ISSN 2079-0597, Russian Journal of Genetics: Applied Research, 2013, Vol. 3, No. 2, pp. 102–113. © Pleiades Publishing, Ltd., 2013.
Original Russian Text © V.S. Zotov, N.V. Punina, S.A. Khapchaeva, S.V. Didovich, T.N. Melnichuk, A.F. Topunov, 2012, published in Ekologicheskaya Genetika, 2012, Vol. 10, No. 2, pp. 50–63.

A New Taxonomic Marker of Nodule Bacteria of the *Rhizobium* Genus and its Evolution

V. S. Zotov[a], N. V. Punina[a,b], S. A. Khapchaeva[a], S. V. Didovych[c], T. N. Melnichuk[c], and A. F. Topunov[a]

[a] *FSBI of Science, Bakh Institute of Biochemistry, Russian Academy of Sciences, Moscow, Russia*

[b] *FSBI Medical Genetic Research Center, Russian Academy of Medical Sciences, Moscow, Russia*

[c] *Crimean Institute of Agriculture, Ukrainian National Acadamy of Sciences, Simferopol, Crimea, Ukraine*

Received November 29, 2011; in final form, June 5, 2012

Abstract—A new taxonomic marker (the *hin*-region) for study of the diversity of *Rhizobium* bacteria at the species–groups of strains level was proposed. Groups of *Rhizobium* strains were isolated with its use that could not be detected by other methods that correlated with the evolutionary proximity of bacteria. This approach to the creation of marker systems allows one to effectively describe the intra- and inter-species genetic diversity of nodule bacteria and to provide an assessment of the prospects of their use in agricultural practices. The proposed marker system was used to describe *Rhizobium* genus bacterial samples that were isolated in different ecological and geographical regions of Ukraine.

Keywords: *Rhizobium*, phylogeny, taxonomy, detection, diversity, saAFLP, *hin*-region PCR, PCR–RFLP, *16S* ribosomal RNA, *16S–23S* ribosomal RNA

DOI: 10.1134/S2079059713020111

INTRODUCTION

It is known that nodule bacteria are divided into two large groups, viz., fast- and slow-growing, which differ in growth rate, as well as carbohydrate and nitrogen compound metabolism (*Rhizobiaceae*. Molecular biology of bacteria that interact with plants, 2002) and the specificity of interactions with various bean cultures (broadly and narrowly specific) (Martínez–Romero, Caballero–Mellado, 1996).

At present, as a rule, the taxonomy and evolutionary relationships of the order *Rhizobiales* are studied by comparing the nucleotide sequences of *16S* rRNA. However, such an assessment does not always correspond to the data that are obtained for other loci (Gaunt et al., 2001). Due to frequent interspecies recombination between different alleles of taxonomically important genes (such as *16S* rRNA, *dnaK*, and *nod*), the rhizobia isolates are not always classified adequately, that fact leads to the need for additional genetic data (Suominen et al., 2001; Eardly et al., 2005; Vinuesa et al., 2008). The genus *Rhizobium* (Frank, 1889), has a complex taxonomic structure that is difficult to study by current methods, and is of particular interest as a model object of a given order.

The purpose of this research was to develop a specific genetic marker system for nodule bacteria of the genus *Rhizobium* in order to study them at the species–groups of strains level.

In order to create a fast, inexpensive, and reliable method for detecting and studying the biodiversity of the population structure of nodule bacteria, we proposed methods, viz., saAFLP (Zotov et al., 2010) and a comparative analysis of nucleotide sequences of the *hin*-region (Zotov et al., 2011).

MATERIALS AND METHODS

Bacterial strains. We used 52 isolates of nodule bacteria of the genus *Rhizobium* from the collections of beneficial microorganisms of various scientific institutions of Ukraine and Russia: the Institute of Agriculture of Crimea of the National Academy of Agrarian Sciences of Ukraine (IAC, NAAS of Ukraine), Simferopol; the Institute of Agricultural Microbiology and Agricultural Production of the National Academy of Agricultural Sciences, Chernigov; Zabolotny Institute of Microbiology and Virology, National Academy of Sciences of Ukraine, Kiev; and the All Russia Research Institute for Agricultural Microbiology of the Russian Academy of Agricultural Sciences, ARRIAM RAAS, St. Petersburg. Nine strains from the collection from the ARRIAM RAAS were used as a reference. The isolates from the collection of the Department of Microbiology of IAC NAAS were isolated in different eco-geographical zones of Ukraine: the steppe, forest steppe, and Polissya (Table 1).

Identification. The cultural, physiological, and biochemical properties of *Rhizobium* sp. isolates were studied according to generally accepted methods of microbiology and biochemistry (Methods of Soil Microbiology and Biochemistry, 1991; Bergey's Manual..., 1993)

Table 1. Test strains of rhizobia and their grouping using a variety of molecular techniques

Strain	Genus/species	Similarity by *16S* rRNA, %	Host plant (source)	Geographical origin*	ITS PCR, RFLP genotype	saAFLP group	*hin*-region (class/genotype)	*hin*-region length, bp
U-1	*R. leguminosarum*	100	pea	Ukraine (VR)	6	IA.1	I/A	270 + 410
28		100	pea	Ukraine (LR)	1	IA.1	I/A	270 + 410
P-2		100	pea	Ukraine (ARC)	1	IA.1	I/A	270 + 410
Ch-14		100	pea	Ukraine (LR)	1	IA.1	I/A	270 + 410
K-29		100	pea	Ukraine (LR)	2	IA.1	I/A	270 + 410
34		100	pea	Ukraine (ARC)	1	IA.5	I/A	270 + 410
65		100	pea	Ukraine (OR)	1	IA.6	I/A	270 + 410
U-2		100	pea	Ukraine (VR)	1	IA.7	I/A	270 + 410
U-4		100	pea	Ukraine (VR)	1	IA.3	I/A	270 + 410
U-5		100	pea	Ukraine (VR)	1	IA.3	I/A	270 + 410
32		100	pea	Ukraine (LR)	1	IA.3	I/A	270 + 410
32-2		100	pea	Ukraine (LR)	1	IA.3	I/A	270 + 410
682		100	beans	Ukraine	1	IA.8.1	I/A	270 + 410
31		99.93	pea	Ukraine (LR)	2	IA.4	I/A	270 + 410
261bst		100	pea	Ukraine	3	IA.2	I/A	270 + 410
G-8		100	pea	Ukraine (KhR)	1	IA.2	I/A	270 + 410
24		100	pea	Ukraine (LR)	1	IA.2	I/A	270 + 410
U-3		100	pea	Ukraine (VR)	1	IA.2	I/A	270 + 410
U-11		100	pea	Ukraine (VR)	1	IA.2	I/A	270 + 410
U-12		100	pea	Ukraine (VR)	1	IA.2	I/A	270 + 410
FA-4		100	beans	Ukraine (KhR)	1	IA.8.2	I/A	270 + 410
F-1		100	beans	Ukraine (KhR)	1	IA.9	I/A	270 + 410
FA-34		100	beans	Ukraine (KhR)	1	IA.8.3	I/A	270 + 410
1326st		99.93	beans	Russia (LR)	13	LA10	I/A	270 + 410
245ast		100	pea	Russia (MR)	4	IB.1	I/B	320 + 470
248b		100	pea	Ukraine	3	IB.2	I/B	320 + 470
U-9		100	pea	Ukraine (VR)	3	IB.2	I/B	320 + 470
U-7		100	pea	Ukraine (VR)	5	IB.3	I/B	320 + 470
96st		100	broad beans	Lithuania	7	IB.1	I/B	320 + 470
97st		100	broad beans	Latvia	8	IB.2	I/B	320 + 470
B-6		100	broad beans	Ukraine (VR)	9	IB.2	I/B	320 + 470
B-8	*R. leguminosarum*	100	broad beans	Ukraine (VR)	9	IB.2	I/B	320 + 470
B-17		100	broad beans	Ukraine (VR)	9	IB.2	I/B	320 + 470
B-18		100	broad beans	Ukraine (VR)	9	IB.2	I/B	320 + 470
B-9		100	broad beans	Ukraine (VR)	10	IB.3	I/B	320 + 470
B-15		100	broad beans	Ukraine (VR)	10	IB.3	I/B	320 + 470
B-16		100	broad beans	Ukraine (VR)	10	IB.3	I/B	320 + 470
B-25		100	broad beans	Ukraine (VR)	10	IB.3	I/B	320 + 470
KP15		100	clover	Ukraine (ARC)	11	IB.3.2	I/B	320 + 470
MK		100	clover	Ukraine (ARC)	12	IB.2	I/B	320 + 470
KK		100	clover	Ukraine (ARC)	12	IB.2	I/B	320 + 470
KP011		100	clover	Ukraine (ARC)	13	IB.3.1	I/B	320 + 470
KP012		100	clover	Ukraine (ARC)	13	IB.3.1	I/B	320 + 470
KL-90		100	clover	Ukraine (ARC)	14	IB..3.3	I/B	320 + 470
KL-91		100	clover	Ukraine (ARC)	14	IB.3.1	I/B	320 + 470

