352973	T	C	MEDIUM	0.985	coding	0	0	0.34	0.499	0.329	2649	1514	NON	V	A	344	Cthe_2836	3352630	3353019	1	hypothetical protein	
3353012	A	G	LOW	0.93	coding	0	0	-0.99	0.822	0.768	708	1413	NON	D	G	383	Cthe_2836	3352630	3353019	1	hypothetical protein	
3353550	T	C	MEDIUM	0.99	coding	6.58	0	1.4	0.971	0.985	6697	13726	SYN	G	G	21	Cthe_2837	3353530	3354081	1	prophage Lp4 protein 7, DNA replication	KEGG: lpl:lp_3383 prophage Lp4 protein 7, DNA replication
3353578	T	G	LOW	0.99	coding	6.86	0	10005.99	0.531	0.651	2235	1790	NON	S	A	49	Cthe_2837	3353530	3354081	1	prophage Lp4 protein 7, DNA replication	KEGG: lpl:lp_3383 prophage Lp4 protein 7, DNA replication
3353592	G	A	LOW	0.99	coding	6.58	0	10005.99	0.376	0.392	951	1143	SYN	K	K	63	Cthe_2837	3353530	3354081	1	prophage Lp4 protein 7, DNA replication	KEGG: lpl:lp_3383 prophage Lp4 protein 7, DNA replication
3353670	T	C	MEDIUM	0.96	coding	0	0	0.97	0.375	0.548	384	373	SYN	I	I	141	Cthe_2837	3353530	3354081	1	prophage Lp4 protein 7, DNA replication	KEGG: lpl:lp_3383 prophage Lp4 protein 7, DNA replication
3353692	C	T	MEDIUM	0.985	coding	4.61	0	0.97	0.873	0.97	3447	4384	NON	P	S	163	Cthe_2837	3353530	3354081	1	prophage Lp4 protein 7, DNA replication	KEGG: lpl:lp_3383 prophage Lp4 protein 7, DNA replication
3353756	A	G	LOW	0.975	coding	9.86	0	-0.91	0.879	0.824	4130	2489	NON	D	G	227	Cthe_2837	3353530	3354081	1	prophage Lp4 protein 7, DNA replication	KEGG: lpl:lp_3383 prophage Lp4 protein 7, DNA replication
3353877	C	A	LOW	0.98	coding	9.86	0	-0.53	0.404	0.25	525	678	NON	S	R	348	Cthe_2837	3353530	3354081	1	prophage Lp4 protein 7, DNA replication	KEGG: lpl:lp_3383 prophage Lp4 protein 7, DNA replication
3354749	G	T	MEDIUM	1	coding	24.97	0	0.27	0.993	0.988	4721	12714	SYN	V	V	273	Cthe_2838	3354477	3356318	1	conserved hypothetical protein precursor	TIGRFAM: Protein splicing (intein) site SMART: Hedgehog/intein hint-like KEGG: bce:BC3021 hypothetical cytosolic protein
3355085	A	C	LOW	0.985	coding	2.29	0	-0.13	0.647	0.762	1804	8484	SYN	G	G	609	Cthe_2838	3354477	3356318	1	conserved hypothetical protein precursor	TIGRFAM: Protein splicing (intein) site SMART: Hedgehog/intein hint-like KEGG: bce:BC3021 hypothetical cytosolic protein
3355091	G	A	LOW	0.98	coding	0	0	-0.13	0.328	0.671	440	2359	SYN	L	L	615	Cthe_2838	3354477	3356318	1	conserved hypothetical protein precursor	TIGRFAM: Protein splicing (intein) site SMART: Hedgehog/intein hint-like KEGG: bce:BC3021 hypothetical cytosolic protein
3355100	G	A	LOW	0.935	coding	0	0	-0.13	0.734	0.949	453	4012	SYN	A	A	624	Cthe_2838	3354477	3356318	1	conserved hypothetical protein precursor	TIGRFAM: Protein splicing (intein) site SMART: Hedgehog/intein hint-like KEGG: bce:BC3021 hypothetical cytosolic protein
3355577	C	T	LOW	0.955	coding	3.29	0	-1	0.766	0.904	871	3838	SYN	I	I	1101	Cthe_2838	3354477	3356318	1	conserved hypothetical protein precursor	TIGRFAM: Protein splicing (intein) site SMART: Hedgehog/intein hint-like KEGG: bce:BC3021 hypothetical cytosolic protein
3355583	C	A	LOW	0.96	coding	3.29	0	-1	0.873	0.946	10694	2771	SYN	G	G	1107	Cthe_2838	3354477	3356318	1	conserved hypothetical protein precursor	TIGRFAM: Protein splicing (intein) site SMART: Hedgehog/intein hint-like KEGG: bce:BC3021 hypothetical cytosolic protein
3355670	A	G	LOW	1	coding	4.58	0	10002.21	0.559	0.728	1601	2108	SYN	K	K	1194	Cthe_2838	3354477	3356318	1	conserved hypothetical protein precursor	TIGRFAM: Protein splicing (intein) site SMART: Hedgehog/intein hint-like KEGG: bce:BC3021 hypothetical cytosolic protein
3355683	A	G	LOW	0.95	coding	0.98	0	10002.21	0.496	0.45	361	754	NON	R	G	1207	Cthe_2838	3354477	3356318	1	conserved hypothetical protein precursor	TIGRFAM: Protein splicing (intein) site SMART: Hedgehog/intein hint-like KEGG: bce:BC3021 hypothetical cytosolic protein
3355719	G	A	MEDIUM	0.98	coding	1.95	0	1.49	0.487	0.303	1881	457	NON	E	K	1243	Cthe_2838	3354477	3356318	1	conserved hypothetical protein precursor	TIGRFAM: Protein splicing (intein) site SMART: Hedgehog/intein hint-like KEGG: bce:BC3021 hypothetical cytosolic protein
3356114	A	G	LOW	0.99	coding	4.26	0	-0.69	0.989	0.983	18434	11879	SYN	K	K	1638	Cthe_2838	3354477	3356318	1	conserved hypothetical protein precursor	TIGRFAM: Protein splicing (intein) site SMART: Hedgehog/intein hint-like KEGG: bce:BC3021 hypothetical cytosolic protein
3356163	A	C	MEDIUM	0.99	coding	2.29	0	0.74	0.593	0.511	2413	1503	NON	K	Q	1687	Cthe_2838	3354477	3356318	1	conserved hypothetical protein precursor	TIGRFAM: Protein splicing (intein) site SMART: Hedgehog/intein hint-like KEGG: bce:BC3021 hypothetical cytosolic protein
3356226	A	G	MEDIUM	0.99	coding	0	0	10001.27	0.391	0.432	550	2330	NON	I	V	1750	Cthe_2838	3354477	3356318	1	conserved hypothetical protein precursor	TIGRFAM: Protein splicing (intein) site SMART: Hedgehog/intein hint-like KEGG: bce:BC3021 hypothetical cytosolic protein
3356246	A	G	MEDIUM	0.99	coding	3.29	0	10001.27	0.955	0.928	3578	4630	SYN	R	R	1770	Cthe_2838	3354477	3356318	1	conserved hypothetical protein precursor	TIGRFAM: Protein splicing (intein) site SMART: Hedgehog/intein hint-like KEGG: bce:BC3021 hypothetical cytosolic protein
3356255	A	G	LOW	0.99	coding	2.93	0	10008.09	0.527	0.367	1901	688	SYN	K	K	1779	Cthe_2838	3354477	3356318	1	conserved hypothetical protein precursor	TIGRFAM: Protein splicing (intein) site SMART: Hedgehog/intein hint-like KEGG: bce:BC3021 hypothetical cytosolic protein
3356650	T	C	LOW	0.985	coding	9.86	1	-1	0.985	0.986	6510	13768	SYN	D	D	306	Cthe_2839	3356345	3357583	1	conserved hypothetical protein	KEGG: bce:BC3020 hypothetical protein
3358094	C	A	MEDIUM	0.93	coding	0	0	0.08	0.348	0.474	387	883	NON	S	Y	8	Cthe_2840	3358087	3358431	1	conserved hypothetical protein	KEGG: btk:BT9727_2794 hypothetical protein
3358227	A	G	LOW	0.985	coding	4.26	0	-0.83	0.991	0.982	6878	5189	SYN	K	K	141	Cthe_2840	3358087	3358431	1	conserved hypothetical protein	KEGG: btk:BT9727_2794 hypothetical protein
3358248	T	C	MEDIUM	0.985	coding	3.29	0	10002.03	0.747	0.817	4867	5900	SYN	L	L	162	Cthe_2840	3358087	3358431	1	conserved hypothetical protein	KEGG: btk:BT9727_2794 hypothetical protein
3358395	G	A	LOW	0.995	coding	3.29	0	10003.28	0.811	0.578	2854	2723	SYN	L	L	309	Cthe_2840	3358087	3358431	1	conserved hypothetical protein	KEGG: btk:BT9727_2794 hypothetical protein
