AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology®

PLANT MICROBIOLOGY

Biogeography of a Novel *Ensifer meliloti* Clade Associated with the Australian Legume *Trigonella suavissima*

Bertrand Eardly,[a] Patrick Elia,[b] John Brockwell,[c] Daniel Golemboski,[d] Peter van Berkum[b]

Biology Department, Penn State University, Berks College, Reading, Pennsylvania, USA[a]; Soybean Genomics and Improvement Laboratory, Agriculture Research Service, U.S. Department of Agriculture, Beltsville, Maryland, USA[b]; Commonwealth Scientific and Industrial Research Organization (CSIRO) Agriculture and Food, Black Mountain Laboratories, Canberra, ACT, Australia[c]; Medical Laboratory Science Department, Bellarmine University, Louisville, Kentucky, USA[d]

ABSTRACT Here, we describe a novel clade within *Ensifer meliloti* and consider how geographic and ecological isolation contributed to the limited distribution of this group. Members of the genus *Ensifer* are best known for their ability to form nitrogen-fixing symbioses with forage legumes of three related genera, *Medicago* L., *Melilotus* Mill., and *Trigonella* L., which are members of the tribe Trifolieae. These legumes have a natural distribution extending from the Mediterranean Basin through western Asia, where there is an unsurpassed number of species belonging to these genera. *Trigonella suavissima* L. is unusual in that it is the only species in the tribe Trifolieae that is native to Australia. We compared the genetic diversity and taxonomic placement of rhizobia nodulating *T. suavissima* with those of members of an *Ensifer* reference collection. Our goal was to determine if the *T. suavissima* rhizobial strains, like their plant host, are naturally limited to the Australian continent. We used multilocus sequence analysis to estimate the genetic relatedness of 56 *T. suavissima* symbionts to 28 *Ensifer* reference strains. Sequence data were partitioned according to the replicons in which the loci are located. The results were used to construct replicon-specific phylogenetic trees. In both the chromosomal and chromid trees, the Australian strains formed a distinct clade within *E. meliloti*. The strains also shared few alleles with *Ensifer* reference strains from other continents. Carbon source utilization assays revealed that the strains are also unusual in their ability to utilize 2-oxoglutarate as a sole carbon source. A strategy was outlined for locating similar strains elsewhere.

IMPORTANCE In this study, we employed a biogeographical approach to investigate the origins of a symbiotic relationship between an Australian legume and its nitrogen-fixing rhizobia. The question of the ancestral origins of these symbionts is based on the observation that the legume host is not closely related to other native Australian legumes. Previous research has shown that the legume host *Trigonella suavissima* is instead closely related to legumes native to the Mediterranean Basin and western Asia, suggesting that it may have been introduced in Australia from those regions. This led to the question of whether its rhizobia may have been introduced as well. In this study, we were unable to find persuasive evidence supporting this hypothesis. Instead, our results suggest either that the *T. suavissima* rhizobia are native to Australia or that our methods for locating their close relatives elsewhere are inadequate. A strategy to investigate the latter alternative is proposed.

KEYWORDS *Ensifer*, rhizobium, *Sinorhizobium meliloti*, *Trigonella*, biogeography, clade, symbiotic nitrogen fixation

Received 20 December 2016 **Accepted** 2 March 2017

Accepted manuscript posted online 10 March 2017

Citation Eardly B, Elia P, Brockwell J, Golemboski D, van Berkum P. 2017. Biogeography of a novel *Ensifer meliloti* clade associated with the Australian legume *Trigonella suavissima*. Appl Environ Microbiol 83:e03446-16. https://doi.org/10.1128/AEM.03446-16.

Editor Emma R. Master, University of Toronto

Address correspondence to Bertrand Eardly, bde1@psu.edu.

In this study, we describe a novel clade of *Ensifer meliloti* and consider how geographic and ecological isolation may have contributed to the evolution of this group of bacteria. Members of the genus *Ensifer* (formerly *Sinorhizobium*) are best known for their ability to form nitrogen-fixing symbioses with forage legumes of three related genera, *Medicago* L., *Melilotus* Mill., and *Trigonella* L. These legumes, which are members of the tribe Trifolieae, have a natural distribution that extends from the Mediterranean Basin through western Asia (1, 2). This region contains an unsurpassed number of legume species belonging to these genera, and it also contains a rich diversity of microsymbionts that nodulate these legumes. In the study in which this extensive rhizobial diversity was first identified, a strongly divergent genotype within the species *E. meliloti* was noted (3). The strain possessing this genotype was USDA6670 (syn., CC2013), a strain symbiotically effective with the native Australian legume *Trigonella suavissima* Lindl. (4). In subsequent studies, four other *T. suavissima* symbionts were also described as having multilocus genotypes similar to that of USDA6670 (5, 6).

Trigonella suavissima is a herbaceous annual legume that is widely distributed in the arid interior of Australia and is well adapted to low-lying areas of flood plains, particularly those in the Channel Country of South West Queensland (7). The presence of *T. suavissima* on that continent was first documented during a collection expedition there in 1835 (8). In that report, the similarity of *T. suavissima* to other legume species of southern Europe was noted. Unlike its widely cultivated relative *Medicago sativa* L. (alfalfa or lucerne), which is native to western Asia and which can be nodulated by at least two *Ensifer* species, *T. suavissima* is more specific in its requirements for nitrogen-fixing microsymbionts (4, 9). *Trigonella suavissima* is also the only member of the tribe Trifolieae that is endemic to Australia (7). This observation and the observation that most other species of *Trigonella* are native to the Mediterranean region and western Asia (2, 10) suggest that *T. suavissima* may have been introduced into Australia from those regions. Although the rhizobial populations in Australia that nodulate *T. suavissima* have been characterized with regard to their abundance, symbiotic effectiveness, and host range (7), there is little information on the range and extent of their genetic diversity or their phylogenetic relatedness to *Ensifer* strains from other continents.

The objective of the current study was to assess the genetic diversity within a collection of 61 rhizobia isolated from 29 localities in southeastern Australia. Forty-seven of the strains were isolated from *T. suavissima* and 14 of the strains were isolated from *M. sativa*. Although both hosts are members of the same cross-inoculation group, the symbiotic effectiveness of a strain may differ between the host species. An additional 32 *Ensifer* sp. reference strains from other continents were included for genotype and phenotype comparisons. These comparisons were intended to aid in the taxonomic placement of the Australian *T. suavissima* rhizobial strains and to provide evidence as to whether they might be closely related to native rhizobial lineages found elsewhere. For the purposes of this study, a *T. suavissima* strain is defined as a rhizobial strain that either was isolated from a *T. suavissima* nodule or can fix nitrogen with *T. suavissima* but was isolated from a related legume species, such as *M. sativa*. In our comparative studies, we included representatives of three other *Ensifer* species (*E. medicae*, *E. arboris*, and *E. numidicus*) that were previously shown to be closely related to *E. meliloti* (11–13).

Currently, the two species of *Ensifer* having the greatest agricultural significance are *E. meliloti* and *E. medicae* (14), which are capable of symbiotic nitrogen fixation with a broad range of perennial and annual forage legumes, respectively (13). Their genomes consist primarily of a single circular chromosome of approximately 3.65 Mb and two large accessory replicons (or megaplasmids) (14, 15). *Ensifer* spp. may also contain smaller auxiliary plasmids, the number and identity of which vary widely among strains (16) and the functional importance of which is largely unknown. Of the two megaplasmids, one is referred to as a symbiotic megaplasmid and is approximately 1.3 Mb in size (pSymA and pSMED02 in *E. meliloti* and *E. medicae*, respectively). It carries genes important in the development and function of the nitrogen-fixing symbiosis. The other

TABLE 1 Distance matrix showing the number of pairwise nucleotide sequence differences in *atpD* and 16S rRNA alleles between the Australian *Trigonella suavissima* strain USDA6670 (CC2013) and *Ensifer* (*Sinorhizobium*) type strains

Ensifer species and strain	Strain designation	No. of sequence differences with strain designation[a]:						
		1	2	3	4	5	6	7
E. meliloti USDA6670[b]	1		**12**	**38**	**39**	**47**	**44**	**47**
E. meliloti USDA1002^T	2	1		**32**	**33**	**43**	**36**	**41**
E. numidicus ORS1407^T	3	4	4		**32**	**40**	**35**	**42**
E. medicae A321^T	4	5	4	6		**42**	**42**	**41**
E. arboris HAMBI 1552^T	5	5	6	7	6		**45**	**42**
E. psoraleae CCBAU 65732^T	6	10	9	11	9	11		**42**
E. fredii ATCC 35243^T	7	15	10	17	18	18	23	

[a]Differences are based on a 420-bp segment of *atpD* (shown in boldface type) and a 1,404-bp segment of the 16S rRNA gene (shown in lightface type). The accession numbers for the *atpD*/16S rRNA sequences for the type strains are AM418760/X67222 (*E. meliloti*), AM946551/AY500254 (*E. numidicus*), AM418754/L39882 (*E. medicae*), AM418767/Z78204 (*E. arboris*), EU617988/EU618039 (*E. psoraleae*), and AM418761/D14516 (*E. fredii*).

