AGRICULTURAL BIOLOGY,
2020, V. 55, № 3, pp. 489-498
(SEL'SKOKHOZYAISTVENNAYA BIOLOGIYA)

ISSN 2412-0324 (English ed. Online)
ISSN 0131-6397 (Russian ed. Print)
ISSN 2313-4836 (Russian ed. Online)

Evolution of plant-microbe symbiosis

UDC 631.461.52:575.22:575.85

doi: 10.15389/agrobiology.2020.3.489eng
doi: 10.15389/agrobiology.2020.3.489rus

IDENTIFICATION OF THE ANCESTRAL CHARACTERISTICS IN THE GENOME OF *Rhizobium leguminosarum* bv. *trifolii*

T.S. AKSENOVA[1], E.R. CHIRAK[1], O.P. ONISHCHUK[1], O.N. KURCHAK[1], A.M. AFONIN[1], A.G. PINAEV[1], E.E. ANDRONOV[1, 2, 3], N.A. PROVOROV[1]

[1]*All-Russian Research Institute for Agricultural Microbiology,* 3, sh. Podbel'skogo, St. Petersburg, 196608 Russia, e-mail tsaksenova@mail.ru, chirak.elizaveta@gmail.com, olony@yandex.ru, okurchak@yahoo.com, aafonin@arriam.ru, agpinaev@gmail.com, eeandr@gmail.com (✉ corresponding author), provorovnik@ya.ru;
[2]*Dokuchaev Soil Science Institute,* 7/str. 2, Pyzhyovskii per., Moscow, 119017 Russia;
[3]*Saint-Petersburg State University,* 7-9, Universitetskaya nab., St. Petersburg, 199034 Russia

ORCID:
Aksenova T.S. orcid.org/0000-0002-7294-8410
Chirak E.R. orcid.org/0000-0002-1610-8935
Onishchuk O.P. orcid.org/0000-0002-5378-7826
Kurchak O.N. orcid.org/0000-0003-3555-7426
Afonin A.M. orcid.org/0000-0002-8530-0226
Pinaev A.G. orcid.org/0000-0001-8272-9679
Andronov E.E. orcid.org/0000-0002-5204-262X
Provorov N.A. orcid.org/0000-0001-9091-9384

The authors declare no conflict of interests
Acknowledgements:
Supported financially by Russian Foundation for Basic Research, the project No. 18-34-00839 (collection of nodules and isolation of strains), and Russian Science Foundation, the project No. 19-16-00081 (genome sequencing and bioinformatics data analysis)
Received March 12, 2020

Abstract

Nodule bacteria of the species *Rhizobium leguminosarum* are symbiotic N_2-fixers that divide into two biotypes: *viciae* and t*rifolii* (D.C Jordan. et al., 1984). Symbiotic genes, the evolution of which depends on host plants, are responsible for the function of symbiotic nitrogen fixation (J.P.W. Young et al., 1989). Recently it was shown that according to the type of organization of the symbiotic regions of the genomes, rhizobia isolated from the *Vavilovia formosa* (Stev.) Fed. are close to the protosymbiont of the tribe *Fabeae R. leguminosarum bv. viceae* (E.R. Chirak et al., 2019). However, in the evolution of *R. leguminosarum*, there was another earlier divergence between the biotypes *viceae* and *trifolii*, the starting point of which was the protosymbiont of the entire species *R. leguminosarum*, which existed before its separation into biovars. In this work we present the results of genomes sequencing of a group of *Rhizobium leguminosarum* bv. *trifolii* and comparisons of the structure of their symbiotic regions with the corresponding regions of the genomes of *Rhizobium leguminosarum* bv. *viciae*, related to the ancestral and "advanced" types. In the program CLC Genomics Workbench 7.5.1, we compared the obtained genome-wide sequences of the strains *R. leguminosarum* bv. *trifolii* (Oxford Nanopore sequencing technique) with reference strains of *R. leguminosarum* bv. *viceae*, related to ancestral and "advanced" types. It was shown that in the genomes of strains of clover symbionts, four of five ancestral characters are found: an increased size of intergenic regions in the symbiotic region, the presence of the *nodX* gene in the nod-operon, the absence of the *nodT* gene in the *sym*-region, and only one copy of the *fixNOPQ* operon located on the pSym. Based on the results obtained, we suggest that the protosymbiont *R. leguminosarum* could be close to clover rhizobia.

Keywords: *Rhizobium leguminosarum* bv. *trifolii*, symbiosis evolution, symbiotic genes, protosymbiont, genome-wide sequences

Nodule bacteria *Rhizobium leguminosarum*, the most widespread symbiotic nitrogen fixators in temperate latitudes, comprise two biotypes contrasting in host specificity, bv. *viciae* (symbionts of vetch, pea, rank, lentil and vavilovia) and bv. *trifolii* (clover symbionts) [1]. Symbiotic genes (*sym*-genes) are responsible for the function of symbiotic nitrogen fixation, the evolution of which is largely determined by the host plants [2].

