Current Microbiology (2025) 82:284
https://doi.org/10.1007/s00284-025-04265-3

Characterization and De Novo Genome Assembly for New *Rhizobium Ruizarguesonis* Rhizobial Strain Vst36-3 Involved in Symbiosis with *Pisum* and *Vicia* Plants

Alexandra V. Dolgikh[1] · Elizaveta A. Salnikova[1] · Alina M. Dymo[1] · Elizaveta S. Kantsurova[1] · Tatyana S. Aksenova[1] · Oleg S. Yuzikhin[1] · Oksana N. Kurchak[1] · Olga P. Onishchuk[1] · Elena A. Dolgikh[1]

Received: 2 January 2025 / Accepted: 26 April 2025
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2025

Abstract
Pea and vetch are the important legume crops used as food, forage, and green manure in agriculture. Several new rhizobial isolates were obtained from vetch *Vicia sativa* root nodules. For one of them, Vst36-3, the nodulation test showed various specificity in relation to plant hosts from the *Fabeae* tribe, such as pea and vetch. It is in contrast to typical strains of the *Rhizobium leguminosarum* species complex (*Rlc*), which formed effective nodules as in pea and vetch. Here, whole genome sequencing was performed followed by de novo genome assembly for Vst36-3 strain. As a result of de novo genome assembly, seven contigs were generated using Oxford Nanopore Technology long reads and subsequently Illumina short reads. Phylogenetic analysis allowed us to identify this strain as *Rhizobium ruizarguesonis* Vst36-3. Analysis of the Sym plasmid containing the *nod* and *nif* genes revealed that *R. ruizarguesonis* Vst36-3 has a complete suite of essential genes for the development of symbiosis. Nevertheless, this new strain forms ineffective nodules in pea. This makes *Rhizobium ruizarguesonis* Vst36-3 attractive for the search and investigation of new factors of host specificity in future.

Abbreviations

Rlc	*Rhizobium leguminosarum* Species complex
ONT	Oxford Nanopore Technology
PGAP	Prokaryotic Genome Annotation Pipeline

Introduction

Rhizobium leguminosarum is a widespread N_2-fixing legume symbiont that belongs to the class *Alphaproteobacteria.* Three symbiovars are known to be distinguished in *Rhizobium leguminosarum* differing in their host specificity—*viciae*, *trifolii*, *phaseoli*. Symbiovar *viciae* nodulates legumes such as *Pisum*, *Vicia, Lathyrus, Lens*, and *Vavilovia* of the *Fabeae* (*Vicieae*) tribe [1, 2]. Symbiovar *trifolii* participate in symbiosis with plants from the *Trifolieae* tribe (genus *Trifolium*), while symbiovar *phaseoli* includes rhizobia that nodulate plants from the *Phaseoleae* tribe [1]. However, comprehensive genomic analysis indicates that *Rhizobium leguminosarum* can be considered as a group of related species (*Rhizobium leguminosarum* species complex, *Rlc*) that includes 18 distinct genospecies or probably more [3, 4]. Among them, *R. leguminosarum* sensu stricto (belongs to genospecies E, gsE) [3], *Rhizobium laguerreae* (gsR) [5], *Rhizobium ruizarguesonis* (gs C) [6], *Rhizobium sophorae* (gsG) [7], *Rhizobium indicum* (gsI) [8], *Rhizobium changzhiense* (gsS) [9], *Rhizobium brockwellii* (gsA), *Rhizobium johnstonii* (gsB) and *Rhizobium beringeri* (gsD) [4] and others. Members of the *Rlc* contain a complex genome with a circular chromosome and several large plasmids (from one to six), but the number and size of the plasmids vary between strains. The presence of a large number of plasmids distinguishes *Rlc* from other rhizobia for which complete genome sequences have been determined. Host specificity genes are known to be carried on large plasmids that also contain genes for nitrogen fixation and metabolism that allow these bacteria to grow under different conditions. Therefore, exploring the genomes of new *Rlc* strains is important to verify their specificity and symbiotic and metabolic characteristics.

Despite the fact that a large number of *Rlc* genomes are publicly available, only a few of them are fully assembled.

Alexandra V. Dolgikh and Elizaveta A. Salnikova are contributed equally to this work.

