Genetic resources, introgression, immunity (towards 130th Anniversary of N.I. Vavilov)

UDC 633.31:631.461.52:577.21

doi: 10.15389/agrobiol.2017.5.928rus

doi: 10.15389/agrobiol.2017.5.928eng

A COMPARATIVE ANALYSIS OF GENOMIC CHARACTERS OF REFERENCE *Sinorhizobium meliloti* STRAINS, THE ALFALFA SYMBIONTS (review)

M.L. ROUMIANTSEVA¹, V.S. MUNTYAN¹, M.E. CHERKASOVA¹,
E.E. ANDRONOV¹, A.S. SAKSAGANSKAYA¹, E.A. DZYUBENKO²,
N.I. DZYUBENKO², B.V. SIMAROV¹

¹All-Russian Research Institute for Agricultural Microbiology, Federal Agency of Scientific Organizations, 3, sh. Podbel'skogo, St. Petersburg, 196608 Russia, e-mail mroumiantseva@yandex.ru (corresponding author), vucovar@yandex.ru, mariacherkasova@mail.ru;

²Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources, Federal Agency of Scientific Organizations, 42-44, ul. Bol'shaya Morskaya, St. Petersburg, 190000 Russia, e-mail elena.dzyubenko@gmail.com, n.dzyubenko@vir.nw.ru

ORCID:

Roumiantseva M.L. orcid.org/0000-0001-5582-6473

Muntyan V.S. orcid.org/0000-0002-1979-0853

Cherkasova M.E. orcid.org/0000-0003-1873-9674

Andronov E.E. orcid.org/0000-0002-5204-262X

The authors declare no conflict of interests

Acknowledgements:

Supported by Russian Science Foundation (grant № 17-06-01095)

Received August 2, 2017

Saksaganskaya A.S. orcid.org/0000-0002-8547-4904

Dzyubenko N.I. orcid.org/0000-0003-0250-5814

Simarov B.V. orcid.org/0000-0002-6893-557X

Abstract

Plant-microbial symbiotic interaction is a unique highly specific biological system for fixing atmospheric nitrogen and its transformation into compounds accessible to living organisms. A fundamentally new approach may be the creation of a system for genetic monitoring of the stability of economically valuable strains in microbioms of agroecosystems. By comparison of the genomic characteristics of symbiotically active strains functionally significant marker sequences could be identified and the basis for genetic monitoring system will be created. In the review, we compared genomic characteristics of the symbiotic active strains obtained on the basis of *Sinorhizobium meliloti* 425a and SU47 strains, used for production of biologicals. Strain 425a was isolated from alfalfa nodules in the mid-1970s in the Almaty region (Kazakhstan), which is a part of the Central Asian primary center of cultivated plant origin designated by N.I. Vavilov. Strain SU47 was isolated from alfalfa nodules in the late 1930s in Australia, which is the secondary center of the diversity of cultivated plants. Strains CXM1-105 (CXM1) and Rm1021 (Rm2011) are widely applied as a reference strain as they have been used to develop or adapt a wide range of symbiogenetics methods. Genomes of both the original strains and their derivatives consist of a chromosome (SMc) and two megaplasmids (SMa, SMb), and do not contain cryptic plasmids. However CXM1-105 genome, unlike Rm1021 genome, did not contain 508 protein encoding ORFs (open reading frames) of which 242 are located on SMa, 121 are on SMb and 145 are on SM, as it was found using DNA biochips. This indicates significant differences in the structure of all three replicons in the reference strains CXM1-105 and Rm1021. Chromosome of CXM1-105 (CXM1), as well as of 425a did not contain sequences of phage origin (the «genomic islands») described in Rm1021. It was found that 62 ORFs of genomic islands are similar or homologous to those of the members of the same species or genus, as well as of phylogenetically distant bacterial classes. However, in the CXM1-105 chromosome there are sites for the integration of genomic islands (EU196757, EU196758 and EU196759), which are 99-100 % homologous to appropriate sequences of Rm1021 (Rm2011). We studied the occurrence of *S. meliloti* strains harboring type SMc_{Rm1021} or SMc_{CXM1-105} chromosome (presence or absence of genomic islands, respectively) in native populations. The significant prevalence of strains inherited SMc_{Rm1021} was shown for the area, belonging to the Middle-Asian gene center of cultivated plants, while strains harboring the chromosomal type SMc_{CXM1-105} were dominant in the area of extremely saline soils next to Aral Sea area ($P < 0.05$). Consequently, the presence of additional «foreign sequences», which could participate in horizontal gene transfer, is typical of native *S. meliloti* isolates abundant in

the primary center of the diversity of their host plants whereas those sequences are lost under an abiotic stress (salinity) impact. In addition, strains harboring different chromosome types, according to structural differences in the intergenic sequence of *rrn-rrl* operons, can be referred to divergent clonal lineages. According to the discussed data, it was suggested to consider strain 425a and its derivatives as the model *S. meliloti* strains to create a system for genetic monitoring of practically valuable strains in agroecosystems.