[b]Homologous *atpD*/16S rRNA sequences were obtained from the GenBank draft genome sequence, accession number NZ_ATWE00000000.

megaplasmid has been referred to as a chromid (17) and is approximately 1.6 Mb in size (pSymB and pSMED01 in *E. meliloti* and *E. medicae*, respectively). This replicon carries genes that are important for other accessory functions, including the import of small molecules and polysaccharide biosynthesis (15). For taxonomic placements, we relied primarily on comparisons of the chromosomal genotypes across the strains. However, we were also interested in the evolutionary histories of their large extrachromosomal replicons because of their agricultural and ecological significance.

In the first part of the study, chromosomal genotypic profiles of 56 *T. suavissima* strains were compared to those of related species through the use of a genomic indexing method called multilocus sequence typing (MLST) (18, 19). Previous whole-genome sequence analyses of diverse *E. meliloti* and *E. medicae* strains have shown that chromosomal MLST analyses can provide a robust description of the taxonomic boundaries between *Ensifer* species (20). In the second part of this study, we used a similar MLST approach to examine the range of megaplasmid genotypes in a subsample of the *T. suavissima* strains used for chromosomal profiling. In addition to the genotypic analyses, we also compared phenotypic characteristics of the strains to explore possible genotype-by-phenotype correlations.

RESULTS

Taxonomic placement of the Australian *Trigonella* strain USDA6670 (syn., CC2013). To obtain a preliminary taxonomic placement for the Australian isolates, we compared the previously published nucleotide sequences of two evolutionarily conserved loci (the 16S rRNA gene and *atpD*) in *T. suavissima* USDA6670 with the homologous sequences in six *Ensifer* type strains (Table 1). The sequence comparisons revealed that the 16S and *atpD* alleles of USDA6670 are most similar to their respective homologs in *E. meliloti* USDA1002^T and that the USDA6670 alleles are less similar to their respective homologs in *E. numidicus* and *E. medicae*. The corresponding alleles in the other species, particularly *E. psoraleae* and *E. fredii*, share the smallest number of nucleotides with their corresponding alleles in USDA6670.

Two criteria that are commonly used to determine species affiliations are average genomic nucleotide identity (ANI) and 16S rRNA gene sequence similarity. It has been estimated that most bacterial species encompass organisms that have an ANI of greater than 95% (21). This threshold has been shown to correspond to a level of 16S rRNA gene sequence similarity of approximately 98.7% (22). The percentage of sequence similarity between the 16S gene of USDA6670 and that of the *E. meliloti* type strain USDA1002^T is greater than 99.9% (Table 1). A two-way ANI comparison was calculated for the draft genome sequence of USDA6670 and the chromosomal sequence of the *E.*

meliloti genome reference strain USDA1021 (23). The comparison revealed a shared ANI of 95.2%. Thus, both the 16S rRNA gene sequence comparisons in Table 1 and the shared ANI estimate indicate that based solely on genotype, the Australian *T. suavissima* strain USDA6670 is either at or within the generally accepted species limits for *E. meliloti*.

Chromosomal genetic divergence and allelic diversity among the *T. suavissima* strains and related species. The nucleotide sequences of 10 chromosomal loci in the *T. suavissima* strains and corresponding loci in the four related *Ensifer* species were screened for polymorphic nucleotide sequence positions. The results revealed that each locus was represented by multiple alleles. A comparison of the multilocus allelic profiles for each of the 84 strains revealed 61 distinct chromosomal sequence types, or CSTs (see Table S1 in the supplemental material). Table 2 lists the 61 CSTs, the 84 strains associated with each of the individual CSTs, the geographical origins of the strains, and the respective strain phenotypes. Fifty-six of the 61 CSTs were represented by only one or two strains, while the other five CSTs were represented by multiple strains of differing geographic origins. For example, two CSTs (CST 2 and CST 46) were represented both by a strain from Israel and by two strains from another country. In the case of CST 2, the two additional strains were from Australia, while in the case of CST 46, the two additional strains were from the United States.

Nucleotide sequence-based phylogenetic analysis of the 61 CSTs by the maximum likelihood (ML) method revealed that the CSTs were clustered on four primary branches in an ML tree. Each of these branches corresponds to a named *Ensifer* species (Fig. 1). The largest cluster, consisting of CST 1 through CST 50, will be referred to as the *E. meliloti* cluster. The *E. meliloti* cluster is subdivided into two distinct clades, the *T. suavissima* strain clade (CST 1 through CST 39) and the *E. meliloti* reference strain clade (CST 40 through CST 50). With one exception, the 54 strains in the *T. suavissima* strain clade are of Australian origin. The single exception is strain CC2283c (CST 2), a strain isolated from *T. arabica* in Israel in 1965. That particular CST is also shared by two *T. suavissima* strains that were isolated from field soils in Australia (Table 2). The *E. meliloti* reference strain clade, which is also within the *E. meliloti* cluster, contains 11 CSTs that represent seven Australian field soil isolates and 12 reference strains from other countries (Table 2). The second largest cluster, which will be referred to as the *E. medicae* cluster (Fig. 1), consists of nine CSTs representing nine diverse *E. medicae* strains (CST 52 through CST 60). Only one of these strains (CC5051) was isolated from Australian soils, while the other seven were isolated from locations in the Mediterranean region (3, 13). The two remaining major branches in Fig. 1 represent the type strains for *E. arboris* (CST 51) and *E. numidicus* (CST 61), which were isolated from soils in Sudan and Tunisia, respectively.

The numbers of alleles observed for each locus within the *T. suavissima* strain clade (CST 1 through CST 39) are summarized in Table 3. The number of alleles per locus ranged from seven for *nuoE1* to 17 for *edd*, with an average of 12.4 alleles per locus. The corresponding estimates of allelic diversity reported previously for a global collection of 230 non-Australian *Medicago* sp. *Ensifer* strains (18) are also shown in Table 3. It is evident that the numbers of alleles observed for most of the *T. suavissima* strain loci are roughly equivalent to, and in some cases greater than, the numbers of alleles observed for the corresponding loci in the global collection of *Medicago* sp. *Ensifer* strains. It is also evident that only a small number of alleles are shared between the two geographically divergent groups of strains. The extensive allelic diversity among the loci of the *T. suavissima* strains is noteworthy, considering that the corresponding data for the global collection of *Medicago* sp. *Ensifer* strains were obtained from a much larger number of strains, collectively representing two genetically divergent species of *Ensifer* (*E. meliloti* and *E. medicae*) (Fig. 1). Furthermore, the global strains were isolated from a wide variety of geographic locations, primarily from the Mediterranean Basin and western Asia.

Estimated phylogenetic relationships among the extrachromosomal replicons. The phylogenetic relationships among the extrachromosomal replicons in four *T.*

TABLE 2 CSTs, strain origins, and phenotypes of the 84 strains in the MLST analysis

