APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Feb. 2003, p. 1067–1074 Vol. 69, No. 2
0099-2240/03/$08.00+0 DOI: 10.1128/AEM.69.2.1067–1074.2003

Symbiotic and Genetic Diversity of *Rhizobium galegae* Isolates Collected from the *Galega orientalis* Gene Center in the Caucasus

E. E. Andronov,[1]* Z. Terefework,[2] M. L. Roumiantseva,[1] N. I. Dzyubenko,[3] O. P. Onichtchouk,[1] O. N. Kurchak,[1] A. Dresler-Nurmi,[2] J. P. W. Young,[4] B. V. Simarov,[1] and K. Lindström[2]

Research Institute of Agricultural Microbiology, St. Petersburg, Pushkin 196608,[1] and Department of Forage Sciences, N. I. Vavilov Russia Research Institute of Plant Genetic Resources, St. Petersburg 190000,[3] Russia; Department of Applied Chemistry and Microbiology, Biocenter 1, Fin-0014 University of Helsinki, Finland[2]; and Department of Biology, University of York, York YO10 5YW, United Kingdom[4]

Received 10 June 2002/Accepted 13 November 2002

This paper explores the relationship between the genetic diversity of rhizobia and the morphological diversity of their plant hosts. *Rhizobium galegae* strains were isolated from nodules of wild *Galega orientalis* and *Galega officinalis* in the Caucasus, the center of origin for *G. orientalis*. All 101 isolates were characterized by genomic amplified fragment length polymorphism fingerprinting and by PCR-restriction fragment length polymorphism (RFLP) of the rRNA intergenic spacer and of five parts of the symbiotic region adjacent to *nod* box sequences. By all criteria, the *R. galegae* bv. officinalis and *R. galegae* bv. orientalis strains form distinct clusters. The *nod* box regions are highly conserved among strains belonging to each of the two biovars but differ structurally to various degrees between the biovars. The findings suggest varying evolutionary pressures in different parts of the symbiotic genome of closely related *R. galegae* biovars. Sixteen *R. galegae* bv. orientalis strains harbored copies of the same insertion sequence element; all were isolated from a particular site and belonged to a limited range of chromosomal genotypes. In all analyses, the Caucasian *R. galegae* bv. orientalis strains were more diverse than *R. galegae* bv. officinalis strains, in accordance with the gene center theory.

Rhizobium galegae (14) is a species that forms an effective symbiosis with plants of *Galega orientalis* and *Galega officinalis*, the only species in the genus *Galega* (Fabaceae) that have been studied for symbiosis. This symbiotic system represents a sharply defined cross-inoculation group. However, there are some differences in symbiotic performance. *R. galegae* strains are able to infect both *Galega* species, but strains isolated from *G. officinalis* form effective nodules on that plant and ineffective nodules on *G. orientalis*, while the converse is true for strains from *G. orientalis* (16). This finding, along with a wide range of phenotypic and genotypic approaches using numerical taxonomy (15), phage typing, DNA homology (13, 40), lipopolysaccharide and protein patterns (16), plasmid profiling (29), phylogeny of ribosomal genes (25, 35), randomly amplified polymorphic DNA, and repetitive PCR (21, 28), led to the proposal of two biovars for strains forming an effective symbiosis with *G. officinalis* (*R. galegae* bv. officinalis) and *G. orientalis* (*R. galegae* bv. orientalis) (24). Symbiosis-related genetic traits were found to be the main factor in genetic divergence between the biovars (24). *G. orientalis* is a good fodder plant with some widely used cultivars, whereas *G. officinalis* is rather poisonous because of its high alkaloid content and has no agricultural importance.

The interesting symbiotic properties, together with the extensive taxonomic knowledge about *R. galegae*, prompted us to use the species to answer important questions related to the diversity and evolution of rhizobia and their symbiosis with legumes. A recent study indicated that *R.galegae* bv. orientalis strains have lower genetic diversity than *R. galegae* bv. officinalis strains (37). However, very few biovar orientalis strains were included in that study. In order to obtain a larger sample, *R. galegae* strains belonging to both biovars were collected during an expedition to the Caucasus, one of the centers of diversity of *G. orientalis* but not of *G. officinalis*. The collection was used for two different purposes.

First, we wanted to investigate the influence of the host plant on the genetic diversity of rhizobial populations. The host plant is an important factor shaping the genetic structure of a natural population of rhizobia (1, 2, 4, 9, 42, 45). The genomic diversity of the collected rhizobia was assessed by amplified fragment length polymorphism (AFLP) fingerprinting and by restriction fragment typing of the rRNA gene internal transcribed spacer (ITS). It is generally believed that insertion sequence (IS) elements in rhizobia promote genetic diversification through genomic rearrangements and recombination (8, 11, 17). The use of IS elements as probes for Southern hybridization can provide high-resolution fingerprints of rhizobial strains (26, 27, 31) and some information for understanding the function of these elements in rhizobial populations (2). An IS-related sequence, closely related to an *Agrobacterium* transposase from the widely studied plasmid pTiA6NC, has been found in the symbiotic region of *R. galegae* bv. orientalis strain HAMBI540

* Corresponding author. Mailing address: Russia Research Institute of Agricultural Microbiology, Podbelsky, 3, St. Petersburg, Pushkin 196608, Russia. Phone: 812 4762802. Fax: 812 4704362. E-mail: eeandr@yandex.ru.

