



## OPEN ACCESS

## EDITED BY

Guillaume J. Bilodeau,  
Canadian Food Inspection Agency (CFIA),  
Canada

## REVIEWED BY

Gonzalo Arturo Torres Tejerizo,  
CONICET Institute of Biotechnology and  
Molecular Biology (IBBM), Argentina  
Mariano Pistorio,  
National Scientific and Technical Research  
Council (CONICET), Argentina

## \*CORRESPONDENCE

Eden S. P. Bromfield  
✉ eden.bromfield@agr.gc.ca

RECEIVED 28 March 2023

ACCEPTED 09 May 2023

PUBLISHED 14 June 2023

## CITATION

Bromfield ESP, Cloutier S and Hynes MF (2023)  
*Ensifer canadensis* sp. nov. strain T173<sup>T</sup> 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*.  
*Front. Microbiol.* 14:1195755.  
doi: 10.3389/fmicb.2023.1195755

## COPYRIGHT

© 2023 Michael F. Hynes and His Majesty the  
King in Right of Canada, as represented by the  
Minister of Agriculture and Agri-Food Canada  
for the contribution of Eden S. P. Bromfield and  
Sylvie Cloutier. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The  
use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in this  
journal is cited, in accordance with accepted  
academic practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# *Ensifer canadensis* sp. nov. strain T173<sup>T</sup> 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<sup>1\*</sup>, Sylvie Cloutier<sup>1</sup> and Michael F. Hynes<sup>2</sup>

<sup>1</sup>Ottawa Research and Development Centre, Agriculture and Agri-Food Canada, Ottawa, ON, Canada,

<sup>2</sup>Department of Biological Sciences, University of Calgary, Calgary, AB, Canada

A bacterial strain, designated T173<sup>T</sup>, was previously isolated from a root-nodule of a *Melilotus albus* plant growing in Canada and identified as a novel *Ensifer* lineage that shared a clade with the non-symbiotic species, *Ensifer adhaerens*. Strain T173<sup>T</sup> was also previously found to harbour a symbiosis plasmid and to elicit root-nodules on *Medicago* and *Melilotus* species but not fix nitrogen. Here we present data for the genomic and taxonomic description of strain T173<sup>T</sup>. Phylogenetic analyses including the analysis of whole genome sequences and multiple locus sequence analysis (MLSA) of 53 concatenated ribosome protein subunit (*rps*) gene sequences confirmed placement of strain T173<sup>T</sup> in a highly supported lineage distinct from named *Ensifer* species with *E. morelensis* Lc04<sup>T</sup> as the closest relative. The highest digital DNA–DNA hybridization (dDDH) and average nucleotide identity (ANI) values of genome sequences of strain T173<sup>T</sup> compared with closest relatives (35.7 and 87.9%, respectively) are well below the respective threshold values of 70% and 95–96% for bacterial species circumscription. The genome of strain T173<sup>T</sup> has a size of 8,094,229bp with a DNA G+C content of 61.0mol%. Six replicons were detected: a chromosome (4,051,102bp) and five plasmids harbouring plasmid replication and segregation (*repABC*) genes. These plasmids were also found to possess five apparent conjugation systems based on analysis of TraA (relaxase), TrbE/VirB4 (part of the Type IV secretion system (T4SS)) and TraG/VirD4 (coupling protein). Ribosomal RNA operons encoding 16S, 23S, and 5S rRNAs that are usually restricted to bacterial chromosomes were detected on plasmids pT173d and pT173e (946,878 and 1,913,930bp, respectively) as well as on the chromosome of strain T173<sup>T</sup>. Moreover, plasmid pT173b (204,278bp) was found to harbour T4SS and symbiosis genes, including nodulation (*nod*, *noe*, *nol*) and nitrogen fixation (*nif*, *fix*) genes that were apparently acquired from *E. medicae* by horizontal transfer. Data for morphological, physiological and symbiotic characteristics complement the sequence-based characterization of strain T173<sup>T</sup>. The data presented support the description of a new species for which the name *Ensifer canadensis* sp. nov. is proposed with strain T173<sup>T</sup> (= LMG 32374<sup>T</sup>=HAMBI 3766<sup>T</sup>) as the species type strain.

## KEYWORDS

*Ensifer canadensis* sp. nov. T173<sup>T</sup>, complete genome sequence, phylogenetics, symbiosis plasmid, horizontal gene transfer, type IV secretion system (T4SS), plasmid conjugation systems

## Introduction

The genus *Ensifer* consists of diverse species of soil bacteria and includes bacterial predators and agriculturally important species capable of fixing nitrogen in symbiotic association with legume plants such as alfalfa (*Medicago sativa*). This bacterial genus is currently divided into two major clades based on phylogenetic analysis of multiple housekeeping (core) genes (Martens et al., 2008; Kumar et al., 2017; Fagorzi et al., 2020; Kuzmanović et al., 2022). One of these clades is represented by the non-symbiotic species, *Ensifer morelensis* (Wang et al., 2002) and the predatory species, *Ensifer adhaerens* (Casida, 1982), and the other clade by symbiotic nitrogen fixing species such as *Ensifer meliloti* (de Lajudie et al., 1994) and *Ensifer medicae* (Rome et al., 1996). Based on comprehensive genomic and phenotypic data, Kuzmanović et al. (2022) proposed the recognition of these two clades as separate genera named *Ensifer* and *Sinorhizobium*, respectively. However, this proposal was not adopted by the Judicial Commission of the International Committee on Systematics of Prokaryotes (Oren and Garrity, 2022) and the genus *Ensifer* (Casida, 1982) remains intact.

In previous studies (Bromfield et al., 2001, 2010), bacteria were isolated from root-nodules of alfalfa and sweet clover (*Melilotus albus*) plants grown at a Canadian field site without a history of cultivation. Characterization of these bacterial isolates by phylogenetic analysis of four house-keeping gene sequences (Bromfield et al., 2010) resulted in the identification of several lineages that were assigned to the genus *Ensifer*. One of these lineages, represented by a bacterial strain designated T173<sup>T</sup>, was found to be novel and closely related to the type strain of *E. morelensis*. Strain T173<sup>T</sup> was also found to possess a plasmid harboring nodulation (*nod*) and nitrogen fixation (*nif*) genes and to be able to elicit nodules on the roots of several *Medicago* and *Melilotus* species but not fix nitrogen.

In the present study our purpose was to use genomic, phylogenetic and phenotypic analyses to further characterize and verify the taxonomic status of strain T173<sup>T</sup>. Based on the results presented here a novel bacterial species named *Ensifer canadensis* sp. nov. is proposed.

## Materials and methods

### Bacteria

Novel strain T173<sup>T</sup> was isolated from a root-nodule of a *Melilotus albus* plant grown at a field site in Ottawa, Ontario, Canada that had no known history of agriculture as described by Bromfield et al. (2001). Reference strains used in phylogenetic and genomic analyses are listed in Supplementary Table S1; reference strains used in different phenotypic tests are listed in relevant Tables and Figures in the main text or Supplementary material section. Bacterial strains were routinely grown on yeast extract-mannitol (YEM) agar medium (Tang et al., 2012) or modified tryptone yeast-extract (TY) agar medium with the following composition (g/L):

tryptone (Oxoid, United States), 0.5; yeast-extract (Oxoid, United States), 1.0; calcium chloride dihydrate, 0.1; Bacteriological agar no. 1 (Oxoid, United States), 15.0. Bacteria were maintained at  $-80^{\circ}\text{C}$  in 20% w/v glycerol. Strain T173<sup>T</sup> was deposited in the BCCM/LMG Bacteria Collection, University of Ghent, Belgium as LMG 32374<sup>T</sup> and in the HAMBI Microbial Culture Collection, University of Helsinki, Finland as HAMBI 3766<sup>T</sup>.

### Phenotypic characterization

Assessment of Gram-stain reaction of bacteria was done using the KOH method of Buck (1982).

Multiple tests including carbon source utilization and chemical sensitivity assays were carried out using BIOLOG GEN III MicroPlates (Biolog, United States) according to manufacturer's instructions.

For analysis of fatty acids, bacteria were grown for 2–4 days on TY agar medium (Beringer, 1974) at 28°C. Bacteria were harvested and fatty acids extracted as described by Sasser (1990). Fatty acid identification was carried out using the Sherlock Microbial Identification System (MIDI) version 6.0 and the RTSBA6 database.

Electron microscopy was used to investigate cell morphology employing bacteria grown in modified TY broth at 28°C. For transmission electron microscopy (model, H-7000; Hitachi), bacterial samples were adsorbed to formvar coated copper grids and negatively stained with 1% phosphotungstic acid (pH 7.0). For scanning electron microscopy (model, Hitachi SU7000 FESEM), bacteria were adsorbed to Poly-L-Lysine coated silicon wafers (Ted Pella Inc., CA, United States). Samples were fixed using 3% glutaraldehyde in 0.1 M sodium Cacodylate buffer pH 7.2, dehydrated in a graded ethanol series and critical point dried in an Autosamdri-931 critical point dryer (Tousimis Research Corp, MD, United States). Wafers were mounted on aluminum stubs coated with a 6.5 nm layer of platinum in an Emitech K550V sputter coater (EM Technologies Ltd., Kent, United Kingdom).

Tests of ability to “track” bacterial prey cells to detect potential bacterial predatory activity of strain T173<sup>T</sup> were carried out using plates of 0.1x Heart Infusion (BD Difco, United States) agar medium supplemented with 0.1% glucose as detailed by Martin (2002). Plates were incubated at 28°C for 7 days and any spreading growth of predator strains over cells of the prey strain was recorded. Bacterial test strains were *Micrococcus luteus* JCM 1464<sup>T</sup> (as prey), *Escherichia coli* (EZ competent cells, Qiagen, Netherlands) (resistant to predation; negative control), *Ensifer adhaerens* Casida A<sup>T</sup> (as predator) and *E. morelensis* (as predator).