Table 1. (Contd.)

Strain	Genus/species	Similarity by 16S rRNA, %	Host plant (source)	Geographical origin*	ITS PCR, RFLP genotype	saAFLP group	*hin*-region (class/genotype)	*hin*-region length, bp
FK-0		100	beans	Ukraine (KhR)	15	Ib.1.1	I/b	320
FK-4		100	beans	Ukraine (KhR)	15	Ib.1.2	I/b	320
FK-6		100	beans	Ukraine (KhR)	15	Ib.1.3	I/b	320
108*		100	beans	Ukraine (KhR)	15	Ib.1.4	I/b	320
105		100	beans	Armenia	16	II.1.1	II/C	470 + 655
108**		100	beans	Ukraine (KhR)	16	II.1.2	II/C	470 + 655
FN		100	beans	Ukraine (KhR)	16	II.1.3	II/C	470 + 655
FN-6		99.09	beans	Ukraine (KhR)	17	III.1	III/D	250
FS	*R. giardinii*	99.72	beans	Ukraine (KhR)	18	IV.1	IV/E	260
700st	*R. etli*	99.93	beans	Mexico	19	V.1	V/F	370 + 520
912st	*R. galegae*	99.86	milk vetch	Russia (LR)	20	VI.1	VI/G	230
913st		99.86	milk vetch	Estonia	20	VI.1	VI/G	230
916st		99.86	milk vetch	Russia (LR)	20	VI.1	VI/G	230
926st		99.86	milk vetch	Ukraine	20	VI.1	VI/G	230
K-3		99.86	milk vetch	Ukraine	20	VI.1	VI/G	230
K-18		99.86	milk vetch	Ukraine	20	VI.1	VI/G	230

Notes: * Vinnitsa region (VR); Odessa region (OR); AR of Crimea (ARC); Moscow region (MO); Leningrad region (LR); Lugansk region (LR); Kharkov region (KhR).

** Genotypes by RFLP data set using restriction endonuclease *Alu*I, *Msp*I, and *Hae*III).

and methodical recommendations of ARRIAM RAAS (Voznyakovskaya, Popova, 1985).

DNA isolation. To isolate total cellular DNA preparations, the strains and isolates were cultured on TY agar medium: yeast extract: 1 g/L; peptone, 10 g/L; $CaCl_2$, 0,4 g/L; and agar, 20 g/L (Beringer, 1974). DNA was isolated from fresh cultures on the first or second day of growth using sorption on magnetic particles (Miniprep and Silex, Russian Federation).

Single adapter AFLP analysis (saAFLP) (Zotov et al., 2010). Restriction enzyme analysis was performed simultaneously with ligation in 10 μL of a mixture containing 80 ng of template DNA, 1 × ligase buffer (Fermentas, United States), 10 pmol of single-stranded adapter (Ad.CTAG1: 5'-ctagCTGGAATCGATTCCAG-3'), 5 units of T4 DNA ligase (Fermentas, United States), and 1 unit of the *Xma*JI endonuclease (*Xba*I). The mixture was incubated at 37°C for 2 hours, after that the reaction volume was adjusted to 100 μL. PCR was performed on an Eppendorf Mastercycler gradient thermocycler in 25 μL of a mixture containing 1 × PCR buffer, 2.8 mM $MgCl_2$, 0.2 mM dNTP, 2 μL of the restriction–ligase mixture as a DNA template, 0.4 μM of primer (Pr.CTAG1: 5'-CTGGAATCGATTCCAGctag-3') complementary to the adapter, and 1 unit of DNA polymerase BioTaq (Dialat Ltd, Russian Federation). For PCR amplification we used the following temperature–time profile: initial denaturation at 94°C, 2 min; the following 30 cycles: 94°C, 30 s; 40°C, 30 s; 72°C, 3 min; and a final elongation, 5 min at 72°C.

PCR amplification of *16S* rRNA and of the *hin*-region. PCR and sequencing (Sanger et al., 1977) of *16S* rRNA gene nucleotide sequences were performed using the FD1: 5'-AGAGTTTGATCCTGGCTCAG-3' and RD1: 5'-AAGGAGGTGATCCAGCC-3' (Weisburg et al., 1991) primers. *Hin*-region amplification was carried out using primers we developed that were specific to the *Rhizobium* genus, and the following protocol: initial denaturation at 94°C, 2 min; the following 30 cycles: 94°C, 30 s; 55°C, 30 s; 72°C, 1 min, and a final elongation for 5 min at 72°C. The amplified fragments were detected by electrophoresis on a 1.5% agarose gel. Nucleotide sequences were determined on a Genetic Analyzer 3130xl automatic sequencer (Applied Biosystems, United States).

Restriction fragment length polymorphism analysis (RFLP). Amplification of the intergenic region of the ribosomal cluster (ITS) for subsequent restriction analysis was carried out using the FGPS1490, 72: 5'-TGCGGCTGGATCCCCTCCTT-3' (Normand et al., 1992) and ITS_R: 5'-CGCAGCGTATCACGTCCTTC-3' primers, which were adapted for bacteria of the genus *Rhizobium*. Restriction enzyme analysis was performed using the *Alu*I, *Hae*III, and *Msp*I restriction endonucleases (Fermentas, United States)

according to the manufacturer's instructions. Restriction-enzyme treated DNA was analyzed by electrophoresis in a 4.0% agarose gel.

Nucleotide sequence analysis. A primary comparative analysis of the obtained sequences with sequences of GenBank database was carried out using an NCBI Blast search (Altschul et al., 1990). Sequence alignment was performed using CLUSTALW 1.75v. (Thompson et al., 1994). Verification and editing were carried out using the BioEdit 7.0.5.3 editor (Hall, 1999). Phylogenetic trees were built in the Mega 3.1 program (Kumar et al., 2004) using the Neighbor-Joining NJ algorithm (Nei and Kumar, 2000). The pairwise genetic distances between the sequences were determined by the two-parameter model of Kimura (Kimura, 1980). The level of nucleotide polymorphism was calculated in the DnaSP 5.10 program (Rozas et al., 2010).

RESULTS AND DISCUSSION

The Physiological and Biochemical Properties of Bacteria of the Rhizobium Genus

The *Rhizobium* genus is traditionally classified on the basis of phenotypic characteristics, such as the ability to form nodules and physiological and biochemical properties. According to results of analysis, all the studied isolates belonged to the genus *Rhizobium*, which has the specificity to infection; the group of plants includes peas, beans, vetch, and lentils (bv. *viciae*), as well as haricot plants (bv. *phaseoli*), clover (bv. *trifolii*), and milk vetch (*R. galegae*).