3362839	G	A	MEDIUM	0.965	coding	11.82	0	0.77	0.957	0.892	15846	9961	SYN	A	A	18	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3363091	T	C	LOW	0.99	coding	1.95	0	10004.18	0.778	0.708	3284	5602	SYN	G	G	270	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3363104	C	A	LOW	0.995	coding	4.61	0	10004.18	0.613	0.784	2887	11174	NON	L	I	283	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3363118	G	A	LOW	1	coding	4.58	0	10004.18	0.667	0.718	2336	3093	SYN	E	E	297	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3363169	A	C	MEDIUM	0.995	coding	3.29	0	0.15	0.976	0.795	3025	5127	SYN	G	G	348	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3363173	C	T	MEDIUM	1	coding	1.66	0	0.15	0.78	0.564	2934	740	NON	R	W	352	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3363233	C	T	MEDIUM	0.985	coding	4.61	0	0.18	0.438	0.504	1933	3593	NON	P	S	412	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3363246	T	C	MEDIUM	0.98	coding	5.59	0	0.18	0.712	0.827	5278	17497	NON	M	T	425	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3363526	G	A	MEDIUM	0.985	coding	0	0	1.59	0.582	0.523	980	1330	SYN	L	L	705	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3363579	G	A	LOW	0.99	coding	3.29	0	-0.89	0.919	0.392	4469	1482	NON	R	K	758	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3363655	G	T	MEDIUM	0.97	coding	6.58	1	1.21	0.988	0.857	6984	4627	SYN	G	G	834	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3363677	A	C	LOW	0.995	coding	0	0	10009.14	0.705	0.458	1896	825	SYN	R	R	856	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3363694	T	A	LOW	0.99	coding	0	0	10009.14	0.395	0.343	784	424	SYN	I	I	873	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3363944	A	C	MEDIUM	1	coding	3.29	0	0.59	0.994	0.967	7960	14405	NON	I	L	1123	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3363973	T	C	LOW	0.935	coding	0	0	10008.25	0.533	0.328	353	622	SYN	C	C	1152	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3363975	G	A	LOW	0.905	coding	0.24	0	10008.25	0.473	0.363	499	342	NON	R	K	1154	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3364013	C	T	MEDIUM	0.985	coding	3.29	1	10001.4	0.68	0.825	1848	4234	NON	P	S	1192	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3364159	G	A	MEDIUM	0.93	coding	0	0	0.55	0.341	0.409	349	379	SYN	S	S	1338	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3364222	A	C	LOW	0.985	coding	3.29	0	10003.58	0.981	0.927	19363	7636	SYN	G	G	1401	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3364244	A	G	LOW	1	coding	3.29	0	10003.39	0.823	0.944	1548	1263	NON	I	V	1423	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3364249	T	C	MEDIUM	0.995	coding	3.29	0	10001.45	0.521	0.732	869	668	SYN	S	S	1428	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3364706	A	G	MEDIUM	0.995	coding	0	0	0.32	0.405	0.306	1118	460	NON	K	E	1885	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3364942	T	C	LOW	0.985	coding	4.26	0	10007.29	0.964	0.986	5181	7337	SYN	G	G	2121	Cthe_2847	3362822	3364975	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3365321	G	T	LOW	0.935	coding	3.29	0	10005.89	0.973	0.912	2090	2040	SYN	A	A	75	Cthe_2848	3365247	3366560	1	phage major capsid protein, HK97	PFAM: phage major capsid protein, HK97 KEGG: det:DET1085 major capsid protein, HK97 family
3365330	A	G	LOW	0.995	coding	3.29	0	10011.27	0.984	0.984	6772	12786	SYN	A	A	84	Cthe_2848	3365247	3366560	1	phage major capsid protein, HK97	PFAM: phage major capsid protein, HK97 KEGG: det:DET1085 major capsid protein, HK97 family
3365345	T	A	LOW	0.995	coding	2.29	0	10011.27	0.866	0.622	4278	6715	SYN	T	T	99	Cthe_2848	3365247	3366560	1	phage major capsid protein, HK97	PFAM: phage major capsid protein, HK97 KEGG: det:DET1085 major capsid protein, HK97 family
3365359	A	G	LOW	0.99	coding	0.79	0	10005.3	0.617	0.522	2475	1753	NON	Q	R	113	Cthe_2848	3365247	3366560	1	phage major capsid protein, HK97	PFAM: phage major capsid protein, HK97 KEGG: det:DET1085 major capsid protein, HK97 family
3365448	T	C	LOW	1	coding	0.55	0	10008.41	0.514	0.303	1139	602	SYN	L	L	202	Cthe_2848	3365247	3366560	1	phage major capsid protein, HK97	PFAM: phage major capsid protein, HK97 KEGG: det:DET1085 major capsid protein, HK97 family
3365771	T	C	MEDIUM	0.935	coding	3.29	1	0.24	0.999	0.989	9823	3211	SYN	S	S	525	Cthe_2848	3365247	3366560	1	phage major capsid protein, HK97	PFAM: phage major capsid protein, HK97 KEGG: det:DET1085 major capsid protein, HK97 family
3365787	T	C	MEDIUM	0.98	coding	0.55	0	0.24	0.753	0.438	2913	1235	NON	S	P	541	Cthe_2848	3365247	3366560	1	phage major capsid protein, HK97	PFAM: phage major capsid protein, HK97 KEGG: det:DET1085 major capsid protein, HK97 family
3365795	T	C	MEDIUM	0.995	coding	0	0	0.24	0.427	0.287	489	771	SYN	F	F	549	Cthe_2848	3365247	3366560	1	phage major capsid protein, HK97	PFAM: phage major capsid protein, HK97 KEGG: det:DET1085 major capsid protein, HK97 family
3365960	T	G	MEDIUM	0.98	coding	0.98	0	0.03	0.797	0.936	433	595	SYN	G	G	714	Cthe_2848	3365247	3366560	1	phage major capsid protein, HK97	PFAM: phage major capsid protein, HK97 KEGG: det:DET1085 major capsid protein, HK97 family
3365972	G	A	MEDIUM	0.99	coding	0.98	0	0.03	0.362	0.407	451	1658	SYN	E	E	726	Cthe_2848	3365247	3366560	1	phage major capsid protein, HK97	PFAM: phage major capsid protein, HK97 KEGG: det:DET1085 major capsid protein, HK97 family
3366080	A	G	LOW	0.985	coding	4.83	0	10002.81	0.96	0.916	7347	3731	SYN	A	A	834	Cthe_2848	3365247	3366560	1	phage major capsid protein, HK97	PFAM: phage major capsid protein, HK97 KEGG: det:DET1085 major capsid protein, HK97 family
3366119	C	T	MEDIUM	0.975	coding	3.29	0	10001.53	0.741	0.637	2629	671	SYN	A	A	873	Cthe_2848	3365247	3366560	1	phage major capsid protein, HK97	PFAM: phage major capsid protein, HK97 KEGG: det:DET1085 major capsid protein, HK97 family
3366152	C	T	LOW	0.945	coding	3.29	0	-0.51	0.989	0.986	7427	13018	SYN	F	F	906	Cthe_2848	3365247	3366560	1	phage major capsid protein, HK97	PFAM: phage major capsid protein, HK97 KEGG: det:DET1085 major capsid protein, HK97 family
3366170	A	C	LOW	1	coding	3.29	0	-0.51	0.975	0.742	7230	6345	SYN	S	S	924	Cthe_2848	3365247	3366560	1	phage major capsid protein, HK97	PFAM: phage major capsid protein, HK97 KEGG: det:DET1085 major capsid protein, HK97 family
3366899	T	C	MEDIUM	0.975	coding	3.29	0	1.02	0.44	0.561	771	1240	SYN	S	S	81	Cthe_2849	3366819	3367121	1	hypothetical protein	
3366911	C	A	MEDIUM	0.945	coding	3.29	0	1.02	0.682	0.968	1511	2264	SYN	A	A	93	Cthe_2849	3366819	3367121	1	hypothetical protein	
3367733	A	C	LOW	0.99	coding	0	0	-0.21	0.908	0.512	6194	3072	SYN	V	V	285	Cthe_2850	3367449	3368705	1	transposase IS116/IS110/IS902	PFAM: transposase, IS111A/IS1328/IS1533; transposase IS116/IS110/IS902 KEGG: btk:BT9727_3066 transposase
3367751	A	C	LOW	0.985	coding	0	0	-0.22	0.549	0.299	867	755	SYN	A	A	303	Cthe_2850	3367449	3368705	1	transposase IS116/IS110/IS902	PFAM: transposase, IS111A/IS1328/IS1533; transposase IS116/IS110/IS902 KEGG: btk:BT9727_3066 transposase
3367851	C	T	LOW	0.995	coding	3.29	0	10006.39	0.913	0.98	3256	7483	SYN	L	L	403	Cthe_2850	3367449	3368705	1	transposase IS116/IS110/IS902	PFAM: transposase, IS111A/IS1328/IS1533; transposase IS116/IS110/IS902 KEGG: btk:BT9727_3066 transposase
3367854	C	A	LOW	0.95	coding	0.24	0	10006.39	0.92	0.985	467	958	NON	Q	K	406	Cthe_2850	3367449	3368705	1	transposase IS116/IS110/IS902	PFAM: transposase, IS111A/IS1328/IS1533; transposase IS116/IS110/IS902 KEGG: btk:BT9727_3066 transposase
3367925	A	G	MEDIUM	1	coding	0	0	0.27	0.561	0.737	1286	2621	SYN	K	K	477	Cthe_2850	3367449	3368705	1	transposase IS116/IS110/IS902	PFAM: transposase, IS111A/IS1328/IS1533; transposase IS116/IS110/IS902 KEGG: btk:BT9727_3066 transposase
3367928	T	C	MEDIUM	0.995	coding	0	0	0.27	0.584	0.789	2198	7931	SYN	S	S	480	Cthe_2850	3367449	3368705	1	transposase IS116/IS110/IS902	PFAM: transposase, IS111A/IS1328/IS1533; transposase IS116/IS110/IS902 KEGG: btk:BT9727_3066 transposase
3368231	A	G	MEDIUM	0.99	coding	1.95	0	0.03	0.443	0.441	495	687	SYN	R	R	783	Cthe_2850	3367449	3368705	1	transposase IS116/IS110/IS902	PFAM: transposase, IS111A/IS1328/IS1533; transposase IS116/IS110/IS902 KEGG: btk:BT9727_3066 transposase