Tests of acid production by bacteria grown on YEM agar medium for 21 days at 28°C were carried out as described previously (Bromfield et al., 2010).

The ability of bacteria to grow in LB broth medium at 30°C was assessed using 96-well microplates as detailed by Fagorzi et al. (2020). Growth was estimated (over a 48 h period) on the basis of optical

density (595 nm) by taking readings at hourly intervals using a FLUOstar OPTIMA plate reader (BMG LABTECH, Germany).

Plant tests of nodulation and nitrogen fixation using seedlings from surface sterilized seed were done using Leonard jar assemblies (Vincent, 1970) (three replicate jars, two plants per jar) as described previously (Bromfield et al., 2010).

## Analyses of partial gene sequences

Sequences of 16S rRNA, *atpD*, *glnII*, *gyrB*, *recA*, and *rpoB* housekeeping genes were extracted from genome sequences and used for phylogenetic analyses. Sequence accession numbers are shown in [Supplementary Table S1](#). Alignment of 16S rRNA gene sequences was done using the fast, secondary-structure aware Infernal aligner implemented in the online Ribosomal Database Project version 11.5 (Cole et al., 2014). Sequences of protein-encoding housekeeping gene sequences (*atpD*, *glnII*, *gyrB*, *recA*, and *rpoB*) were aligned using MUSCLE (Edgar, 2004).

Best fit substitution models were selected using ModelTest-NG (Darriba et al., 2020) implemented in CIPRES Science Gateway version 3.3 (Miller et al., 2010). Bayesian phylogenetic analyses were performed using MrBayes version 3.2.1 with default priors (Altekar et al., 2004) as previously described (Yu et al., 2014). Maximum-likelihood (ML) phylogenetic analyses (Guindon et al., 2010) were carried out using 1,000 non-parametric bootstrap replications to assess support as described by Tang et al. (2012). Phylogenetic trees from Bayesian and ML analyses exhibited similar topologies and therefore only Bayesian trees are shown in this work.

## Genome analyses

Bacterial cells for genomic DNA preparation were grown on modified TY agar plates for 2 days at 28°C. Genomic DNA was extracted from bacterial cells (washed twice in sterile water) using the Promega Wizard SV Genomic DNA Purification System (Promega, United States). The genomic DNA was purified using the DNeasy PowerClean Pro Cleanup Kit (Qiagen, Netherlands) according to manufacturer's instructions.

Complete genome sequencing of strain T173<sup>T</sup> was carried out at the Genome Quebec Innovation Centre, Montreal, Canada, using the Pacific Biosciences (PacBio) Sequel single-molecule real-time (SMRT) platform (Ardui et al., 2018) as described previously (Nguyen et al., 2018). Assembly of sequence reads was done using SMRTLink (v.7.0.0.63985) software (Pacific Biosciences of California, Inc.) and circularization was carried out using Circlator v.1.5.5 (Hunt et al., 2015).

Methods for estimating overall genome relatedness such as digital DNA–DNA hybridization (dDDH) and average nucleotide identity (ANI) have replaced the outdated and error prone DNA–DNA hybridization (DDH) method for bacterial species delineation (Chun et al., 2018; Meier-Kolthoff and Göker, 2019; Meier-Kolthoff et al., 2022). We calculated dDDH values for strain T173<sup>T</sup> and reference strains of the genus *Ensifer* using the suite of algorithms implemented in the web-based Type Strain Genome Server (TYGS) (Meier-Kolthoff et al., 2013; Meier-Kolthoff and Göker, 2019; Meier-Kolthoff et al., 2022). The established dDDH threshold of 70% was used to delineate

species boundaries (Meier-Kolthoff and Göker, 2019). ANI values were estimated using the FastANI method (Jain et al., 2018) implemented in the K base web server (Arkin et al., 2018). The established ANI threshold of 95–96% was employed for species circumscription (Richter and Rosselló-Móra, 2009; Lee et al., 2016; Ciuffo et al., 2018).

To investigate horizontal transfer of symbiosis and related genes to strain T173<sup>T</sup>, comparative analysis of the complete sequence of the symbiosis plasmid (pT173b; accession no., CP083372) with the symbiosis megaplasmid (pWSM1115\_2; accession no. CP088111) of *E. medicae* strain WSM1115 (Reeve et al., 2014), was done using GenomeMatcher (Ohtsubo et al., 2008) and Geneious Prime 2023.0.4<sup>1</sup> software.

The BV-BRC web-based platform (Olson et al., 2023) was used to search for antibiotic resistance genes in the genomes of strain T173<sup>T</sup> and reference strains.

Ribosomal multilocus sequence typing employing 53 full-length gene sequences encoding bacterial ribosome protein subunits (*rps*) was used to assess phylogenetic relationships between novel strain T173<sup>T</sup> and 17 type strains of *Ensifer* species (Jolley et al., 2012). The Genome Comparator tool in the bacterial domain genome database of the BIGSdb software platform (Jolley and Maiden, 2010) was used to retrieve aligned concatenated sequences of *rps* genes from the genome sequences of novel strain T173<sup>T</sup> and reference strains. The best-fit substitution model was selected using ModelTest-NG (Darriba et al., 2020).

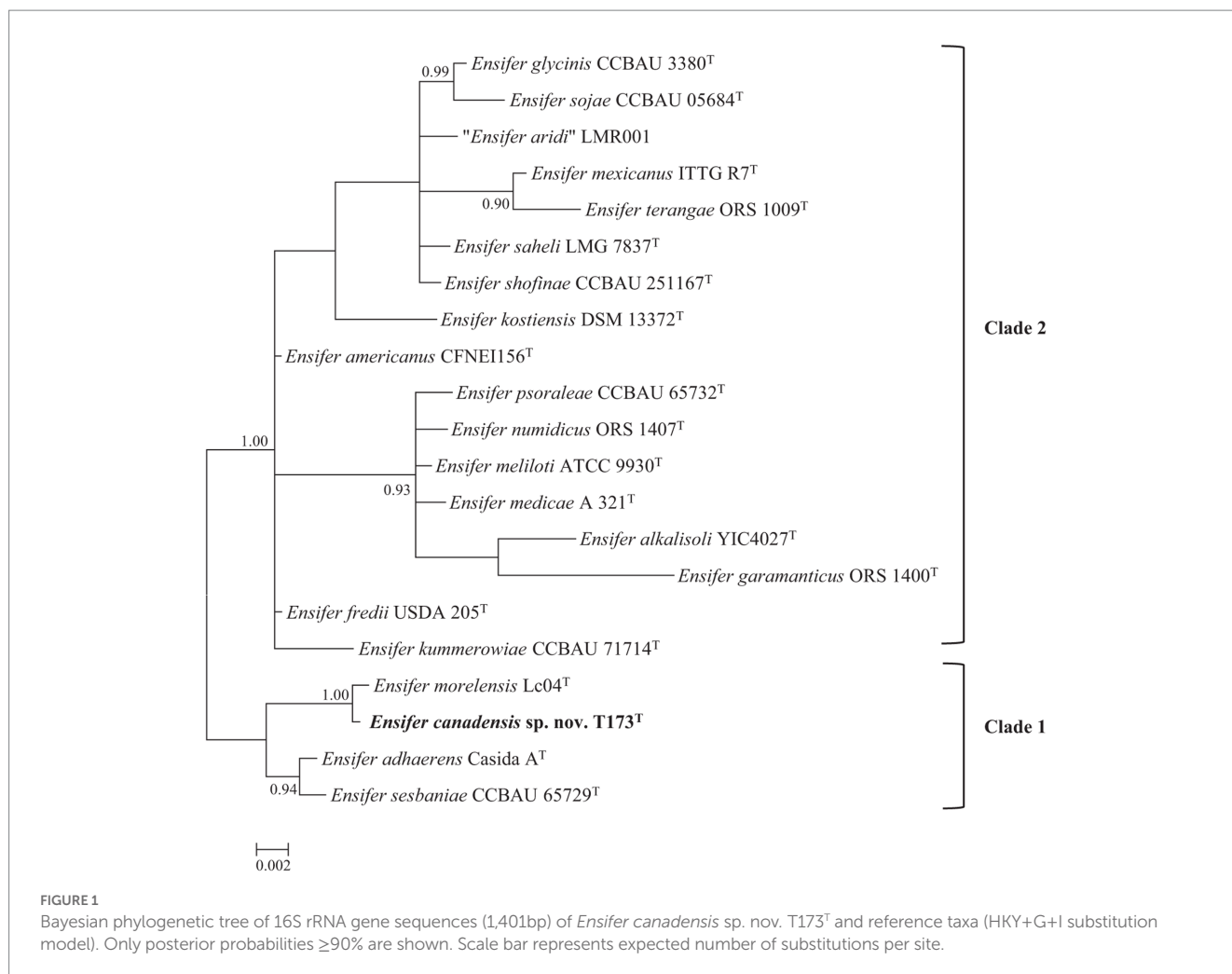
A phylogenomic tree of strain T173<sup>T</sup> and species type strains of the genus *Ensifer* was inferred with FastME 2.1.6.1 (Lefort et al., 2015) from Genome Blast Distance Phylogeny (GBDP) distances calculated from whole genome sequences using the suite of algorithms implemented in the TYGS (Type Strain Genome Server) web-based server (Meier-Kolthoff and Göker, 2019).

## Results and discussion

### Analyses of partial gene sequences

The 16S rRNA gene is universally present in all bacteria and sequences of this gene represent the most common house-keeping genetic marker that has been used in studies of bacterial taxonomy (Janda and Abbott, 2007). To reconstruct a 16S rRNA gene tree of type strains of all described *Ensifer* species ([Supplementary Table S1](#)) it was necessary to trim aligned sequence lengths to 1,401 bp. The Bayesian phylogenetic tree of 16S rRNA gene sequences ([Figure 1](#)) confirms placement of novel strain T173<sup>T</sup> in the genus *Ensifer* and shows division of species type strains into two highly supported clades (labelled 1 and 2) represented by the type strains of *E. adhaerens* and *E. meliloti*, respectively. These two clades correspond to the “non-symbiotic” and “symbiotic” clades defined by Fagorzi et al. (2020) and Kuzmanović et al. (2022) on the basis of genomic and phenotypic analyses. [Figure 1](#) further shows that novel strain T173<sup>T</sup> is placed in the clade represented by *E. adhaerens* (clade 1), with the type strain of *E. morelensis* as closest relative.

