UNIVERSITY *of York*

This is a repository copy of *Diversity of Sinorhizobium meliloti from the Central Asian Alfalfa Gene Center*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/295/

Article:

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Diversity of *Sinorhizobium meliloti* from the Central Asian Alfalfa Gene Center

Marina L. Roumiantseva,[1]* Evgeny E. Andronov,[1] Larissa A. Sharypova,[1]
Tatjana Dammann-Kalinowski,[2] Mathias Keller,[2] J. Peter W. Young,[3]
and Boris V. Simarov[1]

*Research Institute for Agricultural Microbiology, St.-Petersburg-Pushkin 8, 196608, Russia[1]; Lehrstuhl für Genetik,
Fakultät für Biologie, Universität Bielefeld, D-33501 Bielefeld, Germany[2]; and Department of Biology,
University of York, York YO10 5YW, United Kingdom[3]*

***Sinorhizobium meliloti* was isolated from nodules and soil from western Tajikistan, a center of diversity of the host plants (*Medicago*, *Melilotus*, and *Trigonella* species). There was evidence of recombination, but significant disequilibrium, between and within the chromosome and megaplasmids. The most frequent alleles matched those in the published genome sequence.**

Bacteria that nodulate alfalfa form effective symbioses with three related genera: *Medicago* (alfalfa and perennial and annual medics), *Melilotus* (sweet clover), and *Trigonella* (fenugreek). The bacteria fall into two closely related species, *Sinorhizobium meliloti* and *Sinorhizobium medicae* (16, 17), the latter mainly associated with annual medics around the Mediterranean. The alfalfa-*Sinorhizobium* symbiosis is one of the best-studied plant-microbe associations, and the complete genome sequence of *S. meliloti* strain 1021 has been determined (10). Central Asia was recognized by N. I. Vavilov (19), the pioneer of plant biogeography, as a gene center of alfalfa diversity, where it is believed that *Medicago sativa* plants were first cultivated by humans and the tetraploid alfalfa forms arose. We examined rhizobia from this region in the expectation that they might also be very diverse and shed light on the natural gene pool of *S. meliloti*.

Rhizobia were obtained from nodules and soil collected during an expedition to Tajikistan (Fig. 1) in early summer. Their genetic diversity was characterized by plasmid profiling (11), *Rsa*I digestion of amplified 16S ribosomal DNA (1, 17, 20), and restriction fragment length polymorphism (RFLP) of 10 single-copy loci and four insertion sequence (IS) elements (Table 1). All 27 isolates were identified by ribosomal DNA RFLP as *S. meliloti* rather than *S. medicae*, and all formed effective nodules on *M. sativa* cv. Europe. Isolates trapped from soil differed significantly from those from field nodules in genotype frequencies at the *recA*, *exo*, and *exp* loci; the presence of a 200-kb plasmid; and distribution of three of the four IS elements.

If Central Asia is the center of origin of alfalfa rhizobia as well as of their host plants, a wider range of genetic variants would be expected there than elsewhere. In fact, though, the level of polymorphism among these isolates was surprisingly modest, despite the variety of hosts and sites sampled. Brom-

field et al. (7) studied the diversity in strains isolated from alfalfa in Canada, using RFLP at a different but comparable set of loci. They found 22 chromosomal types, 33 pSymA types, and 18 pSymB types, which contrasts with 7, 9, and 3, respectively, in our Central Asian isolates. Admittedly, they examined many more isolates, but this comparison certainly does not support the idea that Central Asia is a repository of alfalfa symbionts with many divergent endemic genotypes.

At each locus that we examined (four chromosomal, four on pSymA, and two on pSymB), the most common variant among the isolates was the "a" type, which by our definition was the type found in the standard laboratory strain 2011. Nevertheless, there was sufficient polymorphism that only one isolate, CA67, had the "a" genotype at every locus. It seems, therefore, that the choice of 1021 (a derivative of 2011) as a representative of the species for genome sequencing (10) was a remarkably fortunate one. It has the most typical genotype, with alleles that are common not just in agricultural inoculants but also in a population that would be expected to include the breadth of the genetic variation in the species.

Although all the isolates share some RFLP alleles with 2011, there is one strain (CA82) that has a SymA megaplasmid genotype that differs at all four loci examined. In particular, it is the only one that has two novel fragments in place of the 3.9-, 1.8-, and 0.6-kb bands that hybridize to the *nifKDH* probe in all the other strains. This deserves further investigation, as it may have significant functional differences in its symbiosis genes from those of the well-studied alfalfa-nodulating rhizobia. Curiously, the chromosomal and pSymB markers of this isolate are indistinguishable from those of Rm2011, so it seems that the "exotic" pSymA has been transferred into a very typical genetic background.