3368372	T	C	MEDIUM	0.98	coding	0	0	0.72	0.471	0.286	851	381	SYN	R	R	924	Cthe_2850	3367449	3368705	1	transposase IS116/IS110/IS902	PFAM: transposase, IS111A/IS1328/IS1533; transposase IS116/IS110/IS902 KEGG: btk:BT9727_3066 transposase
3368662	A	G	MEDIUM	0.99	coding	0	0	10001.95	0.443	0.377	539	572	NON	K	R	1214	Cthe_2850	3367449	3368705	1	transposase IS116/IS110/IS902	PFAM: transposase, IS111A/IS1328/IS1533; transposase IS116/IS110/IS902 KEGG: btk:BT9727_3066 transposase
3368666	T	G	LOW	0.99	coding	0	0	10002.58	0.439	0.383	581	427	SYN	T	T	1218	Cthe_2850	3367449	3368705	1	transposase IS116/IS110/IS902	PFAM: transposase, IS111A/IS1328/IS1533; transposase IS116/IS110/IS902 KEGG: btk:BT9727_3066 transposase
3368688	A	G	LOW	0.99	coding	0	0	10004.84	0.355	0.406	951	1524	NON	T	A	1240	Cthe_2850	3367449	3368705	1	transposase IS116/IS110/IS902	PFAM: transposase, IS111A/IS1328/IS1533; transposase IS116/IS110/IS902 KEGG: btk:BT9727_3066 transposase
3370211	G	A	MEDIUM	0.99	coding	9.5	0	0.05	0.989	0.989	9839	2416	SYN	A	A	18	Cthe_2853	3370194	3370994	1	prophage Lp4 protein 7, DNA replication	KEGG: lpl:lp_3383 prophage Lp4 protein 7, DNA replication
3370223	T	C	MEDIUM	0.98	coding	4.26	0	0.05	0.992	0.985	1110	1423	SYN	Y	Y	30	Cthe_2853	3370194	3370994	1	prophage Lp4 protein 7, DNA replication	KEGG: lpl:lp_3383 prophage Lp4 protein 7, DNA replication
3370605	C	T	LOW	0.99	coding	0	0	-0.53	0.826	0.258	478	444	NON	P	S	412	Cthe_2853	3370194	3370994	1	prophage Lp4 protein 7, DNA replication	KEGG: lpl:lp_3383 prophage Lp4 protein 7, DNA replication
3370669	A	G	LOW	0.98	coding	9.86	0	0	0.738	0.682	2925	2650	NON	D	G	476	Cthe_2853	3370194	3370994	1	prophage Lp4 protein 7, DNA replication	KEGG: lpl:lp_3383 prophage Lp4 protein 7, DNA replication
3371662	G	T	MEDIUM	1	coding	24.97	0	1.6	0.991	0.99	6456	17553	SYN	V	V	273	Cthe_2854	3371390	3373237	1	conserved hypothetical protein precursor	TIGRFAM: Protein splicing (intein) site SMART: Hedgehog/intein hint-like KEGG: btk:BT9727_2794 hypothetical protein
3373049	G	A	MEDIUM	1	coding	3.29	1	1.3	0.462	0.822	1792	4838	NON	V	M	1660	Cthe_2854	3371390	3373237	1	conserved hypothetical protein precursor	TIGRFAM: Protein splicing (intein) site SMART: Hedgehog/intein hint-like KEGG: btk:BT9727_2794 hypothetical protein
3373208	A	G	LOW	0.965	coding	0	0	10003.08	0.354	0.382	354	356	NON	N	D	1819	Cthe_2854	3371390	3373237	1	conserved hypothetical protein precursor	TIGRFAM: Protein splicing (intein) site SMART: Hedgehog/intein hint-like KEGG: btk:BT9727_2794 hypothetical protein
3377773	G	A	MEDIUM	1	coding	9.86	0	0.55	0.668	0.644	3349	3969	NON	S	N	74	Cthe_2859	3377700	3378047	1	hypothetical protein	
3383308	A	C	LOW	0.985	coding	0	0	-0.97	0.88	0.79	981	435	NON	E	A	233	Cthe_2865	3383076	3385229	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3383311	A	G	LOW	0.98	coding	0	0	-0.97	0.878	0.875	2714	4169	NON	K	R	236	Cthe_2865	3383076	3385229	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3383317	A	G	LOW	0.98	coding	0.24	0	-0.97	0.923	0.977	1317	1453	NON	D	G	242	Cthe_2865	3383076	3385229	1	phage / plasmid primase, P4 family	TIGRFAM: phage / plasmid primase, P4 family PFAM: Poxvirus D5 protein KEGG: bld:BLi03629 hypothetical protein
3391700	C	T	HIGH	0.955	coding	0	1	-0.27	0.895	0.995	3341	4557	SYN	F	F	114	Cthe_2871	3391587	3392099	1	TipAS antibiotic-recognition	PFAM: TipAS antibiotic-recognition KEGG: dsy:DSY4240 hypothetical protein
3400528	A	G	HIGH	0.99	coding	0	1	0.29	0.835	0.718	2151	1509	NON	Q	R	1505	Cthe_2876	3399024	3401249	1	ATP-dependent DNA helicase PcrA	TIGRFAM: ATP-dependent DNA helicase PcrA PFAM: UvrD/REP helicase KEGG: tte:TTE0604 Superfamily I DNA and RNA helicases
3420674	A	G	HIGH	0.99	coding	0	1	10001.72	0.76	0.576	4340	2266	NON	I	V	271	Cthe_2894	3420404	3420934	1	conserved hypothetical protein	KEGG: tte:TTE1989 hypothetical protein
3440083	C	T	HIGH	0.985	coding	0	1	10001.25	0.888	0.837	3369	2629	NON	A	V	50	Cthe_2917	3440034	3440432	1	SSU ribosomal protein S8P	PFAM: ribosomal protein S8 KEGG: tte:TTE2277 ribosomal protein S8
3445873	G	A	HIGH	0.985	coding	0	1	-0.45	0.505	0.32	920	964	NON	D	N	151	Cthe_2927	3445723	3445941	1	bacterial translation initiation factor 1 (bIF-1)	TIGRFAM: translation initiation factor IF-1 PFAM: RNA binding S1; S1, IF1 type KEGG: cpe:CPE2381 translation initiation factor IF-1
3458321	G	A	HIGH	0.98	coding	0	1	10001.76	0.977	0.986	2473	13014	NON	A	T	625	Cthe_2943	3457697	3458914	1	ABC-2 type transporter	PFAM: ABC-2 type transporter KEGG: tte:TTE0411 ABC-type Na+ efflux pump, permease component
3467210	A	G	HIGH	0.975	coding	0	1	10001.44	0.988	0.948	6822	4101	NON	V	A	1289	Cthe_2949	3466795	3468498	-1	Pectinesterase	PFAM: Pectinesterase; cellulosome enzyme, dockerin type I KEGG: bld:BLi03498 hypothetical protein
3485280	T	G	HIGH	1	coding	0	1	1.34	0.636	0.977	545	5331	NON	F	V	670	Cthe_2966	3484611	3485357	1	hypothetical protein	
3491205	G	T	HIGH	0.97	coding	0	1	10001.68	0.928	0.988	837	2472	NON	P	H	938	Cthe_2972	3490091	3492142	-1	glycoside hydrolase, family 11	PFAM: glycoside hydrolase, family 11; cellulosome enzyme, dockerin type I; polysaccharide deacetylase; Carbohydrate binding family 6 SMART: Cellulose binding, type IV KEGG: sde:Sde_3061 xylanase/chitin deacetylase-like
3491215	C	T	HIGH	0.99	coding	0	1	10002.35	0.89	0.983	8575	19751	NON	A	T	928	Cthe_2972	3490091	3492142	-1	glycoside hydrolase, family 11	PFAM: glycoside hydrolase, family 11; cellulosome enzyme, dockerin type I; polysaccharide deacetylase; Carbohydrate binding family 6 SMART: Cellulose binding, type IV KEGG: sde:Sde_3061 xylanase/chitin deacetylase-like
3494548	C	T	HIGH	0.965	coding	0	1	10002.25	0.493	0.989	2119	8722	SYN	K	K	675	Cthe_2975	3494458	3495222	-1	DNA-directed RNA polymerase sigma factor	PFAM: sigma-70 region 2 KEGG: tte:TTE0118 DNA-directed RNA polymerase specialized sigma subunit
3507977	G	A	HIGH	0.985	coding	0	1	10001.58	0.971	0.596	5992	2547	SYN	K	K	1770	Cthe_2989	3506208	3509162	1	glycosyltransferase 36	PFAM: glycosyltransferase 36 KEGG: tma:TM1848 cellobiose-phosphorylase
3513247	C	T	HIGH	0.965	coding	0	1	1.31	0.603	0.621	1213	1648	NON	Q	*	265	Cthe_2992	3512983	3513507	1	RNA polymerase, sigma-24 subunit, ECF subfamily	PFAM: sigma-70 region 2; Sigma-70, region 4 type 2 KEGG: dsy:DSY2654 hypothetical protein
3526732	C	A	HIGH	0.975	coding	0	1	10001.61	0.987	0.985	22125	17210	NON	C	F	230	Cthe_3003	3525027	3526961	-1	hydrogenase, Fe-only	KEGG: dsy:DSY0936 hypothetical protein TIGRFAM: hydrogenases, Fe-only PFAM: 4Fe-4S ferredoxin, iron-sulfur binding; iron hydrogenase, small subunit; hydrogenase large subunit-like
3526997	A	G	HIGH	0.99	coding	0	1	-0.25	0.906	0.874	5676	2528	NON	*	Q	1480	Cthe_3004	3526995	3528476	-1	ferredoxin	PFAM: ferredoxin; FAD-dependent pyridine nucleotide-disulphide oxidoreductase KEGG: dsy:DSY0936 hypothetical protein
3532242	A	G	HIGH	0.995	coding	0	1	-0.68	0.936	0.721	3797	2076	NON	F	S	347	Cthe_3008	3531917	3532588	-1	Superoxide dismutase	PFAM: manganese and iron superoxide dismutase KEGG: mta:Moth_1916 superoxide dismutase