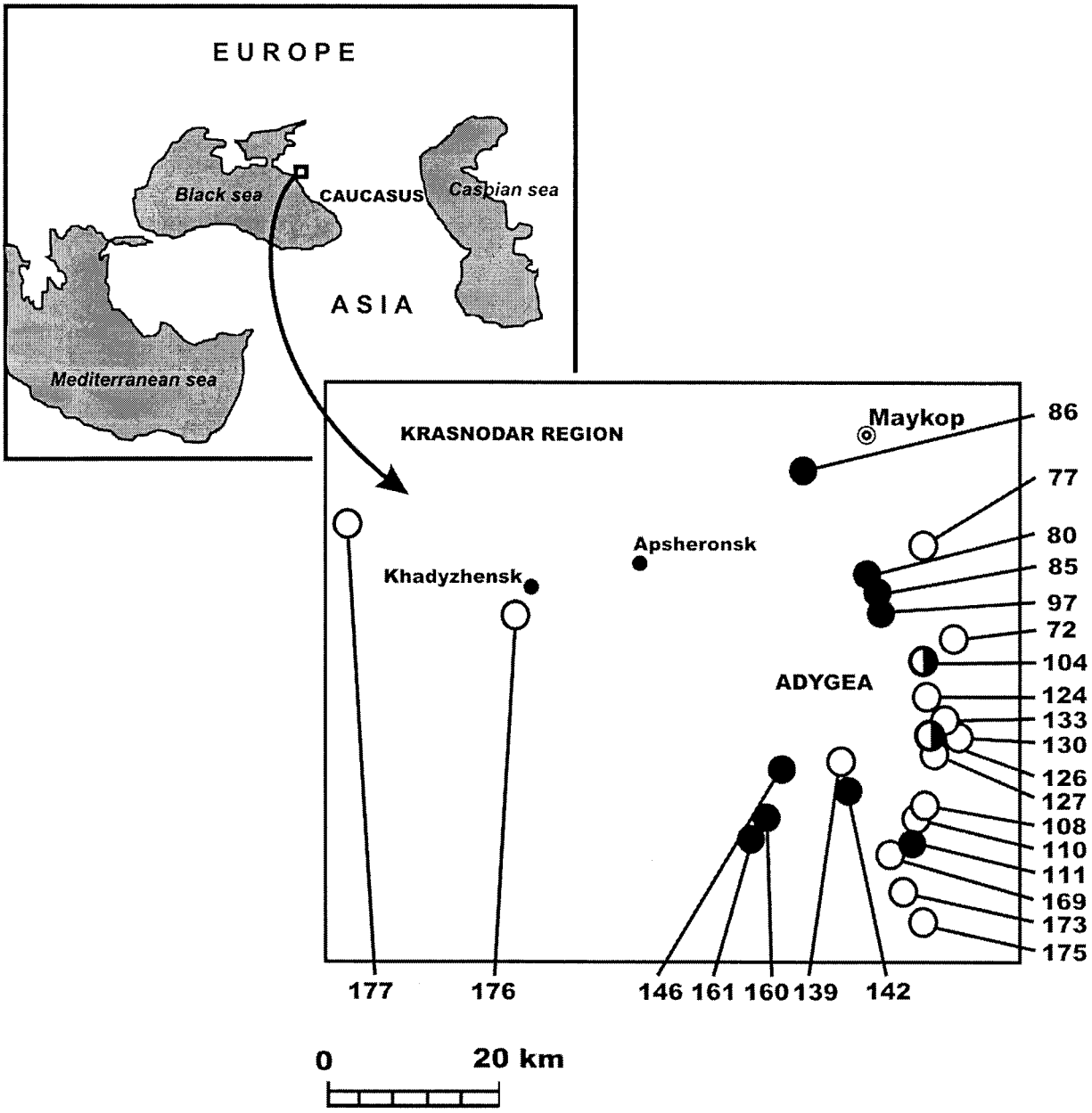

FIG. 1. Sampling locations. Solid circles, sites where *G. orientalis* grew; open circles, sites where *G. officinalis* grew; split circles, sites where both host plants grew.

(24, 33), so we also used this IS sequence for fingerprinting the *R. galegae* strains.

Second, the collection was used to get a picture of the roles that the different parts of the symbiotic genome play in processes related to the evolutionary adaptation of rhizobia to different host plants. The genetic diversity in the symbiotic gene regions of *R. galegae* strains belonging to the two biovars was studied by PCR-restriction fragment length polymorphism (RFLP) analysis of five different symbiotic loci containing previously identified *nod* box sequences (32).

MATERIALS AND METHODS

Expedition and plant sampling. *G. officinalis* and *G. orientalis* plants were sampled during the INTAS expedition to the Caucasus (Russia) in the summer of 1999 (Fig. 1). The plants, with roots, were taken from 25 sites (~1 ha each) along a 300-km route through a mountainous part of the Krasnodar region in the northwest of the Caucasus (Fig. 1). An interesting observation was that the *Galega* species were found growing separately, with the exception of only two sites (Fig. 1, sites 104 and 126). *Galega* species in the area surveyed were readily distinguished by morphological traits of inflorescences and of the root system. *G. orientalis* was always part of well-established plant associations, whereas *G. officinalis* preferred disturbed or destroyed ecotopes and was part of the pioneer flora. The *G. officinalis* population was homogeneous morphologically, while the *G. orientalis* population was differentiated into distinct ecotypes, as is typical for the species. At least four well-differentiated ecotypes of *G. orientalis* were present in the region: mountain ecotypes A, from the Lagonaki tract (site 161), and B, from the mountain Otrub (site 97), were characterized by great height, good foliation, and a large inflorescence (up to 40 to 50 cm). The mountain ecotype C from Mezmay (site 160) was distinguished by dark foliage, and ecotype D from Khadyzhensk (site 176) and from Kamenoostrovsky (sites 104, 124, 126,

and 127) was not as tall as the mountain ecotypes; had paler, shorter inflorescences; and sometime grew together with *G. officinalis* plants (sites 104 and 126).

Bacterial strains. The plants sampled were brought to the laboratory and nodules were cut from the roots and washed in sterile water for 2 min in 70% ethanol, 4 min in 4% $Ca(OCl)_2$, and three times in sterile water. The nodules were crushed, suspended in 100 μl of liquid TY medium (3), and streaked out on TY agar. Single colonies were picked and restreaked. In total, 47 strains were isolated from *G. orientalis* sampled from 10 sites and 54 strains were isolated from *G. officinalis* sampled from 17 sites (Table 1). Two reference strains, HAMBI540 (*R. galegae* bv. orientalis) and HAMBI1141 (*R. galegae* bv. officinalis), were included in the study. All strains are available from the HAMBI collection, University of Helsinki.

Plant test. The symbiotic properties of the five *R. galegae* bv. officinalis isolates recovered from *G. orientalis* nodules were studied on *G. orientalis* cv. Nadezhda and *G. officinalis* var. E-103 (Maykop region) from the VIR (N. I. Vavilov Russia Research Institute of Plant Genetic Resources) collection for 30 days under nitrogen-deficient sterile conditions in glass tubes containing vermiculite (30). Each strain was tested in two replicates with each host plant.

DNA isolation. Total genomic DNAs of all *R. galegae* strains were obtained after lysozyme-sodium dodecly sulfate lysis followed by phenol-chloroform extraction and ethanol precipitation (12).

