

Supplementary Material

Ensifer canadensis sp. nov strain T173^T isolated from *Melilotus albus* (sweet clover) in Canada possesses recombinant plasmid pT173b harbouring symbiosis and Type IV secretion system genes apparently acquired from *Ensifer medicae*

Eden S. P. Bromfield*, Sylvie Cloutier and Michael F. Hynes

***Correspondence:** Eden S. P. Bromfield, eden.bromfield@agr.gc.ca

Supplementary Table 1. GenBank sequence accession numbers of *Ensifer canadensis* sp. nov. T173^T and reference taxa used in this study.

Strain	<i>atpD</i>	<i>glnII</i>	<i>gyrB</i>	<i>recA</i>	<i>rpoB</i>	16S rRNA	16S-23S-5S rRNA	<i>nifHDK</i>	<i>nodABC</i>
<i>Ensifer canadensis</i> sp. nov. T173 ^T	CP083370	CP083370	CP083370	CP083370	CP083370	CP083370	CP083370 CP083371 CP083372	CP083372	CP083372
<i>Ensifer adhaerens</i> Casida A ^T	CP015880	CP015881	CP015880	CP015880	CP015880	CP015880	CP015880 CP015881 CP015882	na	na
<i>Ensifer alkalisoli</i> YIC4027 ^T	CP034909	CP034910	CP034909	CP034909	CP034909	CP034909	CP034909	CP034911	CP034911
<i>Ensifer americanus</i> CFNEI156 ^T	LNQC01	LNQC01	LNQC01	LNQC01	LNQC01	LNQC01	LNQC01	LNQC01	LNQC01
<i>Ensifer arboris</i> LMG 14919 ^T	ATYB01	HM997093	ATYB01	ATYB01	ATYB01	‡ATYB01	ATYB01	ATYB01	ATYB01
" <i>Ensifer aridi</i> " LMR001	LUAV01	LUAV01	LUAV01	LUAV01	LUAV01	LUAV01	LUAV01	KP993264	LUAV01
" <i>Ensifer collicola</i> " Mol12	-	-	-	-	-	*KT724704	-	-	-
<i>Ensifer fredii</i> USDA 205 ^T	WISZ01	WISZ01	WISZ01	WISZ01	WISZ01	WISZ01	WISZ01	WISZ01	WISZ01
<i>Ensifer garamanticus</i> ORS 1400 ^T	-	HM997091	-	-	-	AY500255	-	-	-
<i>Ensifer glycinis</i> CCBAU 23380 ^T	LPUX01	LPUX01	LPUX01	LPUX01	LPUX01	LPUX01	LPUX01	LPUX01	LPUX01
<i>Ensifer kostiensis</i> DSM 13372 ^T	JAGILA01	JAGILA01	JAGILA01	JAGILA01	JAGILA01	JAGILA01	JAGILA01	JAGILA01	JAGILA01
<i>Ensifer kummerowiae</i> CCBAU 71714 ^T	-	GU994062	-	-	-	AY034028	-	-	-
<i>Ensifer medicae</i> A 321 ^T	VITA01	VITA01	VITA01	VITA01	VITA01	VITA01	VITA01	VITA01	VITA01
<i>Ensifer medicae</i> WSM1115	na	na	na	na	na	na	CP088109	na	na
<i>Ensifer meliloti</i> ATCC 9930 ^T	BJNJ01	BJNJ01	BJNJ01	BJNJ01	BJNJ01	BJNJ01	BJNJ01	BJNJ01	BJNJ01
<i>Ensifer meliloti</i> 1021	na	na	na	na	na	na	AL591688	na	na
<i>Ensifer mexicanus</i> ITTG R7 ^T	CP041238	CP041241	CP041238	CP041238	CP041238	CP041238	CP041238	CP041240	CP041240
<i>Ensifer morelensis</i> Lc04 ^T	JABEKT01	HM997095	JABEKT01	JABEKT01	JABEKT01	JABEKT01	JABEKT01	-	-
<i>Ensifer numidicus</i> ORS 1407 ^T	-	HM997092	-	-	-	AY500254	-	-	-
<i>Ensifer psoraleae</i> CCBAU 65732 ^T	JABEKV01	JABEKV01	JABEKV01	JABEKV01	JABEKV01	JABEKV01	JABEKV01	JABEKV01	JABEKV01
<i>Ensifer saheli</i> LMG 7837 ^T	LNQB01	LNQB01	LNQB01	LNQB01	LNQB01	LNQB01	LNQB01	LNQB01	LNQB01
<i>Ensifer sesbaniae</i> CCBAU 65729 ^T	JABEKU01	JABEKU01	JABEKU01	JABEKU01	JABEKU01	JABEKU01	JABEKU01	JABEKU01	JABEKU01
<i>Ensifer shofinae</i> CCBAU 251167 ^T †	MBFK01	MBFK01	MBFK01	MBFK01	MG646990	MBFK01	MBFK01	MBFK01	MBFK01
<i>Ensifer sojae</i> CCBAU 05684 ^T	AJQT01	AJQT01	AJQT01	AJQT01	AJQT01	AJQT01	AJQT01	AJQT01	AJQT01
<i>Ensifer terangae</i> ORS 1009 ^T	WITC01	WITC01	WITC01	WITC01	WITC01	WITC01	WITC01	WITC01	WITC01