✉ Elena A. Dolgikh
ea.dolgikh@arriam.ru; dol2helen@yahoo.com

[1] All-Russia Research Institute for Agricultural Microbiology, Podbelsky Chausse 3, 196608 St.-Petersburg, Russia

Published online: 07 May 2025

Sequencing and assembling of complete genomes and their annotation have been done for some widely used *Rlc* strains, such as *R. johnstonii* 3841 (former *R. leguminosarum* bv. *viciae* 3841, genome size is 7.75 Mb) [10] and *R. leguminosarum* bv. *viciae* UMP 791 (genome size is 7.84 Mb) [11]. Analysis has shown that these strains contained a circular chromosome and five-six different-sized plasmids. In *R. johnstonii* 3841 one of these plasmids (pRL10) carries the genes for nodulation and nitrogen fixation and was defined as a symbiotic plasmid. Similarly, in *R. leguminosarum* bv. *viciae* UMP 791 the pRlvC symbiotic plasmid carries the nodulation genes. A comprehensive analysis of other strains of *Rlc* may be important to validate the distribution and function of nodulation genes in 18 new genospecies recently identified.

Recently, new rhizobial isolates were obtained from root nodules of vetch (*Vicia sativa* L.) [12]. The nodulation test showed that one of them, Vst36-3, may demonstrate various efficiency in relation to plant hosts from the *Fabeae* tribe, such as pea and vetch. It is in contrast to typical strains from the *Rlc* complex, which formed effective nodules as in pea and vetch. This makes a new strain attractive for the search and investigation of new factors of host specificity. Here, whole genome sequencing was performed followed by de novo genome assembly for Vst36-3 strain.

Materials and Methods

Plant Material

The *Pisum sativum* cv. Cameor pea seeds (from the collection of the All-Russia Research Institute for Agricultural Microbiology, Saint Petersburg, Russia) and *Vicia villosa* cv. Lugovskaya-2 vetch seeds (from the collection of N. I. Vavilov Institute of Plant Genetic Resources, Saint-Petersburg, Russia) were applied. The seeds were placed in concentrated sulfuric acid for 5 min and washed with sterile water at least 5 times. The pea seeds were transferred to 1% water agar and incubated at 21 °C until germinated (4–5 days). Vetch seeds germinated on moist filter paper at 23 °C.

Bacterial Strains

Rhizobium johnstonii sp. *nov*. 3841 [10], *Rhizobium ruizarguesonis* RCAM 1026 https://patentscope.wipo.int/search/en/detail.jsf?docId=SU28658295 [13], and a new rhizobial isolate Vst36-3 from the Russian Collection of Agricultural Microorganisms (RCAM, Saint-Petersburg, Russia) were used. The strain Vst36-3 was originally isolated from the nodules of *Vicia sativa* L. [12] and later deposited in the RCAM.

Nodulation Tests and Nitrogen-Fixing Activity Measuring

In microvegetative experiments, plants inoculated with rhizobia were grown under gnotobiotic conditions in vermiculite. The pea seedlings were planted in 1-L glass cylinders containing vermiculite, with N-free Krasilnikov-Korenyako liquid growth medium with a set of microelements as previously described [12]. The pea seedlings were inoculated with 10 ml of a suspension of rhizobia containing approximately 10^7 cells. As a negative control, 10 ml of sterile water was added to the vessel. The vetch seedlings were placed in individual 80 ml glass tubes with vermiculite moisturized with Krasilnikov-Korenyako liquid growth medium. Inoculation was performed using 1 ml of rhizobia suspension containing approximately 10^6 cells. The plants were grown for 28–30 days in the growth chamber at 50% relative humidity, with a four-level illumination/temperature mode: night (dark, 18 °C, 8 h), morning (200 μmol m^{-2} s^{-1}, 20 °C, 2 h), day (400 μmol m^{-2} s^{-1}, 23 °C, 12 h) and evening (200 μmol m^{-2} s^{-1}, 20 °C, 2 h).). The illumination was provided by L 36 W/77 Fluora lamps (Osram, Munich, Germany). Four-six replications were carried out for each sample. The results were recorded using a Carl Zeiss Stemi 508 stereo microscope with a Zeiss Axiocam ERc 5S camera (Carl Zeiss Microscopy GmbH, Jena, Germany). Nitrogen-fixing activity was determined using the acetylene method with C_2H_2 as a substrate for nitrogenase. Roots with nodules were placed in hermetically sealed 50 ml Falcon tubes, into which 5 ml of C_2H_2 (10% volume) was introduced and incubated for 24 h. The amount of C_2H_4 per plant was determined on a gas chromatograph GC-2010 (Shimadzu, Tokyo, Japan).