The isolates of nodule bacteria from the collection of the Department of Microbiology of IAC NAAS showed high efficiency in symbiosis with modern varieties of legumes in the soil and the climatic conditions of Ukraine. This allowed us to improve the productivity of plants by 10–30%, to increase the protein content in the grain by 2–6% and in the green mass by 1–3%, even in the presence in the soil of native or previously introduced rhizobia *Rhizobium* sp. populations (Didovich, 2008).

The Genetic Diversity of Bacteria of the Rhizobium Genus

Nucleotide sequence analysis of the *16S* rRNA gene. Complete DNA sequences of *16S* rRNA gene (1445 bp along the genome *R. leguminosarum* bv. *viciae 3841*) were obtained for all tested isolates and strains of *Rhizobium* sp. For each type of nodule bacteria of the genera *Rhizobium*, *Sinorhizobium*, and *Mesorhizobium* we found sequences of the *16S* rRNA gene of the type strains in the NCBI database (Kwon et al., 2005), which were also taken for analysis. The DNA sequence of *16S* rRNA of the typical strain, *Bradyrhizobium japonicum* DSM30131^T, was used as a remote control and a phylogenetic tree was constructed (Fig. 1).

According to the results of analysis all tested isolates were reliably attributed to the four species of the genus *Rhizobium*: *R. leguminosarum* (Frank, 1889), *R. etli* (Segovia et al., 1993), *R. giardinii* (Amarger et al., 1997), and *R. galegae* (Lindström, 1989) (Table 1). The nucleotide sequences of the *R. leguminosarum* species had one intraspecies substitution, G/A/C/T, (1069 bp from the beginning of the gene for the NC_008380 genome). In the *R. giardinii, R. galegae*, and *R. etli* strains intraspecific polymorphism in the *16S* rRNA gene DNA sequences was not detected.

However, due to low resolution, frequent recombination events, and horizontal transfer the *16S* rRNA gene cannot be used to accurately determine the phylogenetic structure of the *Rhizobium* genus at the species level (Willems, Collins, 1993; Gaunt et al., 2001; Eardly et al., 2005). On the results of phylogenetic tree generation one can see the polifilet nature of rhizobia, which has been shown previously (Martinez, 1994; Novikova, 1996).

RFLP analysis of the *16S–23S* rRNA intergenic region (ITS). We carried out an ITS analysis of the bacteria of the *Rhizobium* genus, which showed that 22 of the 61 strains had two PCR amplification products, which differ not only in length, but also in nucleotide composition (Laguerre et al., 1996; Palmer, Young, 2000; Vessey, Chemining'wa, 2006). Using the restriction of the obtained ITS PCR products 20 ITS RFLP genotypes were reliably identified at the intraspecific level (Table 1). As well, sequencing of the ITS was carried out for 20 strains of *Rhizobium* sp. (one strain for each ITS RFLP genotype), and in combination with the nucleotide sequences of the reference strains from the GenBank (Kwon et al., 2005) a phylogenetic tree was constructed (Fig. 2). The high level of variability in the DNA nucleotide sequence of the ITS region allows us to distinguish between closely related strains, but the multicopy nature of the ITS regions complicates the interpretation of phylogenetic relationships (Gürtler and Stanisich, 1996). Thus, according to the comparative analysis of complete nucleotide sequences of the ITS region the 105 and FA-4 strains (bean symbionts) with single ITS PCR products belong to the species *R. etli*. Strains KK, KP012, and KL91 (clover symbionts) belong to the same species. At the same time, according to the DNA sequences of the *16S* rRNA gene all these strains belong to the species *R. leguminosarum*. It is interesting that the genomes of the symbionts of clover, KK, KP012, and KL91, carry two types of ITS-region incorporates, while one belongs to the species *R. leguminosarum* (its-L, ITS PCR product with a high molecular weight) and the other belongs to the *R. etli* (its-S, and ITS PCR product with a lower molecular weight). This probably indicates horizontal transfer or frequent recombination events between the DNA sequences of different species of bacteria of the genus (Terefework et al., 1998).

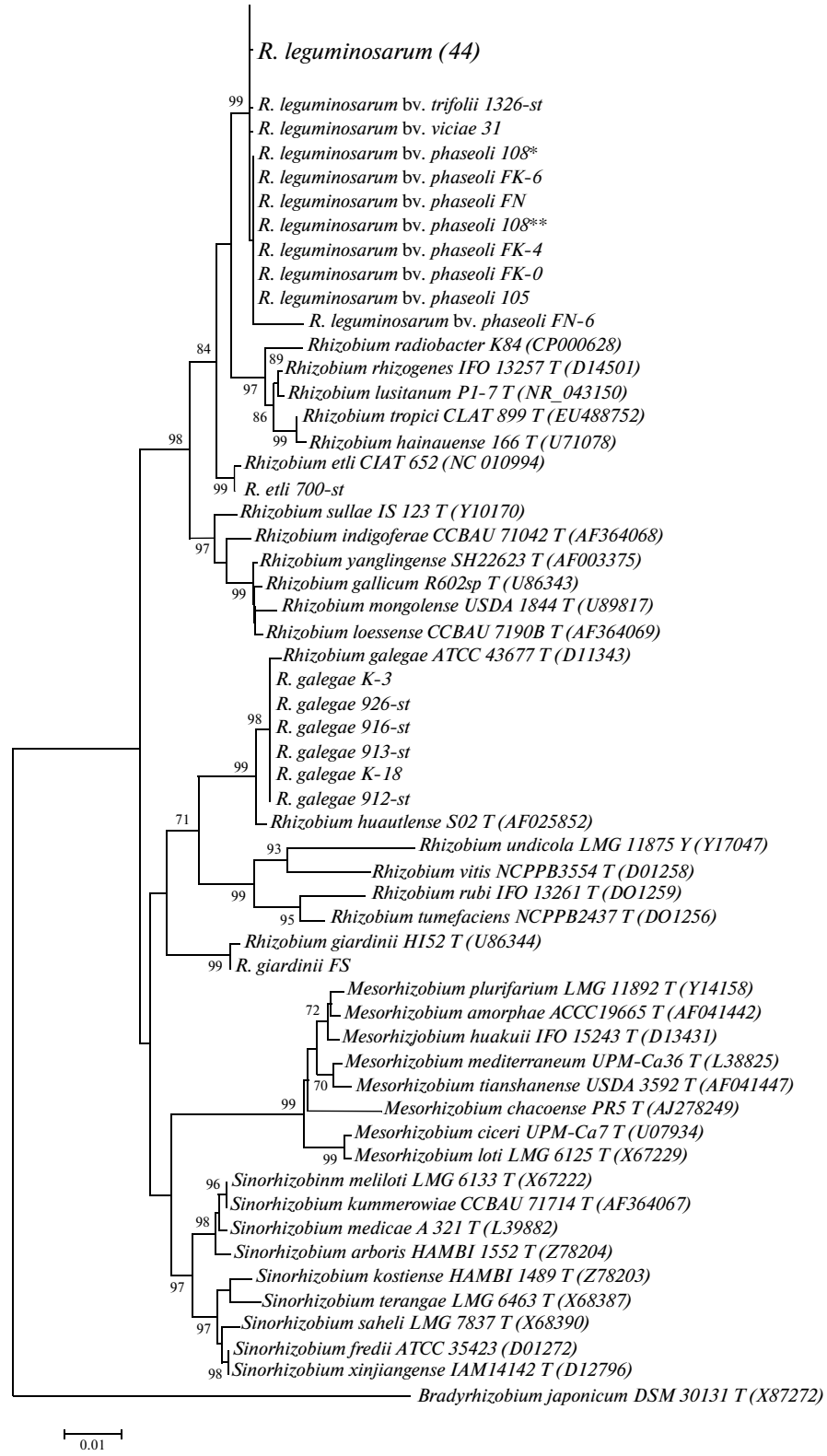

Fig. 1. The phylogenetic tree that was constructed on the basis of the data of comparative analysis of *16S* rRNA gene complete nucleotide sequences of bacteria of the genus *Rhizobium* with other members of *Rhizobiaceae* using the NJ algorithm. The scale corresponds to one substitution per 100 base pairs (genetic distances). The numerals show the statistical reliability of the order of branching (%) defined by bootstrap analysis of 1000 replicas. Bootstrap values less than 70% are not shown. The test strains belong to the genus that have a reduction *R*.