Three main groups of *sym*-genes distinguished in rhizobia are *nod* (syn-

thesis of lipochitooligosaccharide signaling Nod factors that induce nodule development) [3-5], *nif* (synthesis of nitrogenase) [5] and *fix* (energy supply for nitrogenase and regulation of *nif* genes) [5, 6]. The evolution of the symbiotic complex first occurred through the assembly of groups of genes encoding various signaling and metabolic properties that ensure the functioning of symbiosis in primary rhizobia. The primary rhizobia originated from non-symbiotic diazotrophs, followed by the transfer of the assembled symbiotic constructs into the so-called secondary (derivative) species [7]. Primary rhizobia, the relatives of the modern genus *Bradyrhizobium*, were close to the free-living phototroph *Rhodopseudomonas* and acquired the ability to fix nitrogen by recruiting some genes that control photosynthesis into the *fix* system [8]. This reorganization led to the emergence of photosynthetically active Bradyrhizobium spp., nodulating the stems of tropical legumes without the use of *nod* genes. The ability to synthesize Nod factors for which *nod* genes are responsible was probably first acquired by bacteria of the genus *Bradyrhizobium* in which phototrophy was replaced by the ability to use photosynthetic products of plant. These heterotrophic rhizobia usually retain the expression of ex planta *nif* genes, but they are not capable of diazotrophic growth due to low nitrogenase activity [9]. The most studied secondary rhizobia are the members of genera *Rhziobium*, *Sinorhizobium*, *Mezorhizobium*, and *Neorhizobium*. These bacteria are devoid of photosynthetic systems and cannot express ex planta nitrogenase genes; their appearance was the result of horizontal transfer of *sym* genes from primary rhizobia to various soil heterotrophic bacteria [10].

Structural and functional organization of *sym*-gene regions has been studied in detail in *Rhizobium leguminosarum* bv. *viceae*, the strains of which vary significantly in the specificity towards different host plants. It has recently been shown that, according to the type of organization of symbiotic regions (*sym*-regions) of rhizobial genomes, *R. leguminosarum* bv. *viceae* can be divided into two groups [10]. The first group isolated from *Vavilovia formosa* (Stev.) Fed. (a plant, presumably close to the last common ancestor of the entire tribe *Fabeae*) [11] and possessing a complex of ancestral features of the genome, is characterized by i) an extended *sym*-region, sometimes divided between two symbiotic plasmids (pSym), ii) the presence of *nodX* and *fixW* genes in plasmid *sym*-operons, iii) the lack of chromosomal copies of *fixNOPQ*, and iv) the location of *nodT* gene outside the operons of *nod* genes. In the second, derived (or evolutionarily "advanced") group, the *sym*-region is more compact, there are chromosomal copies of *fixNOPQ*, *nodT* is integrated into the *nod* operon between *nodN* and *nodO*, and the *nodX* and *fixW* genes are lost.

The transition from the ancestral form to the advanced one is associated with general (structural and functional) compaction of the genome, an increase in the intensity of nitrogen fixation and a narrowing of the host specificity. Thus, it has been shown that the rhizobia isolated from *V. formosa* are close to *R. leguminosarum* bv. *viceae*, the protosymbiont of the tribe *Fabeae*. However, in the evolution of rhizobia of this group, there was another earlier divergence between *R. leguminosarum* bv. *viceae* and *R. leguminosarum* bv. *trifolii*, the starting point of which was the protosymbiont of the entire species *R. leguminosarum*, which existed before *R. leguminosarum* separation with the formation of biovars *viceae* and *trifolii*. And although today nothing is known either about the host plant of this protosymbiont, much less about the organization of its genome, it is obvious that comparing the symbiotic operons of the *R. leguminosarum* bv. *viceae* ancestral variants and corresponding regions in the *R. leguminosarum* bv. *trifolii* genomes will be very helpful for understanding the evolutionary construction of the protosymbiont.

In this work, we present the first results of sequencing genomes of the *Rhizobium leguminosarum* bv. *trifolii* (Rlt) group of the ancestral type to collate the structure of their symbiotic regions with the corresponding regions in *R. leguminosarum* bv. *viciae* (Rlv) of the advanced type.

The investigation aimed to search for ancestral symbiotic characters in the genome of *Rhizobium leguminosarum* bv. *trifolii.*

Materials and methods. Nodule bacteria were isolated from 50 samples of soil from rhizosphere of three clover plants (species *Trifolium pratense* L., *T. repens* L., and *T. hybridum* L.) collected in the village. Vyritsa (Leningrad Province, Gatchinsky District). To collect samples, we selected areas of compact growth of flowering plants (not farther than 0.2-0.3 m from each other), the distance between the sampling sites was at least 5 m.

Soil suspensions were prepared from each sample, which were used to inoculate sterile seedlings of red clover (*T. pratense*) and white clover (*T. repens*).