AFLP analyses. The AFLP procedure was performed as described by Vos et al. (39), with some modifications. One hundred to 500 ng of DNA was digested with *Eco*RI and *Tru*I (*Mse*I) (MBI Fermentas) restriction enzymes in a PTC-200 thermal cycler (MJ Research, Watertown, Mass.) and, in the same step, ligated with double-stranded adapters specific for each restriction half-site. (7). Two sets of primers were used in separate PCRs: one set with two selective nucleotides (boldface) (*Mse*I-gc [GAT GAG TCC TGA GTA A**GC**] and *Eco*RI-gc [GAC GTC GTA CCA ATT CGA **GC**]) and another set with a mixture of two and three selective nucleotides (*Mse*I-gc and *Eco*RI-gag [GAC GTC GTA CCA ATT CGA **GAG**]). The PCR conditions were as described previously (37). Silver-stained polyacrylamide gels, prepared as described by Dresler-Nurmi et al. (7), were scanned with an AGFA scanner and analyzed using Bionumerics software, version 2.0 (Applied Maths, Kortrijk, Belgium). An unweighted pair group method with averaging (UPGMA) dendrogram was generated from all strains by using Pearson correlation coefficients. For further analyses, including PCR-RFLP and IS fingerprinting, a representative strain (Table 1) was chosen from each AFLP group. Based on AFLP data, a UPGMA dendrogram was constructed for each biovar. We then chose representatives from each AFLP cluster and loaded them on the same gel; there were 20 *R. galegae* bv. orientalis and 21 *R. galegae* bv. officinalis strains with distinct fingerprints. Two different sets of primers were used and two gels, containing all 41 chosen strains, were obtained. The final dendrogram was constructed from the combined gel. The HAMBI1141 and HAMBI540 strains were excluded from the final dendrogram construction to avoid distorting the comparison of the branching depths of the two biovars studied.

PCR amplification. The PCRs were carried out with 5 to 10 ng of total DNA. Dynazyme DNA polymerase (Finnzymes) was used with the supplied buffer (1.5 mm $MgCl_2$) and deoxynucleotide triphosphates. The primers fD1 and rD1 were used to amplify 16S rRNA genes (41). For ITS amplification, the primers FGPS1490-72 and FGPL132′ (22) were used. The locations of primers for amplifying regions containing *nod* boxes are shown in Fig. 2. Primers for *nod* box regions were constructed using Jellyfish software (Biowire): nb1f, TCC ATC TCC TAG ATG CCT CA; nb1r, GGC TAC GCA GTG ATT TGA TG; nodDf, CAG ATG TTG CCT GGG TCG; nodDr, ATC AAC CTC AGT CAA CCG GC; nb2f, TCG ATT GTT TGA ATG CCA TA; nb2r, GGC AAC ACG TCC AAC TTT CT; nb3f, GCG TGA ATT GGT GTC ATC CT; nb3r, AAG CTC GTC GCC AAA ATA GA; nb4f, GCC GCG ATA TAC TGA CCA TT; nb4r, TAT TCG AGG CGT GGA TGT CT; nb5f, GGT CAT GCT CTC GAT GGT TT; and nb5r, CTT CAA ATC CAC GCA ATC AA. All PCRs were performed using a PTC-200 Peltier Thermal Cycler (MJ Research) with the following standard temperature profile: an initial denaturation at 95°C for 3 min; 30 cycles of denaturation (1 min at 94°C), annealing (1 min at 55°C), and extension (1.5 min at 72°C); and final extension at 72°C for 3 min. When amplification was weak, the PCR was repeated with a lower annealing temperature (52 or 53°C). The amplified DNA was examined by electrophoresis in 1% agarose gels.

RFLP analyses. Aliquots of PCR products were digested with the restriction endonucleases *Hae*III, *Msp*I, and *Alu*I (MBI Fermentas) using the Y^+/Tango buffer supplied. The restricted DNA was analyzed by electrophoresis in 4% agarose. pGEM DNA markers (Promega) were used. The gels were stained with ethidium bromide and photographed with the Kodak EDAS290 system. The RFLP type for each particular endonuclease and combined RFLP types were defined for each strain and for each locus. The genetic diversity at each locus was calculated as $H = n(1 - \Sigma x_i^2)/(n - 1)$, where x_i is the frequency of the ith RFLP type and n is the number of RFLP types (5). Similarities between the HAMBI540 and HAMBI1141 strains in symbiotic loci were estimated from the proportion of shared and unshared restriction fragments by using a similarity coefficient (20).

Southern hybridization. An aliquot (0.5 to 1 μg) of the total DNA sample was digested to completion with *Hin*dIII (MBI Fermentas). Restriction fragments were separated by electrophoresis in 1% agarose. DNA of phage λ digested with *Hin*dIII was included on the gel as a size marker. Southern blotting of DNA, hybridization, and detection were done according to instructions provided with the nonradioactive labeling and detection kit (Roche). The digoxigenin (DIG) labeling of the IS probe was performed using PCR amplification with the ISf (CGA TTG CTA CAA TGG CTT CA) and ISr (TCT TTC TTC CAC ATG CAA CC) primers, flanking an inner fragment of the unnamed IS element, constructed from sequence data of *R.galegae* strain HAMBI540 (34).

RESULTS AND DISCUSSION

A number of studies suggest that the host plant is the most important factor that shapes the structure of rhizobial populations (1, 2, 4, 5, 10, 42, 45; M. L. Roumiantseva, N. A. Provorov, and B. V. Simarov, Proc. VIII Eastern Eur. Symp. Biol. Nitrogen Fixation, p. 53, 1992). Moreover, the process of fine tuning symbiosis can be to some extent inferred from the contemporary diversity of both host plants and their rhizobial microsymbionts. It appears that the greater the genetic diversity of the host plant population, the more diverse is the rhizobial population associated with it. This can best be shown by comparing rhizobia belonging to the same species or even biovar but isolated from different host plants (4, 5, 42). Additional evidence comes from the higher diversity of rhizobia in gene centers of host plants (46).

One hundred and one *R. galegae* strains were isolated from nodules of wild-growing *G. officinalis* and *G. orientalis* in this work. All the strains had 16S rRNA *Msp*I restriction profile identical to that of the reference *R. galegae* strains HAMBI540 and HAMBI1141 (data not shown). AFLP fingerprinting (Fig. 3), ITS PCR-RFLP, and *nod* box PCR-RFLP (Table 1) clearly distinguished the strains isolated from each host plant. However, five strains originally isolated from *G. orientalis* nodules belong by their AFLP and ITS and *nod* box PCR-RFLP patterns to the *G. officinalis* cluster (Table 1). The symbiotic performances of these five isolates were studied. They produced relatively large pink nodules on *G. officinalis* plants and small white or light-green nodules on *G. orientalis* plants. Thus, we can conclude that these strains are *R. galegae* bv. officinalis even though they were isolated from nodules of the nontypical host plant *G. orientalis*.