<sup>1</sup> <https://www.geneious.com>



As a cautionary note, it should be pointed out that the 16S rRNA gene is highly conserved and as such has limited usefulness as a taxonomic tool for delineating bacterial species (Richter and Rosselló-Móra, 2009; de Lajudie et al., 2019).

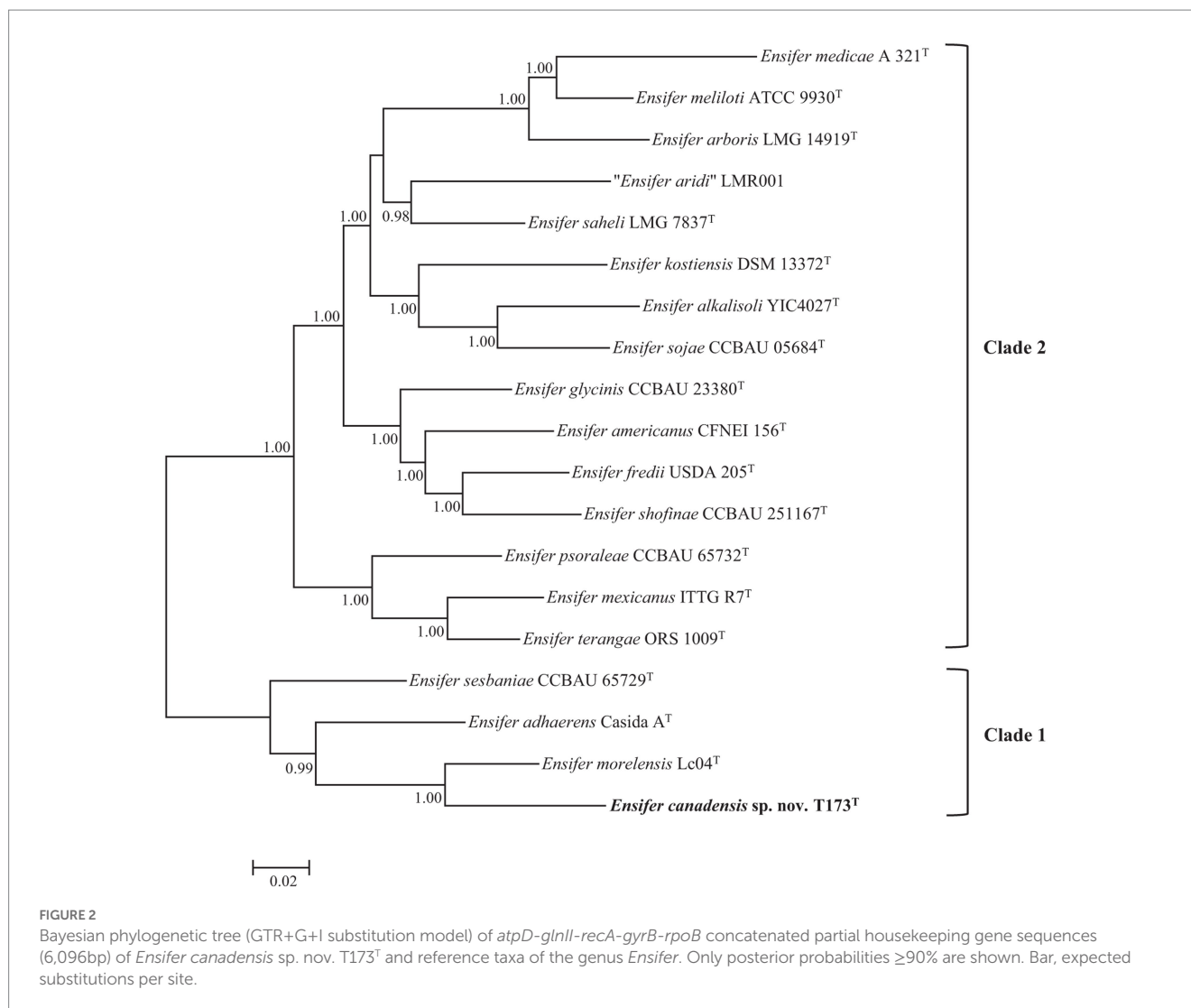
Phylogenetic analysis of multiple protein encoding (house-keeping) partial gene sequences (Multiple Locus Sequence Analysis, MLSA) has often been used to facilitate bacterial species delineation (e.g., Martens et al., 2008; de Lajudie et al., 2019). The topology of the Bayesian tree of five concatenated protein encoding (*atpD*, *glnII*, *gyrB*, *recA*, and *rpoB*) gene sequences (Figure 2), confirms the division of 18 species type strains of the genus *Ensifer* into two highly supported clades. Figure 2 also confirms placement of strain T173<sup>T</sup> in a highly supported lineage that is distinct from type strains of *Ensifer* species with *E. morelensis* as closest relative.

As one or more housekeeping gene sequences of several type strains of *Ensifer* species are not available in public databases, we used sequences of the *glnII* gene (the only protein encoding house-keeping gene for which sequences are available for type strains of 21 *Ensifer* species) in a supplementary phylogenetic analysis to verify the uniqueness of strain T173<sup>T</sup>. To include all 21 type strains in the phylogenetic tree it was necessary to trim the aligned lengths of *glnII* gene sequences to 615 bp. Consistent with the topology of trees of 16S rRNA (Figure 1) and five protein encoding (Figure 2) gene sequences, the tree of partial *glnII* sequences (Supplementary Figure S1) shows

placement of strain T173<sup>T</sup> in a highly supported lineage distinct from all 21 described species of the genus *Ensifer*.

## Genome analyses

A complete circularized genome sequence of strain T173<sup>T</sup> was generated in this study using PacBio Sequel technology. Estimated genome coverage was 419-fold with 96,490 polymerase reads and an average read length of 37,345 bp. Data for genome characteristics of strain T173<sup>T</sup> and reference strains of *Ensifer* for which complete genome sequences are available are shown in Table 1. The size of the complete genome sequence of strain T173<sup>T</sup> is 8,094,229 bp with an average DNA G+C content of 61.0 mol%. Six circularized replicons corresponding to a chromosome and five plasmids (pT173a through pT173e) were detected in the genome sequence of T173<sup>T</sup>. All plasmids harboured *repABC* genes encoding proteins involved in plasmid replication and segregation (Pinto et al., 2012); plasmid pT173c harboured two complete *repABC* gene copies. The six replicons of strain T173<sup>T</sup> have the following sizes (bp): 195,834 (pT173a); 204,278 (pT173b); 782,207 (pT173c); 946,878 (pT173d); 1,913,930 (pT173e), and 4,051,102 bp (chromosome). These data for plasmid number and size are consistent with the results of plasmid profile analysis (horizontal agarose gel electrophoresis) of strain T173<sup>T</sup> in our earlier



study (Bromfield et al., 2010): four plasmid bands were clearly resolved with the fourth band (exhibiting greatest mobility) likely representing a doublet consisting of co-migrating plasmids pT173a and pT173b with closely similar sizes.

In many bacterial genomes the ribosomal RNA (*rrn*) operon (encoding 16S, 23S, and 5S rRNAs) exists in multiple copies that are usually restricted to the chromosome (Acinas et al., 2004; Pei et al., 2010; Espejo and Plaza, 2018). Consistent with these reports multiple *rrn* copies were found on the chromosome of strain T173<sup>T</sup> and all *Ensifer* reference strains listed in Table 1. Of the five *rrn* copies detected in strain T173<sup>T</sup>, three were on the chromosome and single copies were on the large plasmids (megaplasmids), pT173d and pT173e. The closely related strain *E. adhaerens* Casida A<sup>T</sup> also possessed single *rrn* copies on megaplasmids (pCasidaAA and pCasidaAB) as well as three copies on the chromosome. The presence of *rrn* copies on both bacterial chromosomes and plasmids is not without precedence and has been found in *Vibrio parahaemolyticus* (Yamaichi et al., 1999), *Bacillus megaterium* (Kunnimalaiyaan et al., 2001) and soil isolates of *Paracoccus* species (Battermann et al., 2003). Moreover, in one unusual case a strain of *Aureimonas* sp. was reported

to have its sole *rrn* operon on a small (9.4kb) plasmid instead of the chromosome (Anda et al., 2015).

The occurrence of genetic heterogeneity among multiple *rrn* operons within genomes is well documented in bacteria (Yap et al., 1999; Tourova et al., 2001; Acinas et al., 2004).

Of the seven *Ensifer* strains in Table 1, only strain T173<sup>T</sup> and *E. medicae* strain WSM115 were found to exhibit intra-genomic variation in *rrn* copies. For strain T173<sup>T</sup> the two *rrn* copies on megaplasmids pT173d and pT173e differed (at the 23S rRNA locus) from the three copies on the chromosome whereas of the three *rrn* copies on the chromosome of *E. medicae* WSM115, one copy differed at the 16S rRNA locus and a second copy differed at the 23S rRNA locus. In all cases there were <1% nucleotide differences between operons which is consistent with the levels of intra-genomic divergence reported for most bacteria (Acinas et al., 2004).

The finding that the very large megaplasmid, pT173e, of strain T173<sup>T</sup> exhibited a G+C content value (61.2%) that was similar to the chromosome (61.8%), together with the presence of *repABC* genes and housekeeping genes (*rrn* operon) (Table 1) suggests that this

TABLE 1 Characteristics of complete genome sequences of *Ensifer canadensis* sp. nov. T173<sup>T</sup> and reference strains.