CST[a]	Strain	Origin					Symbiotic effectiveness		2-oxo growth[f]	Reference(s)
		Country[b]	State	Locality	Year	Trap host[c]	*Ts*[d]	*Ms*[e]		
1	CC2137	Au	NSW	Bootingee[g]	1960	*Ts*	ND[h]	I	+	
2	CC2283c	Is	NA[i]	Tel Sharuhen	1965	*Ta*	I, E	I	+/−	9
	CC5033	Au	NSW	North Bourke	2004	*Ts*	I/E	I/E	+	
	CC5043	Au	Qld	Jackson	2004	*Ts*	I/E	I	+	
3	CC2155b	Au	Qld	Malagarga[g]	1963	*Ts*	E	I	+	6, 7
4	CC2129	Au	Qld	Wyandra	1960	*Ts*	ND	I	+	6
5	CC2341	Au	Qld	Currawilla[g]	ND	*Ts*	E	I	ND	
6	CC5015	Au	SA	Etadunna[g]	1999	*Ts*	ND	ND	+	
	CC5022	Au	NSW	Cuttabura Creek	2000	*Ts*	ND	ND	+	
	CC5035	Au	Qld	Birdsville	2004	*Ts*	ND	ND	+	
	CC5036	Au	Qld	Birdsville	2004	*Ts*	ND	ND	+	
7	CC5019	Au	NSW	Louth	2000	*Ts*	ND	ND	+	
8	CC5021	Au	NSW	Cuttaburra Creek	2000	*Ts*	ND	ND	+	
9	CC5038	Au	Qld	Innamincka station[g]	2004	*Ts*	ND	ND	+	
	CC5049	Au	Qld	Birdsville	1999	*Ts*	ND	ND	0	
10	CC5037	Au	Qld	Birdsville	2004	*Ts*	E	I/E	+	
	CC5040	Au	Qld	Innamincka station[g]	2004	*Ts*	ND	ND	+	
11	CC5039	Au	Qld	Innamincka station[g]	2004	*Ts*	ND	ND	ND	
12	CC5044	Au	Qld	Jackson	2004	*Ts*	ND	ND	+	
13	CC5053	Au	NSW	Menindee	2000	*Ts*	ND	ND	+	
14	CC5031	Au	NSW	North Bourke	2004	*Ts*	ND	ND	+	
15	CC5032	Au	NSW	North Bourke	2004	*Ts*	ND	ND	+	
16	CC5028	Au	NSW	Menindee	2000	*Ts*	ND	ND	+	
17	CC5034	Au	Qld	Glengyle[g]	2004	*Ts*	ND	ND	0	
18	CC2325	Au	NSW	Fords Bridge	1995	*Ts*	I/E	I/E	+/−	
	CC5020	Au	NSW	Louth	2000	*Ts*	ND	ND	+	
	CC5027	Au	SA	Marree	2000	*Ts*	ND	ND	+	
	CC5029	Au	NSW	Bourke	2004	*Ts*	ND	ND	+	
	CC5030	Au	NSW	Bourke	2004	*Ts*	ND	ND	+/−	
19	CC5045	Au	NSW	Wilcannia	2004	*Ts*	ND	ND	+	
20	CC5024	Au	SA	Margaret Creek	2000	*Ts*	ND	ND	+	
21	CC2324	Au	NSW	North Bourke	1995	*Ts*	ND	I	+	
22	CC2160	Au	NSW	Birriegoolpa Bore	1967	*Ts*	E	I	+	5, 6, 7
23	CC5023	Au	SA	Strzelecki Creek	2000	*Ts*	ND	ND	+	
24	CC5026	Au	SA	Marree	2000	*Ts*	ND	ND	+	
25	CC2017	Au	Qld	Roma	1958	*Ms*	E	E	+	4, 7, 9
26	CC5016	Au	Qld	Marion Downs[g]	1999	*Ts*	ND	ND	0	
	CC5017	Au	Qld	Marion Downs[g]	1999	*Ts*	ND	ND	0	
27	CC2154	Au	Qld	Malagarga[g]	1960	*Ts*	E	I/E	+	7, 9
28	CC5006	Au	NSW	Menindee	2006	*Ts*	E	I/E	+	7
29	CC2153	Au	Qld	Malagarga[g]	1960	*Ts*	E	I	+	5, 7
	CC2156b	Au	Qld	Malagarga[g]	1963	*Ts*	E	I	+	7
30	USDA6670 (CC2013)	Au	NSW	Moonie River	1966	*Ms*	E	E	+	4, 5
31	CC2157	Au	NSW	Teurika[g]	1967	*Ts*	I, I/E	I, I/E	+	5, 7
	CC2152	Au	Qld	Malagarga[g]	1960	*Ms*	E	I	+/−	7
32	CC2326	Au	Qld	Windorah	1995	*Ts*	ND	I	+	
33	CC5011	Au	NSW	Lake Tandou	2006	*Ms*	E	I	+	7

(Continued on following page)

TABLE 2 (Continued)

CST[a]	Strain	Origin					Symbiotic effectiveness		2-oxo growth[f]	Reference(s)
		Country[b]	State	Locality	Year	Trap host[c]	*Ts*[d]	*Ms*[e]		
34	CC5002	Au	NSW	Lake Tandou	2006	*Ms*	I/E	I	+	7
	CC5003	Au	NSW	Lake Tandou	2006	*Ts*	E	I	+	7
35	CC5009	Au	NSW	Lake Tandou	2006	*Ms*	I/E	I	+	7
36	CC5025	Au	SA	Margaret Creek	2000	*Ts*	ND	ND	+	
37	CC5001	Au	NSW	Lake Tandou	2006	*Ts*	I/E	I	+	7
38	CC5007	Au	NSW	Lake Tandou	2006	*Ms*	E	I	+	7
39	CC5008	Au	NSW	Lake Tandou	2006	*Ms*	E	I	+	7
40	102F51	US	ND	ND	1927	*Ms*	ND	E	0	3, 18
41	CC5012	Au	NSW	Menindee	2006	*Ms*	I	E	0	7
42	15A6	Pa	NA	ND	ND	*Ms*	ND	E	0	3, 18
43	15B4	Pa	NA	ND	ND	*Ms*	ND	ND	0	3, 18
44	74B12	Pa	NA	ND	ND	*Ms*	ND	ND	0	3, 18
45	CC5052	Au	NSW	Wentworth	2000	*Ts*	ND	ND	0	
46	CC2282b	Is	NA	Shikmaa	1965	*Ta*	I	I	0	9
	USDA1114	US	CA	ND	ND	*Tb*	ND	ND	0	6
	USDA1115	US	CA	ND	ND	*Tc*	ND	ND	0	6
47	M119	Sy	NA	ND	ND	*Mp*	ND	ND	0	3, 18
48	CC164	ND	NA	ND	ND	ND	ND	ND	0	
	CC2019	Au	Qld	Bollon	1958	*Ms*	I	E	0	4, 7, 9
	CC2053	Au	Qld	Mitchell	1958	*Ms*	I	E	+/−	4, 7, 9
	CC5013	Au	NSW	Menindee	2006	*Ms*	I	E	0	7
	USDA1176	ND	NA	NA	ND	ND	ND	ND	0	
	USDA1177	US	ND	NA	ND	*Tf*	ND	ND	0	6
	CC2003	Au	Qld	Narromine	1958	*Ms*	I	I	0	4, 7, 18, 19, 26
49	CC5010	Au	NSW	Lake Tandou	2006	*Ms*	I	I/E	0	7
50	USDA1002^T	US	VA	Washington, DC	1919	*Ms*	ND	E	0	3, 18, 19
51	LMG 14919^T	Su	NA	Kosti	1987	*Pc*	ND	ND	+	29, 30
52	M254	Jo	NA	ND	ND	*Mr*	ND	ND	0	3, 18
53	CC5051	Au	NSW	Wentworth	2000	*Ts*	ND	ND	0	
54	A321^T	Fr	NA	Aude	ND	*Mt*	ND	ND	0	13
55	M3	Sy	NA	ND	ND	*Mo*	ND	ND	0	3, 18
56	M102	Sy	NA	ND	ND	*Mt*	ND	ND	0	3, 18
57	M161	Sy	NA	ND	ND	*Mn*	ND	ND	0	3, 18
58	M205	Tu	NA	ND	ND	*Mt*	ND	ND	0	3, 18
59	M1	Sy	NA	ND	ND	*Mo*	ND	ND	0	3, 18
60	M280	Jo	NA	ND	ND	*Mo*	ND	ND	0	3, 18
61	LMG 24690^T (ORS 1407)	Tu	NA	Infra-arid zone	ND	*Au*	ND	ND	0	12, 31

[a]CST, multilocus chromosomal sequence type. The corresponding allelic profiles are listed in Table S1 in the supplemental material.
[b]Au, Australia; Is, Israel; Fr, France; Jo, Jordan; Pa, Pakistan; Sy, Syria; Su, Sudan; Tu, Tunisia; US, United States.
[c]*Au*, *Argyrolobium uniflorum*; *Mn*, *Medicago noeana*; *Mo*, *M. orbicularis*; *Mp*, *Medicago* spp.; *Mr*, *M. rotata*; *Mt*, *M. truncatula*; *Ms*, *M. sativa*; *Pc*, *Prosopis chilensis*; *Ta*, *Trigonella arabica*; *Tb*, *T. balansae*; *Tc*, *T. corniculata*; *Tf*, *T. foenum-graecum*; *Ts*, *T. suavissima*.
[d]Symbiotic N-fixing effectiveness with *T. suavissima*: I, ineffective; I/E, partly effective; E, effective.
[e]Symbiotic N-fixing effectiveness with *M. sativa*: I, ineffective; I/E, partly effective; E, effective.
[f]Ability to utilize 2-oxoglutarate (alpha-ketoglutarate) as a sole carbon source after 72h. +, strong growth; +/−, moderate-to-weak growth; 0, poor growth.
[g]Name of the property from which the strain originated.
[h]ND, not determined.
[i]NA, not applicable.

suavissima strains and in 18 reference strains were also estimated using the MLST method. The nucleotide sequences of three concatenated chromid loci and three concatenated symbiotic megaplasmid loci were aligned for these comparisons. The multilocus sequence profiles for each of the replicons in the 22 strains revealed 16 distinct sequence types (STs) for their chromids and 20 distinct STs for the symbiotic megaplasmids in the same strains (Table 4). Some strains share the same ST for one extrachromosomal replicon but not the other (e.g., *E. meliloti* strains 15B4 and 74B12), while other strains share STs for both extrachromosomal replicons (e.g., *E. medicae* strains M254 and A321^T) but do not share their chromosomal profiles (e.g., CST 52 and