3537881	G	A	HIGH	0.985	coding	0	1	-0.3	0.72	0.671	1815	2710	NON	A	V	290	Cthe_3013	3537193	3538170	-1	hydrogenase expression/formation protein HypE	TIGRFAM: hydrogenase expression/formation protein HypE PFAM: AIR synthase related protein; AIR synthase related protein-like KEGG: cac:CAC0809 hydrogenase formation factor (hypE)
3573194	C	T	HIGH	0.96	coding	0	1	10002.93	0.847	0.991	2898	8457	SYN	G	G	834	Cthe_3047	3572361	3573773	1	Peptidoglycan glycosyltransferase	PFAM: penicillin-binding protein, transpeptidase KEGG: tte:TTE1967 cell division protein FtsI/penicillin-binding protein 2
3594982	A	G	HIGH	0.995	coding	0	1	10000.69	0.647	0.862	4146	4431	NON	H	R	374	Cthe_3059	3594609	3595157	1	phosphotransferase KptA/Tpt1	PFAM: phosphotransferase KptA/Tpt1 KEGG: btk:BT9727_2777 probable RNA 2'-phosphotransferase
3608159	G	A	HIGH	0.98	coding	0	1	10002.97	0.924	0.912	4246	5064	NON	G	D	500	Cthe_3066	3607660	3608385	1	ABC transporter related protein	PFAM: ABC transporter related SMART: ATPase KEGG: ctc:CTC01493 transporter
3626155	A	G	MEDIUM	0.99	coding	13.15	1	0.11	0.969	0.993	9713	12426	SYN	E	E	927	Cthe_3078	3625229	3632170	1	cellulosome anchoring protein, cohesin region	PFAM: S-layer-like region; cellulosome anchoring protein, cohesin region KEGG: mhu:Mhun_1291 mucin 2, intestinal/tracheal
3626328	T	C	MEDIUM	0.985	coding	6.58	0	1.55	0.726	0.799	2341	17806	NON	V	A	1100	Cthe_3078	3625229	3632170	1	cellulosome anchoring protein, cohesin region	PFAM: S-layer-like region; cellulosome anchoring protein, cohesin region KEGG: mhu:Mhun_1291 mucin 2, intestinal/tracheal
3626384	T	C	MEDIUM	0.97	coding	0	0	1.81	0.969	0.983	3538	754	NON	S	P	1156	Cthe_3078	3625229	3632170	1	cellulosome anchoring protein, cohesin region	PFAM: S-layer-like region; cellulosome anchoring protein, cohesin region KEGG: mhu:Mhun_1291 mucin 2, intestinal/tracheal
3626388	T	C	MEDIUM	0.985	coding	3.29	0	1.81	0.968	0.985	32592	11445	NON	V	A	1160	Cthe_3078	3625229	3632170	1	cellulosome anchoring protein, cohesin region	PFAM: S-layer-like region; cellulosome anchoring protein, cohesin region KEGG: mhu:Mhun_1291 mucin 2, intestinal/tracheal
3626599	C	T	MEDIUM	0.99	coding	9.86	1	0.17	0.991	0.983	36112	31603	SYN	S	S	1371	Cthe_3078	3625229	3632170	1	cellulosome anchoring protein, cohesin region	PFAM: S-layer-like region; cellulosome anchoring protein, cohesin region KEGG: mhu:Mhun_1291 mucin 2, intestinal/tracheal
3626934	T	C	MEDIUM	0.97	coding	6.58	0	1.02	0.744	0.907	5457	19458	NON	V	A	1706	Cthe_3078	3625229	3632170	1	cellulosome anchoring protein, cohesin region	PFAM: S-layer-like region; cellulosome anchoring protein, cohesin region KEGG: mhu:Mhun_1291 mucin 2, intestinal/tracheal
3626990	T	C	MEDIUM	0.955	coding	0	0	1.56	0.974	0.984	7253	686	NON	S	P	1762	Cthe_3078	3625229	3632170	1	cellulosome anchoring protein, cohesin region	PFAM: S-layer-like region; cellulosome anchoring protein, cohesin region KEGG: mhu:Mhun_1291 mucin 2, intestinal/tracheal
3626994	T	C	MEDIUM	0.975	coding	3.29	0	1.56	0.972	0.975	43483	15624	NON	V	A	1766	Cthe_3078	3625229	3632170	1	cellulosome anchoring protein, cohesin region	PFAM: S-layer-like region; cellulosome anchoring protein, cohesin region KEGG: mhu:Mhun_1291 mucin 2, intestinal/tracheal
3627205	C	T	MEDIUM	0.99	coding	9.86	1	0.12	0.986	0.985	24216	36859	SYN	S	S	1977	Cthe_3078	3625229	3632170	1	cellulosome anchoring protein, cohesin region	PFAM: S-layer-like region; cellulosome anchoring protein, cohesin region KEGG: mhu:Mhun_1291 mucin 2, intestinal/tracheal
3630631	A	G	LOW	1	coding	27.89	0	-0.64	0.986	0.993	46945	21533	SYN	P	P	5403	Cthe_3078	3625229	3632170	1	cellulosome anchoring protein, cohesin region	PFAM: S-layer-like region; cellulosome anchoring protein, cohesin region KEGG: mhu:Mhun_1291 mucin 2, intestinal/tracheal
3630655	T	C	LOW	0.975	coding	13.15	0	-0.48	0.996	0.985	44501	19267	SYN	D	D	5427	Cthe_3078	3625229	3632170	1	cellulosome anchoring protein, cohesin region	PFAM: S-layer-like region; cellulosome anchoring protein, cohesin region KEGG: mhu:Mhun_1291 mucin 2, intestinal/tracheal
3639607	C	T	HIGH	0.985	coding	0	1	10000.89	0.91	0.989	3476	6470	SYN	A	A	30	Cthe_3084	3638500	3639636	-1	putative RNA methylase	PFAM: putative RNA methylase; THUMP KEGG: bha:BH1771 hypothetical protein
3653236	T	C	HIGH	0.975	coding	0	1	10006.17	0.889	0.975	4516	10267	NON	T	A	811	Cthe_3096	3653111	3654046	-1	hypothetical protein	KEGG: ddi:DDB0189729 hypothetical protein
3661256	C	T	HIGH	0.99	coding	0	1	10003.6	0.826	0.99	2807	6553	SYN	D	D	297	Cthe_3104	3660960	3661730	1	putative CoA-substrate-specific enzyme activase	TIGRFAM: putative CoA-substrate-specific enzyme activase PFAM: ATPase, BadF/BadG/BcrA/BcrD type KEGG: cac:CAC3623 activator of 2-hydroxyglutaryl-CoA dehydratase (HSP70-class ATPase domain)
3672053	C	T	HIGH	0.975	coding	0	1	10005.38	0.647	0.908	3369	3916	NON	S	L	149	Cthe_3111	3671905	3672384	1	conserved hypothetical protein	KEGG: dsy:DSY4236 hypothetical protein
3676841	C	A	HIGH	0.99	coding	0	1	10004.12	0.676	0.973	3995	5123	SYN	V	V	699	Cthe_3115	3676143	3677156	1	conserved hypothetical protein	KEGG: tte:TTE0067 hypothetical protein
3679049	C	T	HIGH	0.955	coding	0	1	0.33	0.525	0.917	1960	4258	NON	T	M	557	Cthe_3117	3678493	3679248	1	zinc/iron permease	PFAM: zinc/iron permease KEGG: ctc:CTC01320 GufA protein
3710176	G	A	HIGH	0.975	coding	0	1	10002.31	0.803	0.989	4598	8641	NON	A	T	604	Cthe_3141	3709573	3712068	1	lipolytic enzyme, G-D-S-L	PFAM: lipolytic enzyme, G-D-S-L; cellulosome enzyme, dockerin type I; Carbohydrate binding family 6 KEGG: bsu:BG11142 similar to unknown proteins from B. subtilis
3711703	A	G	HIGH	0.995	coding	0	1	1.27	0.678	0.585	1727	2434	NON	K	E	2131	Cthe_3141	3709573	3712068	1	lipolytic enzyme, G-D-S-L	PFAM: lipolytic enzyme, G-D-S-L; cellulosome enzyme, dockerin type I; Carbohydrate binding family 6 KEGG: bsu:BG11142 similar to unknown proteins from B. subtilis
3731348	G	A	HIGH	0.96	coding	0	1	10002.03	0.618	0.57	1056	1491	SYN	V	V	534	Cthe_3158	3730815	3732743	1	aconitase	TIGRFAM: putative aconitate hydratase PFAM: aconitate hydratase-like KEGG: cac:CAC0971 aconitase A
3742942	C	T	HIGH	0.95	coding	0	1	10002.36	0.931	0.988	2592	4538	SYN	V	V	111	Cthe_3167	3742832	3743953	1	glucose-1-phosphate adenylyltransferase, GlgD subunit	KEGG: tma:TM0239 glucose-1-phosphate adenylyltransferase TIGRFAM: glucose-1-phosphate adenylyltransferase, GlgD subunit PFAM: Nucleotidyl transferase
3749144	G	A	HIGH	0.97	coding	0	1	10001.54	0.949	0.989	1628	3958	SYN	V	V	2331	Cthe_3171	3746814	3749261	1	S-layer domain-like protein	KEGG: dsy:DSY0537 hypothetical protein
3752468	A	G	HIGH	0.98	coding	0	1	1.93	0.812	0.933	2857	5525	NON	I	T	53	Cthe_3174	3752131	3752520	-1	conserved hypothetical protein	KEGG: dsy:DSY2102 hypothetical protein
3784408	T	C	HIGH	0.99	coding	0	1	10005.19	0.986	0.962	7429	5094	SYN	F	F	2340	Cthe_3200	3782069	3784711	1	alanyl-tRNA synthetase	KEGG: tte:TTE1248 Alanyl-tRNA synthetase TIGRFAM: alanyl-tRNA synthetase PFAM: alanyl-tRNA synthetase, class IIc; phosphoesterase, DHHA1; Threonyl/alanyl tRNA synthetase, SAD
3834573	A	G	LOW	0.955	coding	3.29	0	-1.03	0.995	0.99	16535	4433	NON	N	S	2042	Cthe_3232	3832532	3837541	1	YD repeat protein	TIGRFAM: YD repeat protein PFAM: YD repeat KEGG: mba:Mbar_A3399 hypothetical protein