The Caucasus region is known as the gene center (or a center of diversity) for many legumes (38, 47) and in particular for *Galega* species (18). The results of our expedition confirm this. However, there is a significant difference between *G. orientalis* and *G. officinalis*. The *G. orientalis* populations in the region explored are exceptionally diverse and have a longer history than the *G. officinalis* populations.

The data obtained in the study by molecular characterization of rhizobial isolates correlates well with host plant diversity. AFLP analysis, which is based on whole-genome variability, shows that the Caucasian populations of *R. galegae* bv. orientalis are more diverse than the *R.galegae* bv. officinalis populations. The branching depth in the UPGMA dendrogram is deeper in the case of *R. galegae* bv. orientalis (Fig. 3). Interestingly, the *R. galegae* bv. orientalis strains group into two

TABLE 1. *R. galegae* strains used in the study

AFLP group[a]	Host plant	Strain(s)[b]	Site(s)	RFLP type revealed in rRNA ITS and symbiotic loci[d]					
				ITS	nb1	nb2	nb3	nb4	nb5
1	*G. officinalis*	G003	72	A	A	–[c]	A	A	A
2	*G. officinalis*	G004	72	C	A	–	A	A	B
3	*G. officinalis*	G005*, G007	72	B	A	–	A	A	B
4	*G. officinalis*	G006	72	B	A	–	A	A	A
5	*G. officinalis*	G008	72	B	A	–	A	A	B
6	*G. officinalis*	G009	72	B	A	–	A	A	A
7	*G. officinalis*	G010*, G011, G027, G028	72, 77, 127	A	A	–	A	A	A
8	*G. officinalis*	G013, G014*, G015, G018, G021	108	B	A	–	A	A	B
9	*G. officinalis*	G017	108	A	A	–	A	A	A
10	*G. officinalis*	G019*, G020	108	A	A	–	A	A	B
11	*G. officinalis*	G022	110	B	A	–	A	A	–
12	*G. officinalis*	G023	110	B	A	–	A	A	A
13	*G. officinalis*	G024	110	D	A	–	A	A	–
14	*G. officinalis*	G025	124	B	A	–	A	A	B
15	*G. officinalis*	G026	127	A	A	–	A	A	A
16	*G. officinalis*	G029	127	B	–	–	A	A	A
17	*G. officinalis*	G030	127	C	A	–	A	A	A
18	*G. officinalis*	G031	127	B	A	–	A	A	A
19	*G. officinalis*	G032*, G035	130	A	A	–	A	A	A
20	*G. officinalis*	G033*, G038	130	B	A	–	A	A	B
21	*G. officinalis*	G034	130	B	A	–	A	A	B
22	*G. officinalis*	G036	130	B	A	–	A	A	A
23	*G. officinalis*	G037	130	B	A	–	A	A	B
24	*G. officinalis*	G039	133	B	A	–	A	A	A
25	*G. officinalis*	G040*, G041, G042, G043, G046, G050	139, 169, 176	B	A	–	A	A	B
26	*G. officinalis*	G045	173	B	A	–	A	A	B
27	*G. officinalis*	G047	173	E	B	–	A	A	B
28	*G. officinalis*	G048	173	E	A	–	A	A	B
29	*G. officinalis*	G049	175	F	C	–	A	A	B
30	*G. officinalis*	G053	176	A	A	–	A	A	A
31	*G. officinalis*	G054	177	B	A	–	A	A	B
32	*G. officinalis*	G055	177	A	A	–	A	A	A
33	*G. officinalis*	G056	177	B	A	–	A	A	A
34	*G. orientalis*	G093*, G097	104	A	A	–	A	A	A
35	*G. orientalis*	G098	126	B	A	–	A	A	A
36	*G. orientalis*	G110*, G112	161	B	A	–	A	A	B
37	*G. orientalis*	**G057**	80	G	D	A	B	B	C
38	*G. orientalis*	G058	80	H	–	A	B	B	C
39	*G. orientalis*	G060	85	I	E	A	B	B	C
40	*G. orientalis*	G061*, G062	85	K	–	A	B	B	C
41	*G. orientalis*	G063	85	L	–	A	B	B	C
42	*G. orientalis*	G064*, G083, G084	85, 97	M	E	A	B	B	C
43	*G. orientalis*	G065	85,	K	E	A	B	B	C
44	*G. orientalis*	G066	85	I	E	A	B	B	C
45	*G. orientalis*	G067	86	N	E	A	B	B	C
46	*G. orientalis*	**G068*, G071, G078, G079**	86	O	D	A	B	B	C
47	*G. orientalis*	**G069*, G070, G076, G077**	86	G	D	A	B	B	C
48	*G. orientalis*	**G072*, G073, G074, G075**	86	G	D	A	B	B	C
49	*G. orientalis*	**G080**	86	G	D	A	B	B	C
50	*G. orientalis*	G081	97	M	E	A	B	B	C
51	*G. orientalis*	G086*, G090	97	M	E	A	B	B	C
52	*G. orientalis*	G087*, G089	97	P	F	A	–	C	D
53	*G. orientalis*	G091	97	Q	E	A	C	B	C
54	*G. orientalis*	G092	97	R	G	A	B	B	C
55	*G. orientalis*	**G094**	111	S	D	A	B	B	C
56	*G. orientalis*	**G095**	111	G	D	A	B	B	C
57	*G. orientalis*	G096	126	T	E	A	B	B	C
58	*G. orientalis*	G099	126	M	E	A	B	B	C
59	*G. orientalis*	G100*, G101	142, 146	I	G	A	B	B	C
60	*G. orientalis*	G102	146	U	H	B	D	B	C
61	*G. orientalis*	G103*, G104	146	W	E	A	B	B	C
62	*G. orientalis*	G105*, G106	146	M	E	A	B	B	C
63	*G. orientalis*	G107	146	X	G	A	B	B	C
64	*G. orientalis*	G108	160	R	G	A	B	B	C
65	*G. orientalis*	G109	160	W	–	A	B	B	C
66	*G. orientalis*	G111	161	M	–	–	–	–	–
	***R. galegae* bv. officinalis**	H1141 (reference strain)		X	A	–	A	A	B
	***R. galegae* bv. orientalis**	**H540** (reference strain)		O	D	A	B	B	C

[a] AFLP groups were revealed in initial AFLP analyses and comprised strains with identical AFLP patterns (see explanations in the text).
[b] Strains printed in boldface contain IS element copies. *, representative strain from each AFLP group that was tested.
[c] No or weakly amplified fragment.
[d] RFLP types are composite types revealed in restriction analyses of PCR-amplified fragments by using three enzymes, *Hae*III, *Msp*I, and *Alu*I.