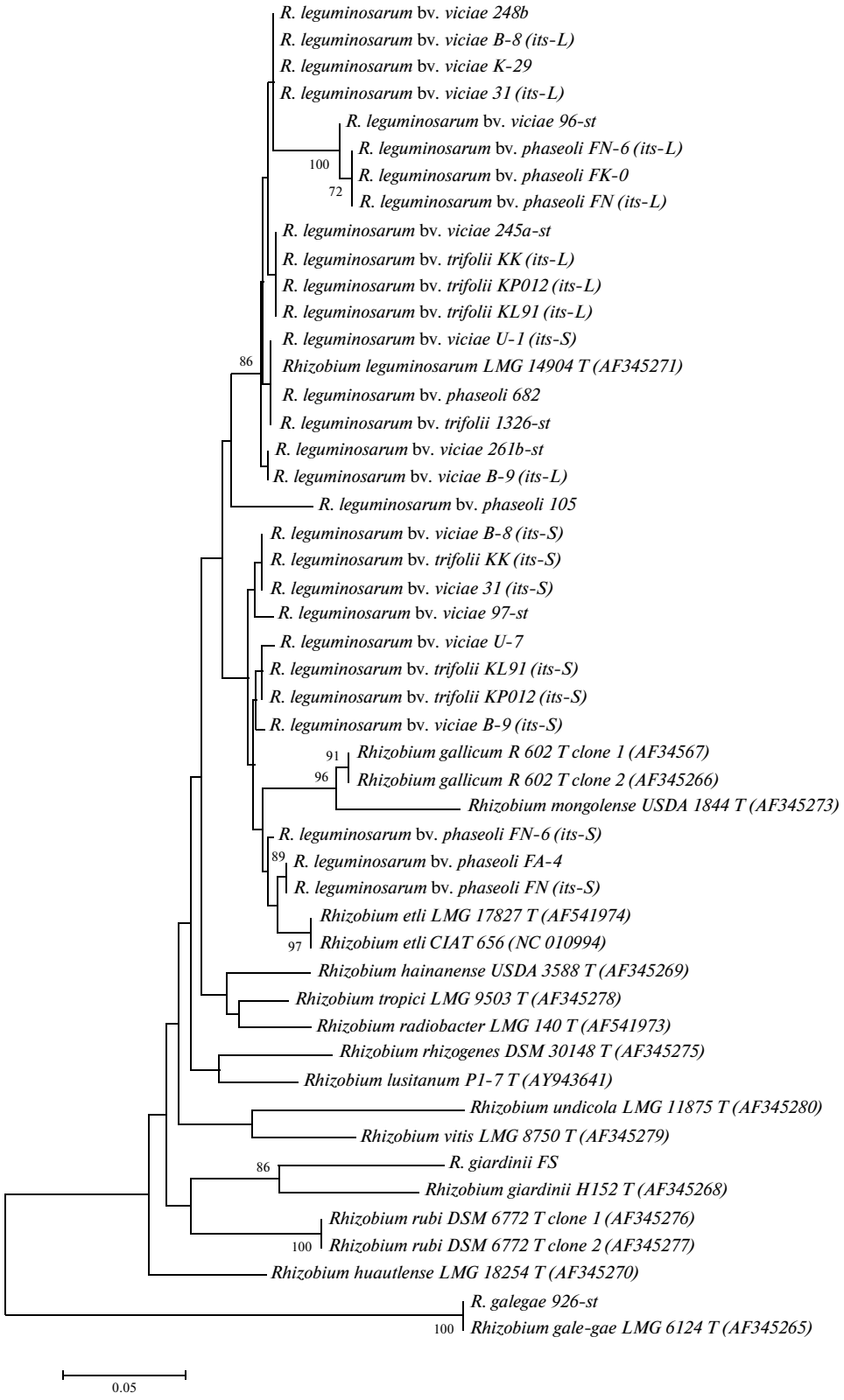

Fig. 2. The phylogenetic tree that was constructed on the basis of data from comparative analysis of the complete ITS nucleotide sequences of bacteria of the genus *Rhizobium* using the NJ algorithm. The scale corresponds to 5 substitutions per 100 base pairs. The numerals show the statistical reliability of the order of branching (%) defined by the bootstrap analysis of 1000 replicas. Bootstrap values less than 70% are not shown. The test strains belong to the genus that have a reduction *R*.

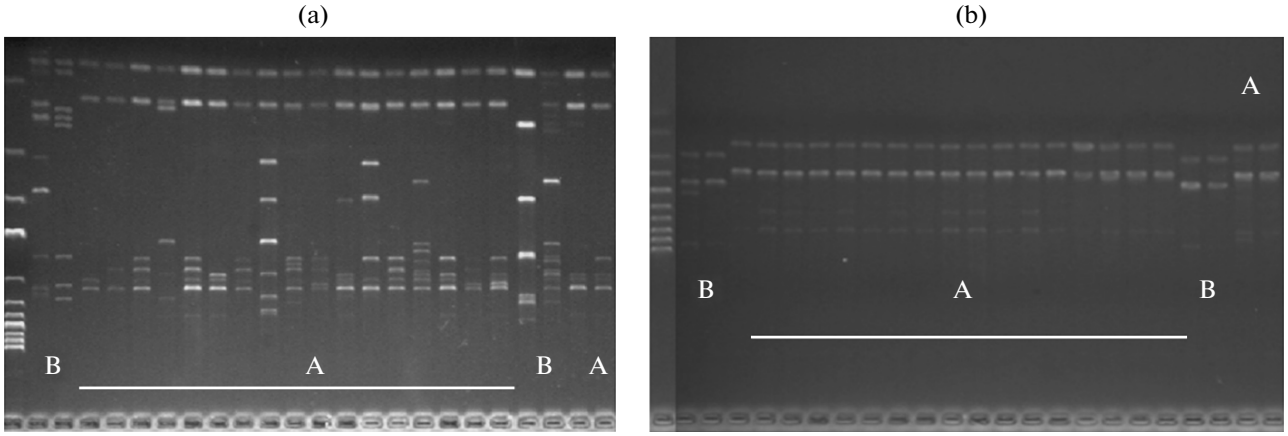

Fig. 3. A. saAFLP analysis using the restriction *Xma*JI endonuclease for bacteria of the genus *R. leguminosarum*. The letters A and B in the figure indicate the genotypes of *R. leguminosarum* by the *hin*-region (I class);

B. Amplification of the *hin*-region of the studied *R. leguminosarum* strains using specific primers for bacteria of the genus *Rhizobium*. The letters A and B in the figure indicate the genotypes of *R. leguminosarum* by the *hin*-region (I class).

Tracks: **1**, molecular weight marker 1kb; **2**, strain 245a-st; **3**, 248b; **4**, 261b-st; **5**, K-29; **6**, P-2; **7**, G-8; **8**, Ch-14; **9**, 24; **10**, 28; **11**, 31; **12**, 32; **13**, 32-2; **14**, 34; **15**, 65; **16**, U-1; **17**, U-2; **18**, U-3; **19**, U-4; **20**, U-5; **21**, U-7; **22**, U-9; **23**, U-11; **24**, U-12.

saAFLP analysis. For a reliable studies of intraspecific polymorphisms of strains of the *Rhizobium* genus saAFLP analysis using the *Xma*JI restriction endonuclease was performed (Zotov et al., 2010).