intergenic regions

GENOME_POSITION	REF	SNP	CONFIDENCE	PROBABILITY	TYPE	NMER_SCORE	UNIQ	HIGHEST_RANK	FORWARD_PBC_SCORE	REVERSE_PBC_SCORE	FORWARD_INTENSITY	REVERSE_INTENSITY	UPSTREAM_LOCUS_TAG	UPSTREAM_DISTANCE	DOWNSTREAM_LOCUS_TAG	INTERGENIC_POSITION
164619	A	G	HIGH	0.99	intergenic	0	1	-0.9	0.41	0.45	754	2548	Cthe_0134	30	Cthe_0135	21
435560	T	A	HIGH	1	intergenic	0	1	0	0.543	0.564	966	606	Cthe_0345	495	Cthe_0346	36
663719	G	A	HIGH	0.98	intergenic	0	1	-1.01	0.823	0.609	895	402	Cthe_0542	371	Cthe_0543	70
770018	C	T	HIGH	0.96	intergenic	0	1	0.81	0.41	0.944	369	1894	Cthe_0624	182	Cthe_0625	384
1425299	T	A	MEDIUM	0.965	intergenic	6.58	1	2.21	0.989	0.994	1675	1903	Cthe_1195	339	Cthe_1196	152
1686740	T	C	HIGH	0.985	intergenic	0	1	0.39	0.889	0.938	6434	5424	Cthe_1381	95	Cthe_1382	191
1855353	T	C	LOW	0.965	intergenic	3.29	1	-0.84	0.798	0.642	3073	2665	Cthe_1528	721	Cthe_1530	711
2370885	T	C	LOW	0.935	intergenic	7.9	0	-0.97	0.983	0.99	9080	2652	Cthe_1988	1194	Cthe_1990	639
2419590	C	T	HIGH	0.99	intergenic	0	1	10000.57	0.629	0.527	2470	995	Cthe_2036	44	Cthe_2038	1922
2509967	G	A	HIGH	0.975	intergenic	0	1	-0.15	0.989	0.986	5665	4156	Cthe_2111	23	Cthe_2112	8
3082151	A	G	HIGH	1	intergenic	0	1	10002.63	0.978	0.985	10142	8918	Cthe_2610	120	Cthe_2611	59
3353047	C	T	LOW	1	intergenic	0	0	10003.93	0.55	0.529	1224	566	Cthe_2836	29	Cthe_2837	484
3353057	A	G	LOW	0.985	intergenic	0.47	0	10003.93	0.46	0.701	1156	2993	Cthe_2836	39	Cthe_2837	474
3353098	T	C	LOW	0.99	intergenic	3.29	0	-0.46	0.989	0.985	5752	9044	Cthe_2836	80	Cthe_2837	433
3353270	G	A	MEDIUM	0.99	intergenic	0.98	0	5.23	0.667	0.472	620	420	Cthe_2836	252	Cthe_2837	261
3353282	T	A	MEDIUM	1	intergenic	0.98	0	5.23	0.718	0.775	2120	3400	Cthe_2836	264	Cthe_2837	249
3353434	G	A	LOW	1	intergenic	7.61	0	10002.94	0.314	0.292	844	879	Cthe_2836	416	Cthe_2837	97
3353449	A	G	LOW	0.995	intergenic	10.65	0	10002.94	0.431	0.742	1767	2689	Cthe_2836	431	Cthe_2837	82
3353523	T	C	MEDIUM	0.995	intergenic	0	0	10001.43	0.342	0.731	410	2114	Cthe_2836	505	Cthe_2837	8
3356341	G	T	MEDIUM	0.97	intergenic	0	0	0.39	0.528	0.522	673	696	Cthe_2838	24	Cthe_2839	5
3356343	A	G	MEDIUM	0.99	intergenic	0	0	0.39	0.515	0.662	949	493	Cthe_2838	26	Cthe_2839	3
3357965	G	T	LOW	0.96	intergenic	3.29	0	-0.98	0.99	0.476	3144	1572	Cthe_2839	383	Cthe_2840	123
3357970	T	C	LOW	0.985	intergenic	2.29	0	-0.98	0.988	0.696	1427	2460	Cthe_2839	388	Cthe_2840	118
3359220	G	A	MEDIUM	0.985	intergenic	4.61	0	0.99	0.465	0.668	760	672	Cthe_2842	10	Cthe_2843	117
3359255	C	T	MEDIUM	0.97	intergenic	0	0	10001.77	0.561	0.355	460	437	Cthe_2842	45	Cthe_2843	82
3360418	G	T	LOW	0.985	intergenic	32.88	1	-0.79	0.988	0.985	13771	41168	Cthe_2843	775	Cthe_2844	49
3362584	C	T	MEDIUM	0.98	intergenic	0	0	0.91	0.706	0.594	1147	655	Cthe_2846	26	Cthe_2847	239
3362716	A	G	MEDIUM	0.95	intergenic	0.55	0	0.27	0.986	0.711	1419	1253	Cthe_2846	158	Cthe_2847	107
3362819	G	A	MEDIUM	0.995	intergenic	3.29	0	2.46	0.993	0.985	13393	8336	Cthe_2846	261	Cthe_2847	4
3365085	A	C	LOW	0.96	intergenic	0.55	0	10003.78	0.506	0.28	370	390	Cthe_2847	111	Cthe_2848	163
3365158	G	A	LOW	0.975	intergenic	0	0	10003.98	0.485	0.387	906	496	Cthe_2847	184	Cthe_2848	90
3365165	G	A	LOW	0.925	intergenic	0	0	10003.98	0.514	0.41	652	369	Cthe_2847	191	Cthe_2848	83
3365167	A	G	LOW	1	intergenic	0	0	10003.24	0.46	0.337	823	590	Cthe_2847	193	Cthe_2848	81
3366759	T	G	LOW	0.995	intergenic	3.29	0	10004.43	0.522	0.982	968	1950	Cthe_2848	200	Cthe_2849	61
3366779	G	C	LOW	0.96	intergenic	0	0	10004.43	0.395	0.301	451	365	Cthe_2848	220	Cthe_2849	41
3369459	G	T	LOW	0.985	intergenic	0	0	10007.23	0.606	0.524	400	557	Cthe_2851	35	Cthe_2852	183
3369463	G	A	LOW	0.95	intergenic	0	0	10007.23	0.538	0.485	422	350	Cthe_2851	39	Cthe_2852	179
3370183	G	A	LOW	0.94	intergenic	0	0	-0.04	0.666	0.726	584	896	Cthe_2852	247	Cthe_2853	12
3370184	A	G	LOW	0.985	intergenic	0	0	-0.04	0.713	0.801	616	1324	Cthe_2852	248	Cthe_2853	11
3378568	G	T	LOW	0.985	intergenic	32.88	1	-0.2	0.983	0.982	17636	39339	Cthe_2859	522	Cthe_2861	1335
3707682	A	G	HIGH	0.985	intergenic	0	1	0.14	0.975	0.974	6333	3385	Cthe_3138	25	Cthe_3139	400
3786978	T	A	MEDIUM	0.94	intergenic	1.98	0	0.14	0.605	0.98	366	1797	Cthe_3200	2268	Cthe_3201	4491
3787345	T	A	MEDIUM	0.905	intergenic	0.24	0	0.71	0.694	0.986	420	2954	Cthe_3200	2635	Cthe_3201	4124
3788963	T	A	MEDIUM	0.945	intergenic	1.27	0	0.2	0.499	0.977	516	2611	Cthe_3200	4253	Cthe_3201	2506

non-called regions of interest

POSITION	RATIO	SNP	NON_CALLED_ROI	AVE_TEST	AVE_REF	RUN_LENGTH
1083481	4.65465	N	Y	1804	4139	1
1329089	4.49615	N	Y	2457	7484	5
1329097	5.984884	N	Y	2555	10086	5
1329105	5.117881	N	Y	3474	11852	5
1329113	4.730067	N	Y	3750	11730	5
1675233	3.56018	N	Y	2104	5057	3
2160377	3.001224	N	Y	9670	19409	1