Characteristic	Strain						
	<i>E. canadensis</i> sp. nov. T173 <sup>T</sup>	<i>E. adhaerens</i> Casida A <sup>T</sup>	<i>E. alkalisoli</i> YIC4027 <sup>T</sup>	<i>E. medicae</i> WSM1115	<i>E. meliloti</i> 1021	<i>E. mexicanus</i> ITTG R7 <sup>T</sup>	<i>E. sojae</i> CCBAU 05684 <sup>T</sup>
Genome assembly quality (no. replicons)	Complete (6)	Complete (3)	Complete (3)	Complete (4)	Complete (3)	Complete (4)	Complete (3)
Genome size (bp)	8,094,229	7,267,502	6,128,433	7,063,185	6,691,694	7,141,863	6,094,027
Chromosome size (bp)	4,051,102	4,071,185	3,690,234	4,106,266	3,654,135	4,316,340	3,672,259
Plasmid size (bp)*	195,834; 204,278; 782,207; 946,878; 1,913,930	1,459,374; 1,736,943	456,454; 1,981,775	276,847; 1,128,391; 1,551,681	1,354,226; 1,683,333	436,172; 455,676; 1,933,675	410,255; 2,011,513
Genes (total)	7,705	6,937	5,797	6,832	6,314	6,641	5,764
CDs (total)	7,625	6,854	5,536	6,763	6,267	6,574	5,699
Plasmids ( <i>repABC</i> copies)	5 (6)	2 (2)	2 (2)	3 (3)	2 (2)	3 (3)	2 (2)
G + C content % (Entire genome sequence)	61.0	62.3	62.2	61.2	62.2	61.4	62.0
G + C content % (Chromosome)	61.8	62.8	62.6	61.5	62.7	61.8	62.4
G + C content % (Megaplasmids*)	60.1; 61.2 <sup>†</sup>	60.3; 62.8	62.3	59.9; 61.6	60.4; 62.4 <sup>†</sup>	61.6	61.9
G + C content % (Plasmids*)	58.0; 58.4; 58.6		59.3	60.6		58.6; 60.2	59.0
Replicon (rRNA operon copies)	Chromosome (3); pT173d (1); pT173e (1)	Chromosome (3); pCasidaAA (1); pCasidaAB (1)	Chromosome (3)	Chromosome (3)	Chromosome (3)	Chromosome (3)	Chromosome (3)
Symbiosis plasmid (pSym)	pT173b (204,278 bp)	none	Accession no: CP034911 (456,424 bp)	pWSM1115_2 (1,128,391 bp)	pSymA (1,354,226 bp)	pEmeITGR7b (455,676 bp)	pSJ05684a (410,255 bp)
Antibiotic Resistance Genes	51	42	39	40	45	43	37
tRNAs	61	64	56	56	55	54	52

\*Data are presented in order of increasing replicon size.

<sup>†</sup>Chromid.

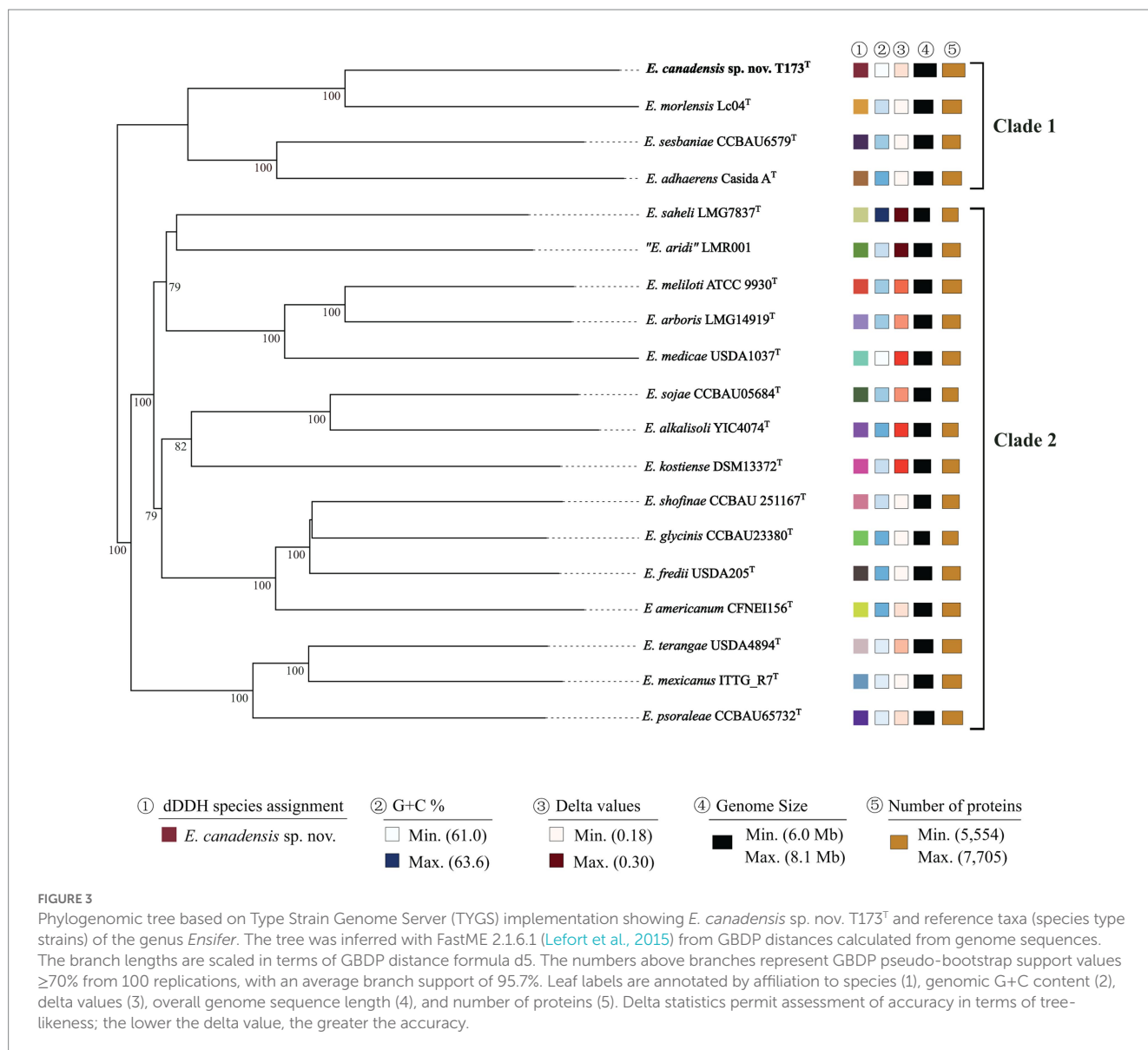
plasmid may represent a “chromid” (secondary chromosome) as defined by Harrison et al. (2010). Consistent with other studies (Nishida, 2012), smaller plasmids in the genome of strain T173<sup>T</sup> (i.e., plasmids other than the “chromid”) as well as those in the genomes of *Ensifer* reference strains (Table 1), consistently show G + C contents (58.0–60.25%) that are lower than that of the respective chromosomes (61.5–62.8%), suggesting that these plasmids may have been acquired from external sources by horizontal genetic exchange (Garcia-Vallvé et al., 2000).

To further verify the taxonomic status of strain T173<sup>T</sup> relative to type strains of *Ensifer* species, we carried out phylogenetic analyses based on: (1) MLSA of 53 concatenated full-length house-keeping gene sequences encoding bacterial ribosome protein subunits (*rps*) (Jolley et al., 2012), and, (2) TYGS analysis of whole genome sequences (Meier-Kolthoff and Göker, 2019). The topology of the phylogenetic tree of *rps* gene sequences (Supplementary Figure S2) as well as the TYGS tree of whole genome sequences (Figure 3) corroborate our findings that strain T173<sup>T</sup> is consistently placed in a highly supported

lineage distinct from type strains of described species of *Ensifer* with *E. morelensis* as closest relative.

Data for dDDH and ANI values for pair-wise comparisons of genome sequences of novel strain T173<sup>T</sup> with the three closest relatives (i.e., all species in the *E. adhaerens* clade) are presented in Supplementary Table S2. The highest dDDH and ANI values (35.7 and 88.6%, respectively) obtained in these comparisons are well below the respective threshold values of 70% and 95–96% for bacterial species circumscription. Based on these data, strain T173<sup>T</sup> is unambiguously classified as a novel *Ensifer* species with *E. morelensis* as closest relative.