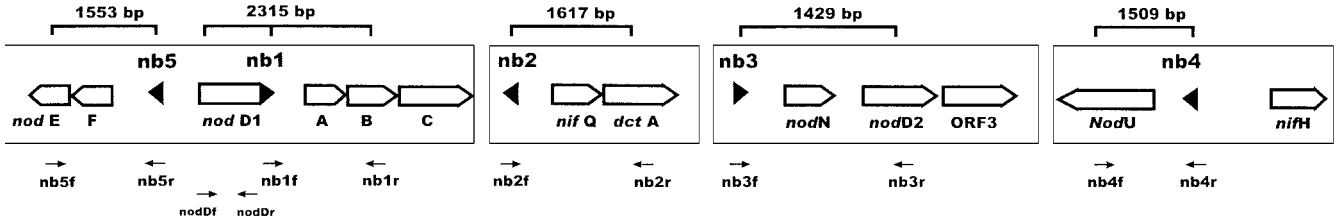

FIG. 2. Maps of symbiotic gene regions of *Rhizobium galegae*. The locations of primers used in the study and of amplified fragments are shown.

different AFLP clusters, A and B, the latter being closer to the *R. galegae* bv. officinalis cluster than to the main *R. galegae* bv. orientalis AFLP cluster (Fig. 3). Another interesting finding is that HAMBI1141 lies outside the main AFLP cluster of *R. galegae* bv. officinalis. This strain also has a unique ITS RFLP type, X (Table 1), and we believe that it is a rare genotype not typical of this region.

The results of the PCR-RFLP analysis, which allows the study of variability at particular loci (Table 1), are in agreement with those of the AFLP analyses. The corresponding values of heterogeneity in the ITS region for the two biovars were 0.87 for *R. galegae* bv. orientalis and 0.56 for *R. galegae* bv. officinalis. The average heterogeneity in the symbiotic regions analyzed was 0.42 for *R. galegae* bv. orientalis and 0.13 for *R. galegae* bv. officinalis. The only symbiotic region in which *R. galegae* bv. officinalis strains are more diverse is *nod* box 5. Overall, the strains of *R. galegae* bv. orientalis are more diverse than the *R. galegae* bv. officinalis strains in symbiotic regions as well as in the ITS region. This finding, together with the AFLP results, provides strong evidence that within the Caucasus region studied *R. galegae* bv. orientalis strains are more diverse than *R. galegae* bv. officinalis strains.

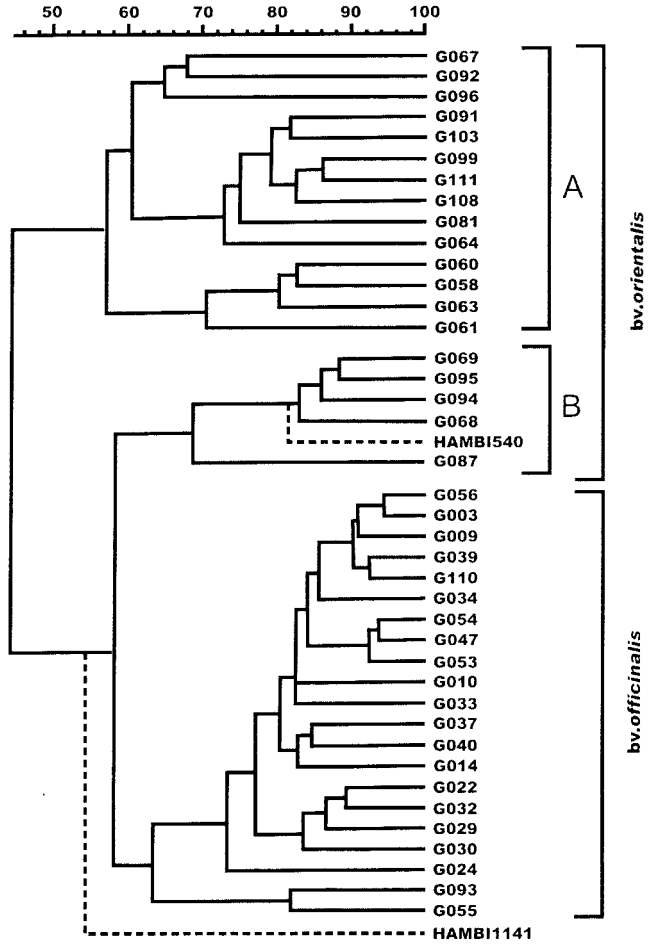

FIG. 3. UPGMA dendrogram constructed from AFLP data. The relative positions of *R. galegae* bv. orientalis HAMBI540 and *R. galegae* bv. officinalis HAMBI1141 are indicated by dashed lines.

Interestingly, the results appear to contradict a previous study (37) in which *G. officinalis* strains were shown to be more diverse than *G. orientalis* strains. However, we believe that there is no contradiction: the population of *G. officinalis* studied by Terefework et al. (37) was geographically heterogeneous, representing five different locations (New Zealand, United Kingdom, Bulgaria, Italy, and Argentina). At least two of them (New Zealand and Bulgaria) are places where native *G. officinalis* is widely represented. At the same time, the population of *G. orientalis* analyzed by Terefework et al. represented two locations, Russia and Finland (where *G. orientalis* for the most part is an introduced plant) and included few strains isolated in the Caucasus. Therefore, the observations of Terefework et al. (37) are in concordance with the results reported here, emphasizing the importance of host plant diversity for the diversity of rhizobia.