The plants were grown in pots in gnotobiotic condition on a nitrogen-free Krasilnikov-Korenyako medium. One nodule was taken from each green plant, which was sterilized in 96% alcohol and washed twice with sterile water. The nodules were destroyed with a glass rod in an eppendorf; 0.1 ml of an aqueous suspension was plated on agar medium 79 [12]. On day 3 of growth, individual colonies were subcultured in tubes on bean agar for storage [12]. A total of 37 clover rhizobia isolates were obtained, of which five were selected and grown in 5 ml of liquid medium 79 for 1 day at 28 °C. The cultures were used to isolate genomic DNA according to a standard technique [13].

For whole genome sequencing, the libraries were constructed according to the 1D native barcoding genomic DNA protocol, recommended by the manufacturer, with EXP-NBD104, EXP-NBD114, and SQK-LSK109 kits (Oxford Nanopore, Great Britain). The libraries were sequenced (a MinION nanopore sequencer, Oxford Nanopore, UK) according to the manufacturer's instructions on well R9.4. Basecalling of fast5 raw files resulted from sequencing was performed with Albacore v. 1 software (https://rubygems.org/gems/albacore/ver-sions/2.3.1). We used Deepbiner v. 0.2.0 software [14] to demultiplex the reads, Porechop v. 0.2.3 software (https://github.com/rrwick/Porechop) for cleaning sequence reads. The reads were assembled in Flye v. 2.6 (https://github.com/fenderglass/Flye). The resulting assemblies were corrected with the use of Racon v. 1.3.2 software (https://github.com/lbcb-sci/racon; -m 8 -x -6 -g -8 -w 500 options), as well as in Medaka v. 0.10.0 software (https://github.com/nanoporetech/medaka). Genome annotation was performed using the Prokka program [15]. Genomes were deposited in GenBank (http://www.ncbi.nlm.nih.gov/bioproject/PRJNA611463, the PRJNA611463 bioproject).

Sequence extraction, concatenation, and other manipulations with genomes during information processing were performed in CLC Genomics Workbench v. 7.5.1 (https://secure.clcbio.com/helpspot/in-dex.php?pg=kb.printer.friendly&id=15). The sequences were aligned using Molecular Evolutionary Genetics Analysis (MEGA) X program (https://www.megasoftware.net) [18]. The construction of phylogenetic trees by the maximum likelihood method with a bootstrap (1000 repeats) was carried out with PhyML v. 3.3 software(http://www.atgc-montpellier.fr/phyml/) [19]. The choice of the distribution model was automatically determined using the least BIC (Bayesian information criterion) method [20]. The resulting dendrograms were visualized in the online application iTOL (https://itol.embl.de) [21].

Results. Genome analysis was performed for five local isolates (3B, 9B,

22B, 23B, and 31B) and five strains of *R. leguminosarum* (Table 1).

1. *Rhizobium leguminosarum* strains used

Strain	Region	Host plant	GenBank accession number	Reference
		Rhizobium leguminosarum bv. *trifolii* (Rlt)		
3B	Pgt Vyritsa, Leningrad Province, Russia	*Trifolium repens* L.	PRJNA611463	This work
9B	Pgt Vyritsa, Leningrad Province, Russia	*Trifolium pratense* L.	PRJNA611463	This work
22B	Pgt Vyritsa, Leningrad Province, Russia	*Trifolium pratense* L.	PRJNA611463	This work
23B	Pgt Vyritsa, Leningrad Province, Russia	*Trifolium pratense* L.	PRJNA611463	This work
31B	Pgt Vyritsa, Leningrad Province, Russia	*Trifolium pratense* L.	PRJNA611463	This work
WSM1689	Greece	*Trifolium uniflorum* L.	CP007045-CP007050	[17]
		Rhizobium leguminosarum bv. *viciae* (Rlv)		
Vaf10	North Ossetia, Russia	*Vavilovia formosa* (Stev.) Fed.	CP016286-CP016293	[10]
Vaf108	Dagestan, Russia	*Vavilovia formosa* (Stev.) Fed.	CP018228-CP018236	[10]
TOM	Turkey	*Pisum sativum* L.	AQUC01000001-AQUC01000006	[16]
248	England	*Vicia faba* L.	ARRT01000001-ARRT01000007	[16]

We sequenced the genomes of five Rlt isolates (3B, 9B, 22B, 23B, 31B) to collate them with the genomes of the Rlv strains (see Table 1). The strains to compare were Vaf108 and Vaf10, the symbionts of *Vavilovia formosa* [22] which is probably the closest living relative of the common ancestor of the tribe *Fabeae* [23], TOM which is a symbiont of pea (*Pisum sativum* L.) Afghan cultivars [24], 248, a symbiont of *Vicia faba* L., and WSM1689, a symbiont of *T. uniflorum* (see Table 1). Since the divergence of *R. leguminosarum* biovars is determined by symbiotic genes, while the chromosome background of these biovars is common [25], we focused on the symbiotic regions of the genomes.