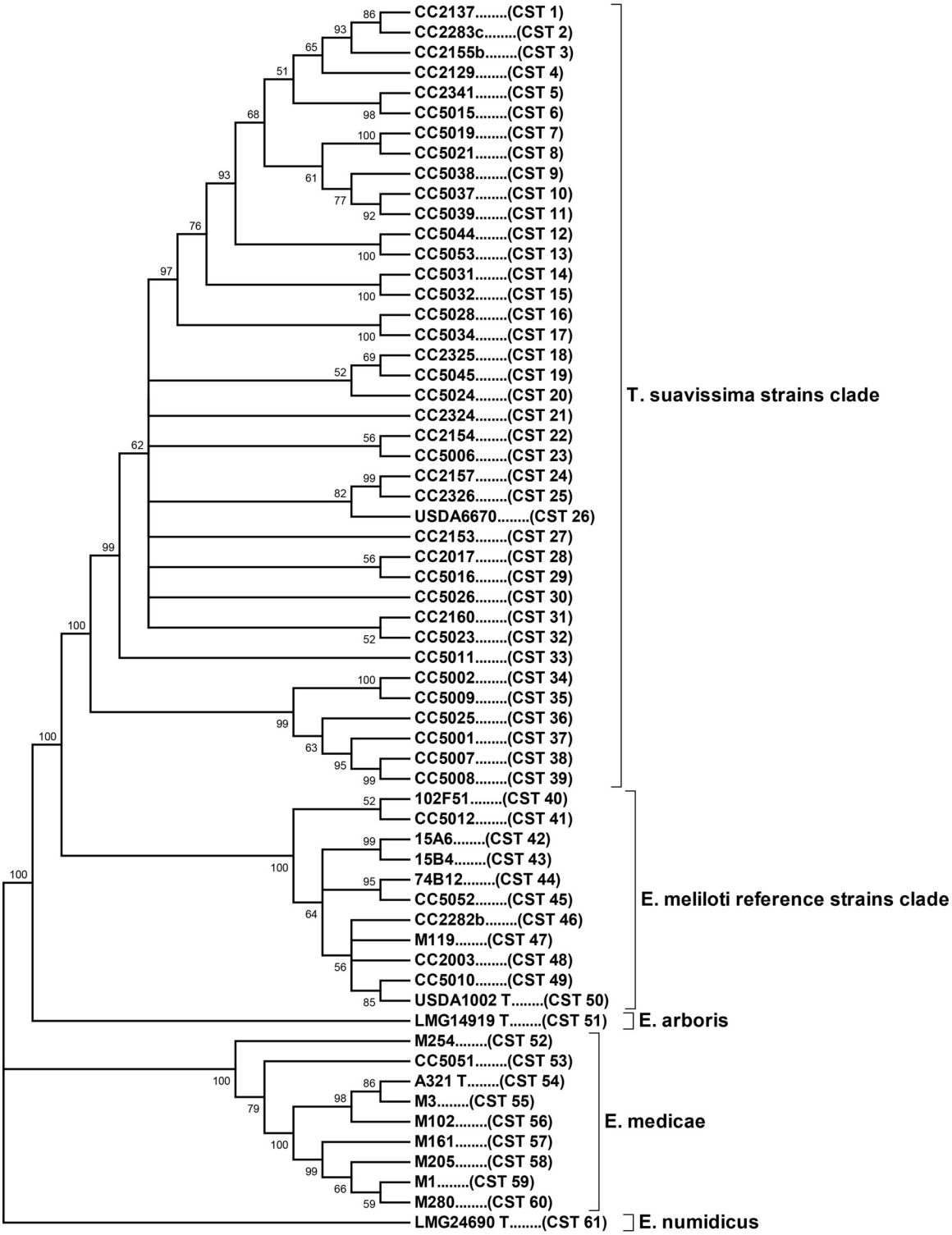

FIG 1 Chromosomal maximum likelihood tree based on concatenated nucleotide sequences of 10 loci in 84 strains. Alignment of the sequences revealed 61 distinct chromosomal sequence types (CSTs); one representative strain for each CST is shown. Bootstrap values of greater than 50% (of 500 replicates) are indicated at nodes. T., *Trigonella*; E., *Ensifer*.

CST 54, respectively). Also, in several cases, the same chromid ST was shared across multiple strains having differing chromosomal genotypes. For example, the *E. medicae* reference strains A321^T, M3, M102, M161, and M254, which were chosen to represent the range of chromosomal genotypes identified in a previous study (18), all have the

TABLE 3 The number of distinct *Ensifer* alleles according to strain geographic origin and trap host

Locus	No. of alleles for: Australian *T. suavissima* strain CSTs[a]	Global (non-Australian) *Medicago* sp. CSTs[b]	No. alleles shared/total
asd	14	15	0/29
edd	17	17	2/32
zwf	16	9	0/25
gap	14	8	1/21
glnD	12	9	0/21
gnd	12	7	1/18
nuoE1	7	11	1/17
ordL2	11	8	0/19
recA	8	9	0/17
sucA	13	11	0/24
	Average = 12.4	Average = 10.4	Sum = 5/223

[a]Thirty-nine chromosomal sequence types (CSTs) reported for the 54 strains in the current study.
[b]Ninety CSTs reported for 230 *Medicago* sp. strains from other continents in a previous study (18).

same chromid ST but have different chromosomal STs. The same is true for the *E. meliloti* reference strains 102F51, CC2003, and 74B12.

The phylogenetic relationships among the extrachromosomal replicons in the 22 strains were also estimated using the maximum likelihood (ML) method. Three-locus concatenated sequences were aligned and compared for each of the replicons in each of the strains. An ML tree illustrating the relationships among the chromid replicons is shown in Fig. 2, and a corresponding ML tree illustrating the relationships among the symbiotic megaplasmid replicons is shown in Fig. 3. Assuming that the strains of the *T.*

TABLE 4 Extrachromosomal replicon sequence types for 22 strains representing four closely related species of *Ensifer* (*Sinorhizobium*)

Species	Strain	Chromid locus[a]			Symbiotic megaplasmid locus[b]		
		dak	*gabT*[c]	*idhA*	*nifD*	*nodC*[c]	sma0198
E. meliloti	USDA1002^T	1	1	10	1	3	13
	102F51	1	1	11	9	3	32
	CC2003	1	1	11	15	32	3
	15A6	5	1	4	3	4	14
	15B4	1	4	10	3	4	6
	M119	3	1	11	12	15	4
	74B12	1	1	11	3	4	6
	CC5052	19	6	11	2	22	3
	USDA6670[d]	9	10	25	6	7	33
	CC2155b[d]	20	25	28	6	7	34
	CC5011[d]	21	26	29	51	7	33
	CC5034[d]	22	10	30	6	7	3
E. medicae	A321^T	2	2	2	2	2	2
	M1	2	9	2	7	11	18
	M3	2	2	2	2	11	2
	M102	2	2	2	2	11	20
	M161	2	2	2	17	17	31
	M205	16	15	24	22	2	22
	M254	2	2	2	2	2	2
	M280	2	20	2	7	2	2
E. arboris	LMG 14949^T	23	27	31	52	52	35
E. numidicus	LMG 24690^T	24	28	32	53	53	36

[a]The concatenated three-locus nucleotide sequence data were used to construct the ML tree in Fig. 2.
[b]The concatenated three-locus nucleotide sequence data were used to construct the ML tree in Fig. 3.
[c]Allele codes.
[d]Members of the *Trigonella suavissima* strain clade in Fig. 1.

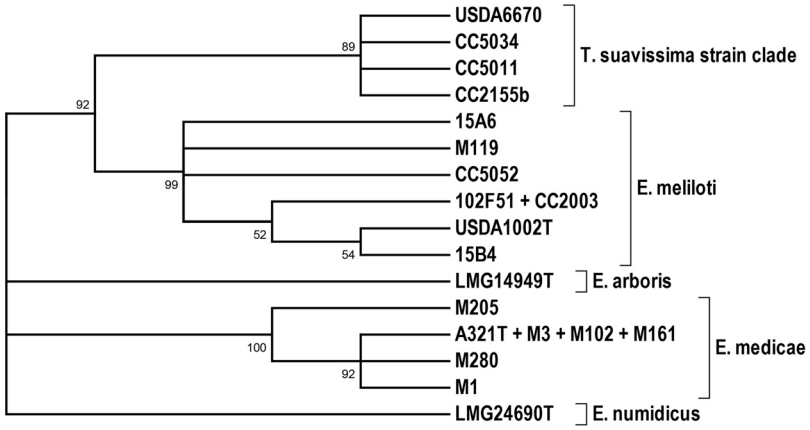

FIG 2 Chromid maximum likelihood tree based on concatenated nucleotide sequences of three loci in 22 strains (listed in Table 4). Alignment of the sequences revealed 16 distinct sequence types; representative strains for each sequence type are shown. Bootstrap values of greater than 50% (of 500 replicates) are indicated at nodes. T., *Trigonella*; E., *Ensifer*.

suavissima clade are indeed conspecific with *E. meliloti*, then a species-level congruence is evident between the topology of the chromosomal tree (Fig. 1) and the topology of the chromid tree (Fig. 2). However, in the symbiotic megaplasmid ML tree, the divergence between the *E. meliloti* STs and the *E. medicae* STs is not well supported (Fig. 3).

Phenotypic characteristics of the *T. suavissima* strains. Transmission electron micrographs of the Australian *T. suavissima* strains CC2017 and CC2155b are shown in Fig. S1 and S2 in the supplemental material, respectively. The ultrastructural characteristics of the cells, including their size, shape, flagellation, and budding morphology, are consistent with the characteristics normally associated with other members of the genus *Ensifer* (*Sinorhizobium*) (24).