2169809	9.715598	N	Y	676	4369	1
2169873	10.28345	N	Y	700	4760	3
2169881	8.996822	N	Y	769	4619	3
2169889	10.88439	N	Y	743	5282	3
2169913	9.961467	N	Y	780	5153	2
2170065	19.27173	N	Y	615	7909	1
2170209	16.72184	N	Y	655	7204	10
2170217	72.93447	N	Y	647	31426	10
2170225	10.88998	N	Y	661	4605	10
2170233	64.65275	N	Y	742	32000	10
2170241	56.56605	N	Y	728	27464	10
2170249	50.59155	N	Y	658	22073	10
2170257	25.55167	N	Y	792	13490	10
2170265	10.43624	N	Y	632	4248	10
2170369	8.688288	N	Y	646	3741	2
2170481	8.375668	N	Y	652	3636	1
2170521	29.08403	N	Y	609	11757	5
2170529	14.84131	N	Y	608	5946	5
2170537	23.77211	N	Y	540	8366	5
2170561	11.02196	N	Y	609	4532	2
2170641	23.53722	N	Y	655	10200	4
2170649	18.32882	N	Y	636	7757	4
2170657	20.60577	N	Y	693	9541	4
2170665	23.43385	N	Y	653	10424	4
2170681	31.05967	N	Y	720	13455	3
2170689	16.65392	N	Y	659	7284	3
2170697	9.959457	N	Y	645	4270	3
2170729	25.68854	N	Y	616	10309	8
2170737	19.18919	N	Y	658	8394	8
2170745	62.04226	N	Y	621	25463	8
2170753	59.21668	N	Y	672	26158	8
2170761	21.03218	N	Y	549	7683	8
2170769	17.05564	N	Y	669	7497	8
2170777	32.41742	N	Y	674	14811	8
2170785	14.63254	N	Y	606	6111	8
2170809	17.66657	N	Y	599	6779	1
2170825	13.03285	N	Y	700	6222	5
2170833	28.53769	N	Y	648	12216	5
2170841	31.8593	N	Y	749	15759	5
2170849	36.91908	N	Y	658	15874	5
2170857	15.05748	N	Y	556	5553	5
2170905	16.98377	N	Y	646	7162	7
2170913	50.87627	N	Y	597	20290	7
2170929	12.40809	N	Y	743	6167	7
2171041	22.3002	N	Y	693	10477	8
2171049	58.62997	N	Y	627	23936	8
2171057	47.72072	N	Y	747	22530	8
2171065	32.69695	N	Y	699	14618	8
2171073	35.27986	N	Y	668	15647	8
2171081	45.58665	N	Y	654	18419	8
2171089	23.94509	N	Y	658	10557	8
2171097	9.133312	N	Y	669	4073	8
2171185	12.13397	N	Y	616	4970	3
2171201	9.4414	N	Y	701	4420	3
2171217	14.59343	N	Y	661	6391	3
2171225	12.17213	N	Y	748	6089	3
2171537	20.47011	N	Y	753	10506	3
2171545	18.40133	N	Y	694	8645	3
2171665	14.48522	N	Y	720	6996	4
2171673	24.2033	N	Y	828	13585	4
2171681	18.70725	N	Y	880	11355	4
2171833	22.78264	N	Y	991	15546	6
2171841	33.26711	N	Y	1714	37452	6
2171849	29.83121	N	Y	1116	22033	6
2171857	32.27158	N	Y	1167	24625	6
2171865	13.78699	N	Y	881	7764	6
2171873	12.15658	N	Y	1045	9086	6
2171961	24.22808	N	Y	967	15454	7
2171969	20.54979	N	Y	1511	20136	7
2171977	14.18353	N	Y	1083	10165	7
2171985	15.12907	N	Y	914	9111	7
2171993	11.53303	N	Y	739	5754	7
2172161	20.46925	N	Y	862	11427	3
2172169	11.99084	N	Y	745	5933	3
2172177	20.36708	N	Y	777	10577	3
2172345	14.92088	N	Y	715	7242	2
2172393	11.04237	N	Y	576	4226	3
2172401	11.91789	N	Y	691	5428	3
2172529	13.24612	N	Y	692	6131	2
2172769	15.47663	N	Y	774	7846	3
2172777	13.92466	N	Y	632	5850	3
2172785	10.64797	N	Y	720	5094	3
2172809	9.110524	N	Y	746	4534	1
2172825	23.1553	N	Y	1119	17088	10
2172833	18.28784	N	Y	858	10291	10
2172841	10.28441	N	Y	783	5361	10
2172849	30.08255	N	Y	838	17087	10
2172857	17.32342	N	Y	843	9746	10
2172865	29.7903	N	Y	866	16870	10
2172873	43.82251	N	Y	1104	33170	10
2172881	31.60136	N	Y	873	17378	10
2172889	14.8368	N	Y	849	8095	10
2173033	12.145	N	Y	957	7489	3
2173393	10.90945	N	Y	1255	8873	1
2173521	9.391908	N	Y	983	6278	3
2173577	9.494155	N	Y	922	5945	1
2174921	22.41699	N	Y	809	11353	1
2174961	9.666076	N	Y	613	3927	1
2175041	28.07293	N	Y	702	13107	3
2175049	21.50342	N	Y	785	11241	3
2175081	10.81938	N	Y	650	4689	1
2175209	10.09006	N	Y	735	4960	2
2175865	10.35486	N	Y	734	4967	4
2175873	12.32421	N	Y	777	6420	4
2175881	24.91574	N	Y	925	15775	4
2175889	28.04487	N	Y	807	15125	4
2175985	11.64151	N	Y	663	5185	1
2176417	16.34477	N	Y	2370	25020	2
2176425	9.109888	N	Y	5258	32423	2
3351529	13.4611	N	Y	2613	17831	10
3351553	17.29159	N	Y	3011	16971	10
3351561	11.6256	N	Y	2328	9721	10
3351569	15.62611	N	Y	1614	12560	10
3351633	9.608298	N	Y	743	3797	7
3351641	9.513971	N	Y	627	3153	7
3351649	11.50774	N	Y	565	3230	7
3351657	24.38932	N	Y	1219	11740	7
3351665	15.34299	N	Y	521	3642	7
3351681	9.821455	N	Y	512	2615	6
3351689	12.41327	N	Y	579	3637	6
3351697	27.20737	N	Y	843	11603	6
3351705	31.87867	N	Y	1417	18171	6
3351713	20.93743	N	Y	634	6217	6
3351721	24.13725	N	Y	764	10793	6
3352017	19.16194	N	Y	413	3877	3
3352025	10.52641	N	Y	373	2063	3
3352033	13.43495	N	Y	426	3181	3
3352281	12.98173	N	Y	465	3099	2
3352289	12.62165	N	Y	442	2839	2
3352329	18.32708	N	Y	708	6670	1
3352353	21.47408	N	Y	637	6395	1
3353393	12.12786	N	Y	1735	10896	6
3353625	9.713994	N	Y	1948	10028	5
3356193	12.4921	N	Y	1318	8446	3
3358337	10.27735	N	Y	1137	6198	2
3359169	9.904162	N	Y	2945	15217	1
3359185	10.38605	N	Y	3128	15691	2
3361449	9.578326	N	Y	2566	12035	6
3361473	10.3509	N	Y	808	4487	2
3361505	18.23812	N	Y	541	5006	11
3361513	16.24131	N	Y	889	6955	11
3361521	15.54013	N	Y	751	5552	11
3361529	9.452328	N	Y	630	3428	11
3361537	10.4136	N	Y	736	4036	11
3361553	17.29115	N	Y	679	6159	11
3361561	41.2036	N	Y	1207	19915	11
3361601	19.688	N	Y	446	4240	7
3361609	25.14331	N	Y	497	5668	7
3361617	40.44778	N	Y	884	18326	7
3361625	19.40042	N	Y	528	5214	7
3361633	21.55608	N	Y	579	7213	7
3361641	37.37577	N	Y	717	13026	7
3361657	14.9416	N	Y	596	4618	3
3361761	24.69814	N	Y	796	9875	1
3361785	19.31248	N	Y	512	4932	2
3361809	14.57735	N	Y	494	3767	2
3361833	21.78641	N	Y	558	6412	2
3361841	26.2002	N	Y	645	7995	2
3361873	10.04596	N	Y	552	2795	4
3361881	12.88106	N	Y	659	4590	4
3361993	12.72459	N	Y	816	4451	1
3366537	8.874688	N	Y	1065	5078	1
3366729	9.905577	N	Y	1227	6366	7
3368793	10.94651	N	Y	1080	5354	2
3368801	35.70805	N	Y	2066	21334	2
3368905	15.88405	N	Y	946	6473	4
3368913	11.59333	N	Y	779	4973	4
3368985	11.51335	N	Y	537	3365	2
3369009	10.85526	N	Y	571	3103	2
3369017	11.93402	N	Y	389	2436	2
3369057	20.93011	N	Y	539	4843	1
3369073	15.56093	N	Y	609	4867	2
3369081	20.5387	N	Y	491	4817	2
3369193	15.40128	N	Y	530	4215	4
3369201	19.17294	N	Y	816	7703	4
3369209	15.55281	N	Y	539	4406	4
3369217	12.30157	N	Y	498	3204	4
3369305	11.66774	N	Y	475	2808	2
3369321	11.17249	N	Y	544	3104	1
3369345	9.19389	N	Y	1167	5788	4

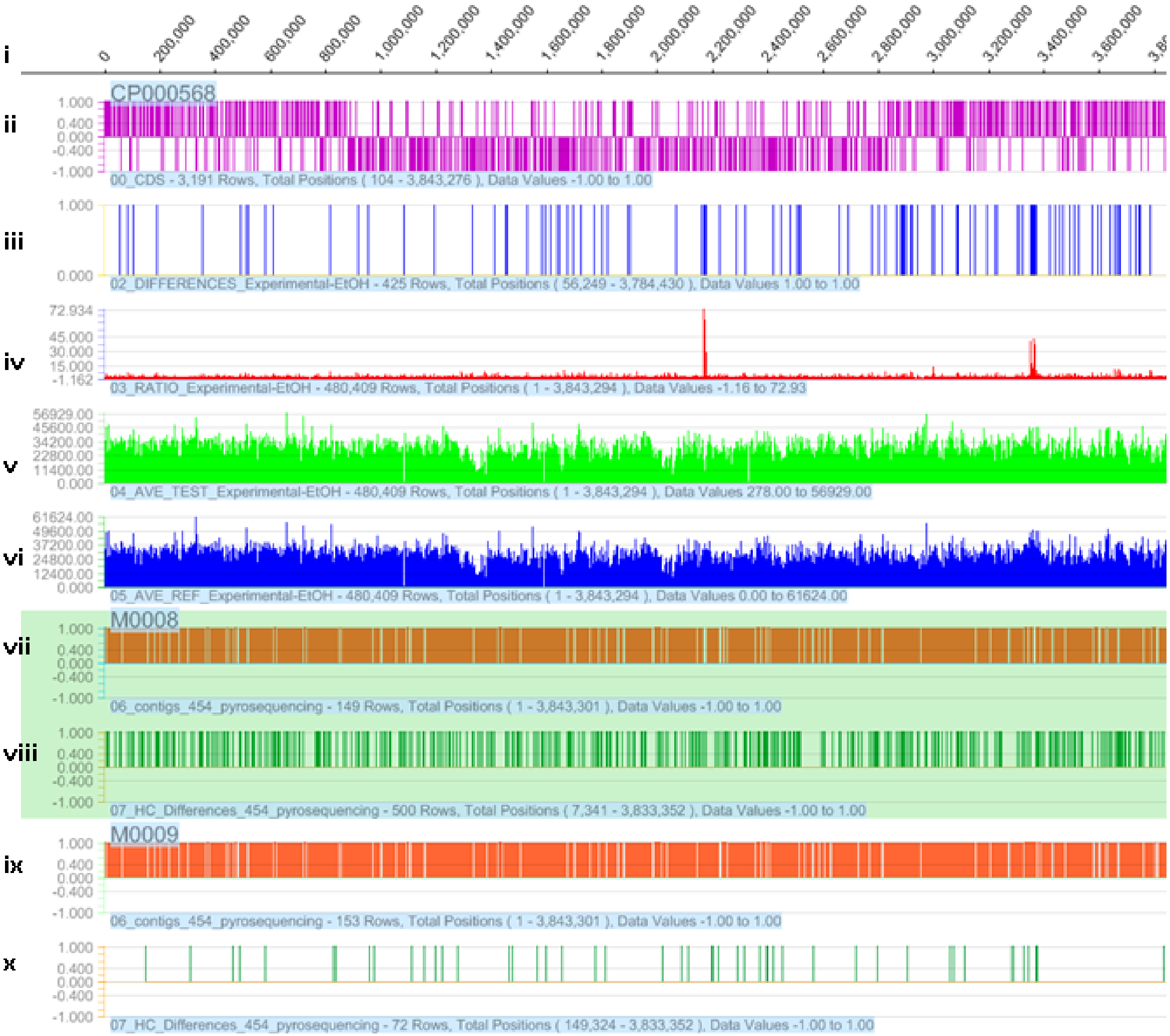
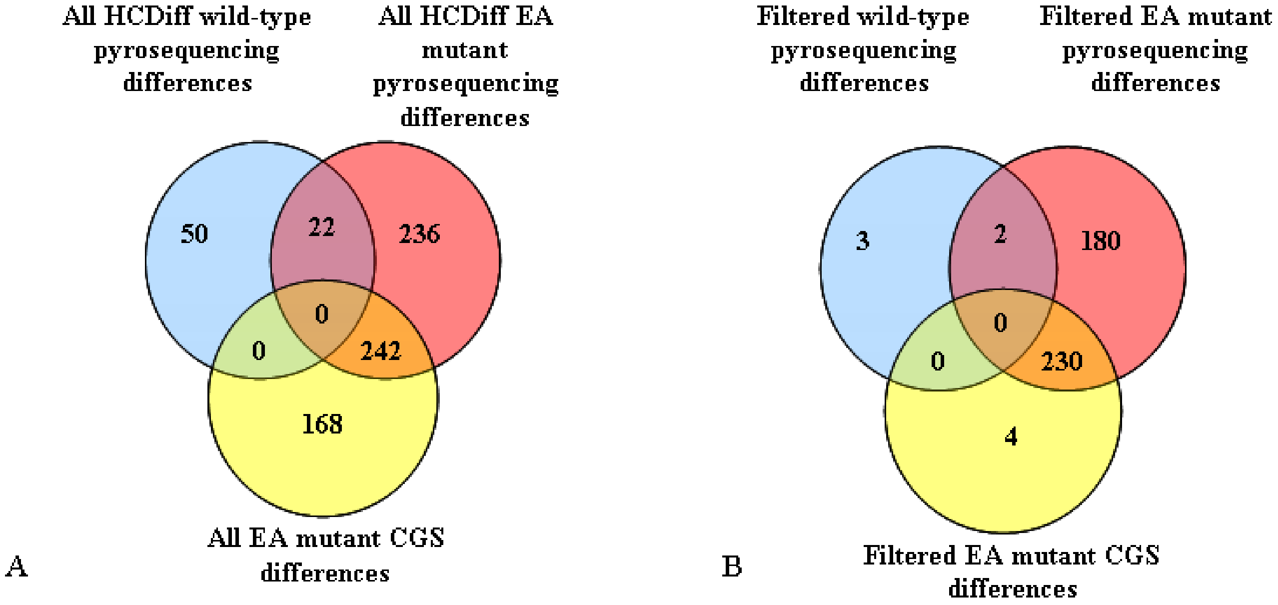
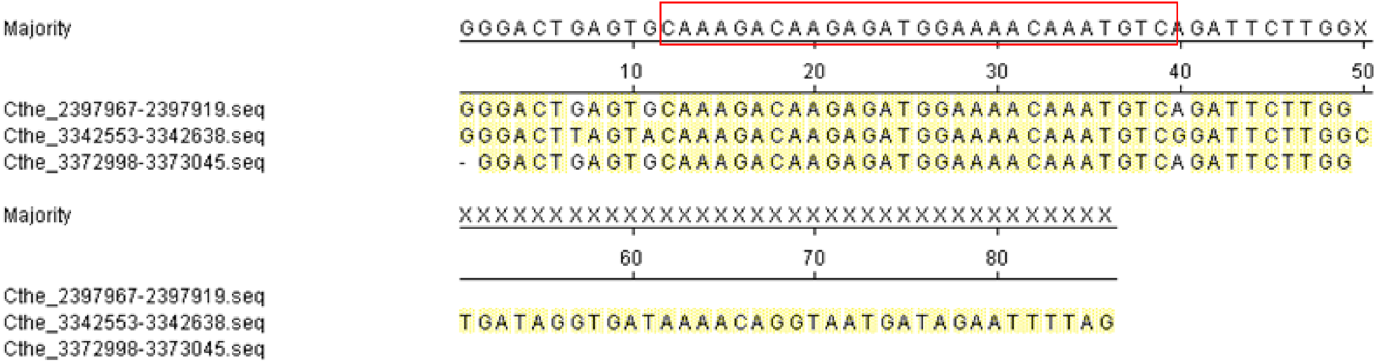