DNA Purification and Quality Controls

For genome sequencing, we used a previously published protocol for DNA extraction with modifications [14]. DNA concentration was measured using the dsDNA Quantitation Broad Range Kit (Thermo Fisher Scientific, Waltham, MA, USA) and NanoDrop (Implen, Germany). DNA quality control was performed by analyzing the sample in different absorbing ratios with the NanoDrop (Implen, Germany). Furthermore, DNA was visualized on a 1% agarose gel using electrophoresis in TAE buffer.

Sequencing, De Novo Genome Assembly, and Annotation

For de novo assembly of the Vst36-3 rhizobial strain genome a combination of Oxford Nanopore Technologies long reads and Illumina short reads were used. Long-read

whole genome sequencing was performed using a MinION sequencer (Oxford Nanopore, UK) at the Center for Genome Technologies, Proteomics and Cell Biology of the ARRIAM by the Nanopore method. Library sample preparation was performed according to the manufacturer's instructions (1D native barcoding genomic DNA protocol) using the SQK-LSK109 ligation sequencing kit and the EXP-NBD104, EXP-NBD114 barcoding kits (Oxford Nanopore, UK). Libraries were sequenced according to the manufacturer's instructions on an R9.4 cell. The basecalling of fast5 raw files obtained as a result of sequencing was performed using the Albacore v. 2.3.1 program (https://rubygems.org/gems/al-bacore/versions/2.3.1). Deepbiner v. 2.3.1 was used to demultiplex the reads. 0.2.0, for cleaning—the program Porechop v. 0.2.3 (https://github.com/rrwick/Porechop). For Illumina sequencing, we used previously published protocol for DNA extraction with modifications [14]. The isolated DNA was sonicated to increase library coverage. Fragments of 200–300 base pairs were extracted from 1% agarose gel. A library was created using the NEBNext® UltraTM II DNA Library Prep Kit for Illumina (New England Biolabs, USA) and AMPure XP beads (Beckman Coulter, USA). Sequencing was performed in the 2× 150 mode on the Illumina NovaSeq 6000 sequencer.

To perform the assembly, we used the Trycycler v0.5.5 tool for the assembly of ONT reads, followed by polishing with Illumina reads [15, 16]. The assembly process involved the use of Flye v2.9.5-b1801, Raven v1.8.3, Miniasm v0.3, and Polypolish v0.6.0 for short read polishing [16–20]. For assembly annotation, the Prokaryotic Genome Annotation Pipeline (PGAP) was used [21]. The Genovi tool was utilized to visualize genome assembly [22]. For comparative analysis, we used genomes of *R. ruizarguesonis* RCAM 1026 (GCF_001927265.3) and *R. johnstonii* 3841 (GCF_000009265.1). The alignment of the whole genome and analysis of orthogroups and COGs was performed using zDB [23].

Raw genome sequencing data and assembled genome are available in the NCBI database under the BioProject accession number PRJNA1169353 (CP171201-CP171207).

Phylogenetic Analysis

Phylogram was created using TYGS server (https://tygs.dsmz.de/) based on whole genome alignment using the Genome BLAST Distance Phylogeny approach. iTOL server was used for visualization of the tree.