The symbiotic nitrogen-fixing effectiveness ratings for 30 of the strains on *T. suavissima* and *M. sativa* are listed in Table 2. These results were compiled from the results of the current study (see Table S3) and also from the results from three previous studies (4, 7, 9). In general, members of the *T. suavissima* strain clade (CST

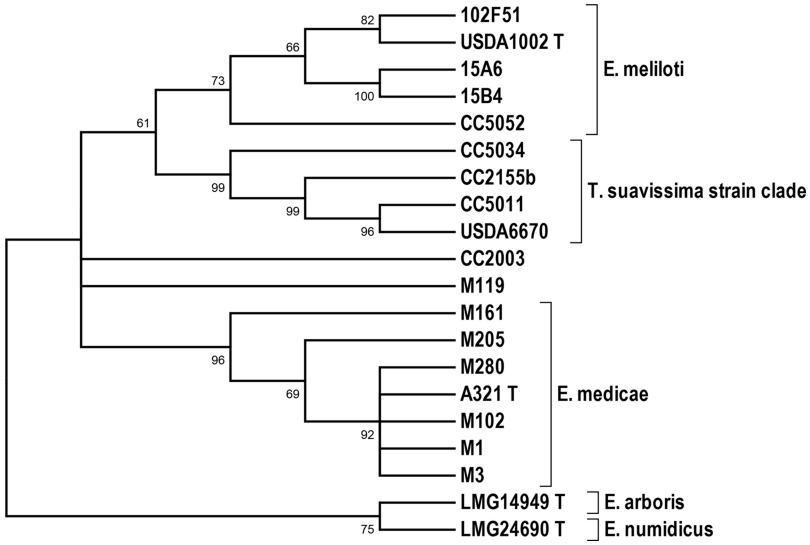

FIG 3 Symbiotic megaplasmid maximum likelihood tree based on concatenated nucleotide sequences of three loci in 22 strains (listed in Table 4). Alignment of the sequences revealed 20 distinct sequence types; representative strains for each sequence type are shown. Bootstrap values of greater than 50% (of 500 replicates) are indicated at nodes. T., *Trigonella*; E., *Ensifer*.

1 through CST 39) are symbiotically more effective on *T. suavissima* than they are on *M. sativa* (Table 2). Conversely, members of the *E. meliloti* reference strain clade (CST 40 through CST 50) are generally more effective on *M. sativa* than they are on *T. suavissima* (Table 2).

Trigonella anguina was included as a potential host in the current effectiveness study because it is a sister species to *T. suavissima* (25) and it was unclear whether the *T. suavissima* strains would be symbiotically effective on *T. anguina*. The results shown in Table S2 indicate that most of the strains that are symbiotically effective on *T. suavissima* are also effective on its sister species, *T. anguina*.

The four strains with the consistently highest levels of symbiotic effectiveness with *T. suavissima* are strains CC2017, CC2155b, USDA6670, and CC5037 (Table 2; see also Table S2). The symbiotic effectiveness results compiled in Table 2 also revealed that different strains sharing the same CST (e.g., CST 2 strains and CST 31 strains) may express significantly different levels of symbiotic effectiveness on the same host.

The results of the initial comparative biochemical profiling of three of the *T. suavissima* strains (CC2017, CC2155b, and USDA6670), two of the *E. meliloti* reference strains (USDA1002^T and USDA1021), and two of the *E. medicae* reference strains (A321^T and CC169) revealed that the only biochemical test that consistently differentiated the *T. suavissima* strains from the other strains was that which showed their ability to utilize 2-oxoglutarate (α ketoglutarate) as a sole carbon source. Neither the two *E. meliloti* reference strains nor the two *E. medicae* reference strains possessed this ability.

To substantiate this observation, an expanded study was initiated to examine the ability of all the strains in the collection to utilize 2-oxoglutarate as a sole carbon source. Of the 54 strains in the *T. suavissima* strain clade (CST 1 through CST 39), 44 showed strong growth on 2-oxoglutarate, 4 showed moderate-to-weak growth, and four showed poor growth (Table 2). The only other strain in the collection showing strong growth on this carbohydrate was the type strain for *E. arboris* (LMG 14949^T). Two other strains in the collection also showed moderate-to-weak growth on 2-oxoglutarate; these were *E. kostiense* (HAMBI 1489^T) and one of the strains in the *E. meliloti* strain clade (CC2053). Although the latter strain is not symbiotically effective with *T. suavissima* (Table 2), it is effective on two other species of *Trigonella* (7, 9). None of the nine representatives of the *E. medicae* strain cluster (CST 52 through CST 60) were able to grow on the 2-oxoglutarate broth media nor were any of the additional *Ensifer* type strains that were examined, which included *E. adhaerens* (ATCC 33499^T), *E. americanum* (CFNEI 156^T), *E. chiapanecum* (IITG S70^T), *E. fredii* (USDA205^T), *E. mexicanus* (IITG R7^T), *E. adhaerens* (Lc04^T), *E. numidicus* (LMG 24690^T), *E. saheli* (HAMBI 215^T), *E. terangae* (HAMBI 220^T), and *E. xingiangensis* (CCBAU 110^T). To determine whether the failure of the strains to grow on 2-oxoglutarate may have been due to an inhibitory effect of that compound on growth, the strains were also grown on 2-oxoglutarate medium supplemented with 5 g liter^{-1} glycerol. The additional carbon source resulted in normal growth of the strains (data not shown), suggesting that their poor growth on 2-oxoglutarate was probably not due to an inhibitory effect of the 2-oxoglutarate.

Since the locus encoding 2-oxoglutarate dehydrogenase (*sucA*) was sequenced in all of the strains in the chromosomal MLST study, it was of interest to determine if there was a correlation between the ability of certain strains to utilize 2-oxoglutarate and the specific *sucA* alleles in those strains. No such correlation was evident, other than the fact that the *T. suavissima* strains, most of which could grow on 2-oxoglutarate, did not share *sucA* alleles with any of the other strains in the study (Table S1).

DISCUSSION

One of the primary goals of this study was to provide a taxonomic placement for the group of Australian *T. suavissima* microsymbionts that are represented by the strain USDA6670 (CC2013). In a previous study which examined the abundance, symbiotic effectiveness, and host range of native populations of rhizobia nodulating *T. suavissima* (7), it was postulated that these rhizobia may be members of a distinct taxon within the genus *Ensifer*. Our initial results, which were based on pairwise 16S RNA gene and *atpD*

nucleotide sequence comparisons and also on a comparative ANI analysis, suggested that USDA6670 is conspecific with the type strain of *E. meliloti*. The subsequent results of our population-scale chromosomal MLST analyses supported this placement; however, they also indicated that the *T. suavissima* strains and the *E. meliloti* reference strains each constitute monophyletic clades within the species *E. meliloti*. The results also confirmed that *E. meliloti* is clearly differentiated from the other closely related species of *Ensifer* (i.e., *E. medicae*, *E. arboris*, and *E. numidicus*).

The distinctiveness of the strains in the *T. suavissima* clade was substantiated by the fact that only five of the 228 allelic variants within the *T. suavissima* strain CSTs were found within the CSTs representing a global collection of 230 *E. meliloti* and *E. medicae* reference strains. The distinctiveness of the *T. suavissima* strains is also supported by the monophyletic clustering of the chromid STs of four representative *T. suavissima* strains (Fig. 2). The topological congruence between chromosomal and chromid ML trees is perhaps not surprising, considering that chromids usually share intrinsic characteristics (e.g., codon usage patterns) with the chromosomes of the cells in which they reside (17). By contrast, the topology of the symbiotic megaplasmid tree differed somewhat from those of the chromosomal and chromid trees, in that two of the *E. meliloti* genotypes were dispersed on divergent nodes of the symbiotic megaplasmid tree, and that these nodes received little bootstrap support (Fig. 3). Previous studies have shown that certain loci (e.g., *nod* genes) that are mapped to the symbiotic megaplasmids in *E. meliloti* often have complex evolutionary histories, apparently because of frequent horizontal transfer and recombination (19, 20, 26).

In addition to their distinct multilocus genotypes, the *T. suavissima* strains also possess some significant differences in phenotype compared with those of most other *Ensifer* strains. For example, the results of the symbiotic nitrogen-fixation effectiveness tests revealed that, even though the *T. suavissima* strains are capable of nodulation and nitrogen fixation with *M. sativa*, they are generally more symbiotically effective with *T. suavissima*. Their modest symbiotic effectiveness with *M. sativa* might be an important agronomic consideration if these strains are shown to be competitive for nodulation in soils where alfalfa is grown.