Figure S1. Summary of pyrosequencing and comparative genome sequencing (CGS) resequencing data for *C. thermocellum* wild-type strain 27405 and the EA mutant strain. Genome positions (i) and genes (ii) are indicated. 425 differences for mutant strain EA were identified by CGS (iii), along with two putative regions where DNA was deleted (iv). CGS microarray signal intensities are represented for the mutant (v) and wild-type (vi) strains. Contiguous regions of DNA that mapped to wild-type reference genome from pyrosequencing data for the mutant (153) and strain 27405 (149) are shown, vii and ix respectively. 500 putative high-confidence differences were determined by pryosequencing for mutant strain EA (viii) and 72 putative high-confidence differences were identified for strain 27405 (x).



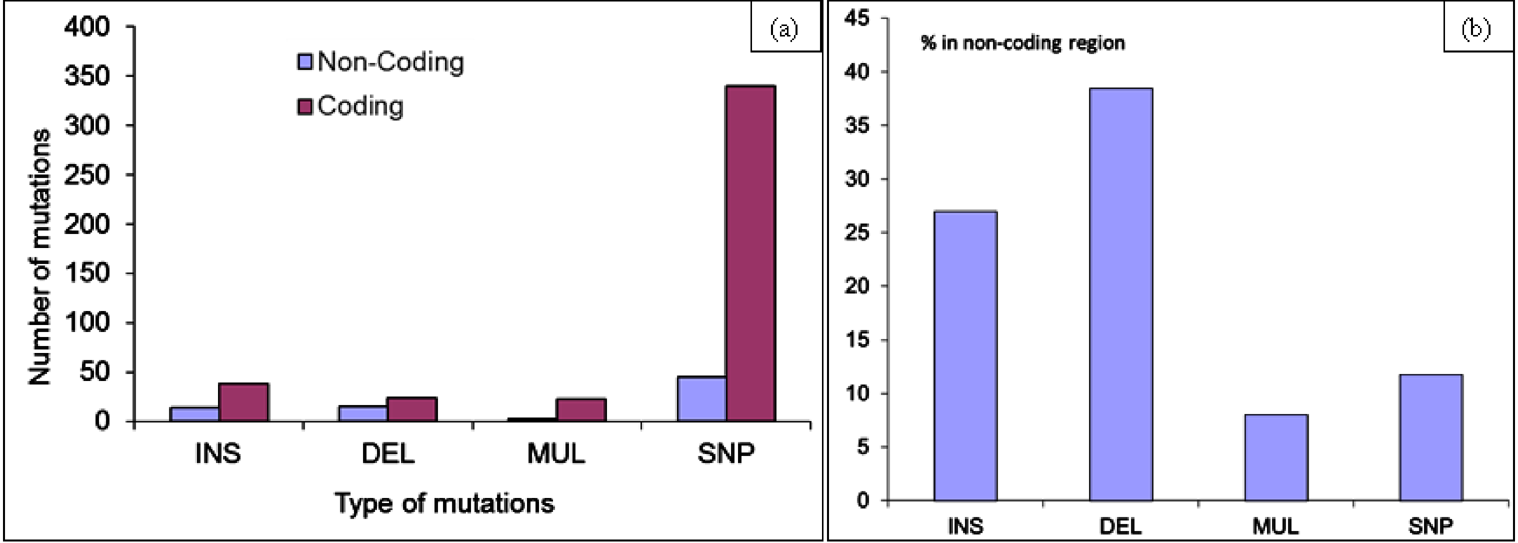
Supplemental Figure S2. Venn analysis of pyrosequencing and CGS resequencing data. 'High confidence' genetic differences identified by pyrosequencing for wild-type and mutant EA strains and all putative CGS differences are shown (A). Pyrosequencing differences with 'variation values' less than 80% and CGS differences defined as either 'low' or 'medium' confidence were filtered out during this analysis (B). All genetic differences and descriptions are shown in Tables Supplemental S1-3.



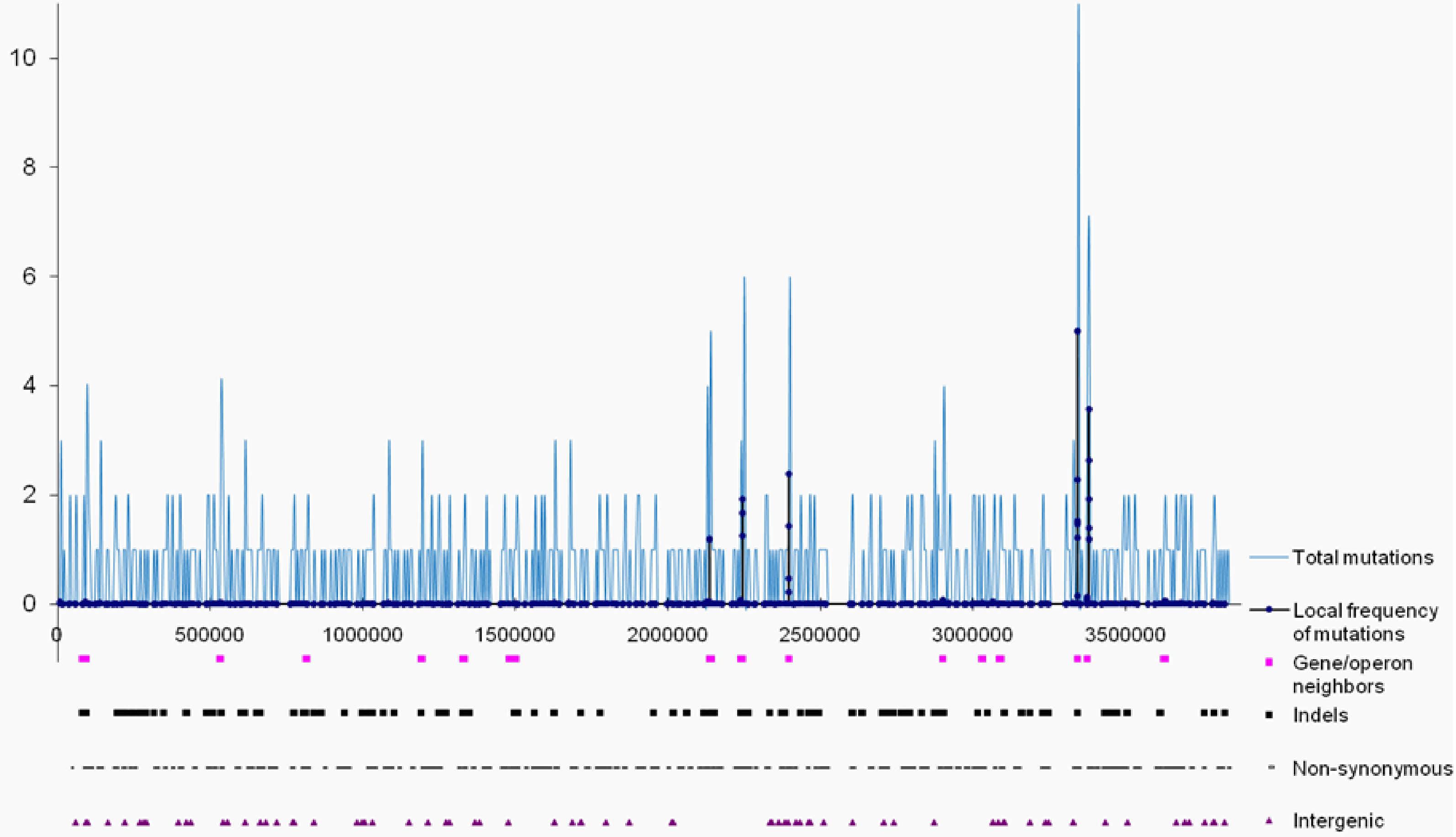
```
Majority                      GGGACTGAGTGCAAAGACAAGAGATGGAAAACAAATGTCAGATTCTTGGX
                                       10        20        30        40        50
Cthe_2397967-2397919.seq      GGGACTGAGTGCAAAGACAAGAGATGGAAAACAAATGTCAGATTCTTGG
Cthe_3342553-3342638.seq      GGGACTTAGTACAAAGACAAGAGATGGAAAACAAATGTCGGATTCTTGGC
Cthe_3372998-3373045.seq      -GGACTGAGTGCAAAGACAAGAGATGGAAAACAAATGTCAGATTCTTGG

Majority                      XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
                                       60        70        80
Cthe_2397967-2397919.seq
Cthe_3342553-3342638.seq      TGATAGGTGATAAAACAGGTAATGATAGAATTTTAG
Cthe_3372998-3373045.seq
```

Supplemental Figure S3. Repetitive DNA detected in "High Confidence Differences (HCDiffs)" regions with low variation values. Twenty five putative HCDiffs were reported by the GSMapper software for conserved hypothetical genes Cthe_2018, Cthe_2825-Cthe_2826, and Cthe_2854-Cthe_2855 for the wild-type strain and the majority were also reported in the mutant strain. A closer examination revealed low variation values for the putative differences (mean of 30) and a CLUSTALW alignment using MegAlign (DNAStar 8.0.2) revealed 28 nucleotides were repeated (red box). A previous study indicated so called HCDiffs with low variation values should be interpreted with caution and are likely false positives. (S. Yang, M. L. Land, D. M. Klingeman, D. A. Pelletier, T.-Y. S. Lu, S. L. Martin, H.-B. Guo, J. C. Smith, S. D. Brown. Paradigm for industrial strain improvement identifies sodium acetate tolerance loci in *Zymomonas mobilis* and *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences*, 2010; 107 (23): 10395 DOI: 10.1073/pnas.0914506107).