Results

Determination of Host Specificity of New Isolate Vst36-3

To verify the host specificity of a new rhizobial isolate obtained from root nodules of *Vicia sativa*, we performed nodulation experiments in vetch *Vicia villosa* cv. Lugovskaya 2 and pea *Pisum sativum* L. cv. Cameor (Fig. 1). It

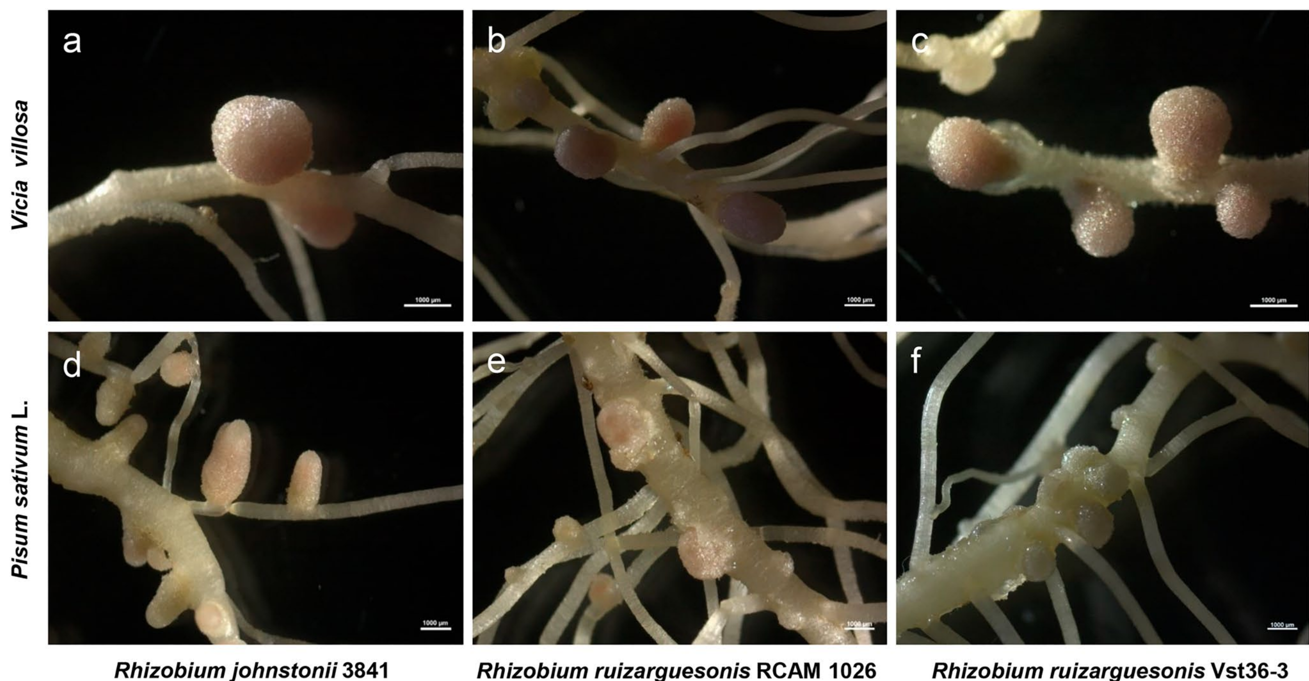


Fig. 1 Nodule formation in vetch *Vicia villosa* cv. Lugovskaya 2 (**a–c**) and pea *Pisum sativum* L. cv. Cameor (**d–f**) with different rhizobial strains

was shown that Vst36-3 strain was able to form small white nodules in pea (Fig. 1b, f), although it formed effective pink nodules in vetch (Fig. 1c). It was in contrast to typical strains *R. johnstonii* 3841 and *R. ruizarguesonis* RCAM 1026, which both form effective nodules in pea (Fig. 1a, b) and vetch (Fig. 1d, e). Subsequently, the number and nitrogen-fixing activity of pea root nodules of 28 dpi grown in N-free medium at 21 °C were determined (Figs. 2, 3).

High activity of nitrogenase in root nodules formed by the strains *R. johnstonii* 3841 and *R. ruizarguesonis* RCAM 1026 indicates the efficiency of biological nitrogen fixation, while a rhizobial isolate Vst36-3 was able to form only ineffective (Fix^-) nodules with no nitrogenase activity (Fig. 2). As a typical ineffective strain, the Vst36-3 forms an increased amount of small ineffective nodules in pea, but not in vetch (Fig. 3). These parameters reflect the efficiency of the symbiotic process that occurs within the legume root nodules. Therefore, we have shown that some strains of *Rlc* complex may demonstrate various specificities in relation to plant hosts of the *Fabeae* tribe. This makes this strain attractive for the search and investigation of new factors of host specificity based on genome sequencing and assembling.