According to the saAFLP data, 23 strains of *R. leguminosarum* bv. *viceae* that inoculate peas were authentically divided into two genotypes related to genetically heterogeneous groups A and B (Fig. 3a), which had less than a 15% level of common fragments. By this analysis we succeeded in establishing the genetic heterogeneity of *R. leguminosarum* strains from the Vinnitsa and Kherson regions of Ukraine (Fig. 3a, Table 1). Thus, within the groups of strains with a *16S* rRNA homogeneous sequence, significant polymorphism of saAFLP products was shown and four genetic groups of nodule bacteria were identified. This variability may be associated with certain genetic changes in the process of adaptation to various eco–geographical areas and/or the host plant (Terefework et al., 2001).

Six isolates, viz., milk vetch symbionts that were referred by the analysis of nucleotide sequences of the *16S* rRNA gene to *R. galegae*, were a genetically homogeneous group of strains with a common characteristic saAFLP profile. Strains classified by DNA sequence analysis of *16S* rRNA to species *R. etli* and *R. giardinii* also had their own unique saAFLP profiles.

As is known, the saAFLP method is a convenient tool for finding taxon-specific markers. The *hin*-region (from the Latin *hin*, unique, unrepeatable), which was found and studied in this paper, became one of these markers (Zotov with et al., 2011. Patent Application Reg. Number of 2011135461 on 25.08.2011). The marker was observed experimentally during the saAFLP analysis of *R. leguminosarum* bv. *viceae* strains (Fig. 3).

***hin*-region analysis.** In the present work the complete nucleotide sequences of the *hin*-region (Zotov et al., 2010) of 61 bacteria strains of the genus *Rhizobium* were amplified and sequenced (Table 1). The obtained sequences were aligned using the CLUSTALW program (Thompson et al., 1994) and compared with five nucleotide sequences of *Rhizobium* sp. strains from the database of the National Center for Biotechnology Information (NCBI) USA. The length of the nucleotide aligned sequences varied among the different species from 230 bp (*R. etli* 926-st) to 650 bp (*R. leguminosarum* bv. *phaseoli* 105). In this case, all the strains, except members of the species *R. galegae* and three groups of isolates of *R. leguminosarum* bv. *phaseoli*, gave three PCR products, thus indicating the presence of three copies of tRNA (Glu) genes (Fig. 4).

Based on our analysis, the 54 studied strains were classified into three biovars: *R. leguminosarum* bv. *viceae*, *R. leguminosarum* bv. *trifolii*, and *R. leguminosarum* bv. *phaseoli*. One strain belonged to the species *R. etli* 700st (the V class of the *hin*-region) and six strains to *R. galegae* (the VI class of the *hin*-region). According to the data, the *hin*-region PCR isolates from the pea nodules were from six groups of strains and the isolates from the beans nodules, were from two groups, which belong to class I of the *hin*-region (Fig. 5). Symbionts of clover were divided into five groups of strains of class I of the *hin*-region, two of which were isolated from soils in Ukraine (Figure 5); the isolates from bean nodules belonged to five genetically remote groups: *R. leguminosarum* (genotypes A, B, b, C, D) and *R. etli* (genotype G) (the V class of the *hin*-region). Thus, on the studied sample strains of the genus *Rhizobium* that were isolated from different eco–geo-

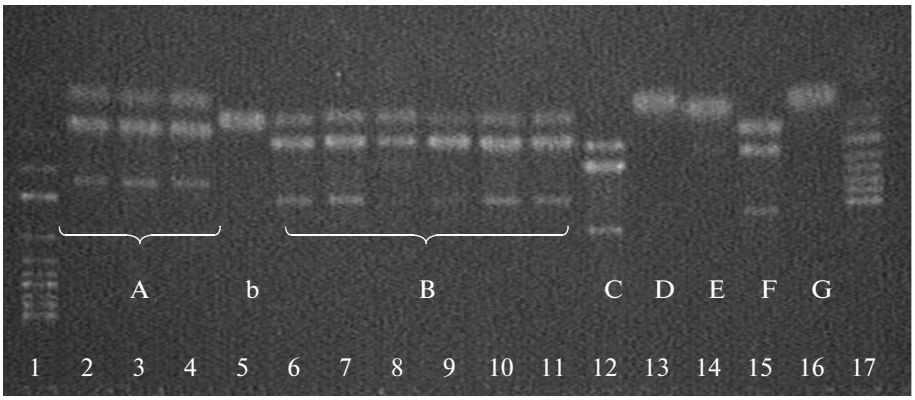

Fig. 4. *Hin*-region amplification using specific primers for bacteria of the genus *Rhizobium*. The letters in the figure indicate the genotypes of bacteria of the genus *Rhizobium* by the *hin*-region PCR (Table 1).

Tracks: **1**, molecular weight marker 100, 1000 bp.;
2, *R. leguminosarum* 28; **3**, *R. leguminosarum* K-29;
4, *R. leguminosarum* 261b-st; **5**, *R. leguminosarum* FK-0;
6, *R. leguminosarum* 96^{st}; **7**, *R. leguminosarum* 245a-st;
8, *R. leguminosarum* U-9; **9**, *R. leguminosarum* KP15;
10, *R. leguminosarum* U-7; **11**, *R. leguminosarum* MK;
12, *R. leguminosarum* FN; **13**, *R. leguminosarum* FN-6;
14, *R. giardinii* FS; **15**, *R. etli* 700^{st}; **16**, *R. galegae K-3*; **17**, molecular weight marker 100 bp.

graphical zones of Ukraine, seven genotypes with different lengths and DNA sequences of the *hin*-region were revealed. In the total strains, eight genotypes were identified according to the DNA sequence of their *hin*-region in the studied sample (Table 1). Division of the study-sample strains into genotypes by the sequences of the *hin*-region coincided with the grouping of strains by saAFLP. However, a correlation between the host plant, the origin of the strain, and genotypes was not found. Some strains, which were combined into one genotype, were symbionts of different biovars of plants (such as bean and clover, genotype B and peas and beans, genotype A). Strains allocated simultaneously on one and the same crop area (for example, symbionts of the pea in the Uladovsky district of the Kherson region) were a genetically remote population according to the *hin*-region PCR and saAFLP data, with the majority of isolates belonging to genotype A, which could be due to both the competitiveness of the strains of this genotype and the number of members of this genotype in the soil. The absence of this correlation was the expected confirmation of the other scientists studies (Wernegreen et al., 1997). This is due to the fact that in bacteria of the genus *Rhizobium* all the genes that determine symbiotic nature are typically located on a plasmid and can be transferred from one species to another, making it saprophytic or capable of symbiosis (Sullivan et al. 1995; Sullivan, Ronson, 1998). However, we are focusing on the study of the taxonomic structure of the genus, i.e., on the study of specific genes and intergenic regions of chromosomal DNA.

Further, for the aligned sequences of *R. leguminosarum* of the I class *hin*-region, a graph was built that showed the level of DNA polymorphism between three *Avr*II sites of the same name with a displacement of a 50-nucleotide window in steps of 1 bp. (Fig. 6). Studies of common polymorphisms of nucleotide sequences based on the example of the *R. leguminosarum* class I *hin*-region revealed that this region had a high level of polymorphism (15%) between strains of different biovars. Thus, the informativeness of the locus we proposed is two orders of magnitude higher than that of *16S* rRNA, is comparable in the resolution with the method of ITS sequencing, amounts to 22 nucleotide substitutions at 100 bp, and allows us to study bacteria at the intraspecific level.