The initial metabolic and biochemical screening of three of the *T. suavissima* strains revealed that, unlike the *E. meliloti* and *E. medicae* reference strains, the *T. suavissima* strains were able to utilize 2-oxoglutarate (α-ketoglutarate) as a sole carbon source. When the entire collection of *T. suavissima* strains, reference strains, and type strains were screened for this phenotype, 48 of the 54 strains in the *T. suavissima* strain clade were able to grow on 2-oxoglutarate as a sole carbon source, while only two of the (twelve) more divergent *Ensifer* type strains (*E. arboris* and *E. kostiense*) were able to do so. Furthermore, only a single strain in the *E. meliloti* clade possessed this phenotype. Although that particular strain is symbiotically ineffective with *T. suavissima* (7, 9) and only marginally effective with *M. sativa* (4), it is effective with two other species of *Trigonella* (9). The presence of the 2-oxoglutarate utilization phenotype in three of the more divergent species examined (*E. meliloti*, *E. arboris*, and *E. kostiense*) suggests that this may be an ancestral trait among members of the genus *Ensifer*.

The limited natural distribution of *T. suavissima* and its symbionts to remote parts of the Australian continent provides an opportunity to address some biogeographical questions on the evolution of the *T. suavissima-Ensifer* symbiosis. For example, did both symbionts coevolve in isolation on the continent, and if so, for how long? Since most species of the genus *Trigonella* are native to the Mediterranean region and to western Asia (2, 10), and since there are no other members of the family Trifolieae that are native to Australia, it would seem reasonable to suggest that the early ancestors of *T. suavissima* may have been introduced into Australia from the Mediterranean region or western Asia. Indeed, the similarity of *T. suavissima* to southern European legumes was first noted in the initial description of the species in 1835 (8).

Whether the rhizobial symbionts of *T. suavissima* may have originated elsewhere is still open to question. However, our observation of the extensive allelic diversity within the Australian *T. suavissima* strains, together with our observation of only very limited

allele sharing between the *T. suavissima* strain CSTs and those observed in a global *Medicago* sp. reference collection, would argue against the proposition that *T. suavissima* strains likely share a very recent common ancestor with strains from the Mediterranean region or western Asia. However, on the contrary, one might reasonably question whether the reference collection in the above comparison can be expected to reflect the full range of allelic diversity in *E. meliloti*. Furthermore, it should be noted that in the development of the global collection, no *Trigonella* species were used as trap hosts. Although species of *Medicago* and *Trigonella* are both members of the same cross-inoculation group (27), the validity of this group has been questioned (28). Thus, it is possible that the *Medicago* species used as trap hosts in the development of the global reference collection may have biased its composition.

If indigenous populations of rhizobia that are closely related to the Australian *T. suavissima* strains do indeed exist naturally in soils outside Australia, then a more targeted isolation strategy might be necessary to expose them. One approach might be to sample field sites at locations where large natural populations of *T. anguina* and other close relatives of *T. suavissima* are known to exist. To avoid any potential trap host bias, it would be advisable to rely on *T. suavissima* as the primary trap host. Subsequent screening and identification of *T. suavissima*-like strains from among the resulting nodule isolates could be facilitated using a two-step approach, involving an initial 2-oxoglutarate enrichment screening, followed by a PCR-restriction fragment length polymorphism (RFLP) screening of the 2-oxoglutarate-positive strains. Diagnostic restriction sites for the latter could be identified among the allelic sequences reported in this study.

MATERIALS AND METHODS

Bacterial strains, DNA samples, and PCR. In this study, a total of 84 strains were examined, 64 of which were from the personal collection of J. Brockwell. The latter strains were isolated mostly from *T. suavissima* trap hosts and several *M. sativa* trap hosts grown in Australian soils (Table 2). Of the remaining 20 strains, 10 were *E. meliloti* reference strains, eight were *E. medicae* reference strains, and the type strains for *E. arboris* and *E. numidicus* were also included (18, 19, 26, 29–31). The reference strains were chosen to represent the range of chromosomal genetic diversity revealed in previous studies examining chromosomal genotypes of *E. meliloti* and *E. medicae* strains from geographically diverse locations (3, 13, 18). Prior to DNA extraction, all strains were plated to check for purity and then were grown for 3 days in modified arabinose-gluconate broth (32). Genomic DNAs were extracted using a QIAamp DNeasy minikit (Qiagen, Valencia, CA). DNA samples were stored at −20°C and the parent cultures were maintained in glycerol suspensions at −70°C.

Loci to be sequenced for chromosomal MLST were selected by referring to the chromosome sequence of *E. meliloti* strain USDA1021 (23). Ten loci distributed around the chromosome of USDA1021 were selected for PCR amplification and sequencing (18). The entire open reading frame for each locus was examined to aid in the design of primers that would amplify a portion of the gene, ranging in size from 200 to 500 bp. Five loci were chosen to represent the first half of the *E. meliloti* chromosome (USDA1021 genome annotation positions 1 to 1,735,000), which contains a higher degree of nucleotide diversity and a more positive GC skew (20). These loci, their products, and the respective fragment sizes that were selected for comparison were as follows: *edd*, phosphogluconate dehydratase, 501 bp; *glnD*, protein-PII uridylyltransferase, 309 bp; *nuoE1*, NADH dehydrogenase I chain E protein, 222 bp; *ordL2*, putative oxidoreductase protein, 432 bp; and *zwf*, glucose-6-phosphate 1-dehydrogenase, 453 bp. Five additional loci were selected to represent the second half of the *E. meliloti* chromosome (USDA1021 genome annotation positions 1,735,000 to 3,654,135), which contains a lower degree of nucleotide diversity and a more negative GC skew (20). These loci, their products, and the respective fragment sizes were as follows: *asd*, aspartate-semialdehyde dehydrogenase, 456 bp; *gap*, glyceraldehyde 3-phosphate dehydrogenase, 414 bp; *gnd*, 6-phosphogluconate dehydrogenase, 303 bp; *recA*, DNA strand exchange and recombination protein, 327 bp; and *sucA*, α-ketoglutarate (2-oxoglutarate) dehydrogenase E1, 429 bp.

Loci to be sequenced for the symbiotic megaplasmid MLST analyses were also selected by referring to the genome sequence of *E. meliloti* strain USDA1021 (15). The three pSymA loci, their products, and the respective fragment sizes were as follows: *nifD*, nitrogenase molybdenum-iron protein alpha chain, 420 bp; *nodC*, acetylglucosaminyltransferase, 348 bp; and SMa0198, ABC transporter permease, 312 bp (19). Similarly, three chromid (pSymB) loci were selected for amplification and analysis. These loci, their products, and the respective fragment sizes were as follows: *dak*, glycerone kinase, 465 bp; *gabT*, 4-aminobutyrate aminotransferase, 522 bp; and *idh*, myo-inositol dehydrogenase, 528 bp.

The details describing the PCR primer sequences and amplification reactions are described in detail elsewhere (18, 19). Briefly, amplification conditions were optimized using the FailSafe PCR PreMix selection kit protocol and reagents (Epicentre, Madison, WI). The PCR thermocycling reaction mixtures were incubated in an MJ Research PTC-225 Peltier thermal cycler (MJ Research, Waltham, MA). Following

amplification, the PCR products were purified to remove PCR primers using the AMPure PCR purification system (Agincourt Bioscience Corp., Beverly, MA).

DNA sequences and data analyses. Initial estimates of the phylogenetic relatedness between the representative *T. suavissima* strain USDA6670 and several *Ensifer* type strains were obtained through pairwise comparisons of published sequences for two conserved "housekeeping" loci (the 16S rRNA gene and *atpD*). The six type strains that were included in these initial comparisons were chosen on the basis of previous studies indicating their close phylogenetic relatedness to *E. meliloti* (11, 12, 31, 33–35). Pairwise similarity matrices were developed for each locus using published sequences that were aligned using the program Muscle, as implemented in MEGA v. 6.0 (36). The numbers of sequence differences were determined for a 1,404-bp core segment of the 16S rRNA locus and for a 420-bp segment of the *atpD* locus. The initial relatedness estimates obtained through these comparisons were then used as a guide to select four reference species for the more in-depth MLST analyses. In a separate comparison, the average nucleotide identity (ANI) between the chromosome of *E. meliloti* genome reference strain USDA1021 (accession no. AL591688) and the draft genome sequence of the *T. suavissima* strain USDA6670 (DOE JGI GEBA Root Nodulating Bacteria project, identification [ID] Ga0002169, accession no. ATWE01) was calculated (37).

The multilocus allelic profiles and corresponding nucleotide sequences for the *E. meliloti* and *E. medicae* reference strains in this study are described elsewhere (18, 19). To obtain sequences for the MLST alleles in the *Trigonella* strains and the other reference strains in the MLST analysis, purified PCR product templates were sequenced using nested PCR primers in both forward and reverse directions. Sequencing reactions were performed using an Applied Biosystems Dye Deoxy terminator cycle sequencing kit in conjunction with an Applied Biosystems 3130 genetic analyzer (Applied Biosystems, Foster City, CA). Homologous chromosomal and extrachromosomal allelic sequences for *E. arboris* LMG 14919^T, *E. meliloti* USDA6670, and *E. numidicus* LMG 27395^T were obtained by a BLAST search of their respective draft genome sequences. Draft sequences for the *E. numidicus* reference strain LMG 27395^T were obtained through a commercial sequencing lab (Molecular Research LP, Shallowater, TX). Briefly, a library pool for strain LMG 27395^T was paired-end sequenced for 300 cycles using the HiSeq 2500 system. Resulting fragments were assembled into 56 contigs.