Fig. 1. Differences in the organization of *sym*-operons in the genomes of *Rhizobium leguminosarum* isolates. 3B, 9B, 22B, 23B, 31B, and WSM1689 are *R. leguminosarum* bv. *trifolii* (Rlt) strains, Vaf10, Vaf108, TOM, and 248 are *R. leguminosarum* bv. *viciae* (Rlv) strains. The *nod*-operons are marked in blue, *nif* is green, and *fix* in yellow. The top scale is the length of sym-regions, bp.

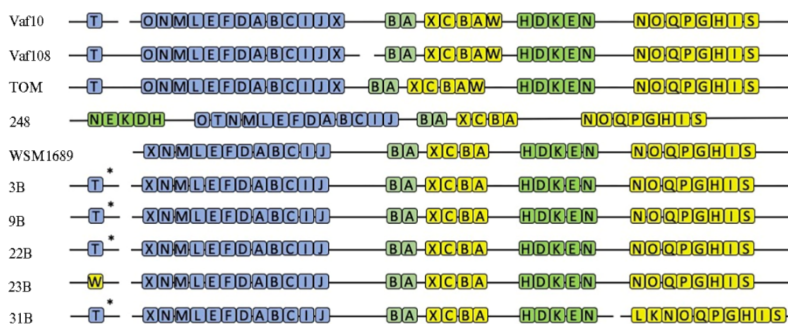

Fig. 2. Schematic structure of *sym*-operons in *Rhizobium leguminosarum* isolates. 3B, 9B, 22B, 23B, 31B, and WSM1689 are *R. leguminosarum* bv. *trifolii* (Rlt) strains, Vaf10, Vaf108, TOM, and 248 are *R. leguminosarum* bv. viciae (Rlv) strains. The *nod* operons are marked in blue, *nif* in green, and fix in yellow. An asterisk (*) marks a gene located on a chromosome.

Structure of *sym*-operons. Comparing *sym*-regions of the genomic sequences of Rlt 3B, 9B, 22B, 23B, and 31B isolates with the *sym*-regions of Rlt WSM1689 and Rlv Vaf10, Vaf108, TOM, and 248 strains revealed differences be-

tween Rlt and Rlv strains in the arrangement and location of *sym* genes (Fig. 1). In strains 3B, 9B, 22B, and 23B, the *sym* genes are organized into *sym* operons located on the pSym. Strain 31B is somewhat different due to the *fixNOQPGHIS* operon location on the chromosome (Fig. 2) and two additional genes (*fixLK*) not detected in other strains (see Fig. 2).

The *nodT* gene was also found on the chromosomes of Rlt strains 3B, 9B, 22B, and 31B, but there was no *fixW* gene. The *nodT* gene was not identified in 23B strain, but of all the studied Rlt strains, only 23B has the *fixW* gene separately located on a plasmid. In the *nod* operons of all Rlt strains, the *nodX* gene was found, but the *nodO* gene was absent.

Genomic distribution of *sym*-regions. In Rlt strains, the structure of *sym*-regions varied (see Fig. 1). In 31B, in contrast to the other strains under consideration, the *fixNOQPGHIS* operon is located on the chromosome. In strains 3B, 9B, and 22B, the distance between *nifHDKE* and *fixNOQPGHIS* operons is greatly increased, while in strain 23B it is noticeably smaller (Fig. 3).

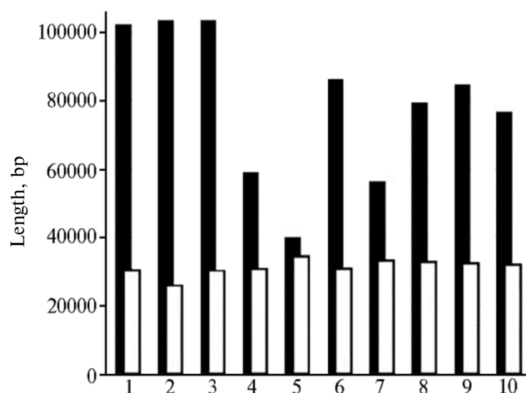

Fig. 3. The proportions between the sizes of *sym*-regions (black bars) **and *sym*-genes** (white bars) **in the genomes of *Rhizobium leguminosarum* bv. *trifolii* (Rlt) strains** (1 — 3B, 2 — 9B, 3 — 22B, 4 — 23B, 5 — 31B, 6 — WSM1689) **and *R. leguminosarum* bv. viciae (Rlv)** (7 — 248, 8 — TOM, 9 — Vaf10, 10 — Vaf108).