It should be noted that the results obtained for the *hin*-region correlate with the results of saAFLP and with the analysis of the nucleotide sequences of *16S* rRNA and ITS. However, the taxonomic marker that we propose has significant advantages: it is chromosomal, has a higher resolution ability and is present in the genome as a single copy. Thus, PCR analysis of *hin*-region is a reliable diagnostic tool for the study of intraspecific genetic polymorphisms of the population of nodule bacteria of the genus *Rhizobium*. As a result of the analysis of the *hin*-region structure we identified specific nucleotide substitutions, including extensive insertions and deletions that were typical for species and groups of strains isolated by saAFLP analysis. These data will allow us to choose taxon-specific primers separately for each group of strains with identical sequences of the *hin*-region.

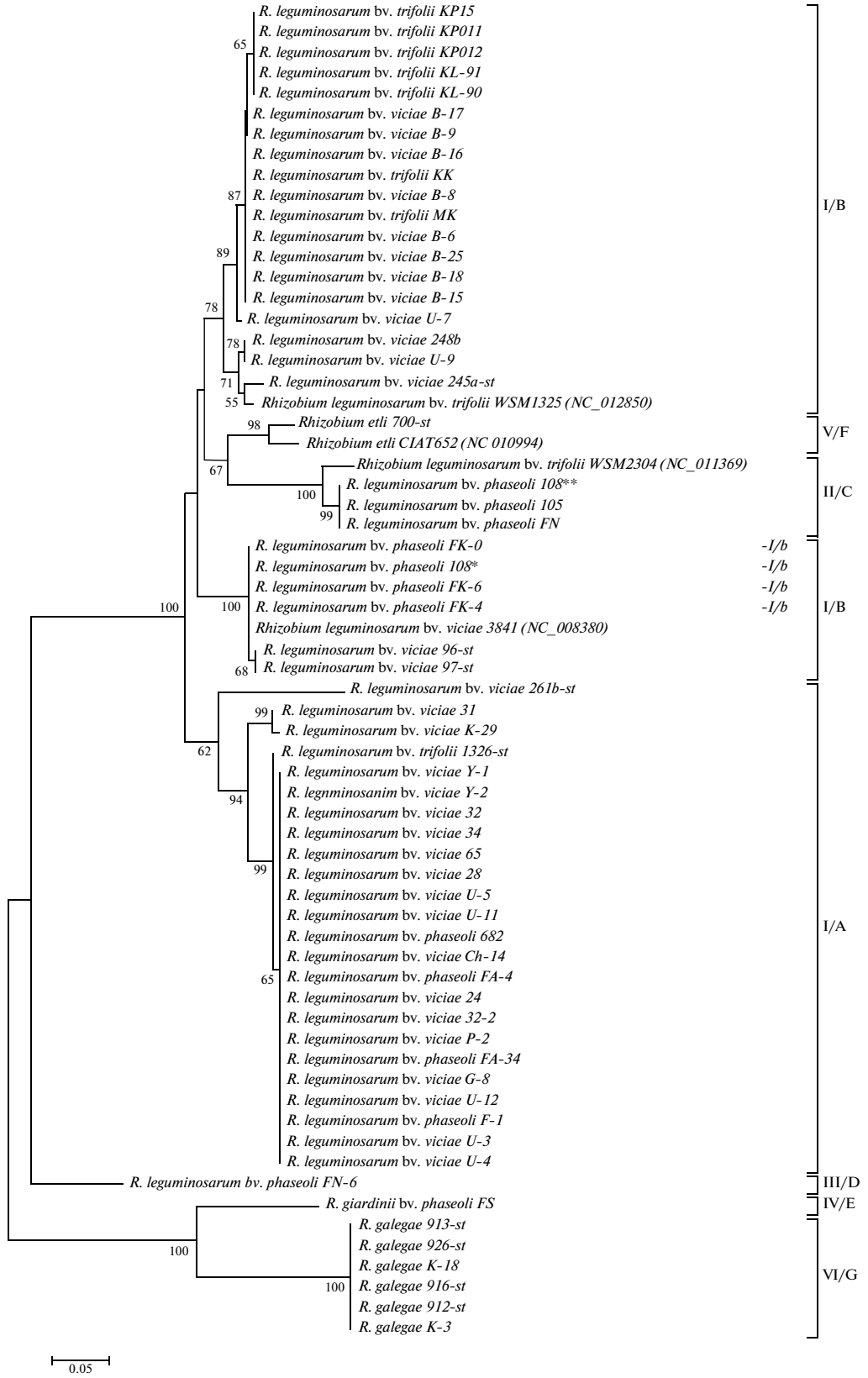

Fig. 5. A phylogenetic tree that was constructed on the basis of comparative sequence analysis of the *hin*-region of *Rhizobium* sp. using the NJ algorithm. The scale corresponds to 5 substitutions per 100 base pairs (genetic distances). The numerals show the statistical reliability of the order of branching (%) defined by bootstrap analysis (1000 replicas). Bootstrap values less than 70% are not shown.

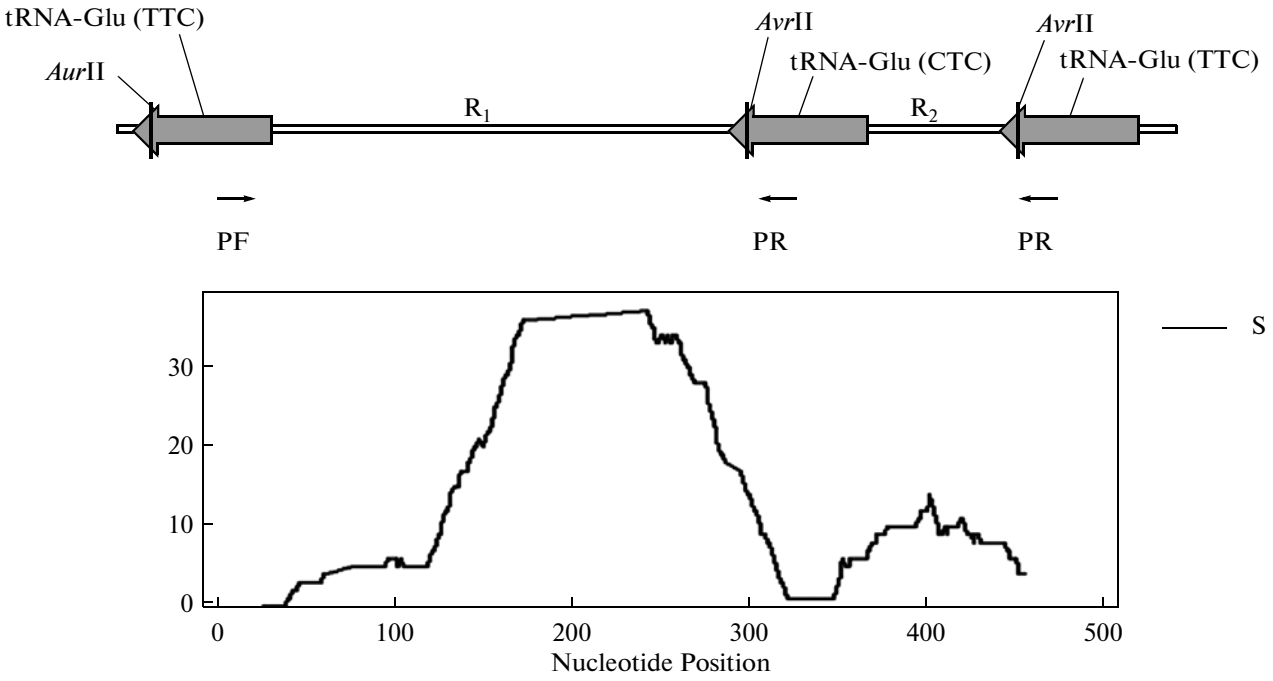

Fig. 6. Graphs drawn using the DnaSP 5.10 software (Rozas et al., 2010) show the level of polymorphism of the class I *hin*-region for the bacteria *R. leguminosarum* (on the x axis, the average position of a 50-nucleotide sliding window; on the y axis, S, the number of segregating sites). The *hin*-region is the sum of the R1 and R2 nucleotide sequences. Specific primers for the genus *Rhizobium* are designated PF and PR. The grey arrows show tRNA(Glu) genes.