There is evidence of recombination between loci, both within replicons and between replicons. In this context, "recombination" between plasmid and chromosomal loci can be interpreted as transfer of plasmids from one chromosomal background to another, whereas recombination between markers linked on the same replicon implies physical breakage and reunion of the DNA. If there are at least two different alleles

* Corresponding author. Mailing address: Research Institute for Agricultural Microbiology, Sh. Podbelsky 3, St.-Petersburg-Pushkin 8, 196608, Russia. Phone: 7 812 470 28 02. Fax: 7 812 470 43 62. E-mail: genet@yandex.ru.

TABLE 1. Origin and features of *S. meliloti* isolates native to the Central Asian center of diversity

Isolate	Source[a]	Site[b]	Plasmid size (kb)[c]	RFLP type(s)[d]			IS element[e]			
				Chr	pSymA	pSymB	i	ii	iii	iv
CA67	Trap	1	—	aa	aaaa	aa	8 (a)	4 (a)	0	2
CA69	Trap	3/d	—	ab	afab	aa	0	0	2 (a)	0
CA72	Trap	4/a	240	aa	agaa	ab	12 (a)	1 (a)	6	5 (a)
CA82	Trap	4/a	200	aa	hbdc	aa	0	5	9	9
CA52	Trap	5/a	100	ba	faba	aa	8 (b)	0	0	0
CA85	Trap	5/a	—	ba	idab	aa	4 (a)	1 (b)	2 (b)	0
CA79	Trap	5/b	—	cc	caaa	aa	7 (a)	0	1 (a)	0
CA96	Trap	5/b	—	cb	dcaa	bb	4 (b)	1 (c)	5 (a)	0
CA97	Trap	6	—	cb	dcaa	bb	4 (b)	1 (c)	5 (a)	0
CA76	Trap	7	—	aa	bebb	ba	3	1 (d)	2 (c)	0
CA107	Trap	7	270	ac	aaaa	aa	12 (b)	1 (e)	2 (d)	0
CA101	Trap	8	—	cb	dcaa	bb	4 (b)	1 (c)	5 (a)	0
CA81	*M. sativa*	3/a	200	ac	aaaa	ac	12 (c)	4 (b)	1 (b)	1 (a)
CA104	*M. sativa*	3/b	200	ac	aaaa	aa	16	2	1 (c)	0
CA110	*M. sativa*	3/b	200	ac	ahaa	aa	11	0	5 (b)	5 (b)
CA112	*M. sativa*	3/c	600, 200, 80	ba	gaba	ca	9	0	0	0
CA114	*M. sativa*	3/d	200	ac	ahaa	aa	11	0	0	1 (b)
CA56	*M. sativa*	5/c	—	ba	caaa	aa	21	0	0	0
CA80	*M. sativa*	5/d	—	bc	ifab	aa	7 (b)	2	4 (a)	3
CA78	*M. sativa*	5/b	360	aa	faca	aa	15 (a)	6	8	8
CA99	*Medicago lupulina*	5/b	300, 200, 80	ac	egaa	aa	13	1 (f)	0	1 (c)
CA98	*Medicago lupulina*	5/b	—	ba	gaba	ca	9	4 (c)	0	0
CA116	*Melilotus officinalis*	4/b	—	aa	egaa	aa	15 (b)	6	4 (b)	4
CA83	*Melilotus officinalis*	4/b	—	ac	aiaa	aa	10	0	5 (c)	0
CA105	*Trigonella germiniflora*	4/a	200	ac	ahaa	aa	11	0	5 (d)	0
CA102	*Trigonella popovi*	3/a	200	ac	ahaa	aa	11	0	0	0
CA108	*Trigonella verae*	2	—	ba	gaba	aa	8 (c)	0	0	0
2011[f]	*M. sativa*		—	aa	aaaa	aa	14	8	0	0

[a] Trap, *M. sativa* cv. Europe inoculated with soil; others are from field nodules.

[b] Sites are indicated in Fig. 1; letters indicate different locations at one site.

[c] All isolates contained two megaplasmids (>1,000 kb), except that CA69, CA76, and CA85 had a single megaplasmid band, like strain MVII (12); only smaller plasmids are listed. —, none were detected.