Keywords: symbiosis, alfalfa, reference strains, *Sinorhizobium meliloti*, molecular genetic analysis, genomic islands, sites for specific integration, accessory genome

Plant-microbial symbiosis is a unique highly specific biological system for fixing atmospheric nitrogen and its transformation into compounds available for living organisms. An important role in increasing the productivity of plants is performed by the microbial component [1], the importance of which is often underestimated [1-3]. It is known that the microorganisms that form the basis of biopharmaceuticals must have certain properties, including virulence, specificity, competitiveness, symbiotic activity and efficiency (productivity), and also must meet a number of technological requirements [4]. In order to comply with these requirements and to preserve the properties, strains should be subjected to supporting selection, because their economically valuable properties deteriorate or are lost over time [5-6]. One of the reasons may be that genes that determine the symbiotic properties of nodule bacteria are localized, as a rule, on plasmids, the availability and structure of which may be unstable (for example, under the influence of an abiotic stress factor), which, in turn, can lead to a decrease or loss of efficiency (productivity) of inoculated plants [7, 8].

A fundamentally different approach is to create a system for genetic monitoring of the stability of economically valuable strains and to study the pan-genome of nodule bacteria of alfalfa and the microbiome of agroecosystems [9-11]. The comparison of the genomic characteristics of symbiotic active strains will allow identifying functionally important marker sequences and can become the basis for creating such a system [12]. One of the stages of this study is a comparison of the properties of production strains that can be recommended as model in the development of a system for monitoring the stability of agroecosystems.

This review includes the first analysis of the strains of *Sinorhizobium meliloti*, obtained on the basis of the commercial strains 425a and SU47, which have been studied for a long time in domestic and foreign laboratories dealing with the issues of symbiogenetics.

Origin of the 425a, SU47 strains and their derivatives. A modern taxonomic name for the strain *Rhizobium meliloti* 425a is *Sinorhizobium (Ensifer) meliloti*. The strain was isolated as highly active from the nodules of alfalfa in the Almaty Region (Kazakhstan, AS No. 549454 dated May 25, 1977). In 1986, on its basis, a highly efficient, resistant to streptomycin (StrR) strain 425a-str-6, or CXM1 [13] was obtained, from which subsequently, as a result of UV mutagenesis, the strain CXM1-105 [14] was obtained. The strain *R. meliloti* SU47 was isolated from the nodules of *Medicago varia* in New South Wales (Australia) in 1937 [15]. On the basis of SU47, the strain Rm2011 was created. The strain Rm1021 is its StrR derivative. Also, the derivatives of Rm2011 include the strain RCR2011. All strains obtained from SU47 were maintained in different laboratories of the world, which caused differences in their names.

Alfalfa is one of the oldest cultures used by mankind. It is mentioned in the Babylonian texts relating to 700 BC [16]. The main centers of the formation and distribution of the more ancient diploid species of blue and yellow alfalfa are Asia and Middle Asia, included, according to the theory of N.I. Vavilov [17-19], in different primary centers of origin of cultivated plants (gene centers). These centers, in which natural and artificial selections are intensively and jointly act-

ing, serve as a source of the diversity of natural plant genotypes [20-21]. The Almaty Region in Kazakhstan, where the 425a strain was isolated from the nodules, is adjacent to the Central Asian gene center, while Australia (the place of isolating the SU47 strain) does not belong to the primary center of diversity of perennial alfalfa species; however, a secondary gene center of the genera *Medicago* L. was formed in this continent. Perennial alfalfa first arrived in the territory of Australia from France (at the end of the 18th century), and then from the US [19].

Taking into account the historical ways of distributing alfalfa as a host plant, it can be assumed that strains-microsymbionts have also undergone similar distribution. It is possible to assume that the strain SU47 could be introduced with seeds or soil from the territory of the Central Asian center. This does not contradict the conclusions of the researchers that the chromosomal genotype of the strain RCR2011 (derived from SU47) has a wide geographical prevalence [22]. Thus, it is possible to assume that the strains 425a and SU47 considered could have common historical origin.

Symbiotic effectiveness of the strains 425a, SU47 and their derivatives. The SU47 strain was available as a symbiotic effective inoculant for di- and tetraploid forms of alfalfa in Australia for farms since 1955 and was used there on an industrial scale since the 1960s, and in New Zealand since 1973-1975 [23]. On the basis of the strain Rm2011, the biopreparation Nitrogen [24] was created. The symbiotic properties of Rm1021 and Rm2011 have been studied during long-term model microvegetation and field experiments, which served as a rationale for the classification of these strains as referent strains [25-26]. The strain 425a is used to prepare the Rhizotorphin biopreparation. The average increase in the yield of alfalfa during inoculation with this strain, as it was stated (AC No. 549454 of May 25, 1977), was 14.5%. Field and plot experiments conducted using different varieties of alfalfa (Yakutskaya, Agniya, Pastbishchnaya 88, Vega) in the continental and transitional conditions of the temperate climatic belt (Tyumen, Vladikavkaz, Leningrad Region), as well as in the Far North in the Republic of Sakha (Yakutia) in the period from 1999 to 2016, showed an increase in the yield of green alfalfa mass by 9.8-48.4% when inoculated with the strain 425a ($P < 0.0015$) [27]. Its derivatives – the strains CXM1 and CXM1-105, have also been studied for economically valuable properties, such as productivity (symbiosis efficiency), nitrogen-fixing (acetylene-reductase) activity, host specificity, competitiveness in model and plot experiments [13, 14, 28]. According to the results of independent micro vegetation and vegetation experiments performed in different years, it should be concluded that the strains CXM1 and CXM1-105 stably retain high symbiotic activity. The average increase in dry weight in inoculated plants was 90.8-100.2% ($P < 0.0015$) with respect to the control without inoculation and 28.6-32% ($P < 0.0015$) compared to the strain Rm1021. Differences between the strains under study in terms of the ability to form a symbiosis with the model diploid species of alfalfa – *M. truncatula* of the Jemalong variety in model micro vegetation experiments was identified. The strain CXM1-105 formed effective pink nodules, and the increase in the dry weight of plants with respect to control without inoculation was 31.5% higher compared to similar plants inoculated with the strain Rm1021. The fact that Rm1021 does not form an effective symbiosis with *M. truncatula* A17 was shown by Australian researchers [29]. At the same time, in the conditions of poor salinity (0.3% NaCl), significant differences ($P < 0.05$) between the strains under study in terms of symbiotic efficiency were not found with either *M. truncatula* or *M. varia* (V.S. Muntyan, personal message). Therefore, the strains obtained on the basis of 425a (CXM1, CXM1-105) are used as reference