na, Not applicable; - Sequence not available or too short.

* The 16S rRNA gene sequence (KT724704) of “*E. collicola*” Mol 12 was an outlier and excluded from phylogenetic analyses.

† *E. shofinae* CCBAU 251167^T genome sequence MBFK01: NCBI notes "anomalous assembly"; *rpoB* (housekeeping gene) not found in genome sequence.

‡ Three copies of the 16S rRNA gene present in the genome of *Ensifer arboris* LMG 14919^T (Accession no. ATYB01) were found to be anomalous and excluded from the 16S rRNA phylogenetic analysis.

Supplementary Table 2. Average Nucleotide Identity (ANI) and digital DNA–DNA hybridization (dDDH) values for pair-wise comparisons of genome sequences of *Ensifer canadensis* sp. nov. T173^T (Accession no. CP083370-CP083375) with closest relatives.

Reference Strain (Sequence Accession no.)	Fast ANI*	dDDH % [C.I.] [†]
	T173 ^T	T173 ^T
<i>Ensifer morelensis</i> Lc04 ^T (JABEKT01)	88.6	35.7 [33.3-38.2]
<i>Ensifer adhaerens</i> Casida A ^T (CP015880 - CP015882)	83.8	25.5 [23.2-28.0]
<i>Ensifer sesbaniae</i> CCBAU 65729 ^T (JABEKU01)	83.7	25.8 [23.5-28.3]

*Values represent averages of reciprocal comparisons.

[†]dDDH values and confidence intervals [C.I.] based on Genome BLAST Distance Phylogeny (GBDP) formula 4 implemented in the Type Strain Genome Server (TYGS); formula 4 is independent of genome length and is robust against use of incomplete draft genomes.

Supplementary Table 3. BLAST results for T4SS (*virB1* through *virB11*) and *repABC* genes of pT173b (CP083372) as query versus *Rhizobiales* whole-genome shotgun contigs and standard nr/nt databases in NCBI.

Query genes (co-ordinates on pT173b sequence)	Top BLAST hits				
	NCBI Database	ID (%)	Coverage (%)	Strain (Accession no.)	Comments
<i>virB1</i> - <i>virB11</i> (23916 - 35397 bp)	Whole- Genome Shotgun Contigs	92.48	100	<i>Neorhizobium galegae</i> bv. orientalis strain HAMBI 2605 (CCRM01)	NCBI notes: anomalous assembly
		91.17	100	<i>Neorhizobium galegae</i> strain NG_87_Ori (VZUL01)	Classification as <i>N. galegae</i> confirmed by phylogenomic analysis implemented in the Type Strain Genome Server (TYGS)
		90.80	99	<i>Rhizobiales</i> bacterium isolate AFS086343 (UCKS01)	
		90.80	100	<i>Rhizobium</i> sp. strain Root483D2 (LMFB01)	
		90.80	100	<i>Rhizobium</i> sp. strain Root268 (LMJD01)	
	Standard nr/nt	79.44	99	<i>Shinella</i> sp. PSBB067 plasmid unnamed (CP069305)	
		79.35	99	<i>Rhizobium daejeonense</i> strain KACC 13094 unnamed plasmid (CP048428)	
		79.14	99	<i>Shinella</i> sp. HZN7 plasmid pShin-06 (CP015742)	
		78.79	99	<i>Shinella zoogloeoides</i> strain ATCC 19623 unnamed plasmid (CP086613)	
<i>repABC</i> (13014 -16648 bp)	Whole- Genome Shotgun Contigs	97.88	100	<i>Neorhizobium galegae</i> strain NG_87_Ori (VZUL01)	Classification as <i>N. galegae</i> confirmed by phylogenomic analysis implemented in the Type Strain Genome Server (TYGS)
		96.75	100	<i>Neorhizobium galegae</i> bv. orientalis HAMBI 2605 (CCRM01)	NCBI notes: anomalous assembly
		98.05	65	<i>Rhizobiales</i> bacterium isolate R129_J (UCML01)	
		84.46	99	<i>Pararhizobium</i> sp. YC-54 (JAOYTH01)	
		76.36	100	<i>Rhizobiaceae</i> bacterium AM104-69 (JAQDQG01)	
	Standard nr/nt	76.43	99	<i>Shinella</i> sp. HZN7 plasmid pShin-06 (CP015742)	
		79.70	59	<i>Rhizobium daejeonense</i> strain KACC 13094 unnamed plasmid (CP086613)	