Analysis of the symbiotic region is of prime interest when subtle genetic adaptation of rhizobia to particular host plants is under investigation. Five different fragments containing *nod* box sequences were studied in this work. The presence of *nod* boxes itself points to the great importance of these regions in symbiotic interaction. We constructed five pairs of primers (see Materials and Methods) to amplify the region adjacent to each of five *nod* boxes, which were identified by Suominen et al. (32) in the symbiotic regions of the *R. galegae* bv. orientalis reference strain HAMBI540 (Fig. 2). Regions nb3-*nodD2*, nb4-*nodU*, and nb5-*nodE* were amplified successfully for almost all *R. galegae* bv. officinalis and *R. galegae* bv. orientalis strains (Table 1). However, the nb2-*dctA* region was successfully amplified only for *R. galegae* bv. orientalis and not for *R. galegae* bv. officinalis strains (Table 1). The same was true for the nb1-*nod*B region (not shown), but when the nb1f primer was replaced by the nodDf primer, located inside the *nodD1* gene (Fig. 2), the corresponding fragment comprising the *nodD1-nodB* region (which includes nb1) was amplified successfully

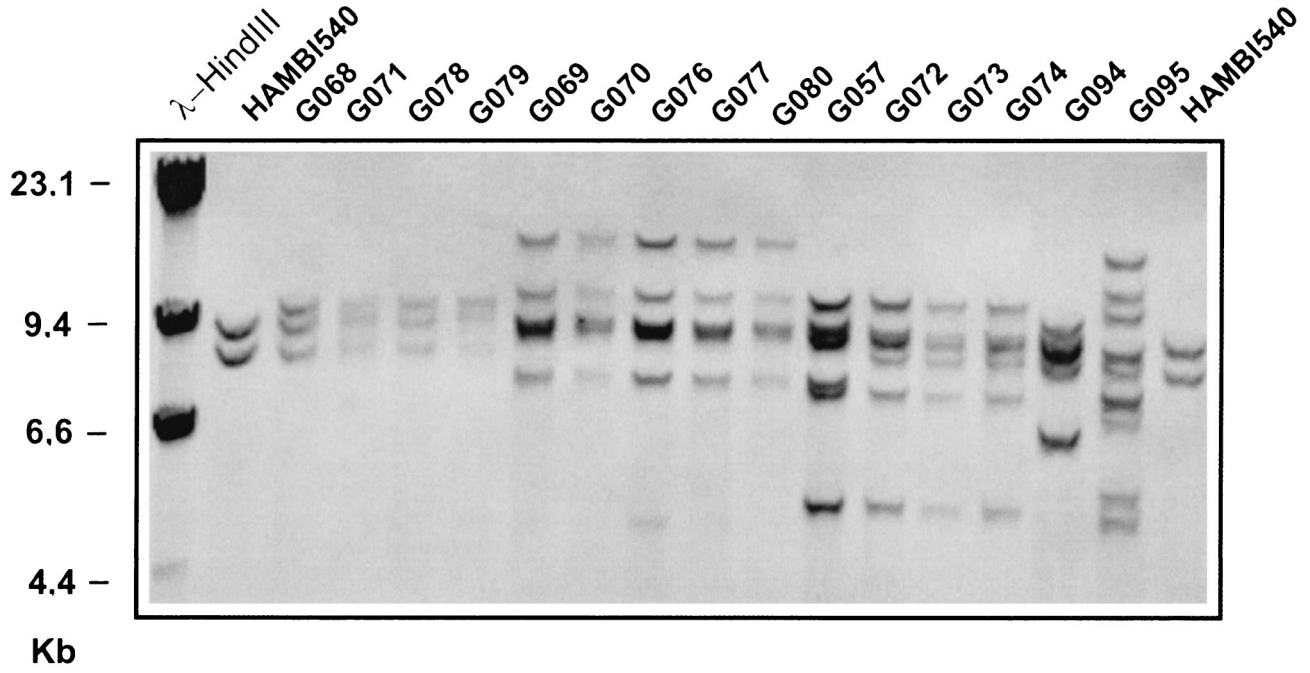

FIG. 4. IS fingerprinting by Southern hybridization of *R. galegae* strains containing the IS element.

for most of the strains of both biovars (Table 1). The last finding means that there is some difference between the nucleotide sequences of *nod* box 1 in strains belonged to different biovars.

The most conserved symbiotic genes seem to be in the *nod* box 3 and *nod* box 4 regions. The *nod* box 1 region is highly conserved among *R. galegae* bv. officinalis strains, and the *nod* box 5 region is highly conserved among *R. galegae* bv. orientalis strains. We expected to find high levels of divergence in the *nod* box regions because of the high proportion of noncoding sequence. However, they proved to be highly conserved despite the high noncoding content, e.g., 56% of noncoding sequence in the case of the most conserved *nod* box 3 region (unpublished sequence data for strain HAMBI540). We assume that there may be strict functional constraints on the noncoding regions adjacent to *nod* boxes (6).

The Nei similarity coefficient was calculated by comparing the restriction patterns of strains HAMBI540 and HAMBI1141 to give a measure of differences between different biovars in particular symbiotic loci. The coefficients were 0.77 for the *nod* box 3 region, 0.76 for *nod* box 4, 0.52 for *nod* box 5, and 0.47 for nod box 1. Thus, the greatest differences were found in the *nod* box 1 and *nod* box 5 regions, and the smallest differences were found in the *nod* box 3 and *nod* box 4 regions. The *nod* box 2 region has not been taken into account, as there was no amplified fragment for the *R. galegae* bv. officinalis strains. Thus, we can suggest that the *nod* box 1 and 5 regions with adjacent genes might play a more important role in host plant adaptation and symbiotic tuning. This suggestion is quite plausible because we know that the *nodA* and *nodEF* genes (located adjacent to nod boxes 1 and 5) are involved in synthesis of unsaturated fatty acids and their attachment to the Nod factor backbone (23). It is established that the unsaturated fatty acid strategy is the main factor for *Rhizobium-Galega* recognition (reviewed in reference 36). The presence and importance of different allelic forms of these host range genes (46) in different biovars of *R. galegae* cannot be excluded. However it seems there is no clear correlation between the genetic background and the Nod factors produced. It was shown that *R. galegae* strains belonging to different biovars produce nearly identical Nod factors, which are a mixture including a series of C_{18} and C_{20} fatty acids with carbonyl-conjugated double bonds (44).

To obtain the whole picture of genome adaptation to particular host plants, this comparative study should be expanded to include more symbiotic genes. Thus, further investigation is needed to complete the picture of symbiotic adaptation in the *Rhizobium-Galega* symbiosis. This work demonstrates an approach to trace evolutionary pressure along the symbiotic genomes of two closely related rhizobium biovars. All the analyses in this study, AFLP and RFLP of ITS and symbiotic regions, delineated the two biovars. The absence of shared restriction patterns or combined profiles from the ITS and symbiotic regions strongly suggests that there has been no recombination between the biovars. Thus, the major factor contributing to the divergence of this part of the symbiotic machinery could be the host plant.