ones: they have higher indicators of symbiotic activity compared with the strain Rm1021 and stably exhibit them under standard (typical) conditions.

Methods of molecular genetic analysis developed on the basis of reference strains *S. meliloti*. Using the reference strains listed above, various models have been developed or adapted to study the formation and functioning of symbiotic systems and for their molecular genetic analysis. The methods of UV and chemical mutagenesis, protoplast fusion, transduction, conjugation, and transformation were developed on the basis of derivatives of the strain 425a [30, 31]. Using the species-specific region of the ISRm2011-2 chromosome of the strain Rm2011, a system for typing genomes (fingerprinting) of natural strains of nodule bacteria of alfalfa has been developed [32]. In the genome Rm2011 (Rm1021) there are 12 copies of ISRm2011-2, in the genome of the strain CXM1 — 11 copies [33]. A significant step in the study of symbiotically important genes was the development of systems of general en masse and directed Tn5-Mob and mini-Tn5 mutagenesis. A collection of 12,000 mini-Tn5 mutants was obtained on the basis of the reference strain Rm2011, an analysis of more than 9,000 of which showed that insertions of mini-Tn5 into genes encoding peptides occurred at a frequency of 0.6 [34, 35].

More than 20 genes involved in the control of the synthesis of poly- and lipopolysaccharides, symbiotic efficiency, competitiveness, salt tolerance, and acid resistance have been sequenced, and their phenotypic manifestation in the derivatives of the CXM1 strain has been studied [36–41]. The genomes of the strains Rm1021 and Rm2011 were sequenced in 2001 and 2013 respectively and are presented in databases (<https://iant.toulouse.inra.fr/S.meliloti>, <https://iant.toulouse.inra.fr/S.meliloti2011>). Further, this allowed proposing methods of modern genomic analysis using phenotypic, DNA and expression biochips [42–43] to simultaneously study the work of more than 14,000 genes and intergenic regions, and compare their expression and phenotypic manifestations [35, 44].

Comparative genomic analysis of reference strains. The genomes of the considered reference strains are similar; they have three replicons (a chromosome and two megaplasmids) and do not contain additional cryptic plasmids. According to the data of the genome-wide sequencing, in Rm1021 the chromosome size (SMc) is 3.5 million bp, the sizes of the megaplasmids SMA and SMB are 1.35 and 1.68 million bp, respectively, which is accepted as a typical characteristic of the genome of nodule bacteria of the species *Sinorhizobium meliloti*.

The replicon SMB is considered as a small chromosome Rm1021, since it has an average composition of GC pairs of 62.4%, which is 0.3% lower than that of the chromosome. 1,570 genes are located on this replicon (open reading frame, ORF), which are functionally related to 20 different COG-groups (clusters of orthologous groups) [45]. However, mainly these are the genes responsible for the carbohydrate metabolism and synthesis of polysaccharides necessary for successful microbial-plant interaction and for the saprophytic existence of bacteria in the soil or in the rhizosphere. The second mega-replicon Rm1021-SMA contains 1,293 genes (ORF), and its nucleotide composition contains 60.4% of GC pairs on average, which is significantly lower than in SMc and SMB. In this replicon, there is a zone of 90 thousand bp including clusters of *nod*, *nif*, *fix* genes that determine the process of formation and functioning of nitrogen-fixing symbiosis with the host plant.

Comparative analysis of the genomes of the reference strains CXM1-105 and Rm1021 using the DNA biochips SM6kOligo showed that 242 ORFs, localized on SMA, and 121 ORFs, localized on SMB, were changed or absent in CXM1-105 [46]. This indicates significant differences in the structural organiza-

tion of mega-replicons in the reference strains CXM1-105 and Rm1021.

We identified 69 genes belonging to five different groups that are involved in the control of such features of symbiosis formation as virulence (*nod*), nodule formation and its specificity (*nol*, *noe*), and the process of nitrogen fixation (*fix* and *nif*), and analyzed them using the technique of DNA-biochips [46]. As a result, it was found that in both model strains 65 ORFs are similar, including *nodD1*, *nodABC*, *nodEFGH*, which also had similar PCR-RFLP types (the analysis of polymorphism of the length of restriction DNA fragments) [46]. The remaining four ORFs were not detected in CXM1-105 [46], which meant the absence or significant alteration of these sequences (divergent sequences). Three of them are localized on SMA and belong to the fix-genes group, and the fourth is localized on SMb and presumably encodes the acyltransferase (EC 2.3.1.-), that belongs to the proteins gene family *cysElacA/lpxA/nodL*.