Consistent with our earlier study (Bromfield et al., 2010), key nodulation (*nod*) and nitrogen fixation (*nif*) genes were detected on plasmid pT173b, representing the symbiosis plasmid (or pSym). To investigate the evolutionary history of symbiosis genes on pT173b we reconstructed Bayesian phylogenetic trees of concatenated *nodABC* and *nifHDK* full length gene sequences of strain T173<sup>T</sup> and 16 species type strains of the genus *Ensifer*. The topologies of the *nodABC* and *nifHDK* gene trees (Figures 4A,B) are closely similar with strain T173<sup>T</sup>

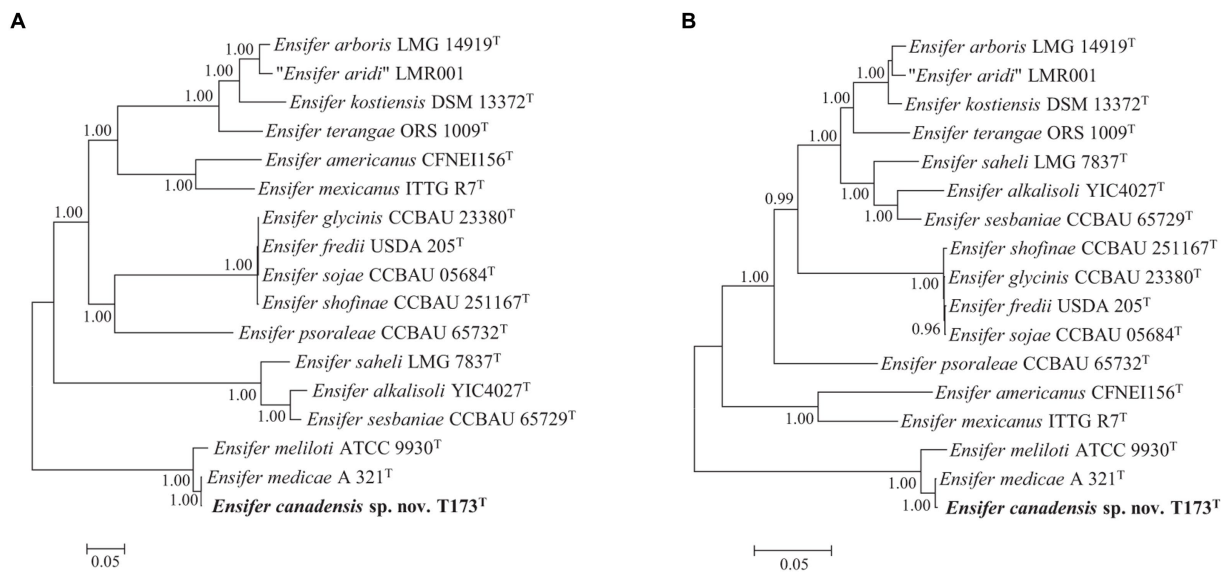


and the type strain of *E. medicae* (possessing almost identical *nod* and *nif* gene sequences) placed together in the same lineage. These results suggest that strain T173<sup>T</sup> placed in the *E. adhaerens* clade in the house keeping gene and genome trees (e.g., Figures 2, 3), has acquired its symbiosis genes by horizontal transfer from *E. medicae* placed in the phylogenetically distant *E. meliloti* clade in Figures 2, 3.

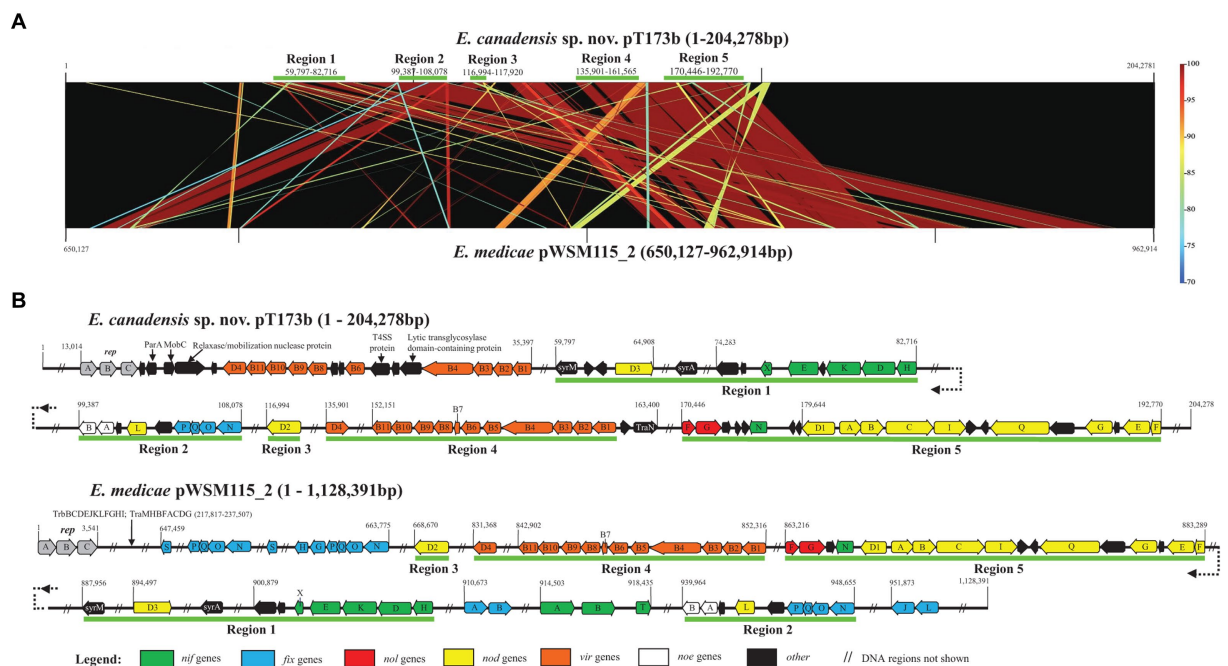
To further investigate the apparent horizontal acquisition of symbiosis and related genes by strain T173<sup>T</sup> we compared the complete sequence of symbiosis plasmid pT173b with the symbiosis gene region of the pSymA megaplasmid (=pWSM1115\_2) of *E. medicae* WSM1115, for which a complete genome sequence is available; pWSM1115\_2 was selected based on the best BLAST hit using the complete sequence of pT173b as query. Figure 5A shows that five regions of the pT173b sequence contain symbiosis and related genes that exhibit high sequence similarity (>98%) to genes on pWSM1115\_2. Regions 1, 2, 3, and 5 consist of nodulation (*nod*, *noe*, and *nol*) and nitrogen-fixation (*nif* and *fix*) genes that were apparently acquired from *E. medicae* (Figure 5B). However, key genes required for nitrogen-fixation such as *nifA*, *nifB*, *fixLJ*, *fixK*, and *fixGHIS*

(Lindström and Mousavi, 2020) present in the sequence of pWSM1115\_2 (*E. medicae*) were not found in the sequence of pT173b. Therefore, it was not unexpected to find that that strain T173<sup>T</sup> did not fix nitrogen in association with any of the host legumes tested in this and in the previous study (Bromfield et al., 2010). Region 4 of the pT173b sequence (Figure 5B), apparently acquired from *E. medicae*, contains genes encoding a conjugative type IV secretion system (T4SS) (Alvarez-Martinez and Christie, 2009; Rudder et al., 2014) that is incomplete and lacks a relaxase gene (Trokter and Waksman, 2018) necessary for conjugation. A second conjugative T4SS as well as plasmid replication and segregation (*repABC*) genes were detected on pT173b that are not shared with pWSM1115\_2. BLAST searches of the T4SS gene segment, *virB1* through *virB11* (pT173b co-ordinates: 23916–35,397 bp) and the *repABC* gene segment (pT173b co-ordinates: 13,014–16,648 bp) against NCBI genome sequence databases (Supplementary Table S3) suggest that these genes are most closely related to replication and conjugation systems in *Neorhizobium galegae*.

Collectively, these findings suggest that symbiosis plasmid pT173b is a recombinant plasmid with plasmid replication and conjugative



**FIGURE 4** Bayesian phylogenetic trees (HKY+G+I substitution model) of (A) *nodABC* concatenated gene sequences (2,352bp), and (B) *nifHDK* concatenated gene sequences (3,828bp) of *Ensifer canadensis* sp. nov. T173T and reference taxa of the genus *Ensifer*. Only posterior probabilities  $\geq 90\%$  are shown. Bar, expected substitutions per site.



**FIGURE 5** Comparison of the sequence of symbiosis plasmid pT173b with the symbiosis gene region of the pSymA megaplasmid (=pWSM1115\_2) of *Ensifer medicae* WSM1115 showing (A) five regions of the pT173b sequence that exhibit high sequence similarity (>98%) with genes on pWSM1115\_2 and (B) detail of regions 1 through 5 harbouring nodulation (*nod*, *noe*, and *nol*), nitrogen-fixation (*nif* and *fix*) and type IV secretion system (T4SS) genes.

T4SS gene regions that have their closest relatives in species like *N. galegae* and large segments of DNA containing symbiosis and T4SS genes that apparently originated from a symbiosis megaplasmid of *E. medicae*. It is noteworthy that host plants of *E. medicae* (*Medicago* and *Melilotus* species) and *N. galegae* (*Galega officinalis* and *Galega orientalis*) are native to the same region of Eurasia (Mediterranean

basin/Caucasus) (Quiros and Bauman, 1988; Rome et al., 1996; Steele et al., 2010; Ramirez-Bahena et al., 2015; Darbyshire et al., 2021) suggesting that the hybrid symbiosis plasmid, pT173b, may have had its origins in Eurasia.

Plasmid conjugation systems represent important mechanisms for the horizontal transmission of genetic information such as symbiotic



lifestyle and antibiotic resistance traits that facilitate the rapid evolution and adaptation of bacteria to new or changing environments (Chen et al., 2022). We detected genes involved in conjugation on four of the five plasmids in strain T173<sup>T</sup> (plasmids pT173a, b, c, and e). The plasmid conjugation systems in species of *Rhizobium*, *Agrobacterium* and *Ensifer*, as well as other members of the *Rhizobiaceae*, are classified into at least four different types (Ding and Hynes, 2009; Ding et al., 2013). Sequence comparisons of the conjugation systems and analysis of gene synteny in plasmids of strain T173<sup>T</sup> with well characterized members of each type of conjugation system show that pT173a carries a typical type I system, such as is found in *Agrobacterium* Ti plasmids, and is most likely regulated by quorum sensing as indicated by the presence of an autoinducer synthase gene (*traI*) and a *traR* orthologue (Gordon and Christie, 2014). The large plasmid or “chromid,” pT173e, was found to encode a complete type II system, similar to those found on plasmid pCFN42d of *Rhizobium etli* strain CFN42 and the pSymA plasmid of *Ensifer meliloti* strain 1021; it is common for this type of system, or a reduced version of it, to be present on megaplasmids or chromids of *Ensifer* species (Pérez-Mendoza et al., 2005; Blanca-Ordóñez et al., 2010). Two conjugation systems were detected on the symbiosis plasmid, pT173b, but one of these lacks a relaxase gene (as described above) and may not be functional. The second is similar to the system encoded on pT173c, and most closely resembles type IVb systems as described by Ding et al. (2013). We compared the relaxase (MobZ or TraA), coupling protein (usually designated TraG or VirD4) and best conserved protein from the T4SS (designated TrbE or VirB4, depending on the system) between these conjugation gene clusters on pT173b and pT173c and found that there is 62% identity between the relaxases (J3R84\_28450 on pT173b and J3R84\_31165 on pT173c), 77% identity between the coupling proteins (J3R84\_28460 and J3R84\_30945, respectively), and 91% identity between the VirB4/TrbE proteins (J3R84\_28515 and J3R84\_30995, respectively). These conjugation gene clusters on pT173b and pT173c are organized like those on plasmid pAtS4a of *Agrobacterium vitis* strain S4 and plasmid pSmed03 of *Ensifer medicae* strain WSM419 (Ding et al., 2013), which were among the earliest recognized members of the type IVb system (Giusti et al., 2012).