In addition, the *sym*-region clusters in strain 23B is the most compact and comparable in size to that in strain 248, a *V. faba* symbiont. The sizes of the *sym*-region in strains 23B and 31B correspond to those characteristic of the evolutionarily advanced group, for which a compact arrangement of *sym*-genes is typical, while strains 3B, 9B, and 22B, in which the *sym*-region is expanded, can be attributed to the ancestral evolutionary group.

Phylogenetic analysis of *sym*-genes. In strains Rlt and Rlv, in addition to the revealed structural features of the symbiotic region, we analyzed the nucleotide polymorphism of three gene groups, *fix*, *nif*, and *nod*. Figure 4 shows the phylogenies of the corresponding concatenates. The grouping of clover symbionts in a relatively compact cluster occurred in two gene groups, *nif* and *nod*, while Rlv strains grouped in a compact cluster only for *fix* genes. Noteworthy is the fact that, in the phylogeny for the *nif* and *nod* genes, the advanced symbionts Rlv (248 and TOM) appear, with reliable statistical support, in one cluster with clover rhizobia, while for *fix* genes, there are two Rlt strains (WSM1689 and 31B) fall into a relatively compact cluster with the Rlv group, which includes both advanced and ancestral Rlv strains.

The biovars of *R. leguminosarum* are represented by symbionts of two very different leguminous tribes. The *R. leguminosarum* separation into *viciae* and *trifolii* biovars has a long evolutionary history, and strains of these biovars do not nodulate legumes from the tribes *Trifolieae* and *Fabeae* upon cross-inoculation. In biovar *viciae*, symbionts of Vavilovia are distinguished, possessing a number of ancestral characters, and it is assumed that they are closest to the protosymbiont of the tribe *Fabeae*, the common ancestor of the biovar *R. leguminosarum* bv. *viceae* [10]. In the presented study, when comparing the Vavilovia rhizobia genomes sequenced earlier and the genomes of clover rhizobia studied in this work, we obtained data concerning the protosymbiont common for the entire

species *R. leguminosarum*, which existed evolutionary earlier, i.e. before the separation into biovars *viceae* and *trifolii.*

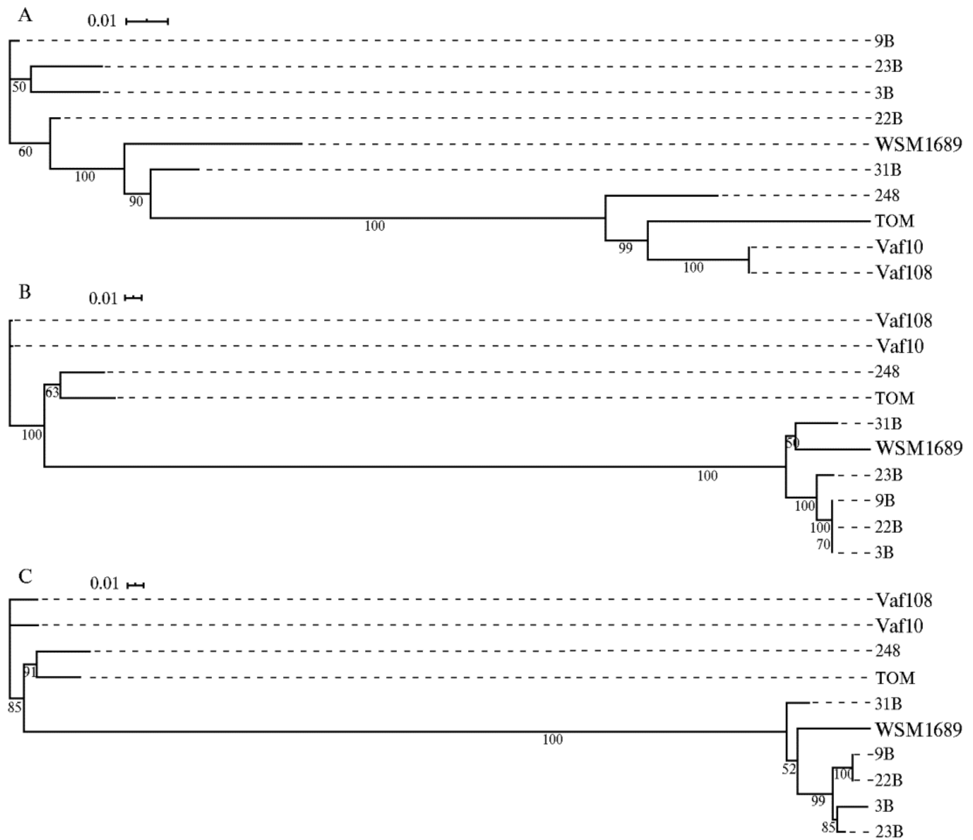

Fig. 4. Dendrogram of *Rhizobium leguminosarum* strains based on concatenated sequences of genes *fixABCGHINOPQ* (A), ***nifABDEHKN*** (B), and ***nodABCDEFIJLMN*** (C). 3B, 9B, 22B, 23B, 31B, and WSM1689 are *R. leguminosarum* bv. *trifolii* (Rlt) strains, Vaf10, Vaf108, TOM, and 248 are *R. leguminosarum* bv. *viciae* (Rlv) strains.