Statistically significant differences in the number of nodules for plants inoculated with effective *Rj* 3841, Rr 1026,

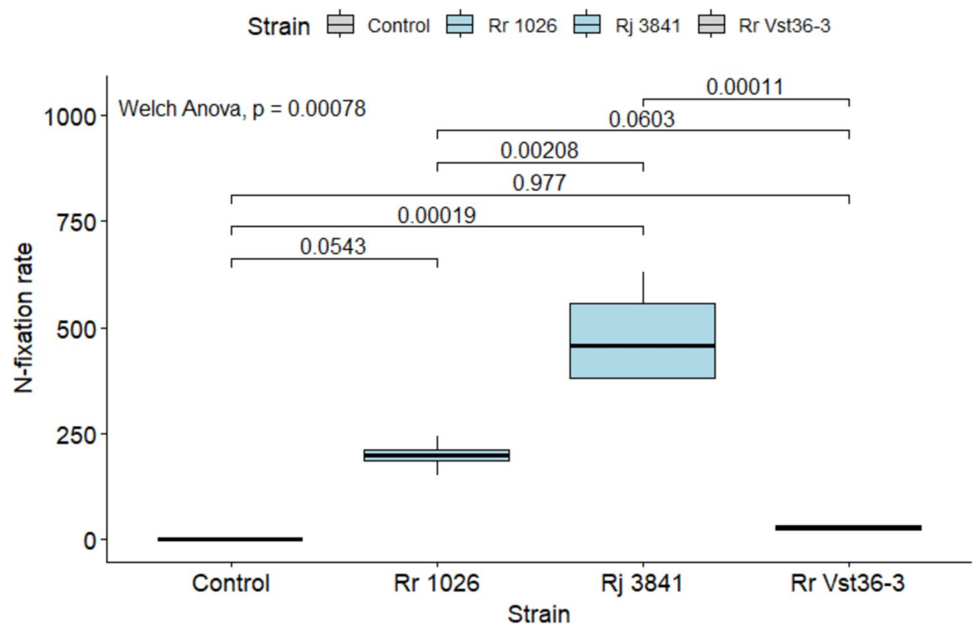
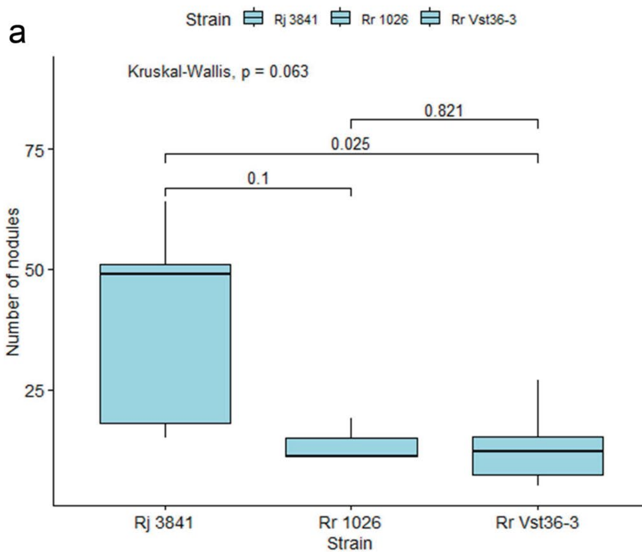


Fig. 2 Nitrogen-fixing activity of cv. Cameor pea plants inoculated with *Rhizobium ruizarguesonis* RCAM 1026 (Rr 1026), *Rhizobium johnstonii* 3841 (Rj 3841) and *Rhizobium* Vst36-3 (Rr Vst36-3) at 28 dpi (C_2H_4 µl per plant), control – noninoculated cv. Cameor pea plants. Error bars indicate standard error of the mean of four-five replicates