The Rm1021 chromosome contains 3,341 open reading frames, whose products are involved in the functioning of information systems – replication, transcription, translation, and are also responsible for the key pathways of metabolism and the formation of cellular structures [45]. Therefore, genes located on the chromosome are often called core.

Comparative analysis of the chromosome structure in the strains *Sinorhizobium meliloti* Rm1021 (A, B) and CXM1-105 (C): A — the content (%) of GC pairs in Rm1021 chromosome determined by the moving average method (10,000 bp window, 5,000 bp step, Unipro UGENE program); B and C — graphics images of the chromosome structure in Rm1021 and CXM1-105; *oriC* — the point of the origin of autonomous replication of the chromosome; I, II, III — the genomic islands Sme21T, Sme19T, Sme80S, respectively; ■ — the projection of the regions of localization of genomic islands on the graphic image of chromosomes Rm1021 and CXM1-105; □ — sequences that are homologous in Rm1021 and CXM1-105, ▤ — sequences, that are not detectable in the CXM1-105 genome; ▥ — *rrn*-operons; ▦ — loci *betABCI* and SMc04407-SMc04881 (for details, see the text of the article).

We analyzed 5 loci in chromosomes Rm1021 and CXM1-105, which are located in remote regions of the chromosome and can give an idea of the structure of a vital replicon. This is the intergenic sequence SMc04407-04881 with a length of 1,280 bp between the genes SMc04407 and SMc04881, which is located to the left of the origin of replication of *oriC* (Fig.). The second sequence is the locus, which includes the part of the *bet* operon located at a distance of 1,039 million bp to the right of the *oriC* (see Fig.). This sequence of 1,544 bp included 1,400 bp of the *betC* gene, an intergenic region of 1 bp and 143 bp of the *betB* gene, which are involved in the synthesis of osmoprotectant glycine-betaine. PCR-RFLP analysis did not identify differences in loci SMc04407-04881 and *betCB* in the referent strains under consideration [47]. Three other regions correspond to taxonomically significant intergenic sequences of ribosomal operons (*rrn*, see Fig.). According to the data of PCR-RFLP analysis, all three intergenic sequences of *rrn* operons in the Rm1021 strain, 1,307 bp long each, are of the a-type [48], and similar regions of CXM1-105 are b-type, which

may indicate that the genomes of the strains under consideration belong to phylogenetically remote divergent clonal lines [48, 49].

The regions in which the content of GC pairs is 6-8% lower than the average content (62.73 %) are of particular interest in the structure of the chromosome Rm1021. Such regions have less structural rigidity [50] and are evolutionarily younger in comparison with the core part of the chromosome. In the chromosome Rm1021 (Rm2011) there are three such sequences (Sme80S, Sme21T and Sme19T), which are considered as so-called genomic islands (see Fig.). These structures are atypical extended mobile elements that contain sequences of phage origin, IS elements and functionally significant genes, as well as more than seven dozen sequences from which non-coding RNAs are transcribed. The islands can actively participate in the horizontal transfer of genes [51]. It is necessary to consider in detail the structural arrangement of the islands in Rm1021 relative to the origin of replication of the chromosome *oriC* (see Fig.) The two islands (Sme21T and Sme19T) are located adjacent to the right of the *oriC* and have a similar length (20.7 and 18.6 thousand bp). Sme21T is located at a distance of 541 thousand bp from the *oriC*, Sme19T — at a distance of 98.6 thousand bp on the right of it (see Fig.). The third island (Sme80S, length 80.2 thousand bp) is located to the left of the *oriC* and is 216.4 thousand bp away. Such an arrangement of islands in regions close to *oriC*, as well as to ribosomal operons, indicates that these sequences can have important functional significance, since they are replicated in the first place. Structures analogous to genomic islands were not detected by us in the chromosomes of the strains 425a, CXM1 and CXM1-105. In addition, according to data obtained using DNA-biochips, 145 protein-coding ORFs [46] are absent in the structure of the chromosome CXM1-105. The figure above illustrates the differences between the chromosome structure in the reference strains Rm1021 and CXM1-105.

All three genomic islands, like the so-called pathogenicity islands, which are most well studied, have site-specific embedding. The places of integration of the Sme21T and Sme19T islands are direct nucleotide repeats located at the 3'-ends of the two isoacceptor tRNA of threonine (tRNA-Thr), and in the case of Sme80S at the 3'-end of the tRNA of serine (tRNA-Ser) [52]. Based on the complex in silico analysis of the site-specific integration of genomic islands in *S. meliloti* Rm1021, a system for their in vivo detection by PCR and original pairs of primers was developed [52]. According to the technical parameters of the PCR method, the PCR products can be synthesized with primers for amplification of the outer border regions of the islands (regions of integration) only in the absence of islands, as, for example, in the strain CXM1-105 [52]. Sequences obtained as a result of PCR and corresponding to the border regions in the chromosome CXM1-105 are sequenced and deposited in GenBank (EU196757, EU196758 and EU196759). A comparative analysis of these sequences in CXM1-105 and the integration regions of genomic islands in Rm1021 made it possible to study their structure and to estimate the degree of homology. It has been found that the sequence EU196757 (1,230 bp) contains a site for the specific integration of the Sme21T island with a length of 16 bp (direct repetition), to the left and right of which there are sequences of 880 and 334 bp, respectively, which are 99% homologous in Rm1021 and CXM1-105. The sequence EU196758 (561 bp) has one direct repeat of 31 bp in length — the site for the specific integration of the island Sme19T, and to left and right of it there are sequences of 53 and 530 bp with 100 % homology in the reference strains under consideration. The site for the specific integration of the third island Sme80S — a direct repeat of 15 bp in length is identified in the sequence EU196759 with a