Supplementary Table 4. Growth characteristics of (1) *E. canadensis* sp. nov. T173^T, (2) *E. morelensis* Lc04^T, (3) *E. adhaerans* Casida A^T, (4) *E. sesbaniae* CCBAU 65729^T, (5) *E. medicae* A321^T, (6) *E. americanus* CFNEI156^T, (7) *E. fredii* USDA 205^T, and (8) *E. meliloti* ATCC 9930^T.

Characteristic	1	2	3	4	5	6	7	8
Growth on YEM agar medium*								
10 °C	±	±	±	-	-	-	-	-
37 °C	±	±	+	+	+	±	-	±
pH 5	+	+	+	+	±	±	±	±
pH 10	+	+	+	+	±	-	-	+
1% NaCl	+	+	+	+	+	±	±	+
2% NaCl	+	+	+	-	-	-	±	-
3% NaCl	-	-	-	-	-	-	-	-
Growth in LB broth*	+	+	+	+	±	±	±	+
Acid production †	+	ND	ND	+	+	+	+	+

* Positive, +; weak, ± ; negative, - ; not determined, ND. Values are based on three replicates.

† Acid production on YEM agar medium determined as described by Bromfield et al., (2010); pH values varied between pH 5 - 6.

Supplementary Table 5. Phenotypic characteristics (BIOLOG GEN III MicroPlates) of (1) *E. canadensis* sp. nov.T173^T, (2) *E. morelensis* Lc04^T, (3) *E. adhaerans* Casida A^T, (4) *E. sesbaniae* CCBAU 65729^T, (5) *E. medicae* A321^T, (6) *E. americanus* CFNEI156^T, (7) *E. fredii* USDA 205^T, and (8) *E. meliloti* ATCC 9930^T.