[d] Letters indicate distinct *Eco*RI RFLP patterns at each of eight loci, determined with cloned or PCR-amplified probes: *leuB* (SMc04405) (1,2), *recA* (SMc00760) (8) on the chromosome, *nodD1* (SMa0870) (14, 17), *nodD2* (SMa0757) (17), *nodH* (SMa0850) (9), *nifKDH* (SMa0829, SMa0827, and SMa0825) (3, 15) on pSymA, *exoP* to *exoZ* (SMb20943 to SMb20961) (4, 5), and *expA10* to *expE8* (SMb21307 to SMb21327) (6) on pSymB. Accession numbers of genes used as hybridization probes are cited according to sequence data (10). The additional chromosomal locus *mucR* (SMc00058) (13) was invariant. Chr, chromosomal.

[e] Number of copies of IS*Rm2011-2* (i), IS*Rm2011-1* (ii), IS*Rm220-12-3* (iii), and IS*Rm102F34-1* (iv) (18). Letters distinguish different fingerprints with the same copy number.

[f] 2011 is a standard laboratory strain of *S. meliloti* for comparison; it is the parent of 1021, the source of the complete genome sequence (10).

at each of two loci, and they occur in all combinations, this must indicate either recombination or independent parallel mutation to the same allelic state in different lineages. For example, at the two loci *exoP* to *exoZ* and *expA10* to *expE8* on

megaplasmid 2, all four combinations, "aa," "ab," "ba," and "bb," were found. There are many pairs of loci for which this is true ("+" in Table 2), which implies that recombination has been frequent, since so much parallel evolution is implausible.

TABLE 2. Pairwise tests of association and recombination between genetic markers[a]

Genetic marker	Result for genetic marker:												
	leu	*recA*	*nodD1*	*nodD2*	*nodH*	*nifHDK*	*exo*	*exp*	200-kb plasmid[b]	IS*2011-2*	IS*2011-1*	IS*220-12-3*	IS*102F34-1*
leu		0.003	0.000	0.030	NS	NS	0.002	0.012	NS	0.002	NS	NS	NS
recA	+		0.008	0.001	0.021	NS	0.004	0.005	0.020	0.004	0.044	0.012	NS
nodD1		+		0.001	0.000	0.010	0.000	NS	NS	0.000	NS	NS	NS
nodD2	+	+			NS	0.000	0.017	0.038	0.045	0.000	0.025	0.004	NS
nodH	+					NS	NS	NS	NS	NS	NS	NS	NS
nifHDK	+	+			+		NS	NS	NS	NS	NS	0.039	NS
exo	+	+		+		+		0.018	NS	0.033	0.028	NS	NS
exp	+	+		+			+		NS	NS	0.003	0.037	NS
200-kb plasmid	+	+	+	+	+		+			NS	NS	NS	NS
IS*2011-2*											0.001	0.009	NS
IS*2011-1*	+			+		+			+	+		0.009	0.016
IS*220-12-3*													NS
IS*102F34-1*													

[a] The upper half shows probabilities based on the null hypothesis of no association (exact contingency test); NS, *P* > 0.05, i.e., no significant association. The lower half shows evidence (+) for recombination or parallel evolution of loci (see text).

[b] Presence or absence of a 200-kb plasmid.
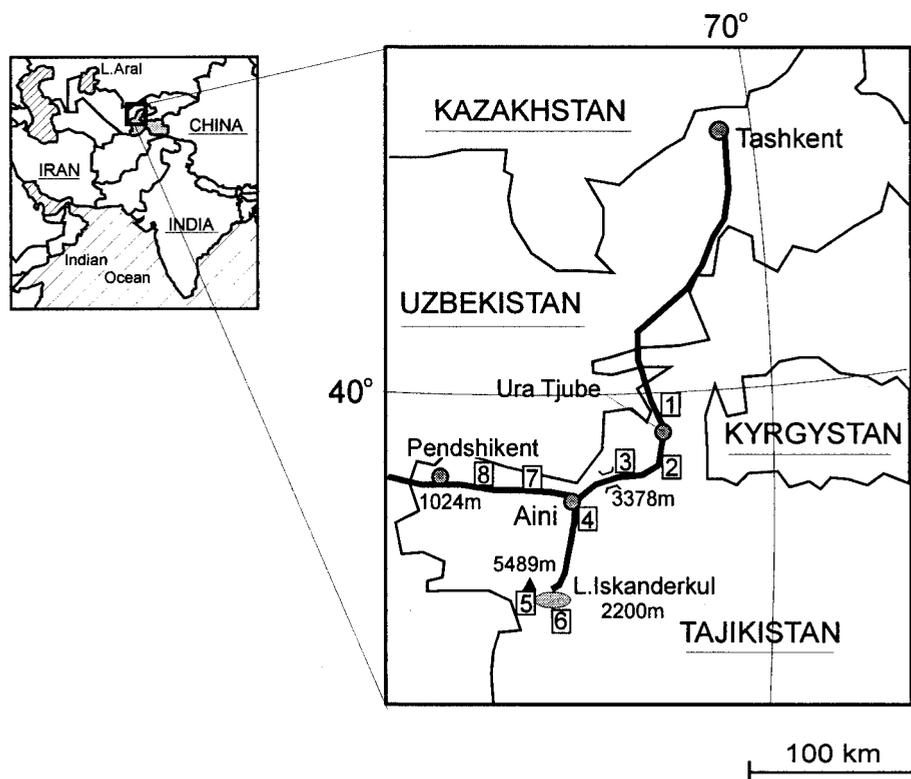
FIG. 1. Map of Central Asia showing the Tajikistan region where nodules from *Medicago*, *Melilotus*, and *Trigonella* species and soil samples were collected. The black line indicates the route of the expedition; numbers from 1 to 8 represent the collection sites.