length of 826 bp. To the left and right of it there are sequences of 397 and 417 bp respectively, which are 100 % homologous in Rm1021 and CXM1-105. On the left in the region between the SMc03748 gene and direct repeat, the Rm1021 also has an IS element of TRm11, which is not present in CXM1-105. An analysis of the border regions of the islands shows that in the CXM1-105 genome, sites for specific integration/insertion of existing genomic islands are preserved. In addition, it is possible to assume that the absence of islands can contribute to a greater structural stability of the chromosome in CXM1-105, as well as in CXM1 and 425a; however, in order to confirm this, appropriate studies are required.

The island detection system was used to identify structural types of chromosomes similar to those of Rm1021 (SMc_{Rm1021}, have islands) or CXM1-105 (SMc_{CXM1-105}, do not have islands) in four geographically remote regions. It was established that the natural strains of *S. meliloti*, isolated in the southern region of Uzbekistan, which is part of the Central Asian gene center, had the chromosome type SMc_{Rm1021}, which has a frequency of 0.72. Strains from the northern region of the Caucasus, adjacent to the Front Asian gene center, which played a leading role in the formation of cultural diploid alfalfa, as well as from the modern center of introgressive hybridization of alfalfa [21] located in the foothills of the Mu-dozhary in the north of Kazakhstan, also predominantly had the chromosome type SMc_{Rm1021}. The frequency of these types was similar, but was lower (0.54) than in the Central Asian center [52]. Only in the territory located at a distance of 500 km from the modern shore of the Aral Sea, which was subjected to extreme salinization, strains isolated from the liquorice rhizosphere or from saline sands predominantly had a chromosome type SMc_{CXM1-105} with a frequency of 0.62, that is, they had no genomic islands. Differences between samples of strains from the Central Asian gene center, in which strains with the chromosome type SMc_{Rm1021} predominated, and from the Aral region where strains with the chromosome type SMc_{CXM1-105} prevailed were significant ($\chi^2 = 4.388$; $P < 0.05$). Consequently, strains that had islands were significantly more frequent in the area of the primary gene center of alfalfa, whereas under the influence of the stress factor (salinization), loss of islands occurred.

The evaluation of the functional significance of genomic islands in Rm1021 was of interest. The islands contain copies of functionally important genes which, for example, can participate in the protection of cells from foreign DNA (locus *hsdRSM*), and can also affect the resistance of cells to the specific soil conditions (involved in the synthesis of the melanin pigment, proline osmoprotectant). In addition, one of the islands contain the *fixT3* gene — a copy of the *fixT* gene [52], which, as previously shown [53], is involved in the control of nitrogen metabolism, is influenced by the two-component global regulation system FixJL and is induced at a low oxygen content, as well as in nodules. However, the functional role of a significant portion of ORFs is not known. We conducted the search for sequences similar (homologous) to those in the strain Rm1021, in the closely related genomes, as well as in the taxonomically distant species. It was found that there are 62 ORFs in the Rm1021 islands, the sequences of which are similar or homologous to those in bacteria — representatives of 22 genera from 4 fillets, as well as from uncultivated bacteria. In most cases (33%), the islands contained OFRs, whose homologues were detected in the α -proteobacteria. They were functionally related to the storage of information (K, L; 21 %), metabolism (E, P, G, M; 33 %) or belonged to the group of poorly characterized (R, S, 17 %). ORFs that are homologous or similar to those in taxonomically distant representatives of the β -, γ - and δ -proteobacteria, were predominantly belonged to the group involved in cellular pro-

cesses and to the signaling group (O, T, V; 17 %). Differences in ORF distribution between the indicated bacterial taxa were significant ($\chi^2 = 11.02$, $P = 0.01$). Thus, the Rm1021 gene has an “additional genome” that includes functionally significant ORFs that are similar or homologous to those in representatives of the predominantly the same species or genera, as well as in phylogenetically remote classes of bacteria. Such diversity of ORFs is the result of high activity of horizontal gene transfer, which could take place both in the soil microbiome and in planta in the nodule. In the latter, according to a recent publication [54], different representatives of phylogenetically remote groups of bacteria can be present simultaneously. Indirect evidence of the possibility of horizontal transfer of genetic determinants in the nodule can be the fact, discovered by us, of the presence of a high phylogenetic diversity of OTF in the genomic islands of the strain Rm1021 – typical representative of saprophytic bacteria that form a non-obligate mutualistic symbiosis with leguminous host plants. Currently, it is not possible to conclude that the genomes of the reference strains Rm1021 (Rm2011) and CXM1-105 (CXM1) differ in terms of the presence of an “additional foreign genome”, since “foreign” genes can be fixed in the core part of the chromosome CXM1-105; however, such a statement can be confirmed with full genome sequencing.