We used a pair of primers which allowed us to amplify the inner fragment of an IS element closely related to an *Agrobacterium* transposase, which was revealed in the symbiotic region of *R. galegae* bv. orientalis strain HAMBI540 (33). Only 16 strains of *R. galegae* bv. orientalis were found to contain copies of the IS sequence (Table 1). Furthermore, we noticed that all IS-positive strains, including the reference strain HAMBI540, had closely similar rRNA ITS types: patterns O and G (Table 1) differ only in their *Hae*III restriction patterns. Strains of

these genotypes formed a clearly separated cluster (B) in the AFLP dendrogram (Fig. 3). It is particularly remarkable that 13 of these strains were isolated from the same site (Table 1). By using a DIG-labeled IS probe for Southern hybridization, it was shown that the strains contain from three to nine copies of the IS element; HAMBI540 contains two copies (Fig. 4). Each hybridizing band corresponds to at least one copy of the IS element, because the IS sequence does not contain *Hin*dIII restriction sites. Strains with identical IS hybridization patterns had identical AFLP fingerprints.

The distribution of the IS-like sequences in the population raises some interesting questions. Why is the IS element found only in genetically similar backgrounds? Why were most strains containing this IS element found in the same site? A likely answer to the first question is a limitation of horizontal transfer between different chromosomal groups, i.e., a clonal structure of the population (19, 43). Our results show no evidence for any recombination event between the biovars. Furthermore, it appears that the strains that belong to AFLP group B and contain the IS element also represent a clonal group. The second question could be explained by limitations on the transport of strains between different locations. The finding that it was rare for the same AFLP type to be found at more than one site provides support for this view. On the other hand, according to Beijerinck's principle that "everything is everywhere, the environment selects," the limitations could be assigned to either specific soil conditions or selection by plant genotype.

Our work shows that the study of rhizobial diversity in gene centers of the host is a way to answer some interesting questions in plant-microbe interaction and can shed some light on patterns and constraints of evolution in rhizobia.

ACKNOWLEDGMENTS

This work was funded by INTAS project 694, INCO-COPERNIGUS II Contr. ICA2-CT-2000-10001, and a grant from the Academy of Finland to E.E.A.

REFERENCES

1. **Andronov, E. E., M. L. Roumyantseva, V. V. Sagoulenko, and B. V. Simarov.** 1999. Effect of the host plant on the genetic diversity of a natural population of *Sinorhizobium meliloti*. Russ. J. Genet. **35:**1169–1176.
2. **Andronov, E. E., M. L. Roumyantseva, and B. V. Simarov.** 2001. Genetic diversity of a natural population of *Sinorhizobium meliloti* revealed in analysis of cryptic plasmids and IS*Rm*2011–2 fingerprints. Russ. J. Genet. **37:** 494–499.
3. **Beringer, J. E.** 1974. R1 transfer in *Rhizobium leguminosarum*. J. Gen. Microbiol. **84:**188–198.
4. **Bromfield, E. S. P., L. R. Barran, and R. Wheatcroft.** 1995. Relative genetic structure of a population of *Rhizobium meliloti* isolated directly from soil and from nodules of alfalfa (*Medicago sativa*) and sweet clover (*Melilotus alba*). Mol. Ecol. **4:**183–188.
5. **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.
6. **Clark, A.** 2001. The search for meaning in noncoding DNA. Genome Res. **11:**1319–1320.
7. **Dresler-Nurmi, A., Z. Terefework, S. Kaijalainen, K. Lindström, and A. Hatakka.** 2000. Silver stained polyacrylamide gels and fluorescence-based automated capillary electrophoresis for detection of amplified fragment length polymorphism patterns obtained from white-rot fungi in the genus *Trametes*. J. Microbiol. Methods **41:**161–172.
8. **Freiberg, C., R. Fellay, A. Bairoch, W. Broughton, A. Rosenthal, and X. Perret.** 1997. Molecular basis of symbiosis between *Rhizobium* and legumes. Nature **387:**394–401.
9. **Hartman, A., J. J. Giraud, and G. Catroux.** 1998. Genotypic diversity of *Sinorhizobium* (formerly *Rhizobium*) *meliloti* strains isolated directly from a soil and from nodules of alfalfa (*Medicago sativa*) grown in the same soil. FEMS Microb. Ecol. **25:**107–116.
10. **Hirsh, P. R.** 1996. Population dynamics of indigenous and genetically modified rhizobia in the field. New Phytol. **133:**159–171.
11. **Laberge, S., A. T. Middleton, and R. Wheatcroft.** 1995. Characterization, nucleotide sequence, and conserved genomic locations of insertion sequence IS*R*m5 in *Rhizobium meliloti.* J. Bacteriol. **177:**3133–3142.
12. **Laguerre, G., S. I. Masurier, and N. Amarger.** 1992. Plasmid profiles and restriction fragment length polymorphism of *Rhizobium leguminosarum* bv. viceae in field populations. FEMS Microb. Ecol. **101:**17–26.
13. **Lindström, K., B. D. V. Jarvis, P. E. Lindström, and J. J. Patel.** 1983. DNA homology, phage-typing, and cross-nodulation studies of rhizobia infecting *Galega* species. Can. J. Microbiol. **29:**781–789.
14. **Lindström, K.** 1989. *Rhizobium galegae*, a new species of legume root nodule bacteria. Int. J. Syst. Bacteriol. **39:**365–367.
15. **Lindström, K., and S. Lehtomäki.** 1988. Metabolic properties, maximum growth temperatures and phage-typing as a means of distinguishing *Rhizobium* sp. (*Galega*) from other fast growing rhizobia. FEMS Microbiol. Lett. **50:**277–287.
16. **Lipsanen, P., and K. Lindström.** 1988. Lipopolysaccharide and protein patterns of *R. galegae*, p. 478. *In* H. Bothe, F. J. De Bruijn, and W. E. Newton (ed.), Nitrogen fixation: hundred years after. VCH Publishers, Stuttgart, Germany.
17. **Martinez, E., D. Romero, and R. Palacios.** 1990. The *Rhizobium* genome. Crit. Rev. Plant Sci. **9:**59–93.
18. **Martinez-Romero, E.** 1996. Comments on *Rhizobium* systematics. Lessons from *R. tropici* and *R. etli*, p. 503–508. *In* G. Stacey, B. Mullin, and P. M. Gresshoff (ed.), Biology of plant-microbe interactions. International Society for Molecular Plant-Microbe Interactions, St. Paul, Minn.
19. **Maynard Smith, J., N. H. Smith, M. O'Rourke, and B. G. Spratt.** 1993. How clonal are bacteria? Proc. Natl. Acad. Sci. USA **90:**4384–4388.
20. **Nei, M., and W. H. Li.** 1979. Mathematical model for studying genetic variations in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA **76:**5269–5273.
21. **Nick, G., and K. Lindström.** 1994. Use of repetitive sequences and the polymerase chain reaction to fingerprint the genomic DNA of *R. galegae* strains and to identify the DNA obtained by sonicating liquid cultures and root nodules. Syst. Appl. Microbiol. **17:**265–273.
22. **Normand, P., B. Cournoyer, S. Nazaret, and P. Simonet.** 1992. Analysis of a ribosomal operon in the actinomycete *Frankia.* Gene **111:**119–124.
23. **Peters, N.** 1997. Nodulation: finding the lost common denominator. Curr. Biol. **7:**223–226.
24. **Radeva, G., G. Jurgens, M. Niemi, G. Nick, L. Suominen, and K. Lindström.** 2001. Description of two biovars in the *Rhizobium galegae* species: biovar orientalis and biovar officinalis. Syst. Appl. Microbiol. **24:**192–205.
25. **Sawada, H., H. Oyaizu, and S. Matsumoto.** 1993. Proposal for rejection of *Agrobacterium tumefaciens* and revised description for the genus *Agrobacterium* and for *Agrobacterium radiobacter* and *Agrobacterium rhizogenes.* Int. J. Syst. Bacteriol. **43:**694–702.
26. **Selbitschka, W., W. Arnold, D. Jording, et al.** 1995. The insertion sequence element IS*Rm*2011–2 belongs to the IS630-Tc1 family of transposable elements and is abundant in *Rhizobium meliloti.* Gene **163:**59–64.
27. **Selbitschka, W., S. Zekri, G. Schröder, A. Pühler, and N. Toro.** 1999. The *Sinorhizobium meliloti* insertion sequence (IS) elements IS*Rm*102F34–1/IS*Rm*7 and IS*Rm*220–13–5 belong to a new family of insertion sequence elements. FEMS Microbiol Lett. **172:**1–7.
28. **Selenska-Pobell, S., L. Gigova, and N. Petrova.** 1995. Strain specific fingerprint of *Rhizobium galegae* generated by PCR with arbitrary repetitive primers. J. Appl. Bacteriol. **79:**425–431.
29. **Selenska-Trajkova, S., G. Radeca, and K. Markov.** 1990. Comparison between *Rhizobium galegae* and *R. meliloti* plasmid contents. Lett. Appl. Microbiol. **10:**123–126.
30. **Sharypova, L. A., O. P. Onishchuk, O. N. Chesnokova, J. G. Fomina-Eschchenko, and B. V. Simarov.** 1994. Isolation and characterization of *Rhizobium meliloti* Tn*5*-mutants showing enhanced symbiotic effectiveness. Microbiology **140:**463–470.
31. **Simon, R., B. Hotte, 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 vector*s*. J. Bacteriol. **173:** 1502–1508.
32. **Suominen, L., A.-M. Saren, L. Paulin, C. Roos, A. Saano, and K. Lindström.** 1999. Identification of the putative host specificity gene region in *R.galegae* using a synthetic nod box probe and phylogeny of *nodD* genes. FEMS Microbiol. Lett. **177:**217–223.
33. **Suominen, L.** 2000. Molecular biology of symbiotic interactions between *Galega orientalis* and *Rhizobium galegae*. Ph.D. thesis. University of Helsinki, Helsinki, Finland.
34. **Suominen, L., C. Roos, G. Lortet, L. Paulin, and K. Lindström.** 2001. Identification and structure of the *Rhizobium galegae* common nodulation genes: evidence for horizontal gene transfer. Mol. Biol. Evol. **18:**907–916.
35. **Terefework, Z., G. Nick, S. Suomalainen, L. Paulin, and K. Lindström.** 1998. Phylogeny of *Rhizobium galegae* with respect to other rhizobia and agrobacteria. Int. J. Syst. Bacteriol. **48:**349–356.
36. **Terefework, Z., G. Lortet, L. Suominen, and K. Lindström.** 2000. Molecular