Characteristic	1	2	3	4	5	6	7	8	Characteristic	1	2	3	4	5	6	7	8
C-source utilization																	
Dextrin	-	+	+	+	-	-	-	+	Glycyl-L-Proline	-	+	±	±	-	-	-	-
D-Maltose	±	+	+	+	±	-	-	±	L-Alanine	-	-	±	-	-	-	-	-
D-Trehalose	±	+	+	+	-	-	-	+	L-Arginine	-	±	±	±	-	-	-	-
D-Cellobiose	±	+	+	+	±	-	±	+	L-Aspartic Acid	-	+	+	+	-	-	-	±
Gentiobiose	±	+	+	+	±	-	-	+	L-Glutamic Acid	±	±	±	+	-	-	-	±
Sucrose	±	+	+	+	-	-	-	+	L-Histidine	-	±	±	±	-	-	-	-
D-Turanose	-	-	-	-	±	-	-	-	L-Pyroglutamic Acid	-	-	-	-	-	-	-	-
Stachyose	-	-	-	-	-	-	-	-	L-Serine	-	±	±	±	-	-	-	-
D-Raffinose	±	±	+	±	-	-	-	-	Pectin	±	+	+	±	+	+	+	±
α-D-Lactose	±	±	+	±	-	-	-	+	D-Galacturonic Acid	-	-	-	-	-	-	-	±
D-Melibiose	±	±	±	-	-	-	-	±	L-Galactonic Acid Lactone	-	-	-	-	-	-	-	-
β-Methyl-DGlucoside	-	±	±	±	-	-	-	±	D-Gluconic Acid	±	-	-	-	-	-	-	-
D-Salicin	-	±	±	-	-	-	-	-	D-Glucuronic Acid	±	-	-	-	-	-	-	-
N-Acetyl-DGlucosamine	±	+	+	+	-	-	-	±	Glucuronamide	-	-	-	-	±	+	+	-
N-Acetyl-β-DMannosamine	±	-	±	-	-	-	-	±	Mucic Acid	-	-	-	-	-	-	-	-
N-Acetyl-DGalactosamine	-	±	±	±	-	-	-	-	Quinic Acid	±	±	-	±	-	-	-	-
N-Acetyl Neuraminic Acid	-	-	-	-	-	-	-	-	D-Saccharic Acid	-	-	-	-	-	-	-	-
α-D-Glucose	±	+	+	+	±	-	-	+	p-HydroxyPhenylacetic Acid	-	-	-	-	-	-	-	-
D-Mannose	±	+	+	+	-	-	-	+	Methyl Pyruvate	-	-	-	-	-	-	-	-
D-Fructose	±	+	+	±	±	-	-	+	D-Lactic Acid Methyl Ester	-	+	-	-	-	-	-	-
D-Galactose	-	±	±	±	±	-	-	±	L-Lactic Acid	-	+	±	±	-	-	-	-
3-Methyl Glucose	±	±	-	+	-	-	-	-	Citric Acid	-	-	-	-	-	-	-	-
D-Fucose	-	-	-	-	-	-	-	-	α-Keto- Glutaric Acid	-	-	-	-	-	-	-	-
L-Fucose	±	+	+	±	-	-	-	+	D-Malic Acid	±	+	±	+	-	-	-	±
L-Rhamnose	-	-	±	±	-	-	-	±	D-Malic Acid	+	+	+	+	-	-	-	+
Inosine	-	±	-	-	-	-	-	-	Bromo-Succinic Acid	-	-	±	-	-	-	-	-
D-Sorbitol	±	+	+	+	+	+	+	+	Tween 40	±	+	±	±	-	-	-	-
D-Mannitol	±	+	+	+	+	+	+	±	γ-Amino-Butyric Acid	-	-	-	-	-	-	-	-
D-Arabitol	+	+	+	+	+	+	+	+	α-HydroxyButyric Acid	-	-	-	-	-	-	-	-
myo-Inositol	+	+	+	+	±	+	+	+	β-Hydroxy-D,LButyric Acid	-	-	-	-	-	-	-	-
Glycerol	±	+	+	+	±	+	+	+	α-Keto-Butyric Acid	-	-	-	-	-	-	-	-
D-Glucose- 6-PO4	±	±	+	+	±	+	+	±	Acetoacetic Acid	+	+	±	±	-	-	-	+
D-Fructose- 6-PO4	±	-	±	-	±	+	+	-	Propionic Acid	±	-	-	+	-	-	-	+
D-Aspartic Acid	±	-	-	-	±	+	+	-	Acetic Acid	+	+	+	+	-	-	-	+
Gelatin	-	-	-	-	-	-	-	-	Formic Acid	±	-	-	-	-	-	-	-
Chemical Sensitivity																	
1% Sodium Lactate	-	+	+	-	-	±	-	-	Vancomycin	-	-	-	-	-	-	-	-
Fusidic Acid	-	-	-	-	±	+	+	-	Tetrazolium Violet	+	+	+	+	+	+	+	+
D-Serine	-	-	-	-	-	-	-	-	Tetrazolium Blue	+	+	+	+	+	+	+	+
Troleandomycin	±	+	+	+	-	+	+	-	Nalidixic Acid	-	-	±	+	±	-	-	-
Rifamycin SV	+	+	+	+	+	+	+	+	Lithium Chloride	-	-	-	-	-	-	-	-
Minocycline	-	-	-	-	-	+	+	-	Potassium Tellurite	-	±	-	+	+	±	-	-
Lincomycin	+	±	+	+	+	-	-	-	Aztreonam	+	+	-	-	+	-	-	-
Guanidine HCl	-	-	-	-	-	-	-	-	Sodium Butyrate	-	-	-	-	-	±	-	-
Niaproof 4	-	-	-	±	-	±	-	-	Sodium Bromate	-	-	-	-	-	±	-	-

Growth after 24 hours incubation at 30 °C: +, positive; ±, weak; −, negative. Values are based on three replicates.