Based on our genome analyses, it seems likely that four of the plasmids of T173<sup>T</sup> are mobile. As such, strain T173<sup>T</sup> represents a valuable resource for studies on plasmid self transmissibility. Indeed, mobile plasmids are highly prevalent in the *Rhizobiaceae*, and there appears to be strong selection for mobility of symbiosis genes between strains (Wardell et al., 2022). This selection for transmission of plasmids harbouring symbiosis genes might explain the common occurrence of mosaic plasmids (Wardell et al., 2022), arising from recombination between segments of multiple plasmids, as seems to have been the case for the symbiosis plasmid, pT173b of strain T173<sup>T</sup>.

## Phenotypic characterization

Colonies of strain T173<sup>T</sup> are raised, mucilaginous, circular and off-white coloured with diameters ~2–3 mm after 3 days growth at 28°C on yeast extract-mannitol (YEM) agar medium. Bacterial cells are Gram-stain-negative rods and based on electron microscopy exhibit multiple (peritrichous) flagella (Supplementary Figure S3).

Data for growth characteristics of strain T173<sup>T</sup> and seven reference strains of the genus *Ensifer* are shown in Supplementary Table S4.

Strain T173<sup>T</sup> produces an acidic reaction on YEM agar after 21 days growth at 28°C typical of other members of the genus *Ensifer*. Strain T173<sup>T</sup>, like close relatives, *E. morelensis* Lc04<sup>T</sup> and *E. adhaerens* Casida A<sup>T</sup>, shows growth in the presence of 2% NaCl, at pH 5 and pH10, and, at temperatures of 10°C and 37°C on YEM agar after 2 days incubation. Strain T173<sup>T</sup> shows good growth in LB broth medium after 48 h at 30°C (Supplementary Table S4; Supplementary Figure S4) which is considered typical of members of the *E. adhaerens* clade (Fagorzi et al., 2020). Unlike *E. adhaerens* Casida A<sup>T</sup>, novel strain T173<sup>T</sup> and closest relative, *E. morelensis* Lc04<sup>T</sup>, did not exhibit tracking activity (i.e., predation) (Martin, 2002) of *Micrococcus luteus* JCM 1464<sup>T</sup> grown on 0.1x heart infusion agar medium supplemented with 0.1% glucose at 28°C after 7 days.

The results of 70 carbon source utilization and 18 chemical sensitivity tests using phenotype microarrays (Biolog) are given in Supplementary Table S5. On the basis of multiple tests, strain T173<sup>T</sup> could be readily differentiated from all seven type strains of *Ensifer* species that we tested. In particular, T173<sup>T</sup> could be distinguished from its closest relative, *E. morelensis* Lc04<sup>T</sup>, based on differential utilization of 20 carbon sources and sensitivity to three chemical compounds. Moreover, in our previous study (Bromfield et al., 2010), strain T173<sup>T</sup> was shown to be unusual in that it was highly resistant to multiple antibiotics including carbenicillin (>1,000 µg ml<sup>-1</sup>), kanamycin (>100 µg ml<sup>-1</sup>) and neomycin (~100 µg ml<sup>-1</sup>), similar to its close relative, *E. morelensis* Lc04<sup>T</sup> (Wang et al., 2002). In this connection, multiple antibiotic resistance genes (Yoneyama and Katsumata, 2006) were detected in the genome of strain T173<sup>T</sup> (Table 1) including genes encoding enzymes that inactivate beta-lactam antibiotics (e.g., carbenicillin) and amino-glycoside antibiotics such as kanamycin and neomycin (Supplementary Table S6).

Data for the fatty acid profiles of T173<sup>T</sup> and reference strains are shown in Supplementary Table S7. Consistent with the results of other studies of the genus *Ensifer* (Tighe et al., 2000), fatty acids 16:0, 18:0, 18:0 3OH, 19:0 cyclo ω8c, 12:0 aldehyde/? (summed feature 2) and 18:1 ω6c/18:1 ω7c (summed feature 8) were common to strain T173<sup>T</sup> and reference strains of seven *Ensifer* species. Fatty acid 18:1 ω7c 11-methyl was predominant (>15%) only in strain T173<sup>T</sup> and close relatives, *E. morelensis* Lc04<sup>T</sup> and *E. adhaerens* Casida A<sup>T</sup>. It is noteworthy that the overall profile (24 fatty acids including minor fatty acids) distinguished strain T173<sup>T</sup> from all seven reference strains of *Ensifer* species.

Plant tests in this work and in the previous study (Bromfield et al., 2010) showed that strain T173<sup>T</sup> elicited numerous small white nodules on roots of *Medicago sativa* (alfalfa), *Melilotus albus* (white sweet clover), *Medicago polymorpha* (burr medic) and *Macroptilium atropurpureum* (siratiro), but did not fix nitrogen.

## Description of *Ensifer canadensis* sp. nov.

*Ensifer canadensis* sp. nov. *ca.na.den'sis*. N.L. masc./fem. Adj. canadensis, of or belonging to Canada, from where the organism was isolated.

Cells are Gram-stain-negative, aerobic, non-spore-forming rods with multiple flagella. The type strain produces colonies that are raised, mucilaginous, circular and off-white coloured with diameters ~2–3 mm after 3 days growth at 28°C on YEM agar medium.

Grows in the presence of 2% NaCl and at pH 5 and pH10 (optimum ~pH 6.0–7.0) after 2 days at 28°C on YEM agar medium.

The type strain grows at temperatures of 10°C and 37°C (optimal at ~28°C) after 2 days on YEM agar. Grows in LB broth medium after 2 days at 30°C. Produces an acidic reaction on YEM agar after 21 days growth at 28°C. Does not show predatory activity against *Micrococcus luteus* after 7 days incubation at 28°C on 0.1x heart infusion agar medium supplemented with 0.1% glucose. Predominant fatty acids are 16:0, 18:0, 18:1 ω7c 11-methyl, 12:0 aldehyde/? (summed feature 2) and 18:1 ω6c/18:1 ω7c (summed feature 8).

The type strain utilizes 35 carbon sources including sucrose, D-raffinose, α-D-lactose, N-acetyl-β-D-mannosamine, D-mannose, L-fucose, glycerol, D-fructose-6-PO<sub>4</sub>, D-aspartic acid, L-glutamic acid, pectin, D-gluconic acid, D-glucuronic acid, propionic acid and formic acid. Does not utilize 35 carbon sources including dextrin, stachyose, β-methyl-D-glucoside, D-salicin, N-acetyl-D-galactosamine, D-galactose, inosine, glycyl-L-proline, L-arginine, L-aspartic acid, L-histidine, L-serine, D-lactic acid methyl ester, L-lactic acid, and citric acid. Resistant to troleandomycin, rifamycin SV, lincomycin, tetrazolium violet, tetrazolium blue and aztreonam. Susceptible to 12 chemical compounds including 1% sodium lactate, guanidine HCl, niaproof 4, vancomycin, nalidixic acid and potassium tellurite. The type strain is highly resistant to carbenicillin (>1,000 μg ml<sup>-1</sup>), kanamycin (>100 μg ml<sup>-1</sup>) and neomycin (~100 μg ml<sup>-1</sup>).

The type strain elicits numerous small white nodules (ineffective for nitrogen fixation) on roots of *Medicago sativa*, *Medicago lupulina*, *Medicago polymorpha*, *Melilotus albus*, and *Macroptilium atropurpureum*.

The type strain, T173<sup>T</sup> (= LMG 32374<sup>T</sup> = HAMBI 3766<sup>T</sup>) was isolated from a root-nodule of a *Melilotus albus* plant grown at a field site in Ottawa, Ontario, Canada. The whole genome shotgun project for *Ensifer canadensis* strain T173<sup>T</sup> was deposited at DDBJ/ENA/GenBank under the accession numbers CP083370–CP083375. Raw PacBio data was deposited in the NCBI Sequence Read Archive under the BioProject accession number PRJNA713338. The genome of the type strain contains a chromosome and five plasmids one of which is a symbiosis plasmid harbouring nodulation and nitrogen fixation genes. The DNA G + C content of the type strain is 61.0 mol% and the genome size is 8,094,229 bp.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repository and accession number(s) can be found in the article/[Supplementary material](#).