Thus, summing up the results of the comparative analysis of chromosome structures in the reference strains obtained on the basis of the production strains 425a and SU47, it should be concluded that these strains can relate to the evolutionarily diverged lines of nodule bacteria that were distributed in the Central Asian center of the origin of cultivated plants. Based on the set of data presented, the derivatives of the strain 425a-CXM1 and CXM1-105 should be considered as model high-performance strains of the *Sinorhizobium meliloti* genera, which are genetically different from Rm1021/Rm2011, which can be used to develop methods for obtaining new genetically stable strains of nodule bacteria that promote the formation of highly productive and stress-resistant plant-microbial symbiosis, and to create an “umbrella” system for genetic monitoring of the stability of economically valuable strains in microbiomes of the agroecocenosis.

REFERENCES

1. Mishustin E.N., Shil'nikova V.K. *Kluben'kovye bakterii i inokulyatsionnyi protsess* [Root-nodule bacteria and inoculation]. Moscow, 1973 (in Russ.).
2. Bianco C., Defez R. Soil bacteria support and protect plants against abiotic stresses. In: *Abiotic stress in plants — mechanisms and adaptations*. A. Shanker, B. Venkateswarlu (eds.), InTech, Croatia, 2011 (doi: 10.5772/23310).
3. Tikhonovich I.A., Andronov E.E., Borisov A.Yu. Dolgikh E.A., Zher-nakov A.I., Zhukov V.A., Provorov N.A., Rumyantseva M.L., Simarov B.V. *Genetika*, 2015, 51(9): 973-990 (doi: 10.7868/S001667581509012X) (in Russ.).
4. Khotyanovich A.V. *Metody kul'tivirovaniya azotfiksiruyushchikh bakterii, sposoby polu-cheniya i primeneniya preparatov na ikh osnove* [Nitrogen fixing bacteria: culture, biological production and use]. Leningrad, 1991 (in Russ.).
5. *Mikrobiologiya* /Pod redaktsiei E.N. Mishustina, V.T. Emtseva [Microbiology. E.N. Mishustin, V.T. Emtsev (eds.)]. Moscow, 1987 (in Russ.).
6. Sadanov A.K., Gavrilova N.N., Dadonova T.N., Ratnikova I.A. *Izvestiya Natsional'noi akademii nauk Respubliki Kazakhstan. Seriya biologicheskaya i meditsinskaya*, 2015, 1: 115-124 (in Russ.).
7. Lopez-Guerrero M.G., Ormeno-Orrillo E., Acosta J.L., Mendoza-Vargas A., Rogel M.A., Ramirez M.A., Rosenblueth M., Martinez-Romero J., Martinez-Romero E. Rhizobial extrachromosomal replicon variability, stability and expression in natural niches. *Plasmid*, 2012, 68(3): 149-158 (doi: 10.1016/j.plasmid.2012.07.002).
8. Kajić S., Hulak N., Sikora S. Environmental stress response and adaptation mechanisms in *Rhizobia*. *Agriculturae Conspectus Scientificus*, 2016, 81(1): 15-19.