At the same time, recombination has not been so frequent as to eliminate linkage disequilibrium, which was significant between many pairs of loci (values of $P < 0.05$ in Table 2). Linkage disequilibrium was analyzed by an extension of Fisher's exact test, suitable for small samples and implemented in Arlequin 2.0 (S. Schneider, D. Roessli, and L. Excoffier, Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland). The estimation was based on sampling the space of possible contingency tables via a random Markov chain of 2 million steps, following 10,000 dememorization steps. All significant P values were within 0.002 in duplicate runs. In part, this linkage disequilibrium may reflect the complex sampling structure, but in any case such disequilibrium is not surprising in a bacterial population, because it can be generated rapidly by "epidemic" reproduction of individual clones. There is, however, no indication of a predominantly epidemic population structure in this case, since almost every isolate was genetically different. The only "clonal" isolates were the three trap isolates CA96, CA97, and CA101. These had a distinctive allele combination on each of the three replicons and clearly represent a well-established clone since they were isolated from three different locations. They form part of the rich genetic structure of the population, which must reflect the joint effects of mutation, recombination, and selection.

REFERENCES

1. **Andronov, E. E., M. L. Roumiantseva, and B. V. Simarov.** 2001. Genetic diversity of a natural population of *Sinorhizobium meliloti* revealed in analysis of cryptic plasmids and ISRm2011–2 fingerprints. Russ. J. Genet. **37:**494–499.
2. **Aronshtam, A. A., B. R. Umarov, V. N. Yerko, E. E. Andronov, and B. V. Simarov.** 1993. The use of the cosmid gene bank of *Rhizobium meliloti* for the molecular cloning of leucine biosynthesis gene involved in genetic control of nitrogen fixing symbiosis with alfalfa. Genetika **29:**235–245. (In Russian.)
3. **Banfalvi, Z., E. Kondorosi, and A. Kondorosi.** 1985. *Rhizobium meliloti* carries two megaplasmids. Plasmid **13:**129–138.
4. **Becker, A., A. Kleickmann, H. Küster, M. Keller, W. Arnold, and A. Pühler.** 1993. Analysis of the *Rhizobium meliloti* genes *exo*U, *exo*V, *exo*W, *exo*T, and *exo*I involved in exopolysaccharide biosynthesis and nodule invasion: *exo*U and *exo*W probably encode glucosyltransferases. Mol. Plant-Microbe Interact. **6:**735–744.
5. **Becker, A., H. Küster, K. Niehaus, and A. Pühler.** 1995. Extension of the *Rhizobium meliloti* succinoglycan biosynthesis gene cluster: identification of the *exsA* gene encoding an ABC transporter protein, and the *exsB* gene which probably codes for a regulator of succinoglycan biosynthesis. Mol. Gen. Genet. **249:**487–497.
6. **Becker, A., S. Rüberg, H. Küster, A. Roxlau, M. Keller, T. Ivashina, H.-P. Cheng, G. C. Walker, and A. Pühler.** 1997. The 32-kilobase *exp* gene cluster of *Rhizobium meliloti* directing the biosynthesis of galactoglucan: genetic organization and properties of the encoded gene products. J. Bacteriol. **179:**1375–1384.
7. **Bromfield, E. S. P., A. M. P. Behara, R. S. Singh, and L. R. Barran.** 1998. Genetic variation in local populations of *Sinorhizobium meliloti*. Soil Biol. Biochem. **30:**1707–1716.
8. **Dammann-Kalinowski, T., S. Niemann, M. Keller, W. Selbitschka, C. C. Tebbe, and A. Pühler.** 1996. Characterization of two bioluminescent *Rhizobium meliloti* strains constructed for field releases. Appl. Microbiol. Biotechnol. **45:**509–512.