## References

- Acinas, S. G., Marcelino, L. A., Klepac-Ceraj, V., and Polz, M. F. (2004). Divergence and redundancy of 16S rRNA sequences in genomes with multiple *rrn* operons. *J. Bacteriol.* 186, 2629–2635. doi: 10.1128/JB.186.9.2629-2635.2004
- Altekar, G., Dwarkadas, S., Huelsenbeck, J. P., and Ronquist, F. (2004). Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. *Bioinformatics* 20, 407–415. doi: 10.1093/bioinformatics/btg427
- Alvarez-Martinez, C. E., and Christie, P. J. (2009). Biological diversity of prokaryotic type IV secretion systems. *Microbiol. Mol. Biol. Rev.* 73, 775–808. doi: 10.1128/MMBR.00023-09
- Anda, M., Ohtsubo, Y., Okubo, T., Sugawara, M., Nagata, Y., Tsuda, M., et al. (2015). Bacterial clade with the ribosomal RNA operon on a small plasmid rather than the chromosome. *Proc. Natl. Acad. Sci. U. S. A.* 112, 14343–14347. doi: 10.1073/pnas.1514326112
- Ardui, S., Ameur, A., Vermeesch, J. R., and Hestand, M. S. (2018). Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. *Nucleic Acids Res.* 46, 2159–2168. doi: 10.1093/nar/gky066
- Arkin, A. P., Cottingham, R. W., Henry, C. S., Harris, N. L., Stevens, R. L., Maslov, S., et al. (2018). KBase: the United States Department of Energy Systems Biology Knowledgebase. *Nat. Biotechnol.* 36, 566–569. doi: 10.1038/nbt.4163
- Battermann, A., Disse-Krömker, C., and Dreiseikelmann, B. (2003). A functional plasmid-borne *rrn* operon in soil isolates belonging to the genus *Paracoccus*. *Microbiology* 149, 3587–3593. doi: 10.1099/mic.0.26608-26600
- Beringer, J. E. (1974). R Factor Transfer in *Rhizobium leguminosarum*. *J. Gen. Microbiol.* 84, 188–198. doi: 10.1099/00221287-84-1-188
- Blanca-Ordóñez, H., Oliva-García, J. J., Pérez-Mendoza, D., Soto, M. J., Olivares, J., Sanjuán, J., et al. (2010). pSymA-dependent mobilization of the *Sinorhizobium meliloti* pSymB megaplasmid. *J. Bacteriol.* 192, 6309–6312. doi: 10.1128/JB.00549-10
- Bromfield, E. S., Butler, G., and Barran, L. R. (2001). Temporal effects on the composition of a population of *Sinorhizobium meliloti* associated with *Medicago sativa* and *Melilotus alba*. *Can. J. Microbiol.* 47, 567–573. doi: 10.1139/w01-034
- Bromfield, E. S. P., Tambong, J. T., Cloutier, S., Prévost, D., Laguerre, G., van Berkum, P., et al. (2010). *Ensifer*, *Phyllobacterium* and *Rhizobium* species occupy nodules

## Author contributions

EB co-ordinated the project, received the funding, and wrote the draft manuscript. SC and EB carried out the experiments. SC, EB, and MH analyzed the data. All authors contributed to the article and approved the submitted version.

## Funding

This research was supported by grants J-002272 and J-002295 from Agriculture and Agri-Food Canada.

## Acknowledgments

The authors are thankful to Keith Hubbard of the Microscopy Centre, Agriculture and Agri-Food Canada, Ottawa, Canada, for preparing electron microscope images.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2023.1195755/full#supplementary-material>

- of *Medicago sativa* (alfalfa) and *Melilotus alba* (sweet clover) grown at a Canadian site without a history of cultivation. *Microbiology* 156, 505–520. doi: 10.1099/mic.0.034058-0
- Buck, J. D. (1982). Nonstaining KOH method for determination of gram reactions of marine bacteria. *Appl. Environ. Microbiol.* 44, 992–993. doi: 10.1128/aem.44.4.992-993.1982
- Casida, L. E. (1982). *Ensifer adhaerens* gen. Nov., sp. nov.: a bacterial predator of bacteria in soil. *Int. J. Syst. Bacteriol.* 32, 339–345. doi: 10.1099/00207713-32-3-339
- Chen, C.-Y., Fuqua, C., Jackson, C. R., Kadlec, K., and Top, E. M. (2022). Editorial: plasmid transfer-mechanisms, ecology, evolution and applications. *Front. Microbiol.* 13:993628. doi: 10.3389/fmicb.2022.993628
- Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, D. R., da Costa, M. S., et al. (2018). Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 68, 461–466. doi: 10.1099/ijsem.0.002516
- Ciufo, S., Kannan, S., Sharma, S., Badretin, A., Clark, K., Turner, S., et al. (2018). Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. *Int. J. Syst. Evol. Microbiol.* 68, 2386–2392. doi: 10.1099/ijsem.0.002809
- Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., et al. (2014). Ribosomal database project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* 42, D633–D642. doi: 10.1093/nar/gkt1244
- Darbyshire, S. J., Francis, A., Bromfield, E. S. P., and Mechanda, S. (2021). The biology of Canadian weeds: 158. *Galega officinalis* L. *Can. J. Plant Sci.* 102, 160–185. doi: 10.1139/cjps-2020-0327
- Darriba, D., Posada, D., Kozlov, A. M., Stamatakis, A., Morel, B., and Flouri, T. (2020). Model test-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Mol. Biol. Evol.* 37, 291–294. doi: 10.1093/molbev/msz189
- de Lajudie, P. M., Andrews, M., Ardley, J., Eardly, B., Jumas-Bilak, E., Kuzmanović, N., et al. (2019). Minimal standards for the description of new genera and species of rhizobia and agrobacteria. *Int. J. Syst. Evol. Microbiol.* 69, 1852–1863. doi: 10.1099/ijsem.0.003426
- de Lajudie, P., Willems, A., Pot, B., Dewettinck, D., Maestrojuan, G., Neyra, M., et al. (1994). Polyphasic taxonomy of rhizobia: emendation of the genus *Sinorhizobium* and description of *Sinorhizobium meliloti* comb. nov., *Sinorhizobium saheii* sp. nov., and *Sinorhizobium teranga* sp. nov. *Int. J. Syst. Bacteriol.* 44, 715–733. doi: 10.1099/00207713-44-4-715
- Ding, H., and Hynes, M. F. (2009). Plasmid transfer systems in the rhizobia. *Can. J. Microbiol.* 55, 917–927. doi: 10.1139/w09-056
- Ding, H., Yip, C. B., and Hynes, M. F. (2013). Genetic characterization of a novel rhizobial plasmid conjugation system in *Rhizobium leguminosarum* bv. *Viciae* strain VF39SM. *J. Bacteriol.* 195, 328–339. doi: 10.1128/JB.01234-12
- Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform.* 5:113. doi: 10.1186/1471-2105-5-113
- Espejo, R. T., and Plaza, N. (2018). Multiple ribosomal RNA operons in bacteria: their concerted evolution and potential consequences on the rate of evolution of their 16S rRNA. *Front. Microbiol.* 9:1232. doi: 10.3389/fmicb.2018.01232
- Fagorzi, C., Ilie, A., Decorosi, F., Cangioli, L., Viti, C., Mengoni, A., et al. (2020). Symbiotic and nonsymbiotic members of the genus *Ensifer* (syn. *Sinorhizobium*) are separated into two clades based on comparative genomics and high-throughput phenotyping. *Genome Biol. Evol.* 12, 2521–2534. doi: 10.1093/gbe/evaa221
- García-Vallvé, S., Romeu, A., and Palau, J. (2000). Horizontal gene transfer in bacterial and archaeal complete genomes. *Genome Res.* 10, 1719–1725. doi: 10.1101/gr.130000
- Giusti, M. A., Pistorio, M., Lozano, J. L., Torres Tejerizo, G. A., Salas, M. E., Martini, M. C., et al. (2012). Genetic and functional characterization of a yet unclassified rhizobial Dtr (DNA-transfer-and-replication) region from a ubiquitous plasmid conjugial system present in *Sinorhizobium meliloti*, in *Sinorhizobium medicae*, and in other nonrhizobial gram-negative bacteria. *Plasmid* 67, 199–210. doi: 10.1016/j.plasmid.2011.12.010
- Gordon, J. E., and Christie, P. J. (2014). The *Agrobacterium* Ti plasmids. *Microbiol. Spectr.* 2, 1–18. doi: 10.1128/microbiolspec.PLAS-0010-2013
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010
- Harrison, P. W., Lower, R. P. J., Kim, N. K. D., and Young, J. P. W. (2010). Introducing the bacterial ‘chromid’: not a chromosome, not a plasmid. *Trends Microbiol.* 18, 141–148. doi: 10.1016/j.tim.2009.12.010
- Hunt, M., Silva, N. D., Otto, T. D., Parkhill, J., Keane, J. A., and Harris, S. R. (2015). Circulator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol.* 16:294. doi: 10.1186/s13059-015-0849-0
- Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9:5114. doi: 10.1038/s41467-018-07641-9
- Janda, J. M., and Abbott, S. L. (2007). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J. Clin. Microbiol.* 45, 2761–2764. doi: 10.1128/JCM.01228-07
- Jolley, K. A., Bliss, C. M., Bennett, J. S., Bratcher, H. B., Brehony, C., Colles, F. M., et al. (2012). Ribosomal multilocus sequence typing: universal characterization of bacteria from domain to strain. *Microbiology* 158, 1005–1015. doi: 10.1099/mic.0.055459-0
- Jolley, K. A., and Maiden, M. C. J. (2010). BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinform.* 11:595. doi: 10.1186/1471-2105-11-595
- Kumar, H. K., Gan, H. M., Tan, M. H., Eng, W. W., Barton, H. A., Hudson, A. O., et al. (2017). Genomic characterization of eight *Ensifer* strains isolated from pristine caves and a whole genome phylogeny of *Ensifer* (*Sinorhizobium*). *J. Genom.* 5, 12–15. doi: 10.7150/jgen.17863
- Kunnimalaiyaan, M., Stevenson, D. M., Zhou, Y., and Vary, P. S. (2001). Analysis of the replicon region and identification of an rRNA operon on pBM400 of *Bacillus megaterium* QM B1551. *Mol. Microbiol.* 39, 1010–1021. doi: 10.1046/j.1365-2958.2001.02292.x
- Kuzmanović, N., Fagorzi, C., Mengoni, A., Lassalle, F., and Dincenzo, G. C. (2022). Taxonomy of Rhizobiaceae revisited: proposal of a new framework for genus delimitation. *Int. J. Syst. Evol. Microbiol.* 72:005243. doi: 10.1099/ijsem.0.005243
- Lee, I., Ouk Kim, Y., Park, S. C., and Chun, J. (2016). OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *Int. J. Syst. Evol. Microbiol.* 66, 1100–1103. doi: 10.1099/ijsem.0.000760
- Lefort, V., Desper, R., and Gascuel, O. (2015). Fast ME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol. Biol. Evol.* 32, 2798–2800. doi: 10.1093/molbev/msv150
- Lindström, K., and Mousavi, S. A. (2020). Effectiveness of nitrogen fixation in rhizobia. *Microb. Biotechnol.* 13, 1314–1335. doi: 10.1111/1751-7915.13517
- Martens, M., Dawyndt, P., Coopman, R., Gillis, M., De Vos, P., and Willems, A. (2008). Advantages of multilocus sequence analysis for taxonomic studies: a case study using 10 housekeeping genes in the genus *Ensifer* (including former *Sinorhizobium*). *Int. J. Syst. Evol. Microbiol.* 58, 200–214. doi: 10.1099/ijse.0.65392-0
- Martin, M. O. (2002). Predatory prokaryotes: an emerging research opportunity. *J. Mol. Microbiol. Biotechnol.* 4, 467–477.
- Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P., and Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinform.* 14:60. doi: 10.1186/1471-2105-14-60
- Meier-Kolthoff, J. P., Carbasse, J. S., Peinado-Olarte, R. L., and Göker, M. (2022). TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res.* 50, D801–D807. doi: 10.1093/nar/gkab902
- Meier-Kolthoff, J. P., and Göker, M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun.* 10:2182. doi: 10.1038/s41467-019-10210-3
- Miller, M. A., Pfeiffer, W., and Schwartz, T. (2010). “Creating the CIPRES science gateway for inference of large phylogenetic trees,” In Proceedings of the Gateway Computing Environments Workshop, New Orleans, LA.
- Nguyen, H. D. T., Cloutier, S., and Bromfield, E. S. P. (2018). Complete genome sequence of *Bradyrhizobium ottawaense* OO99<sup>T</sup>, an efficient nitrogen-fixing symbiont of soybean. *Microbiol. Resour. Announc.* 7, e01477–e01418. doi: 10.1128/MRA.01477-18
- Nishida, H. (2012). Comparative analyses of base compositions, DNA sizes, and dinucleotide frequency profiles in archaeal and bacterial chromosomes and plasmids. *Int. J. Evol. Biol.* 2012:342482. doi: 10.1155/2012/342482
- Ohtsubo, Y., Ikeda-Ohtsubo, W., Nagata, Y., and Tsuda, M. (2008). Genome matcher: a graphical user interface for DNA sequence comparison. *BMC Bioinform.* 9:376. doi: 10.1186/1471-2105-9-376
- Olson, R. D., Assaf, R., Brettn, T., Conrad, N., Cucinell, C., Davis, J. J., et al. (2023). Introducing the bacterial and viral bioinformatics resource center (BV-BRC): a resource combining PATRIC, IRD and ViPR. *Nucl. Acids Res.* 51, D678–D689. doi: 10.1093/nar/gkac1003
- Oren, A., and Garrity, G. M. (2022). Notification that new names of prokaryotes, new combinations, and new taxonomic opinions have appeared in volume 72, part 3 of the IJSEM. *Int. J. Syst. Evol. Microbiol.* 72:005393. doi: 10.1099/ijsem.0.005393
- Pei, A. Y., Oberdorf, W. E., Nossa, C. W., Agarwal, A., Chokshi, P., Gerz, E. A., et al. (2010). Diversity of 16S rRNA genes within individual prokaryotic genomes. *Appl. Environ. Microbiol.* 76, 3886–3897. doi: 10.1128/AEM.02953-09
- Pérez-Mendoza, D., Sepúlveda, E., Pando, V., Muñoz, S., Nogales, J., Olivares, J., et al. (2005). Identification of the *rctA* gene, which is required for repression of conjugative transfer of rhizobial symbiotic megaplasmids. *J. Bacteriol.* 187, 7341–7350. doi: 10.1128/JB.187.21.7341-7350.2005
- Pinto, U. M., Pappas, K. M., and Winans, S. C. (2012). The ABCs of plasmid replication and segregation. *Nat. Rev. Microbiol.* 10, 755–765. doi: 10.1038/nrmicro2882
- Quiros, C. F., and Baughan, G. R. (1988). “The genus *Medicago* and the origin of the *Medicago sativa* complex” in *Alfalfa and alfalfa improvement*. eds. A. A. Hanson, D. K. Barnes and R. R. Hill
- Ramírez-Bahena, H. M., Vargas, M., Martín, M., Tejedor, C., Velázquez, E., and Peix, Á. (2015). Alfalfa microsymbionts from different ITS and *nodC* lineages of *Ensifer meliloti* and *Ensifer medicae* symbiovar *meliloti* establish efficient symbiosis with alfalfa