9. Galardini M., Mengoni A., Brilli M., Pini F., Fioravanti A., Lucas S., Lapidus A., Cheng J.F., Goodwin L., Pitluck S., Land M., Hauser L., Woyke T., Mikhailova N., Ivanova N., Daligault H., Bruce D., Detter C., Tapia R., Han C., Teshima H., Mocali S., Bazzicalupo M., Biondi E.G. Exploring the symbiotic pangenome of the nitrogen-fixing bacterium *Sinorhizobium meliloti*. *BMC Genomics*, 2011, 12: 235 (doi: 10.1186/1471-2164-12-235).
10. Schneiker-Bekel S., Wibberg D., Bekel T., Blom J., Linke B., Neuwe-gger H., Stiens M., Vorhulter F.J., Weidner S., Goesmann A., Pühler A., Schlüter A. The complete genome sequence of the dominant *Sinorhizobium meliloti* field isolate SM11 extends the *S. meliloti* pan-genome. *Journal of Biotechnology*, 2011, 155(1): 20-33 (doi: 10.1016/j.jbiotec.2010.12.018).
11. Ivanov L.A. *Zemledelie*, 2014, 3: 25-29 (in Russ.).
12. Penttinen P., Terefework Z., Lindstroem K., Greco D., Auvinen P., Muntyan V., Roumiantseva M., De Lajudie P., Becker A. Divergent genes in potential inoculant *Sinorhizobium* strains are related to DNA replication, recombination, and repair. *Journal of Basic Microbiology*, 2016, 56(6): 680-685 (doi: 10.1002/jobm.201500592).
13. Zaretskaya A.N. *Mikrobiologiya*, 1976, 45(5): 873-877 (in Russ.).
14. Fedorov S.N. *Poluchenie mutantov kluben'kovykh bakterii lyutsemy s izmenennymi simbiotiches-kimi svoystvami pod deystviem UF-izlucheniya. Avtoreferat kandidatskoi dissertatsii* [UV-induced mutants of alfalfa-nodulating bacteria with specific properties/ PhD Thesis]. Leningrad, 1987 (in Russ.).
15. Geddes B.A., Oresnik I.J. The mechanism of symbiotic nitrogen fixation. In: *The mechanistic benefits of microbial symbionts*. C.J. Hurst (ed.). Springer International Publishing, Switzerland, 2016 (doi: 10.1007/978-3-319-28068-4_4).
16. Bolton J.L. *Alfalfa: botany, cultivation and utilization*. Interscience Publishers, NY, 1962.
17. Vavilov N.I. *Trudy po prikladnoi botanike, genetike i seleksii*, 1926, 16(2): 3-248 (in Russ.).
18. Sinskaya E.N. *Istoricheskaya geografiya kul'turnoi flory* [Historical geography of cultivated flora]. Leningrad, 1969 (in Russ.).
19. Dzyubenko N.I., Dzyubenko E.A. V sbornike: *Problemy evolyutsii i sistematiki kul'turnykh rastenii* [Evolution and systematics of cultivated plants]. St. Petersburg, 2014: 19-20 (in Russ.).
20. Zakharov-Gezekhus I.A. *Nikolai Ivanovich Vavilov: v kontekste epokhi* [Nikolai Vavilov: the epochal context]. Moscow, 2012 (in Russ.).
21. Ivanov A.I. *Lyutserna* [Alfalfa]. Moscow, 1980 (in Russ.).
22. Eardly B.D., Materon L.A., Smith N.H., Johnson D.A., Rumbaugh M.D., Selander R.K. Genetic structure of natural populations of the nitrogen-fixing bacterium *Rhizobium meliloti*. *Appl. Environ. Microbiol.*, 1990, 56(1): 187-194.
23. MacKinnon P.A., Robertson J.G., Scott D.J., Hale C.N. Legume inoculant usage in New Zealand. *Journal of Experimental Agriculture*, 1977, 5(1): 35-39 (doi: 10.1080/03015521.1977.10425930).
24. Smith R.S., Thorne S., Randall R. Rhizobia inoculants for various leguminous species. *Nitragin Brand Inoculants Technical Bulletin*, 1988, Iss. 102: 8.
25. Keller M., Roxlau A., Weng W.M., Schmidt M., Quandt J., Niehaus K., Jording D., Arnold W., Puehler A. Molecular analysis of the *Rhizobium meliloti mucR* gene regulating the biosynthesis of the exopolysaccharides succinoglycan and galactoglucan. *Mol. Plant-Microbe Interact.*, 1995, 8(2): 267-77 (doi: 10.1094/MPMI-8-0267).
26. Bradic M., Sikora S., Redzepovic S., Stafa Z. Genetic identification and symbiotic efficiency of an indigenous *Sinorhizobium meliloti* field population. *Food Technology and Biotechnology*, 2003, 41(1): 69-75.
27. Troyan T.N. *Formirovanie effektivnogo bobovo-rizobial'nogo simbioza i ego rol'*. *Avtoreferat kandidatskoi dissertatsii* [Effective rhizobium-legume symbiosis: formation and the role. PhD Thesis]. Kaliningrad, 2010 (in Russ.).
28. Provorov N.A. *Sel'skokhozyaistvennaya biologiya* [Agricultural Biology], 1985, 20(3): 34-46 (in Russ.).
29. Terpolilli J.J., O'Hara G.W., Tiwari R.P., Dilworth M.J., Howieson J.G. The model legume *Medicago truncatula* A17 is poorly matched for N₂ fixation with the sequenced microsymbiont *Sinorhizobium meliloti* 1021. *New Phytologist*, 2008, 179(1): 62-66 (doi: 10.1111/j.1469-8137.2008.02464.x).
30. Sharypova L.A., Pretorius-Guth I.-M., Simarov B.V., Puehler A. Genetic improvement of *Rhizobium* strains. In: *The nitrogen fixation and its research in China*. G.F. Hong (ed.). Springer-Verlag, Berlin, 1992: 266-285 (doi: 10.1007/978-3-662-10385-2_13).
31. Simarov B.V., Aronshtam A.A., Novikova N.I., Bazhenova O.V., Provorov N.A., Sharypova L.A. *Geneticheskie osnovy seleksii kluben'kovykh bakterii* /Pod redaktsiei B.V. Simarova, A.A. Aronshtama, N.I. Novikovoi. L., 1990 (in Russ.).
32. Selbitschka W., Arnold W., Jording D., Kosier B., Toro N., Puehler A. The insertion sequence element ISRm2011-2 belongs to the IS630-TcI family of transposable