evolution of interactions between rhizobia and their legume hosts, p. 187–206. *In* E. Triplett (ed.), Prokaryotic nitrogen fixation: a model for analysis of a biological process. Horizon Press, Wymondham, United Kingdom.
37. **Terefework, Z., S. Kaijalainen, and K. Lindström.** 2001. AFLP fingerprinting as a tool to study the genetic diversity of *Rhizobium galegae* isolated from *Galega orientalis* and *Galega officinalis*. J. Biotechnol. **91:**169–180.
38. **Vavilov, N. I.** 1926. Centers of origin of cultivated plants. Trends Pract. Bot. Genet. Sel. **16:**3–248. (In Russian.)
39. **Vos, P., R. Hogers, M. Bleeker, M. Rejans, T. Van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kupier, and M. Zabeau.** 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. **23:**4407–4414.
40. **Wedlock, D. N., and B. D. W. Jarvis.** 1986. DNA homologies between *Rhizobium fredii*, rhizobia that nodulate *Galega* sp., and other *Rhizobium* and *Bradyrhizobium* species. Int. J. Syst. Bacteriol. **36:**550–558.
41. **Weisburg, W. G., S. M. Barns, D. A. Pelletier, and D. J. Lane.** 1991. 16S ribosomal DNA amplification for phylogenetic study. J. Bacteriol. **173:**697–703.
42. **Wernegreen, J. J., E. E. Harding, and M. A. Riley.** 1997. *Rhizobium* gone native: unexpected plasmid stability of indigenous *Rhizobium leguminosarum*. Proc. Natl. Acad. Sci. USA **94:**5483–5488.
43. **Wright, S., and D. Finnegan.** 2001. Genome evolution: Sex and the transposable element. Curr. Biol. **11:**296–299.
44. **Yang, G.-P., F. Debellé, A. Savgnac, M. Ferro, O. Schiltz, F. Maillet, D. Promé, M. Treilhou, C. Vialas, K. Lindström, J. Dénarié, and J.-C. Promé.** 1999. Structure of the *Mesorhizobium huakuii* and *Rhizobium galegae* Nod factors: a cluster of phylogenetically related legumes are nodulated by rhizobia producing Nod factors with α,β-unsaturated *N*-acyl substitutions. Mol. Microbiol. **34:**227–237.
45. **Young, J. P. W., L. Demetriou, and R. G. Apte.** 1987. *Rhizobium* population genetics: enzyme polymorphism in *Rhizobium leguminosarum* from plants and soil in a pea crop. Appl. Environ. Microbiol. **53:**397–402.
46. **Young, J. P. W., and A. W. B. Johnston.** 1989. The evolution of specificity in the legume-rhizobium interaction. Trends Ecol. Evol. **4:**341–349.
47. **Zhukovskii, P. M.** 1968. New origin and genetic centers of crops and highly endemic microcenters of related species. Bot. Z. **53:**430–460. (In Russian.)