- in Spanish acid soils. *Appl. Microbiol. Biotechnol.* 99, 4855–4865. doi: 10.1007/s00253-014-6347-6
- Reeve, W., Ballard, R., Howieson, J., Drew, E., Tian, R., Bräu, L., et al. (2014). Genome sequence of *Ensifer medicae* strain WSM1115; an acid-tolerant *Medicago*-nodulating microsymbiont from Samothraki, Greece. *Stand. Genomic Sci.* 9, 514–526. doi: 10.4056/sigs.4938652
- Richter, M., and Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. U. S. A.* 106, 19126–19131. doi: 10.1073/pnas.0906412106
- Rome, S., Fernandez, M. P., Brunel, B., Normand, P., and Cleyet-Marel, J. C. (1996). *Sinorhizobium medicae* sp. nov., isolated from annual *Medicago* spp. *Int. J. Syst. Bacteriol.* 46, 972–980. doi: 10.1099/00207713-46-4-972
- Rudder, S., Doohan, F., Creevey, C. J., Wendt, T., and Mullins, E. (2014). Genome sequence of *Ensifer adhaerens* OV14 provides insights into its ability as a novel vector for the genetic transformation of plant genomes. *BMC Genomics* 15:268. doi: 10.1186/1471-2164-15-268
- Sasser, M. (1990). *Identification of Bacteria by gas Chromatography of Cellular Fatty acids MIDI Technical Note 101*. Newark, DE: MIDI Inc.
- Steele, K. P., Ickert-Bond, S. M., Zarre, S., and Wojciechowski, M. F. (2010). Phylogeny and character evolution in *Medicago* (Leguminosae): evidence from analyses of plastid trnK/matK and nuclear GA3ox1 sequences. *Am. J. Bot.* 97, 1142–1155. doi: 10.3732/ajb.1000009
- Tang, J., Bromfield, E. S. P., Rodrigue, N., Cloutier, S., and Tambong, J. T. (2012). Microevolution of symbiotic *Bradyrhizobium* populations associated with soybeans in East North America. *Ecol. Evol.* 2, 2943–2961. doi: 10.1002/ece3.404
- Tighe, S. W., de Lajudie, P., Dipietro, K., Lindstrom, K., Nick, G., and Jarvis, B. D. (2000). Analysis of cellular fatty acids and phenotypic relationships of *Agrobacterium*, *Bradyrhizobium*, *Mesorhizobium*, *Rhizobium* and *Sinorhizobium* species using the Sherlock microbial identification system. *Int. J. Syst. Evol. Microbiol.* 50, 787–801. doi: 10.1099/00207713-50-2-787
- Tourova, T. P., Kuznetsov, B. B., Novikova, E. V., Poltarau, A. B., and Nazina, T. N. (2001). Heterogeneity of the nucleotide sequence of the 16S rRNA genes of the type strain of *Desulfotomaculum kuznetsovii*. *Microbiology* 70, 678–684. doi: 10.1023/A:1013135831669
- Trokter, M., and Waksman, G. (2018). Translocation through the conjugative type IV secretion system requires unfolding of its protein substrate. *J. Bacteriol.* 200, e00615–e00617. doi: 10.1128/JB.00615-17
- Vincent, J. M. (1970). *A manual for the practical study of the root-nodule bacteria*. Blackwell Scientific Publishers, Oxford. 164.
- Wang, E. T., Tan, Z. Y., Willems, A., Fernández-López, M., Reinhold-Hurek, B., and Martínez-Romero, E. (2002). *Sinorhizobium morelense* sp. nov., a *Leucaena leucocephala*-associated bacterium that is highly resistant to multiple antibiotics. *Int. J. Syst. Evol. Microbiol.* 52, 1687–1693. doi: 10.1099/00207713-52-5-1687
- Wardell, G. E., Hynes, M. F., Young, J. P. W., and Harrison, E. (2022). Why are rhizobial symbiosis genes mobile. *Philos. Trans. R. Soc. London Ser. B* 377:20200471. doi: 10.1098/rstb.2020.0471
- Yamaichi, Y., Iida, T., Park, K. S., Yamamoto, K., and Honda, T. (1999). Physical and genetic map of the genome of *Vibrio parahaemolyticus*: presence of two chromosomes in *Vibrio* species. *Mol. Microbiol.* 31, 1513–1521. doi: 10.1046/j.1365-2958.1999.01296.x
- Yap, W. H., Zhang, Z., and Wang, Y. (1999). Distinct types of rRNA operons exist in the genome of the actinomycete *Thermomonospora chromogena* and evidence for horizontal gene transfer of an entire rRNA operon. *J. Bacteriol.* 181, 5201–5209. doi: 10.1128/JB.181.17.5201-5209.1999
- Yoneyama, H., and Katsumata, R. (2006). Antibiotic resistance in bacteria and its future for novel antibiotic development. *Biosci. Biotechnol. Biochem.* 70, 1060–1075. doi: 10.1271/bbb.70.1060
- Yu, X., Cloutier, S., Tambong, J. T., and Bromfield, E. S. P. (2014). *Bradyrhizobium ottawaense* sp. nov., a symbiotic nitrogen fixing bacterium from root nodules of soybeans in Canada. *Int. J. Syst. Evol. Microbiol.* 64, 3202–3207. doi: 10.1099/ijs.0.065540-0