CONCLUSIONS

Thus, we developed a marker system that was specific to the nodule bacteria of the genus *Rhizobium*, and allowed us to separate them at the level of species and groups of strains (Table 2). Genus-specific PCR of the *hin*-region allows one to identify bacteria of the genus *Rhizobium*; the determination of their nucleotide sequences provides information on intraspecific diversity. The application of the *hin*-region analysis together with other methods, which are presented in Table 2, gives a complete picture of the taxonomy of the genus *Rhizobium* at any level. The data obtained by analysis of the *hin*-region are consistent with the traditional techniques discussed above, but far exceed them in their informativeness. Thus, based on our studies, in a comparative analysis of the DNA sequences of the *hin*-regions eight genotypes of rhizobia isolated in different eco-geographical areas of Ukraine were identified. The division of strains into groups coincided with the phylogeny that was obtained by the sequence of *16S* rRNA gene and grouping of strains by saAFLP analysis. However, the variability of the nucleotide sequence of *hin*-regions within certain genotypes was higher and varied from 3.3 to 77.5% (compared to 0.1–1.2% for *16S* rRNA and 3.2–19.2% for saAFLP).

Table 2. The resolutions of the phenotypic and genetic methods that were used in this work to study the biodiversity of bacteria of the genus *Rhizobium*

Method	Genus *Rhizobium*	Species	Biovar	Group of strains
Phenotypic methods				
Biochemical analysis	+	+/–	–	–
Test for inoculation of host plants	+	+/–	+	–
Genotypic methods				
hin-region PCR	+	+	–	+
saAFLP	–	+	+	+
ITS PCR–RFLP	–	+/–	–	+/–
Sequencing of *16S* rRNA	+	+	–	–
Sequencing of *16S–23S* rRNA (ITS)	+	+	–	+

Although, the origin and significance of the *hin*-region for rhizobia remains unclear, on the basis of this study it can be assumed that it is necessary for the bacterial survival. It is known that this sequence, which is unique and is not found in any sequenced pro- and eukaryotes genome, is located between repeats of tRNA (Glu) genes, which, in turn, are the precursors in the initial stage of the tetrapyrrol biosynthesis (hemoglobin and chlorophyll) in plants, Archea, and bacteria (Jahn et. al., 1992).

The biosythesis of legoglobin (leggemoglobin, Lb), symbiotic hemoglobin, occurs in the nodules of leguminous plants, which plays a crucial role in the process

of nitrogen fixation: Lb promotes the transfer of oxygen into the symbiosomes, supplying nitrogen-fixing microsymbionts with bound oxygen. On the other hand Lb performs buffering functions, linking excess oxygen, which inhibits the catalytic activity of nitrogenase (Topunov, Petrov, 2001; Kosmachevskii, Topunov, 2009). The number of tRNA (Glu) genes and the type of anticodon affect the intensity of the Lb biosynthesis and the ability for other protein compounds of the C5 metabolic pathway to participate in biosynthesis (Levican et al., 2005).

We carried out a preliminary search for the presence of transcription factors and promoter areas in this region (Reese, 2001). According to the study, the R2 region contained a conserved area with a length of 46 bp, that was a promoter with a 100% probability for plants and a 90% probability for bacteria; it contained GATA and GCGC motifs which were necessary for recognition of the signaling (immune) system of plants by protein molecules. Studies of the R1 region did not reveal any characteristic promoter regions and motifs, except for the TATA box. However, the secondary structure of this region was the same for all classes of R1 DNA sequences and formed two-hairpin spatial conformation at positions located around –30 and –60 bp.

Baced on the given above, we can assume that the *hin*-region is a promoter of tRNA (Glu) genes and that its unique sequence is probably a consequence of cooperative adaptation (evolution) of bacteria and host plants to each other.

ACKNOWLEDGMENTS

This work was supported by RFBR grant (12-04-01809), a joint RFBR (10-04-90043 Bel_a)-BRFFR (B10 R-211), a special-purpose RAS Presidium program, and the Support for Young Scientists, and the Scientific basis for the management of microbiological processes in high-yield agricultural technologies of agricultural production (Agricultural Microbiology) Program of the Research NAAS of Ukraine

REFERENCES

Altschul, S.F., Gish, W., Miller, W., et al., Basic Local Alignment Search Tool, *J. Mol. Biol.*, pp. 403–410.

Amarger, N., Macheret, V., and Laguerre, G., *Rhizobium gallicum* sp. nov. and *Rhizobium giardinii* sp. nov., from *Phaseolus vulgaris* Nodules, *Int. J. Syst. Bacteriol.*, 1997, vol. 47, no. 4, pp. 996–1006.

Beringer, J.E., R1 Transfer in *Rhizobium leguminosamm*, *J. Gen. Microbiol.*, 1974, vol. 84, pp. 188–198.

Bergey's Manual of Determinative Bacteriology, 9th ed., Holt, J.G., Krieg, N.R., Sneath, P.H.A., Staley, J.T., and Williams, S.T., Eds., 9th ed., Baltimore: Williams and Wilkins, 1993.

Dydovich, S.V., Tolkachov, M.Z., and Butvina, O.Yu., Efficiency of Symbiotic Nitrogen Fixation in Agrocenoses of Ukraine, in *Agricultural Microbiology: Interinstitutional Collection of Scientific Papers,* Chernigiv: ISGM UAAS, 2008, vol. 8, pp. 117–125.

Eardly, B.D., Nour, S.M., van Berkum, P., and Selander, R.K., Rhizobial *16S* rRNA and *dnaK* Genes: Mosaicism and the Uncertain Phylogenetic Placement of *Rhizobium galegae*, *App. Env. Mic.,* 2005, vol. 71, no. 3, pp. 1328–1335.

Frank, B., Ueber die Pilzsymbiose der Leguminosen, *Ber. Deut. Bot. Ges.*, 1889, vol. 7, pp. 332–346.

Gaunt, M.W., Turner, S.L., Rigottier-Gois, L., et al., Phylogenies of *AtpD* and *RecA* Support the Small Subunit rRNA-Based Classification of Rhizobia, *Int. J. Syst. EV. Microbiol.*, 2001, vol. 51, pp. 2037–2048.

Gürtler, V. and Stanisich, V.A., New Approaches to Typing and Identification of Bacteria Using the *16S–23S* rDNA Spacer Region, *Microbiology*, 1996, vol. 142, pp. 3–16.

Hall, T.A., BioEdit: a User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT, *Nucl. Acids. Symp. Ser.,* 1999, vol. 41, pp. 95–98.

Jahn, D., Verkamp, E., and Soll, D., Glutamyltransferase RNA: A Precursor of Heme and Chlorophyll Biosynthesis, *Trends. Biochem. Sci.*, 1992, vol. 17, no. (6), pp. 215–223.

Kimura, M.A., A Simple Method for Estimating Evolutionary Rate at Base Substitudions Tthrough Comparative Studies of Nucleotide Sequences, *J. Mol. EV.,* 1980, vol. 16, pp. 111–120.

Kosmachevskaya, O.V. and Topunov, A., Hemoglobins: Diversity of Structures and Functions, *Appl. Biochem. Microbiol.*, 2009, vol. 45, no. 6, pp. 627–653.