A significantly larger size of intergenic regions in the symbiotic region, due to the primary “rough” assembly at the early stages of evolution, is an ancestral character. Later in evolution, these regions have compacted [10]. An important result of our studies is that we have identified in the Rlt strains 3B, 9B, and 22B a sym-region with the size which is much larger than in symbionts of Vavilovia (see Fig. 1, Table 2).

2. Summarized ancestral genomic characters of in *Rhizobium leguminosarum* biotypes (based on genome-wide sequencing of 10 strains)

Rhizobium biovbar	Extended *sym*-region	Absence of *nodT* in *nod*-operon	Presence of *nodX*	Presence of *fixW*	Absence of chromosomal copy of *fixNOPQ*
R. leguminosarum bv. *trifolii*	+	+	+	–	+
R. leguminosarum bv. *viciae*	+	+	+	+	+

Note. «+» or «–» — the trait is present or absent, respectively. The studied strains are 3B, 9B, 22B, 23B, 31B, and WSM1689 (*R. leguminosarum* bv. *Trifolii*, Rlt), Vaf10, Vaf108, TOM, and 248 (*R. leguminosarum* bv. *viciae*, Rlv).

All Rlt strains are characterized by the presence of the *nodX* gene in *nod*-operon, which in Rlv strains also serves as a trait that marks ancestral genotypes [25]. The significance of *nodX* gene for symbiosis with clover has not been studied, and its loss in advanced Rlv strains is associated with a narrowing of the host specificity and an increase in the activity of nitrogen fixation [26].

Another ancestral feature is the absence of the *nodT* gene which encodes the efflux system ensuring effective release of the nod factor from the rhizobial cell [27]. As shown in Rlv strains, a probable evolutionary scenario is associated with the recruitment of this gene from the chromosome into a symbiotic cluster through duplication, neofunctionalization, and transfer. In Rlt strains, *nodT* gene is present in one copy only on the chromosome (the exception is isolate 31B in which *nodT* was not detected at all). Thus, according to this trait, the genomes of clover rhizobia demonstrate correspondence to even earlier stages of the evolution of the symbiotic gene cluster.

In Rlv strains, ancestral traits associated with functional redundancy of ancestral genotypes also include the presence of *fixW* gene. In the studied clover symbionts, *fixW* gene was found only in strain 31B, but not in the *fix* operon on pSym as in Rlv, but in a separate nonsymbiotic contig. The *fixW* function, as suggested earlier, may be associated with deep differentiation of bacteroids characteristic of rhizobia. The *fixW* manifestation and significance for symbiosis have not been studied in detail, but, most likely, *fixW* does not affect the host specificity [28].

Finally, strains Rlt 3B, 9B, 22B, and 23B, like strains Rlv Vaf10 and Vaf108, have only one copy of *fixNOPQ* genes per pSym (the operon is absent in the chromosome), in contrast to other members of Rlv which have two *fixNOPQ* copies in their genome, i.e. in pSym and in the chromosome. In strain 31B, one copy of *fixNOPQ* was detected, but in the chromosome. The *fixNOQP* genes and their homologues in Gram-negative nitrogen-fixing bacteria encode a high-affinity terminal cytochrome oxidase of the cbb3 type which provides respiration under microaerophilic conditions [29], for example, in nodule symbiosis. Most likely, the duplication of *fixNOQP* cluster and its transfer to the chromosome occurs during the late evolution of *R. leguminosarum* [30].

The differences that we revealed in the phylogenetics topology of *fixABCGHINOPQ*, *nifABDEHKN*, and *nodABCDEFIJLMN* concatenates indicate an independent evolution of the groups of genes that control various functions. The data obtained make it possible to extend the assumptions made earlier about the independent evolution of these groups within the *viceae* biovar [10] to the whole species *R. leguminosarum*.

So, the analysis of whole genome sequencing data showed that at least three (3B, 9B, and 22B) of the studied *Rhizobium leguminosarum* bv. *trifolii* strains possess a large part of the ancestral features (extended *sym*-region, absence of a chromosomal copy of *fixNOPQ* and *nodT* gene in the *nod*-operon, and the presence of *nodX*) found in rhizobia of *Vavilovia formosa* (Stev.) Fed. However, the evolutionary interpretation of the obtained data is complicated by the fact that the mechanisms of *R. leguminosarum* bv. *viciae* and *R. leguminosarum* bv. *trifolii* evolution, undoubtedly, are determined not only by the host plants and their phylogenesis, but also by the history of adaptation of these plants to various ecological and geographical zones. It is possible that characters identified in one group of rhizobia as ancestral in another group may have a different evolutionary meaning. Despite these constraints, one of the significant results of our study, is, in our opinion, the assumption that clover rhizobia, together with *Vavilovia* rhizobia, may be close to the protosymbiont of *R. leguminosarum*. We consider this assumption as one of the working hypotheses for further research.