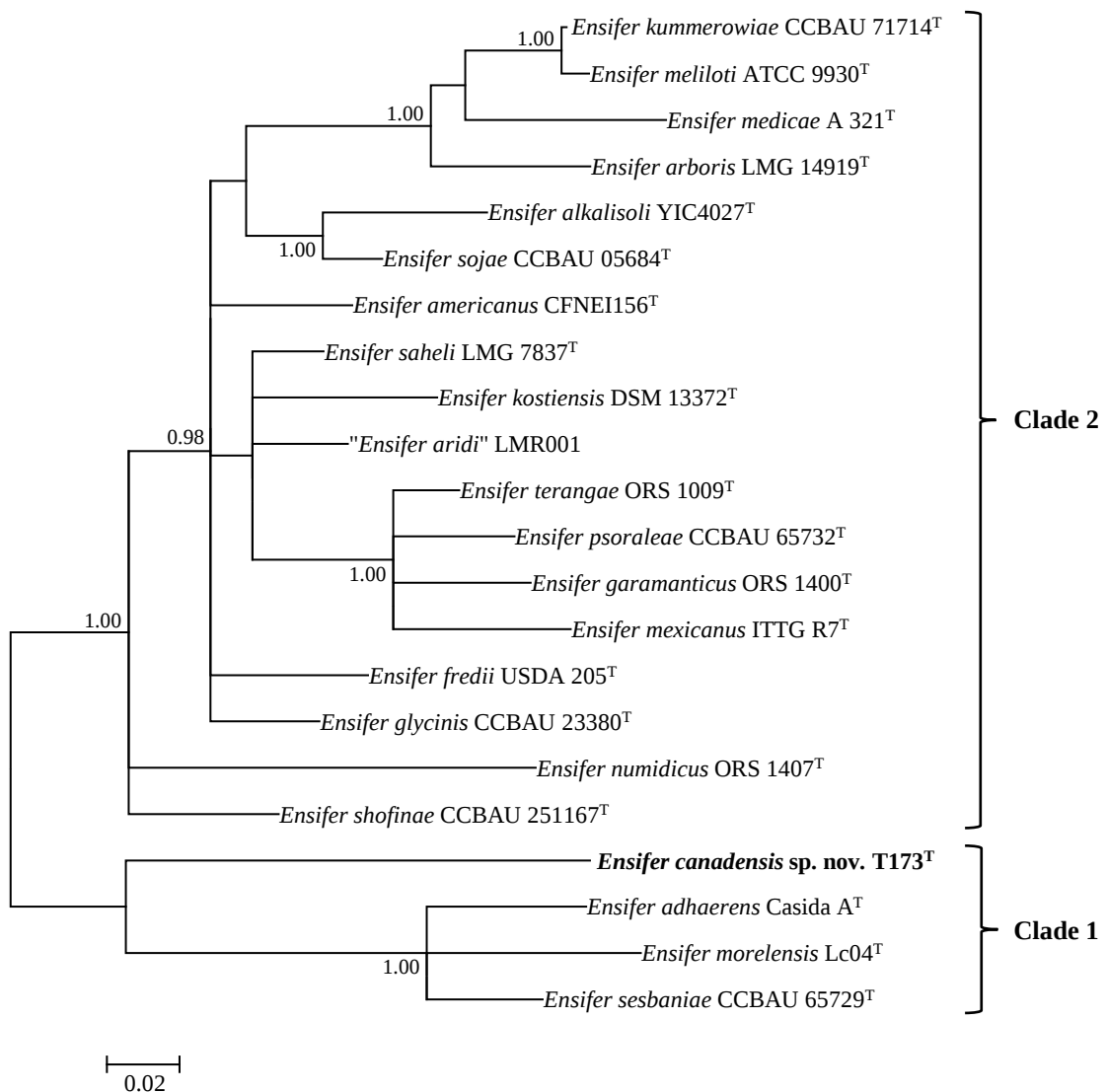
Supplementary Table 6. Examples of antibiotic resistance genes encoding enzymes that inactivate beta-lactam and amino-glycoside antibiotics detected in the genome of *E. canadensis* sp. nov. T173^T.

Gene (location)	Co-ordinates (bp)	Product	Function
bla (Chromosome - CP083370)	698,503-699,387	Class A beta-lactamase	Class A beta-lactamase Classification: antibiotic inactivation enzyme Antibiotic class: beta-lactam antibiotics
bla (pT173e - CP083371)	1,593,270-1,594,130	Subclass B3 beta-lactamase	Subclass B3 beta-lactamase Classification: antibiotic inactivation enzyme Antibiotic class: beta-lactam antibiotics
AAC(3)-II,III,IV,VI,VIII,IX,X (pT173c - CP083373)	24,137-24,955	Aminoglycoside N(3)-acetyltransferase	Aminoglycoside N(3)-acetyltransferase Classification: antibiotic inactivation enzyme Antibiotic class: aminoglycosides Antibiotics: gentamicin C, tobramycin, gentamicin B, amikacin, kanamycin A, apramycin, neomycin.
AAC(6')-Ic,f,g,h,j,k,l,r-z (Chromosome - CP083370)	2,575,834-2,576,295	Aminoglycoside N(6')-acetyltransferase	Aminoglycoside N(6')-acetyltransferase Classification: antibiotic inactivation enzyme Antibiotic class: aminoglycosides Antibiotics: tobramycin, kanamycin A, amikacin, dibekacin, sisomicin, gentamicin B, isepamicin, arbekacin, netilmicin, neomycin.
Unnamed (Chromosome - CP083370)	1,505,251-1,506,051	Aminoglycoside 6-phosphotransferase, (putative)	Aminoglycoside 6-phosphotransferase, (putative). Classification: antibiotic inactivation enzyme Antibiotic class: aminoglycosides
Unnamed (pT173e - CP083371)	3,349,865-3,350,656	Aminoglycoside 3'-phosphotransferase, (putative)	Aminoglycoside 3'-phosphotransferase, putative. Classification: antibiotic inactivation enzyme Antibiotic class: aminoglycosides
APH(3'')-I (pT173e - CP083371)	155,136-155,939	Aminoglycoside 3''-phosphotransferase	Aminoglycoside 3''-phosphotransferase Classification: antibiotic inactivation enzyme Antibiotic class: aminoglycosides Antibiotic: streptomycin