Supplemental Figure S4. Distribution analysis of EA mutant pyrosequencing data. 16 mutational hotspots identified in this analysis are described in detail in Table 3.



Supplemental Figure S5. Types of genomic mutations identified in the EA mutant by pryosequencing. Types and numbers of identified mutations in coding and non-coding regions of the EA genome (a) and percentage of different types of mutations found in non-coding regions of the genome (b).

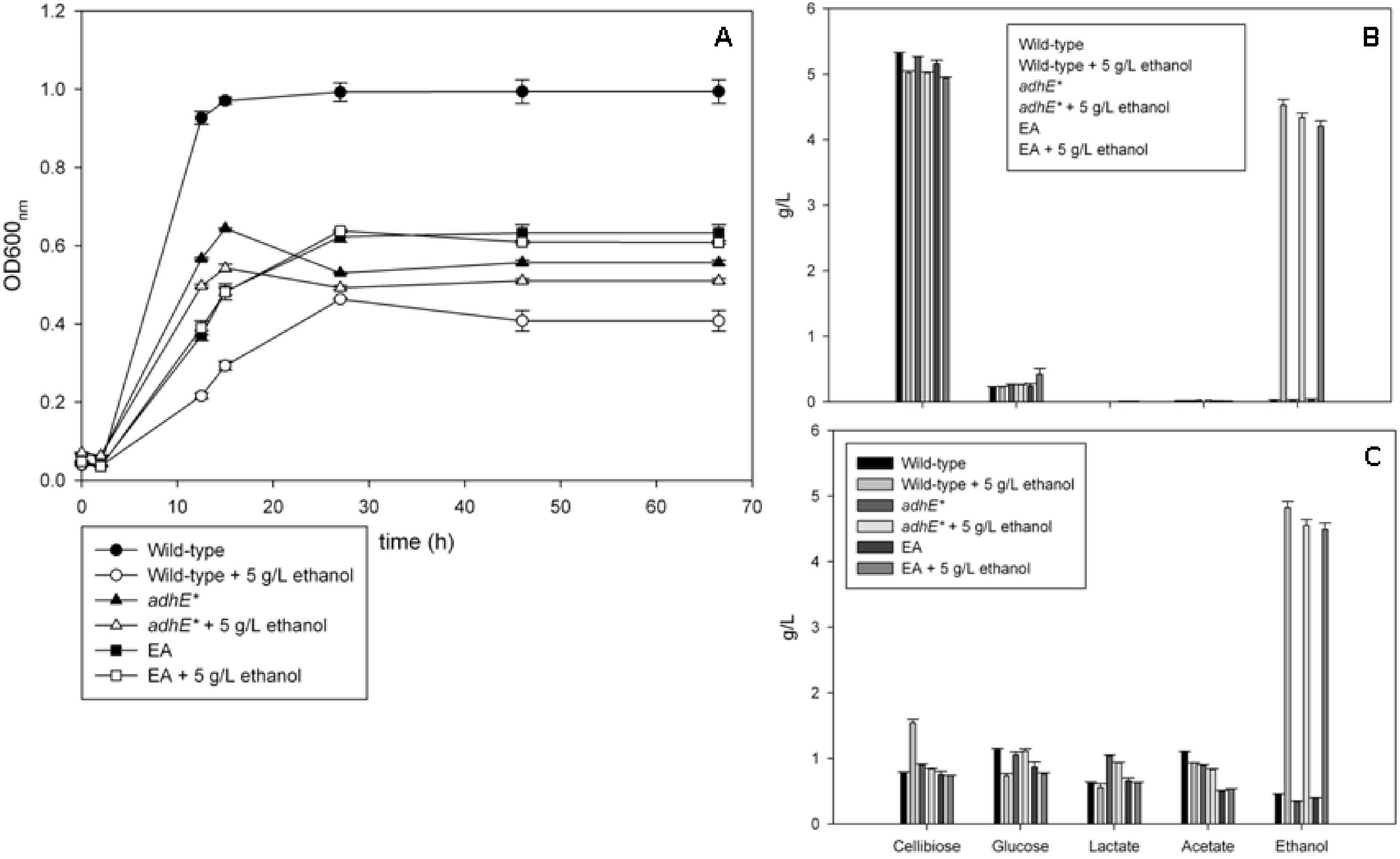
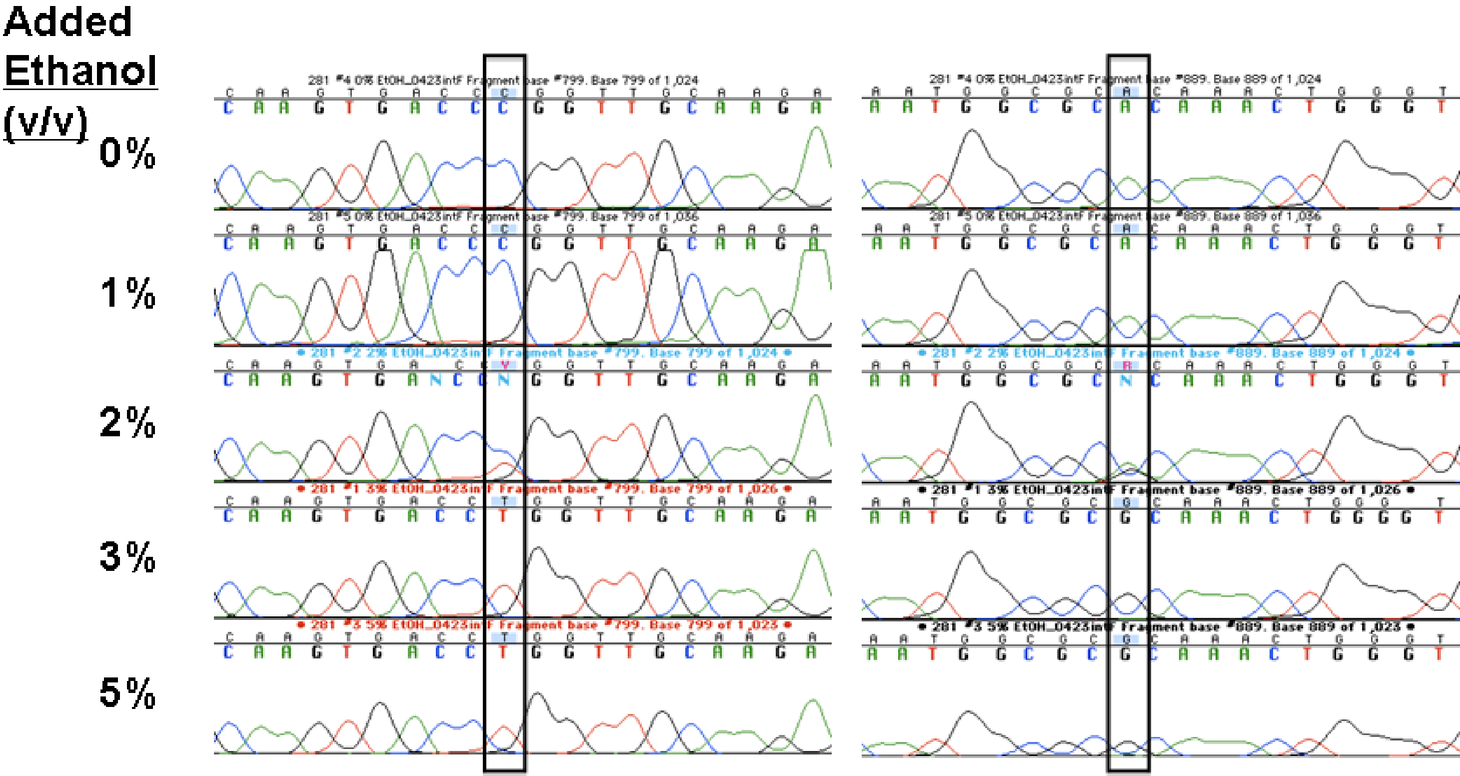


Figure S6. Growth and extracellular metabolite analysis of *C. thermocellum* mutants adhE*, EA and wild-type 27405. All cultures (including inocula) were grown in media for thermophilic clostridia (MTC) containing 5 g/L cellobiose and 2 g/L yeast extract at 55°C and 125 rpm. Anaerobic ethanol was added at a concentration of 5 g/L to select cultures. A 1 mL inoculum from a serum bottle was added to each Balch tube containing 10 mL of medium. (A) Culture growth was monitored by measuring optical densities at 600 nm. Samples were taken for HPLC analysis at the beginning (B) and end (C) of the experiment. For this purpose, samples were filtered through 0.2 µm acrodiscs and acidified with sulfuric acid. Metabolite analysis was performed using a LaChrom Elite system (Hitachi High Technologies America, Inc.) equipped with a refractive index detector (Model L-2490). Metabolites (cellobiose, glucose, lactic acid, acetic acid and ethanol) were separated at a flow rate of 0.5 ml/min in 5 mM H2SO4 using an Aminex HPX- 87H column (Bio-Rad Laboratories, Inc.).



Supplemental Figure S7. Chromosomal wild-type *adhE* gene replacement to *adhE version observed with increasing ethanol concentrations.** Sanger data for SNPs from PCR products amplified using flanking chromosomal primers.