A Microsoft Access database was created to compile the MLST sequences from the various sources. The strategy for the design and manipulation of the database is described elsewhere (18). Briefly, the different alleles were identified using the program Sequence Comparator version 2.0.1 written by Keith Jolley. Alleles from previous studies had already been assigned identification number codes. Alleles that were identified in this study were assigned new allele identification number codes. The number codes for all of the chromosomal alleles included in the study are listed in Table S1 in the supplemental material. Similarly, the number codes for all of the extrachromosomal replicon alleles are listed in Table 4. For each replicon in each strain, a composite multilocus profile was developed. Strains possessing the same multilocus profile for a given replicon were assigned to the same sequence type (ST) for that particular replicon.

To estimate the phylogenetic relatedness between homologous replicons, the allelic sequences for the loci within each ST were aligned and the alleles were concatenated. To estimate chromosomal phylogenetic relatedness through the use of the concatenated sequences, the maximum likelihood (ML) method was implemented using MEGA 6.0 (36). Model selection was facilitated using the corrected Akaike information criterion. The best fits for the chromosomal data were obtained using the Tamura 3-parameter model of DNA substitutions with gamma-distributed rate variations among sites, under the assumption that a certain fraction of sites were evolutionarily invariable (T92 + G + I). The best fits for the chromid and symbiotic megaplasmid data were obtained using the Tamura 3-parameter model of DNA substitutions with gamma-distributed rate variations among sites (T92 + G). The final trees were obtained using the ML heuristic method employing the nearest-neighbor-interchange tree-searching strategy. A multiparametric bootstrap resampling of 500 pseudoreplicates was plotted onto the previously selected best-scored ML tree for each replicon.

Phenotypic characterizations of the strains. In a previous study (7), the symbiotic effectiveness of 20 of the *T. suavissima* strains was evaluated using a plant-tube method. In this study, the effectiveness of five additional strains was examined using a similar method (38). Briefly, axenically grown seedlings of *M. sativa*, *T. suavissima*, and *T. anguina* were inoculated with 100 μl of a dilute broth culture suspension of a single strain. The shoot dry matter production of the seedlings was measured after 35 days of growth. The effectiveness of each strain was assessed by comparing the effect of the strain on shoot dry matter production relative to that observed for nitrate-fertilized control seedlings and uninoculated control seedlings. In this study, the symbiotic effectiveness of four of the isolates from previous studies (7, 9) (CC2017, CC2155b, CC2157, CC2283c) was included, in addition to those of five isolates that were not examined in those studies (CC2325, CC5033, CC5037, CC5043, and USDA6670). *Trigonella anguina* D. was included as a host in this study because it is a sister species to *T. suavissima* (25). However, unlike *T. suavissima*, *T. anguina* is native to the Mediterranean region. The seeds for this study were obtained through the USDA ARS National Genetic Resource Program's Genetic Resources Network. The cultivars used and their accession numbers are as follows: *M. sativa* subsp. *sativa*, PI 673733; *T. anguina*, PI 227394; and *T. suavissima*, PI 198170.

An initial assessment of the metabolic diversity and chemical sensitivity among three representative Australian *T. suavissima* strains and four *Ensifer* reference strains was obtained using an automated Biolog microplate ID system (Biolog, Hayward, CA). A subsample of seven strains was examined, including three Australian *T. suavissima* strains (CC2017, CC2155b, and USDA6670), two *E. meliloti* reference strains (USDA1002^T and USDA1021), and two *E. medicae* reference strains (A321^T and CC169). A dilute suspen-

sion of each strain was inoculated into 96 wells of a GEN III microplate. The results of the 94 biochemical tests were recorded after 6 days.

Based on the results of the microplate biochemical analysis, a subsequent study was undertaken to examine the specific ability of each of the strains in the collection to utilize 2-oxoglutarate (α-ketoglutarate) as a sole carbon source. The following *Ensifer* type strains were also included in this analysis: *E. adhaerens* (ATCC 33499^T), *E. americanum* (CFNEI 156^T), *E. arboris* (HAMBI 1552^T), *E. chiapanecum* (IITG S70^T), *E. fredii* (USDA205^T), *E. kostiense* (HAMBI 1489^T), *E. mexicanus* (IITG R7^T), *E. adhaerens* (Lc04^T), *E. numidicus* (LMG 24690^T), *E. saheli* (HAMBI 215^T), *E. terangae* (HAMBI 220^T), and *E. xingiangensis* (CCBAU 110^T). The inoculants for these assimilation assays were grown for 72 h on tryptone-yeast agar supplemented with 5.3 mM calcium chloride (39). A small portion of a colony was then diluted in 4 ml of sterile phosphate-buffered saline. The resulting suspension was mixed, and an aliquot was collected that was sufficient to provide an optical density at 600 nm (OD_{600}) of 0.010 to 0.015 after addition to a defined broth medium (40) containing 2.5 g liter^{-1} 2-oxoglutarate (α-ketoglutarate) (K1875; Sigma-Aldrich) as the sole carbon source. The cultures were shaken at 200 rpm at 28°C for 72 h, at which time their optical densities were recorded. Strains having OD_{600} values of $>$0.075 were scored as having strong growth, those with values between 0.025 and 0.075 were scored as having moderate-to-weak growth, and those with values $<$0.025 were scored as having poor growth.

The ultrastructure of two of the Australian *T. suavissima* strains (CC2017 and CC2155b) was examined by transmission electron microscopy. Cells were grown for 3 days in modified arabinose-gluconate broth (32). Specimens were then placed on grids and stained with 4% uranyl acetate for 5 min prior to microscopy.

Accession number(s). Accession numbers for the nucleotide sequences in this study are listed alphabetically according to the locus name and increasing allele code numbers in Table S3. Those for newly determined sequences are KX896263 through KX896397.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/AEM.03446-16.

SUPPLEMENTAL FILE 1, PDF file, 0.3 MB.

ACKNOWLEDGMENTS

We thank Charles Murphy (USDA ARS) for electron microscopy and Jeff Newman (Lycoming College) for the initial metabolic and biochemical profiling (Biolog) analyses.

REFERENCES

1. Small E. 2011. Alfalfa and relatives: evolution and classification of *Medicago*. NRC Research Press, Ottawa, ON, Canada.
2. Allen ON, Allen EK. 1981. Papilionoideae: Trifolieae. *Trigonella* L., p 666–668. *In* Allen ON, Allen EK (ed), The Leguminosae, a source book of characteristics, uses, and nodulation. The University of Wisconsin Press, Madison, WI.
3. Eardly BD, Materon LA, Smith NH, Johnson DA, Rumbaugh MD, Selander RK. 1990. Genetic structure of natural populations of the nitrogen-fixing bacterium *Rhizobium meliloti*. Appl Environ Microbiol 56:187–194.
4. Brockwell J, Hely FW. 1966. Symbiotic characteristics of *Rhizobium meliloti*: an appraisal of the systematic treatment of nodulation and nitrogen fixation interactions between hosts and rhizobia of diverse origins. Aust J Agric Res 17:885–899. https://doi.org/10.1071/AR9660885.
5. Gordon DM, Wexler M, Reardon TB, Murphy PJ. 1995. The genetic structure of *Rhizobium* populations. Soil Biol Biochem 27:491–499. https://doi.org/10.1016/0038-0717(95)98624-W.
6. Wernegreen JJ, Riley MA. 1999. Comparison of the evolutionary dynamics of symbiotic and housekeeping loci: a case for the genetic coherence of rhizobial lineages. Mol Biol Evol 16:98–113. https://doi.org/10.1093/oxfordjournals.molbev.a026041.
7. Brockwell J, Evans CM, Bowman AM, McInnes A. 2010. Distribution, frequency of occurrence and symbiotic properties of the Australian native legume *Trigonella suavissima* Lindl. and its associated root-nodule bacteria. Rangeland J 32:395–406. https://doi.org/10.1071/RJ09080.
8. Mitchell TL. 1838. Chapter 2.6, Expedition sent to explore the course of the river Darling, in 1835: Australian shamrock, p 254–255. Three expeditions into the interior of eastern Australia, 2nd ed, vol 1. T. & W. Boone, London, UK. http://gutenberg.net.au/ebooks/e00035.html.
9. Brockwell J. 1971. Patterns of symbiotic behavior in *Trigonella* L. Aust J Agric Res 22:917–921.
10. Petropoulos GA (ed). 2002. Fenugreek: the genus *Trigonella*. Taylor and Francis, New York, NY.
11. Martens M, Delaere M, Coopman R, De Vos P, Gillis M, Willems M. 2007. Multilocus sequence analysis of *Ensifer* and related taxa. Int J Syst Evol Microbiol 57:489–503. https://doi.org/10.1099/ijs.0.64344-0.
12. Merabet C, Martens M, Mahdi M, Zakhia F, Sy A, Le Roux C, Domergue O, Coopman R, Bekki A, Mars M, Willems A, de Lajudie P. 2010. Multilocus sequence analysis of root nodule isolates from *Lotus arabicus* (Senegal), *Lotus creticus*, *Argyrolobium uniflorum* and *Medicago sativa* (Tunisia) and description of *Ensifer numidicus* sp. nov. and *Ensifer garamanticus* sp. nov. Int J Syst Evol Microbiol 60:664–674. https://doi.org/10.1099/ijs.0.012088-0.
13. Rome S, Fernandez MP, Brunel B, Normand P, Cleyet-Marel JC. 1996. *Sinorhizobium medicae* sp. nov., isolated from annual *Medicago* spp. Int J Syst Microbiol 46:972–980. https://doi.org/10.1099/00207713-46-4-972.
14. Reeve W, Chain P, O'Hara G, Ardley J, Nandesena K, Brau L, Tiwari R, Malfatti S, Kiss H, Lapidus A, Copeland A, Nolan M, Land M, Hauser L, Chang Y-J, Ivanova N, Mavromatis K, Markowitz V, Krypides N, Gollagher M, Yates R, Dilworth M, Howieson J. 2010. Complete genome sequence of the *Medicago* microsymbiont *Ensifer* (*Sinorhizobium*) *medicae* strain WSM419. Stand Genomic Sci 2:77–86. https://doi.org/10.4056/sigs.43526.
15. Galibert F, Finan TM, Long SR, Puhler A, Abola P, Ampe F, Barloy-Hubler F, Barnett MJ, Becker A, Boistard P, Bothe G, Boutry M, Bowser L, Buhrmester J, Cadieu E, Capela D, Chain P, Cowie A, Davis RW, Deano S, Federspiel NA, Fisher RF, Gloux S, Godrie T, Goffeau A, Golding B, Gouzy J, Gurjal M, Hernanadez-Lucas I, Hong A, Huizar L, Hyman RW, Jones T, Kahn D, Kahn ML, Kalman S, Keating DH, Kiss E, Komp C, Lelaure V, Masuy D, Palm C, Peck MC, Pohl TM, Portetelle D, Purnelle B, Ramsperger U, Surzycki R, Thebault P, Vandenbol M, et al. 2001. The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science 293:668–672. https://doi.org/10.1126/science.1060966.
16. Kuhn S, Stiens M, Puhler A, Schluter A. 2008. Prevalence of pSmeSM11a-like plasmids in indigenous *Sinorhizobium meliloti* strains isolated in the course of a field release experiment with genetically modified *S. meliloti*