9. **Fisher, R. F., J. A. Swanson, J. T. Mulligan, and S. R. Long.** 1987. Extended region of nodulation genes in *Rhizobium meliloti* 1021. II. Nucleotide sequence, transcription start sites and protein products. Genetics **117:**191–201.

10. **Galibert, F., T. M. Finan, S. R. Long, A. Pühler, P. Abola, F. Ampe, F. Barloy-Hubler, M. J. Barnett, A. Becker, P. Boistard, G. Bothe, M. Boutry, L. Bowser, J. Capela, P. Chain, A. Cowie, R. W. Davis, S. Dreano, N. A. Federspiel, R. F. Fisher, S. Gloux, T. Godrie, A. Goffeau, B. Golding, J. Gouzy, M. Gurjal, I. Hernandez-Lucas, A. Hong, L. Huizar, R. W. Hyman, T. Jones, D. Kahn, M. L. Kahn, S. Kalman, D. H. Keating, E. Kiss, C. Komp, V. Lelaure, D. Masuy, C. Palm, M. C. Peck, T. M. Pohl, D. Portetelle, B. Purnelle, U. Ramsperger, R. Surzycki, P. Thebault, M. Vandenbol, F.-J. Vorholter, S. Weidner, D. H. Wells, K. Wong, K.-C. Yeh, and J. Batut.** 2001. The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science **293:**668–672.

11. **Hynes, M. F., R. Simon, P. Müller, L. Niehaus, M. Labes, and A. Pühler.** 1986. The two megaplasmids of *Rhizobium meliloti* are involved in the effective nodulation of alfalfa. Mol. Gen. Genet. **202:**356–362.

12. **Hynes, M. F., R. Simon, and A. Pühler.** 1985. The development of plasmid-free strains of *Agrobacterium tumefaciens* by using incompatibility with a *Rhizobium meliloti* plasmid to eliminate pAtC58. Plasmid **13:**99–105.

13. **Keller, M., A. Roxlau, W. M. Weng, M. Schmidt, J. Quandt, K. Niehaus, D. Jording, W. Arnold, and A. Pühler.** 1995. Molecular analysis of the *Rhizobium meliloti muc*R gene regulating the biosynthesis of the exopolysaccharides succinoglucan and galactoglucan. Mol. Plant-Microbe Interact. **8:**267–277.

14. **Kondorosi, E., Z. Banfalvi, and A. Kondorosi.** 1984. Physical and genetic analysis of a symbiotic region of *Rhizobium meliloti*: identification of nodulation genes. Mol. Gen. Genet. **193:**445–452.

15. **Reiländer, H.** 1987. Molecular analysis of the *nif* gene region of *Rhizobium meliloti*. Ph.D. thesis. Bielefeld University, Bielefeld, Germany.

16. **Rome, S., M. P. Fernandez, B. Brunel, P. Normand, and J. C. Cleyet-Marel.** 1996. *Sinorhizobium medicae* sp. nov., isolated from annual *Medicago* spp. Int. J. Syst. Bacteriol. **46:**972–980.

17. **Roumiantseva, M. L., V. V. Yakutkina, T. Dammann-Kalinowski, L. A. Sharypova, M. Keller, and B. V. Simarov.** 1999. Comparative analysis of structural organization of the genome in alfalfa nodule bacteria *Sinorhizobium medicae* and *Sinorhizobium meliloti*. Russ. J. Genet. **35:**128–135.

18. **Simon, R., B. Hötte, B. Klauke, and B. Kosier.** 1991. Isolation and characterization of insertion sequence elements from gram-negative bacteria by using new broad-host-range, positive selection vectors. J. Bacteriol. **173:**1502–1508.

19. **Vavilov, N. I.** 1926. Centers of origin of cultivated plants. Trends Pract. Bot. Genet. Sel. **16:**3–248. (In Russian.)

20. **Weidner, S., W. Arnold, and A. Pühler.** 1996. Diversity of uncultured microorganisms associated with the seagrass *Halophila stipulacea* estimated by restriction fragment length polymorphism analysis of PCR-amplified 16S rRNA genes. Appl. Environ. Microbiol. **62:**766–771.