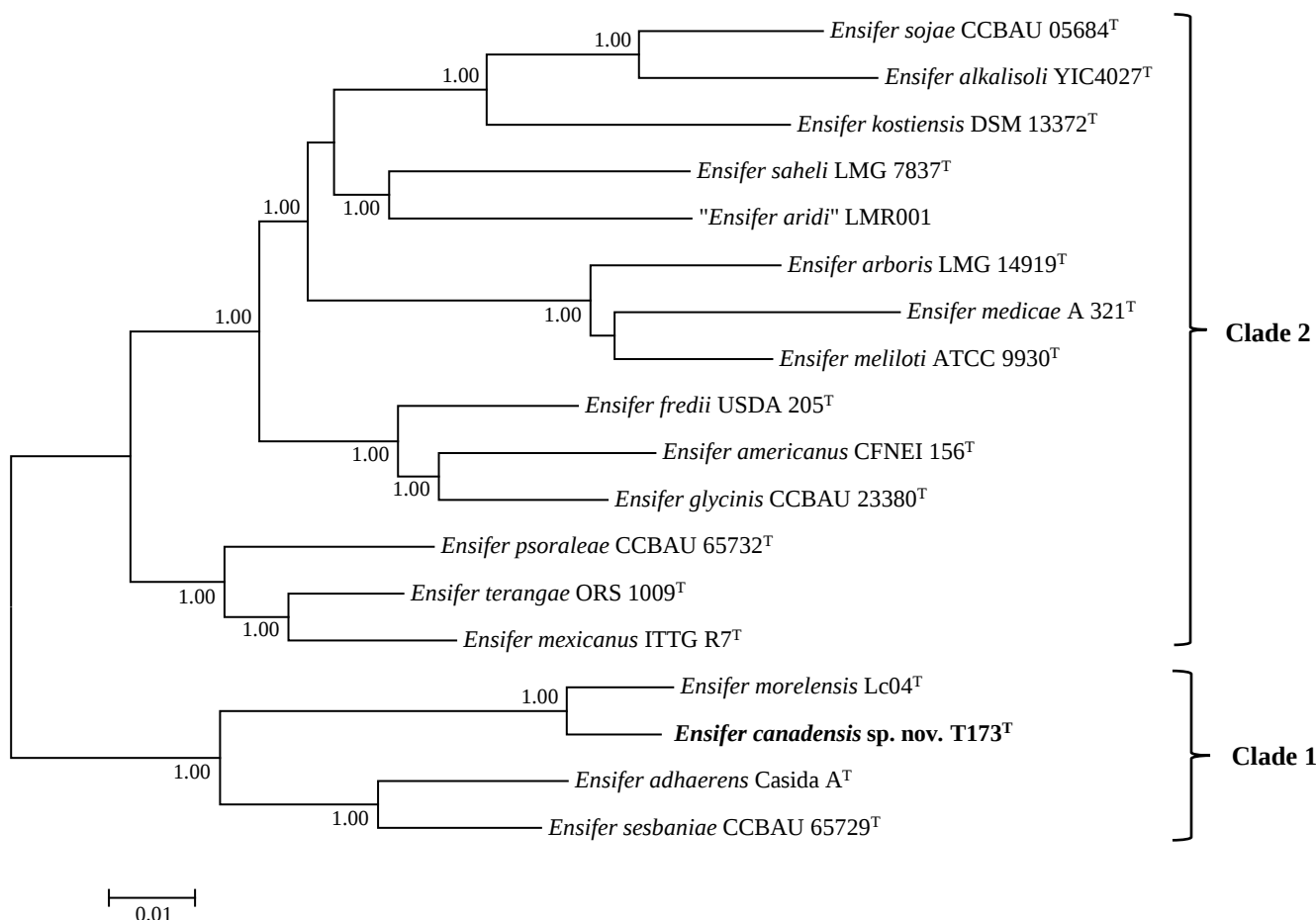
Supplementary Table 7. Fatty acid profiles of (1) *E. canadensis* sp. nov. T173^T, (2) *E. morelensis* Lc04^T, (3) *E. adhaerans* Casida A^T, (4) *E. sesbaniae* CCBAU 65729^T, (5) *E. medicae* A321^T, (6) *E. americanus* CFNEI156^T, (7) *E. fredii* USDA 205^T, and (8) *E. meliloti* ATCC 9930^T.