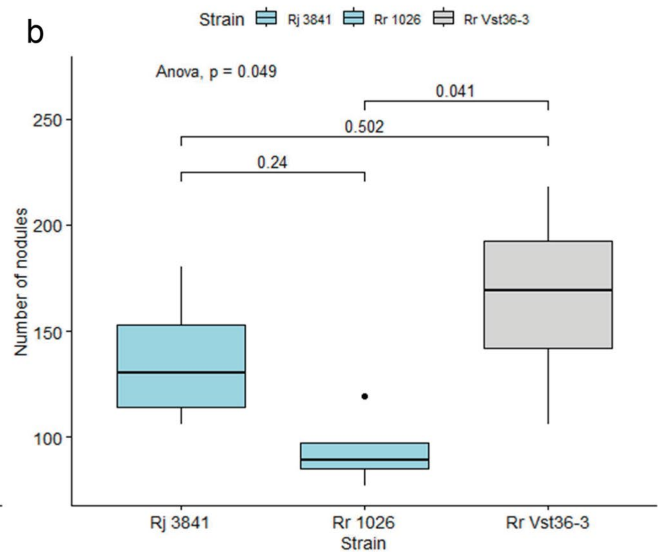


Fig. 3 Statistical analysis of nodule number in vetch (**a**) and pea (**b**)

and ineffective Rr Vst36-3 strains at 28 dpi were assessed. All strains formed effective nodules in vetch (blue boxes), but Rr Vst36-3 formed ineffective nodules in pea (gray box). (a) – Kruskal–Wallis followed by Dunn's post-hoc test was used for analysis. (b)—one-way ANOVA followed by Tukey's post hoc test was applied.

Whole Genome Sequencing of a New Rhizobial Isolate from *Vicia* Nodules and Phylogenetic Analysis

We performed de novo genome assembly for a new rhizobial isolate (Vst36-3 strain) using a combination of Oxford Nanopore Technologies long reads and Illumina short reads. As a result of genome assembly, seven contigs were generated using Oxford Nanopore Technology (ONT) read assembly and subsequently polished with Illumina reads: a circular chromosome and six different size plasmids (Fig. 4, Table S1, Fig. S1).

The summary size of the genome assembly was 8.48 Mb. For primary taxon identification, the TYGS server was used and the resulting assembly was compared with typical strains of different *Rhizobium* species [24]. As a result, by whole genome alignment the Vst36-3 strain was identified as *Rhizobium ruizarguesonis* (Fig. 5).

The taxonomy check performed during the PGAP annotation further validated the assignment of the genomes to these species with average nucleotide identity (ANI) of 98.02% compared to *R. ruizarguesonis* UPM1133 T. 8208 protein coding sequencing was found in the assembly of the Vst36-3 genome during annotation.

Comparative Genome Analysis of a New Isolate with Typical Rhizobial Strains

A comparison of the genome assembly with previously published assemblies of *R. johnstonii* 3841 and *R. ruizarguesonis* bv. *viciae* RCAM 1026 reveals that it possesses the largest genome and the highest number of predicted proteins among the three species. Analysis of the region in the Sym plasmid that contains the *nod* and *nif* genes reveals that *R. ruizarguesonis* Vst36-3 has a complete suite of essential genes for the development of symbiosis. This includes all the necessary *nod* genes involved in nodulation and *nif* genes critical for nitrogen fixation (Fig. 6).

We also performed a comparative analysis of the Clusters of Orthologous Groups of proteins (COGs) across three genomes. This analysis revealed an increase in the number of proteins associated with the mobilome in the *R. ruizarguesonis* Vst36-3 genome. However, in general, the distribution