REFERENCES

1. Jordan D.C. Family III. *Rhizobiaceae*. In: *Bergey's manual of systematic bacteriology*.

N.R. Krieg, J.G. Holt (eds.). Williams and Wilkins, Baltimore, 1984: 234-242.

2. Young J.P.W., Johnston A.W.B. The evolution of specificity in the legume-rhizobium symbiosis. *Trends in Ecology & Evolution*, 1989, 4(11): 341-349 (doi: 10.1016/0169-5347(89)90089-X).
3. Göttfert M. Regulation and function of rhizobial nodulation genes. *FEMS Microbiology Reviews*, 1993, 10(1-2): 39-63 (doi: 10.1111/j.1574-6968.1993.tb05863.x).
4. Debellé F., Moulin L., Mangin B., Dénarié J., Boivin C. Nod genes and Nod signals and the evolution of the *Rhizobium* legume symbiosis. *Acta Biochimica Polonica*, 2001, 48(2): 359-365. (doi: 10.18388/abp.2001_3921).
5. Shamseldin A. The role of different genes involved in symbiotic nitrogen fixation. *Global Journal of Biotechnology & Biochemistry*, 2013, 8(4): 84-94 (doi: 10.5829/idosi.gjbb.2013.8.4.82103).
6. Fischer H.M. Genetic regulation of nitrogen fixation in rhizobia. *Microbiological Reviews*, 1994, 58(3): 352-386.
7. Provorov N.A., Andronov E.E. Evolution of root nodule bacteria: reconstruction of the speciation processes resulting from genomic rearrangements in a symbiotic system. *Microbiology*, 2016, 85: 131-139 (doi: 10.1134/S0026261716020156).
8. Rey F.E., Harwood C.S. FixK, a global regulator of microaerobic growth, controls photosynthesis in *Rhodopseudomonas palustris. Molecular Microbiology*, 2010, 75(4): 1007-1020 (doi: 10.1111/j.1365-2958.2009.07037.x).
9. Wongdee J., Boonkerd N., Teaumroong N., Tittabutr P., Giraud E. Regulation of nitrogen fixation in *Bradyrhizobium* sp. strain DOA9 involves two distinct NifA regulatory proteins that are functionally redundant during symbiosis but not during free-living growth. *Frontiers in Microbiology*, 2018, 9: 1644 (doi: 10.3389/fmicb.2018.01644).
10. Chirak E.R., Kimeklis A.K., Karasev E.S., Kopat V.V., Safronova V.I., Belimov A.A., Aksenova T.S., Kabilov M.R., Provorov N.A., Andronov E.E. Search for ancestral features in genomes of *Rhizobium leguminosarum* bv. *Viciae* strains isolated from the relict legume *Vavilovia formosa. Genes*, 2019, 10(12): 990 (doi: 10.3390/genes10120990).
11. Mikič A., Smýkal P., Kenicer G., Vishnyakova M., Sarukhanyan N., Akopian J., Vanyan A., Gabrielyan I., Smýkalová I., Sherbakova E., Zorić L., Atlagić J., Zeremski-Škorić T., Ćupina B., Krstić P., Jajić I., Antanasović S., Đorđević V., Mihailović V., Ivanov A., Ochatt S., Ambrose M. The bicentenary of the research on ‘beautiful’ vavilovia (*Vavilovia formosa*), a legume crop wild relative with taxonomic and agronomic potential. *Bot. J. Linn. Soc.*, 2013, 172(4): 524-531 (doi: 10.1111/boj.12060).
12. *Biologicheskoe raznoobrazie kluben'kovykh bakterii v ekosistemakh i agrotsenozakh. Teoreticheskie osnovy i metody (monografiya)* /Pod redaktsiei M.L. Rumyantsevoi, B.V. Simarova [The biological diversity of nodule bacteria in ecosystems and agrocenoses. Theoretical foundations and methods (a monograph). M.L. Rumyantseva, B.V. Simarova (eds.)]. St. Petersburg, 2011 (in Russ.).
13. Somasegaran P., Hoben H.J. Isolating and purifying genomic DNA of *Rhizobia* using a large-scale method. In: *Handbook for Rhizobia: methods in legume-rhizobium technology*. Springer, New York, 1994: 279-283 (doi: 10.1007/978-1-4613-8375-8_31).
14. Wick R.R., Judd L.M., Holt K.E. Deepbinner: Demultiplexing barcoded Oxford Nanopore reads with deep convolutional neural networks. *PLoS Computational Biology*, 2018, 14(11): e1006583 (doi: 10.1371/journal.pcbi.1006583)
15. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 2014, 30(14): 2068-2069 (doi: 10.1093/bioinformatics/btu153).
16. Reeve W., Ardley J., Tian R., Eshragi L., Yoon J.W., Ngamwisetkun P., Seshadri R., Ivanova N.N., Kyrpides N.C. A genomic encyclopedia of the root nodule bacteria: assessing genetic diversity through a systematic biogeographic survey. *Standards in Genomic Sciences*, 2015, 10: 14 (doi: 10.1186/1944-3277-10-14).
17. Terpolilli J., Rui T., Yates R., Howieson J., Poole P., Munk C., Tapia R., Han C., Markowitz V., Tatiparthi R., Mavrommatis K., Ivanova N., Pati A., Goodwin L., Woyke T., Kyrpides N., Reeve W. Genome sequence of *Rhizobium leguminosarum* bv. *trifolii* strain WSM1689, the microsymbiont of the one flowered clover *Trifolium uniflorum. Standards in Genomic Sciences*, 2013, 9: 527-539 (doi: 10.4056/sigs.4988693).
18. Kumar S., Stecher G., Li M., Knyaz C., Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, 2018, 35(6): 1547-1549 (doi: 10.1093/molbev/msy096).
19. Guindon S., Dufayard J.F., Lefort V., Anisimova M., Hordijk W., Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 2010, 59(3): 307-321 (doi: 10.1093/sysbio/syq010).
20. Schwarz G. Estimating the dimension of a model. *Annals of Statistics*, 1978, 6(2): 461-464 (doi: 10.1214/aos/1176344136).
21. Letunic I., Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*, 2016, 44(W1): W242-W245 (doi:

10.1093/nar/gkw290).
22. Safronova V.I., Kimeklis A.K., Chizhevskaya E.P., Belimov A.A., Andronov E.E., Pinaev A.G., Pukhaev A.R., Popov K.P., Tikhonovich I.A. Genetic diversity of rhizobia isolated from nodules of the relic species *Vavilovia formosa* (Stev.). Fed. *Antonie van Leeuwenhoek*, 2014, 105: 389-399 (doi: 10.1007/s10482-013-0089-9).
23. Makasheva R.Kh., Drozd A.M., Adamova O.P., Golubev A.A. *Sbornik trudov po prikladnoi botanike, genetike i selektsii*, 1973, 51(1): 44 (in Russ.).
24. Ma S.W., Iyer V.N. New field isolates of *Rhizobium leguminosarum* biovar *Viciae* that nodulate the primitive pea cultivar Afghanistan in addition to modern cultivars. *Applied and Environmental Microbiology*, 1990, 56(7): 2206-2212.
25. Kimeklis A.K., Chirak E.R., Kuznetsova I.G., Sazanova A.L., Safronova V.I., Belimov A.A., Onishchuk O.P., Kurchak O.N., Aksenova T.S., Pinaev A.G., Andronov E.E., Provorov N.A. Rhizobia isolated from the relict legume *Vavilovia Formosa* represent a genetically specific group within *Rhizobium leguminosarum* biovar *viciae*. Genes, 2019, 10(12): 991 (doi: 10.3390/genes10120991).
26. Provorov N., Tikhonovich I. Genetic resources for improving nitrogen fixation in legume-rhizobia symbiosis. *Genetic Resources and Crop Evolution*, 2003, 50: 89-99 (doi: 10.1023/A:1022957429160).
27. Alvarez-Ortega C., Olivares J., Martínez J.L. RND multidrug efflux pumps: What are they good for? *Frontiers in Microbiology*, 2013, 4: 7 (doi: 10.3389/fmicb.2013.00007).
28. Hontelez J.G., Lankhorst R.K., Katinakis P., van den Bos R.C., van Kammen A. Characterization and nucleotide sequence of a novel gene *fixW* upstream of the *fixABC* operon in *Rhizobium leguminosarum*. *Mol. Gen. Genet.*, 1989,218: 536-544 (doi: 10.1007/BF00332421).
29. Talbi C., Sanchez C., Hidalgo-Garcia A., González E.M., Arrese-Igor C., Girard L., Bedmar E.J., Delgado M.J. Enhanced expression of *Rhizobium etli* cbb3 oxidase improves drought tolerance of common bean symbiotic nitrogen fixation. *Journal of Experimental Botany*, 2012, 63(14): 5035-5043 (doi: 10.1093/jxb/ers101).
30. Kopat V.V., Chirak E.R., Kimeklis A.K., Safronova V.I., Belimov A.A., Kabilov M.R., Andronov E.E., Provorov N.A. Evolution of *fixNOQP* genes encoding cytochrome oxidase with high affinity to oxygen in Rhizobia and related bacteria. *Russ. J. Genet.*, 2017, 53: 766-774 (doi: 10.1134/S1022795417070067).
31. Berrada H., Fikri-Benbrahim K. Taxonomy of the rhizobia: current perspectives. *British Microbiology Research Journal*, 2014, 4(6): 616-639 (doi: 10.9734/BMRJ/2014/5635).