Fatty Acid	1	2	3	4	5	6	7	8
13:1 at 12-13	0.07							
14:0	0.13	0.16			0.20			
15 :0 3OH	0.13			0.12				
16:0	7.70	6.40	4.15	4.09	12.40	8.01	7.65	7.61
16:0 3OH	1.56	2.06	2.59	1.73	1.37	0.90	0.81	0.91
17:0 cyclo	0.59	2.12		0.58	1.01	0.97	0.62	0.73
17:0 3OH	0.18			0.19				0.25
17:1 ω8c	0.16			0.14		0.36		0.14
17:0	0.62	0.27		0.37	0.32	0.63	0.72	0.65
18:0	4.34	1.82	2.28	1.72	2.81	2.53	5.60	2.09
18:0 3OH	2.17	1.98	4.24	3.17	3.19	3.50	4.17	4.30
18:1 w5c	0.12			0.22				0.20
18:1 ω7c 11-methyl	18.69	18.33	15.35	1.04	0.74	0.50	0.81	1.86
19:0	0.08							
19:0 cyclo ω8c	2.86	2.02	9.41	22.08	14.18	5.52	9.14	25.01
20:0 iso	0.20							
20:1 w7c	0.16			0.18				
20:2 w6,9c	0.31		0.85	0.81	0.64	0.48	0.77	1.46
Summed feature 1*	0.10							
Summed feature 2*	7.00	6.96	8.12	7.88	7.55	8.68	8.55	7.86
Summed feature 3*	2.44	2.19		0.45	1.03	2.23	0.91	0.51
Summed feature 5*	0.11							
Summed feature 7*	0.18			0.41				0.56
Summed feature 8*	50.07	55.69	53.01	54.82	54.56	65.70	60.24	45.86

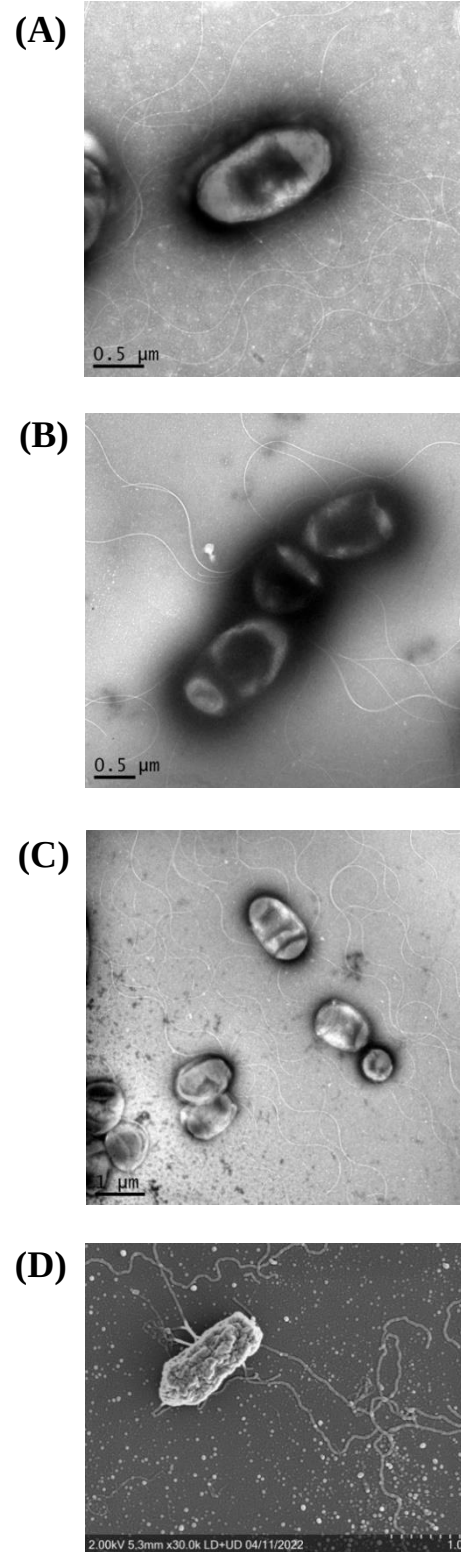
* Summed Features are fatty acids that cannot be resolved reliably from another fatty acid using the chromatographic conditions chosen. The MIDI system groups these fatty acids together as one feature with a single percentage of the total. Summed feature 1, 15:1 iso H/13:0 3OH / 13:0 3OH/15:1 iso H; Summed feature 2, 12:0 aldehyde/?; Summed feature 3, 16:1 ω6c/16:1 ω7c; Summed feature 5, 18:0 ante/18:2 ω6,9c; Summed feature 7, 19:1w7c/19:1 w6c / 19:1 w6c/w7c/19cy; Summed feature 8, 18:1 ω6c/18:1 ω7c.