Kumar, S., Tamura, K., and Nei, M., MEGA3: Integrated Software for Molecular Evolutionary Genetics Analysis and Sequence Alignment, *Brief. Bioinform.*, 2004, vol. 5, pp. 150–163.

Kwon, S.-W., Park, J.-Y., Kim, J.-S., et al., Phylogenetic Analysis of the Genera *Bradyrhizobium, Mesorhizobium, Rhizobium* and *Sinorhizobium* on the Basis of *16S* rRNA Gene and Internally Transcribed Spacer Region Sequences, *Int. J. Syst. EV. Microbiol.,* 2005, vol. 55, pp. 263–270.

Laguerre, G., Mavingui, P., Allard, M.-R., et al., Typing of Rhizobia by PCR DNA Fingerprinting and PCR-RFLP Analysis of Chromosomal and Symbiotic Gene Regions: Application to *Rhizobium leguminosarum* and Its Different Biovars, *App. Env. Microbiol.*, 1996, vol. 62, pp. 2029–2036.

Lindstrom, K., *Rhizobium galegae*, a New Species of Legume Root Nodule Bacteria, *Int. J. Syst. EV. Microbiol.,* 1989, vol. 39, no. 3, pp. 365–367.

Martinez-Romero, E. and Caballero-Mellado, J., *Rhizobium* Phylogenies and Bacterial Genetic Diversity, *Critical Rev. Plant Sci.*, 1996, vol. 15, pp. 113–140.

Martínez, E., Recent Developments in Rhizobium Genome, *Plant Soil*, 1994, vol. 161, pp. 11–20.

Metody pochvennoi mikrobiologii i biokhimii (Methods of Soil Microbiology and Biochemistry) Zvyagintsev, D.G., Ed., Moscow: Izd. Mosk. Gos. Univ., 1991.

Nei, M. and Kumar, S., *Molecular Evolution and Phylogenetics*, New York: Oxford University Press, 2000.

Normand, P., Cournoyer, B., Nazaret, S., and Simonet, P., Analysis of a Ribosomal Operon in the *Actinomycete frankia*, *Gene*, 1992, vol. 111, pp. 119–124.

Novikova, N.I., Modern Concepts of the Phylogeny and Taxonomy of Nodule Bacteria, *Microbiology*, 1996, vol. 65, no. 4, pp. 437–450.

Palmer, K.M. and Young, J.P.W., Higher Diversity of *R. leguminosarum* bv. *viciae* Populations in Arable Soils Than in Grass Soils, *App. Env. Microbiol.*, 2000, vol. 66, pp. 2445–2450.

Reese, M.G., Application of a Time-Delay Neural Network to Promoter Annotation in the *Drosophila melanogaster* Genome, *Comp. Chem.*, 2001, vol. 26, no. 1, pp. 51–56.

Rhizobiaceae. Molekulyarnaya biologiya bakterii, vzaimodeistvuyushchikh s rasteniyami (Rhizobiaceae. Molecular Biology of Bacteria That Interact with Plants), Spaink, G., Kondoroshi, A., and Hukas, P., St. Petersburg: Biont, 2002.

Sanger, F., Air, G.M., Barrell, B.G., et al., Nucleotide Sequence of Bacteriophage Phi X174 DNA, *Nature,* 1977, vol. 265, pp. 687–695.

Segovia, L., Young, J.P.W., and Martinez-Romero, E., Reclassification of American *Rhizobium leguminosarum* Biovar *phaseoli* Type I Strains in a New Species, *Rhizobium etli* sp. nov., *Int. J. Sys. Bacteriol.,* 1993, vol. 43, pp. 374–377.

Sullivan, J.T., Patrick, H.N., Lowther, W.L., Scott, D.B., and Ronson, C.W., Nodulating Strains of *Rhizobium loti* Arise through Chromosomal and Symbiotic Gene Transfer in the Environment, *Proc. Natl. Acad. Sci. U.S.A.*, 1995, vol. 92, pp. 8995–8999.

Sullivan, J.T. and Ronson, C.W., Evolution of Rhizobia by Acquisition of a 500-Kb Symbiosis Island That Integrates into a Phe-tRNA Gene, *Proc. Natl. Acad. Sci. U.S.A.*, 1998, vol. 95, pp. 5145–5149.

Suominen, L., Roos, C., Lortet, G., et al., Identification and Structure of the *Rhizobium galegae* Common Nodulation Genes: Evidence for Horizontal Gene Transfer, *Mol. Biol. EV.*, 2001, vol. 18, pp. 907–916.

Terefework, Z., Kaijalainen, S., and Lindström, K., AFLP Fingerprinting as a Tool to Study the Genetic Diversity of *Rhizobium galegae* Isolated from Galega Orientalis and Galega Officinalis, *J. Biotechnol.*, 2001, vol. 91, pp. 169–180.

Terefework, Z., Nick, G., Suomalainen, S., and Padin, L., Phylogeny of *Rhizobiurn galegae* with Respect to Other Rhizobia and Agrobacteria, *Int. J. Syst. Bacteriol.,* 1998, vol. 48, pp. 349–356.

Thompson, J.D., Higgins, D.G., and Gibson, T.J., CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position Specific Gap Penalties and Weight Matrix Choice, *Nucleic Acids Res.,* 1994, vol. 22, pp. 4673–4680.

Topunov, A.F. and Petrova, N.E., Hemoglobins: Evolution, Distribution, and Heterogeneity, *Usp. Biol. Khim.,* 2001, vol. 41, pp. 199–228.

Vandamme, P., Pot, B., Gillis, M., et al., Polyphasic Taxonomy, a Consensus Approach to Bacterial Systematics, *Microbiol. Rev.,* 1996, vol. 60, pp. 407–438.

Vessey, J.K. and Chemining'wa, G.N., The Genetic Diversity of *Rhizobium leguminosarum* bv. *viciae* in Cultivated Soils of the Eastern Canadian Prairie, *Soil Biol. Biochem.*, 2006, vol. 38, pp. 153–163.

Vinuesa, P., Rojas-Jiménez, K., Contreras-Moreira, B., et al., Multilocus Sequence Analysis for Assessment of the Biogeography and Evolutionary Genetics of Four *Bradyrhizobium* Species That Nodulate Soybeans on the Asiatic Continent, *App. Env. Microbiol.*, 2008, vol. 74, pp. 6987–6996.

Voznyakovskaya, Yu.M. and Popova, Zh.P., *Metodicheskie ukazaniya po identifikatsii nesporovykh bakterii, dominiruyushchikh v rizosfere rastenii* (Guidance on the Identification of Nonsporeformers Bacteria Dominating the Rhizosphere of Plants), Leningrad: VNIISKhM, 1985.

Weisburg, W.G., Barns, S.M., Pelletier, D.A., and Lane, D.J., *16S* Ribosomal DNA Amplification for Phylogenetic Study, *J. Bacteriol.*, 1991, vol. 173, pp. 697–703.

Wernegreen, J.J., Harding, E.E., and Riley, M.A., *Rhizobium* Gone Native: Unexpected Plasmid Stability of Indigenous *Rhizobium leguminosarum*, *Proc. Natl. Acad. Sci. U.S.A.,* 1997, vol. 94, pp. 5483–5488.

Zotov, V.S., Punina, N.V., Ignatov, A.N., et al., Elaboration and Use of New Approach to Species and Strain Identification of Phytopatological and Nitrogenfixing Bacteria, in *Adaptation to Climate Change in the Baltic Sea Region: Contributions from Plant and Microbial Biotechnology. Program and Abstracts, Finland,* 2010, p. 87.

Zotov, V.S., Punina, N.V., and Topunov, A.F., A Method for Identification and Differentiation of Prokaryotic Organisms, RF Patent Appl. No. 2011135461, 2011.