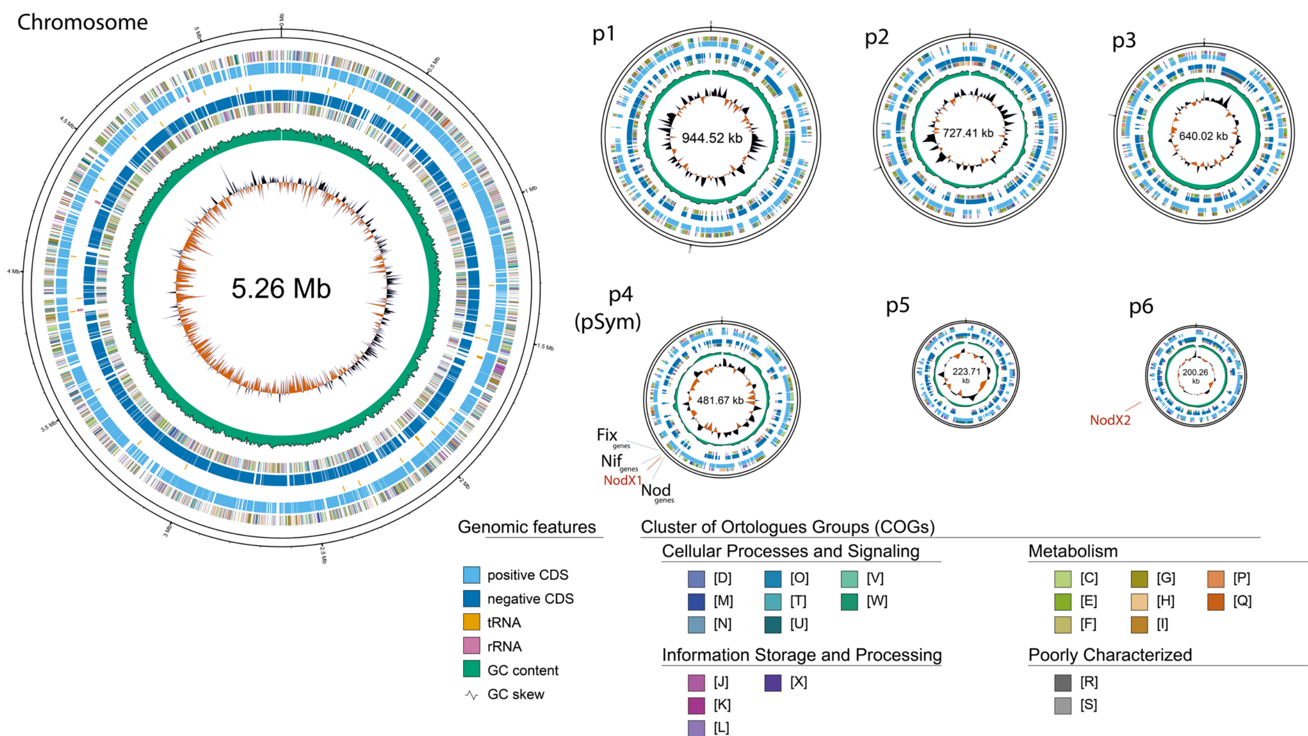


Fig. 4 Visualization of *Rhizobium ruizarguesonis* Vst36-3 genome structure. The description of COG categories represented by characters was provided in Supplementary Table S1

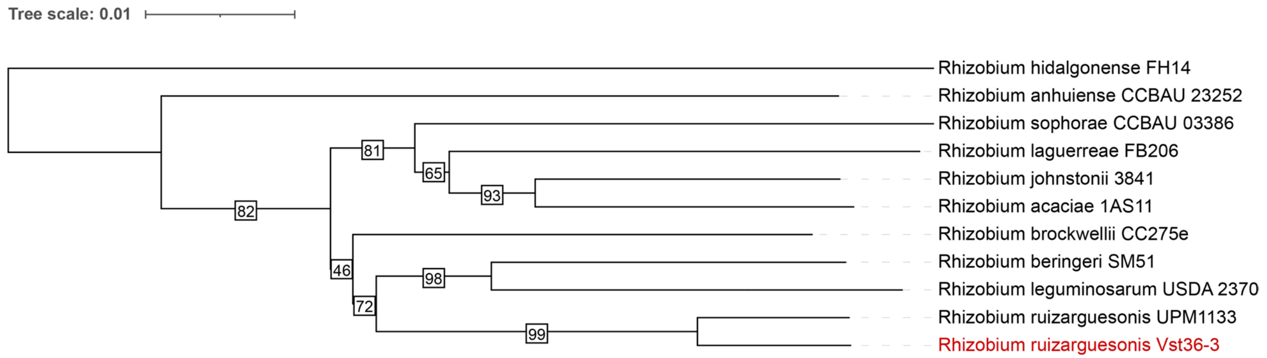


Fig. 5 Phylogenetic tree constructed using Maximum-Likelihood approach based on whole genome alignment for type rhizobial strains and *R. ruizarguesonis* Vst36-3 strain. Values indicate branch support based on 1000 bootstrap replicates