Supplementary Figure 1. Bayesian phylogenetic tree (GTR + G + I substitution model) of *glnII* gene sequences (615 bp) for *Ensifer canadensis* sp. nov. T173^T and reference taxa of the genus *Ensifer*. Only posterior probabilities $\geq 90\%$ are shown. Bar, expected substitutions per site.

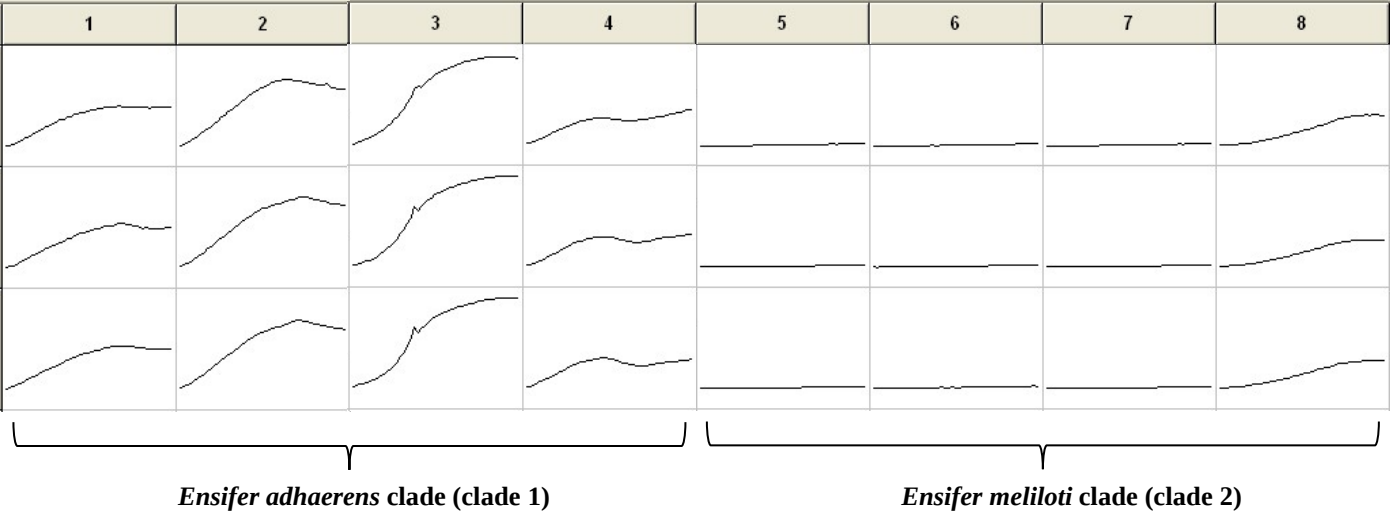


Supplementary Figure 2. Bayesian phylogenetic tree (GTR + G + I substitution model) inferred from 53 full-length concatenated ribosome protein subunit (*rps*) gene sequences of *Ensifer canadensis* sp. nov. T173^T and reference taxa of the genus *Ensifer*. Alignment length, 23067bp. Only posterior probabilities $\geq 90\%$ are shown. Bar, expected substitutions per site.



Supplementary Figure 4. Transmission (A, B and C) and scanning (D) electron microscope images showing cell morphological features of *Ensifer canadensis* sp. nov. strain T173^T.

Supplementary Figure 5. Growth curves (48 hrs at 30 °C) in LB broth medium of (1) *E.canadensis*. sp. nov.T173^T, (2) *E. adhaerans* Casida A^T, (3) *E. morelensis* Lc04^T, (4) *E. sesbaniae* CCBAU 65729^T, (5) *E. americanus* CFNEI156^T, (6) *E. fredii* USDA 205^T, (7) *E. medicae* A321^T, and (8) *E. meliloti* ATCC 9930^T.



Vertical axis: bacterial growth based on OD (595 nm). Horizontal axis: hourly reads for 48 hours at 30 °C (three replicates of each bacterial strain).