strains. FEMS Microbiol Ecol 63:118–131. https://doi.org/10.1111/j.1574-6941.2007.00399.x.
17. Harrison PW, Lower RP, Kim NK, Young JP. 2010. Introducing the bacterial 'chromid': not a chromosome, not a plasmid. Trends Microbiol 18:141–148. https://doi.org/10.1016/j.tim.2009.12.010.
18. van Berkum P, Elia P, Eardly BD. 2006. Multilocus sequence typing as an approach for population analysis of *Medicago*-nodulating rhizobia. J Bacteriol 188:5570–5577. https://doi.org/10.1128/JB.00335-06.
19. van Berkum P, Elia P, Eardly BD. 2010. Application of multilocus sequence typing to study the genetic structure of megaplasmids in *Medicago*-nodulating rhizobia. Appl Environ Microbiol 76:3967–3977. https://doi.org/10.1128/AEM.00251-10.
20. Epstein B, Branca A, Mudge J, Bharti AK, Briskine R, Farmer AD, Sugawara M, Young ND, Sadowsky MJ, Tiffin P. 2012. Population genomics of the facultatively mutualistic bacteria *Sinorhizobium meliloti* and *S. medicae*. PLoS Genet 8:e1002868. https://doi.org/10.1371/journal.pgen.1002868.
21. Rodriguez-R LM, Konstantinidis K. 2014. Bypassing cultivation to identify species. Microbe Wash DC 9:111–118. https://doi.org/10.1128/microbe.9.111.1.
22. Kim M, Oh H-S, Park S-C, Chun J. 2014. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 64:346–351. https://doi.org/10.1099/ijs.0.059774-0.
23. Capela D, Barloy-Hubler F, Gouzy J, Bothe G, Ampe F, Batut J, Boistard P, Becker A, Boutry M, Cadieu E, Dreano S, Gloux S, Godrie T, Goffeau A, Kahn D, Kiss E, Lelaure V, Masuy D, Pohl T, Portetelle D, Puhler A, Purnelle B, Ramsperger U, Renard C, Thebault P, Vandenbol M, Weidner S, Galibert F. 2001. Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. Proc Natl Acad Sci U S A 98:9877–9882. https://doi.org/10.1073/pnas.161294398.
24. Kuykendall LD, Hashem FM, Wang ET. 2005. Genus VII. *Sinorhizobium* Chen Yan and Li 1988b, 396^{VP} emend. de Lajudie, Willems, Pot, Dewettinck, Maestrojuan, Neyra, Collins, Dreyfus, Kersters and Gillis 1994, 732, p 358–362. *In* Brenner DJ, Krieg NR, Staley JT, Garrity GM (ed), Bergey's manual of systematic bacteriology, 2nd ed, vol 2. Springer, New York, NY.
25. Maureira-Butler IJ, Pfeil BE, Muangprom A, Osborn TC, Doyle J. 2008. The reticulate history of *Medicago* (Fabaceae). Syst Biol 57:466–482. https://doi.org/10.1080/10635150802172168.
26. Bailly X, Olivieri I, Brunel B, Cleyet-Marel J-C, Bena G. 2007. Horizontal gene transfer and recombination drive the evolution of the nitrogen-fixing symbionts of *Medicago* species. J Bacteriol 189:5223–5236. https://doi.org/10.1128/JB.00105-07.
27. Fred EB, Baldwin IL, McCoy E. 1932. Root nodule bacteria and leguminous plants. University Wisconsin Press, Madison, WI.
28. Lieberman MT, Mallory LM, Simkins S, Alexander M. 1985. Numerical taxonomic analysis of cross-inoculation patterns of legumes and *Rhizobium*. Plant Soil 84:225–244. https://doi.org/10.1007/BF02143186.
29. Reeve W, Tian R, Brau L, Goodwin L, Munk C, Detter C, Tapia R, Han C, Liolios K, Huntemann M, Pati A, Woyke T, Mavrommatis K, Markowitz V, Ivanova N, Krypides N, Willems A. 2013. Genome sequence of *Ensifer arboris* strain LMG 14919^{T}; a microsymbiont of the legume *Prosopis chilensis* growing in Kosti, Sudan. Stand Genomic Sci 9:473–483. https://doi.org/10.4056/sigs.4828625.
30. Nick G, de Lajudie P, Eardly BD, Suomalainen S, Paulin L, Zhang X, Gillis M, Lindstrom K. 1999. *Sinorhizobium arboris* sp. nov. and *Sinorhizobium kostiense* sp. nov., isolated from leguminous trees in Sudan and Kenya. Int J Syst Bacteriol 49:1359–1368. https://doi.org/10.1099/00207713-49-4-1359.
31. Degefu T, Wolde-Meskel E, Frostegard A. 2012. Phylogenetic multilocus sequence analysis identifies seven novel *Ensifer* genospecies isolated from a less-well-explored biogeographical region in East Africa. Int J Syst Evol Microbiol 62:2286–2295. https://doi.org/10.1099/ijs.0.039230-0.
32. van Berkum P. 1990. Evidence for a third uptake hydrogenase phenotype among the soybean bradyrhizobia. Appl Environ Microbiol 56:3835–3841.
33. Lloret L, Ormeno-Orillo E, Rincon R, Martinez-Romero J, Rogel-Hernandez MA, Martinez-Romero E. 2007. *Ensifer mexicanus* sp. nov. a new species nodulating *Acacia angustissima* (Mill.) Kuntze in Mexico. Syst Appl Microbiol 30:280–290. https://doi.org/10.1016/j.syapm.2006.12.002.
34. Martens M, Dawyndt P, Coopman R, Gillis M, De Vos P, Willems A. 2008. Advantage 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. https://doi.org/10.1099/ijs.0.65392-0.
35. Rincon-Rosales R, Lloret L, Ponce E, Martinez-Romero E. 2009. Rhizobia with different symbiotic efficiencies nodulate *Acaciella angustissima* in Mexico, including *Sinorhizobium chiapanecum* sp. nov. which has common symbiotic genes with *Sinorhizobium mexicanum*. FEMS Microbiol Ecol 67:103–117. https://doi.org/10.1111/j.1574-6941.2008.00590.x.
36. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30:2725–2729. https://doi.org/10.1093/molbev/mst197.
37. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol 57:81–91. https://doi.org/10.1099/ijs.0.64483-0.
38. Vincent JM. 1970. A manual for the practical study of root-nodule bacteria. IBP handbook no. 15. Blackwell Scientific Publications, Ltd., Oxford, United Kingdom.
39. Beringer JE. 1974. R factor transfer in *Rhizobium leguminosarum*. J Gen Microbiol 84:188–198.
40. Manhart JR, Wong PP. 1979. Nitrate reductase activities of rhizobia and the correlation between nitrate reduction and nitrogen fixation. Can J Microbiol 25:1169–1174. https://doi.org/10.1139/m79-181.