- elements and is abundant in *Rhizobium meliloti*. *Gene*, 1995, 163(1): 59-64 (doi: 10.1016/0378-1119(95)00371-C).
33. Roumiantseva M.L., Andronov E.E., Sharypova L.A., Dammann-Kalinowski T., Keller M., Young J.P., Simarov B.V. Diversity of *Sinorhizobium meliloti* from the Central Asian alfalfa gene center. *Appl. Environ. Microbiol.*, 2002, 68(9): 4694-4697. (doi: 10.1128/AEM.68.9.4694-4697.2002).
 34. Pobigaylo N., Wetter D., Szymczak S., Schiller U., Kurtz S., Meyer F., Nattkemper T.W., Becker A. Construction of a large signature-tagged mini-Tn5 transposon library and its application to mutagenesis of *Sinorhizobium meliloti*. *Appl. Environ. Microbiol.*, 2006, 72(6): 4329-4337 (doi: 10.1128/aem.03072-05).
 35. Serrania J., Johner T., Rupp O., Goesmann A., Becker A. Massive parallel insertion site sequencing of an arrayed *Sinorhizobium meliloti* signature-tagged mini-Tn5 transposon mutant library. *J. Biotechnol.*, 2017, 257: 9-12 (doi: 10.1016/j.jbiotec.2017.02.019).
 36. Rumyantseva M.L., Simarov B.V., Onishchuk O.P., Andronov E.E., Chizhevskaya E.P., Belova V.S., Kurchak O.N., Muntyan A.N., Rumyantseva T.B., Zatoevskaya T.V. *Biologicheskoe raznoobrazie kluben'kovykh bakterii v ekosistemakh i agrosenozakh. Teoreticheskie osnovy i metody* /Pod redaktsiei M.L. Rumyantsevoi, B.V. Simarova [Biodiversity of nodulating bacteria in ecosystems and agroecosystems. Theoretical aspects and methodology. M.L. Rumyantseva, B.V. Simarov (eds.)]. St. Petersburg, 2011 (in Russ.).
 37. Onishchuk O.P., Sharypova L.A., Kurchak O.N., Bekker A., Simarov B.V. *Genetika*, 2005, 41(12): 1617-1623 (in Russ.).
 38. Simarov B.V., Sharypova L.A., Chesnokova O.N., Onishchuk O.P., Kuchko V.V. *Genetika*, 1990, 26(4): 630-635 (in Russ.).
 39. Sharypova L.A., Yurgel S.N., Simarov B.V., Keller M., Puehler A., Becker A. The eff-482 locus of *Sinorhizobium meliloti* CXM1-105 that influences symbiotic effectiveness consists of three genes encoding an endoglycanase, a transcriptional regulator and an adenylate cyclase. *Mol. Gen. Genet.*, 1999, 261(6): 1032-1044 (doi: 10.1007/s004380051052).
 40. Chizhevskaya E.P., Krol' E.A., Onishchuk O.P., Sagulenko E.A., Fomina-Eshchenko Yu.G., Simarov B.V., Sharypova L.A. *Genetika*, 1998, 34(9): 1220-1227 (in Russ.).
 41. Rumyantseva M.L., Muntyan V.S. *Mikrobiologiya*, 2015, 84(3): 263-280 (doi: 10.7868/S0026365615030179) (in Russ.).
 42. Krol E., Becker A. Global transcriptional analysis of the phosphate starvation response in *Sinorhizobium meliloti* strains 1021 and 2011. *Mol. Gen. Genet.*, 2004, 272(1): 1-17 (doi: 10.1007/s00438-004-1030-8).
 43. Sharypova L., Andronov E., Krol E., Becker A., Saal B., Wegener C. Untersuchungen der bakteriellen Diversität mittels Oligonukleotid-Microarrays. *BioSpektrum*, 2006, 12(1): 106-108.
 44. Biondi E.G., Tatti E., Comparini D., Giuntini E., Mocali S., Giovannetti L., Bazzicalupo M., Mengoni A., Viti C. Metabolic capacity of *Sinorhizobium (Ensifer) meliloti* strains as determined by phenotype MicroArray analysis. *Appl. Environ. Microbiol.*, 2009, 75(16): 5396-404 (doi: 10.1128/aem.00196-09).
 45. Rumyantseva M.L. *Genetika*, 2009, 45(9): 1157-1172 (in Russ.).
 46. Belova-Muntyan V.S., Simarov B.V., Rumyantseva M.L. *Materialy VII Moskovskogo mezhdunarodnogo kongressa «Biotehnologiya: sostoyaniye i perspektivy razvitiya»*. Moscow, 2013, tom 2: 363-364 (in Russ.).
 47. Saksaganskaya A.S., Subbotina A.R., Cherkasova M.E., Muntyan V.S., Rumyantseva M.L. *Tezisy dokladov IX Mezhdunarodnoi nauchnoi konferentsii «Mikrobynye biotehnologii: fundamental'nye i prikladnye aspekty»* [Proc. IX Int. Conf. «Microbial biotechnologies: fundamental aspects and practice». Minsk, 2015: 48-49 (in Russ.).
 48. Rumyantseva M.L., Muntyan V.S., Mengoni A., Simarov B.V. *Genetika*, 2014, 50(4): 400-412 (doi: 10.7868/S0016675814040109) (in Russ.).
 49. Langille M.G., Hsiao W.W., Brinkman F.S. Detecting genomic islands using bioinformatics approaches. *Nature Rev. Microbiol.*, 2010, 8(5): 373-382 (doi: 10.1038/nrmicro2350).
 50. Hudson C.M., Lau B.Y., Williams K.P. Islander: a database of precisely mapped genomic islands in tRNA and tmRNA genes. *Nucl. Acids Res.*, 2015, 43: 48-53 (doi: 10.1093/nar/gku1072).
 51. Rumyantseva M.L. *Tezisy dokladov Vserossiiskoi konferentsii «50 let VOGiS: uspekhi i perspektivy»*. Moscow, 2016: 52 (in Russ.).
 52. Muntyan V.S., Cherkasova M.E., Andronov E.E., Simarov B.V., Rumyantseva M.L. *Genetika*, 2016, 52(8): 1126-1133 (doi: 10.7868/S0016675816080105) (in Russ.).
 53. Ferrieres L., Francez-Charlot A., Gouzy J., Rouille S., Kahn D. FixJ-regulated genes evolved through promoter duplication in *Sinorhizobium meliloti*. *Microbiology*, 2004, 150(7): 2335-2345 (doi: 10.1099/mic.0.27081-0).
 54. Martinez-Hidalgo P., Hirsch A.M. The nodule microbiome: N₂-fixing rhizobia do not live alone. *Phytobiomes*, 2017, 1(2): 70-82.