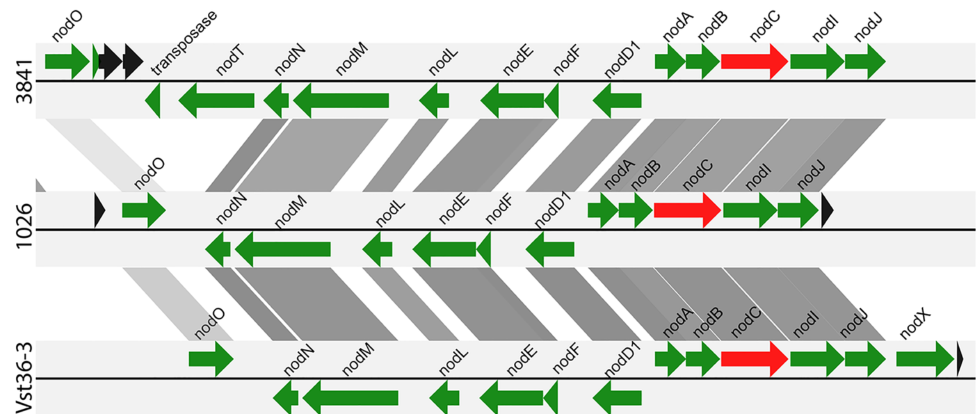


Fig. 6 The analysis of the region on the *Sym* plasmid containing *nod* genes in *R. johnstonii* 3841, *R. ruizarguesonis* RCAM 1026 and *R. ruizarguesonis* Vst36-3 strains. Black arrows indicate pseudogenes. Green arrows indicate functional protein coding genes and red arrow indicate *nodC* gene position. Gray strips connecting green arrows indicate percentage of similarity between orthologs (darker is higher)

of proteins from other COGs was relatively uniform across the genomes of 3 strains (Fig. 7).

Recently, three rhizobial strains UPM1132, UPM1133 T, and UPM1134, which belong to *R. ruizarguesonis*, have been characterized based on genome sequencing [6]. Therefore, we also performed a comparative analysis of these genomes. We found that size of *R. ruizarguesonis* Vst36-3 genome was bigger (8.48 Mbp) than genome sizes of UPM1133 T, UPM1132, and UPM1134 (7.7 Mbp, 7.5 Mbp and 7.5 Mbp, respectively). We also performed comparative analysis of proteins from different COGs between *R. ruizarguesonis* Vst36-3 and these strains (Fig. 8). Analysis has shown that the number of proteins associated with the mobilome was more widely represented in the genome of *R. ruizarguesonis* Vst36-3 compared to other representatives of *R. ruizarguesonis* (Fig. 8a, Table S2). It was related to the presence of additional regulators of the replication, recombination, and repair processes found in a new rhizobial strain. A large number of transposable elements found in the genome of Vst36-3 strain may indicate active acquisition of genetic material through horizontal gene transfer. Moreover, the function of proteins selectively found in *R. ruizarguesonis* Vst36-3 using Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis may be related to more intensive or variable metabolic transformation of amino acids, as well as energy production and conversion (Fig. 8b, Table S3). Additional regulators of quorum sensing were also found in a new rhizobial strain compared to UPM1132, UPM1133 T, and UPM1134.

R. ruizarguesonis Vst36-3 and UPM1133 T strains have the *nodX* gene, which may expand their potential host range [25]. However, in contrast to UPM1133 T, the *R. ruizarguesonis* Vst36-3 has two copies of *nodX,* located in the different plasmids that may also influence on the symbiotic specificity of this strain in pea.

Unlike other strains, the *R. rhuizarguesonis* Vst36-3 possesses type III (T3SS) and IV (T4SS) secretion systems (presented by *virB1-virB11* genes). Although the *virB1-virB11* genes were also found in *R. johnstonii* 3841 [26], our analysis has shown that these proteins demonstrate significant variability between effective *R. johnstonii* 3841 and ineffective *R. rhuizarguesonis* Vst36-3 strains (Fig. 7). It is known that the T3SS and T4SS systems in rhizobial strains may be involved in the symbiosis establishment process, affecting host specificity and the number of formed nodules [27]. Therefore, the behavior of the Vst36-3 strain in nodulation experiments may be associated with the presence of these genes.

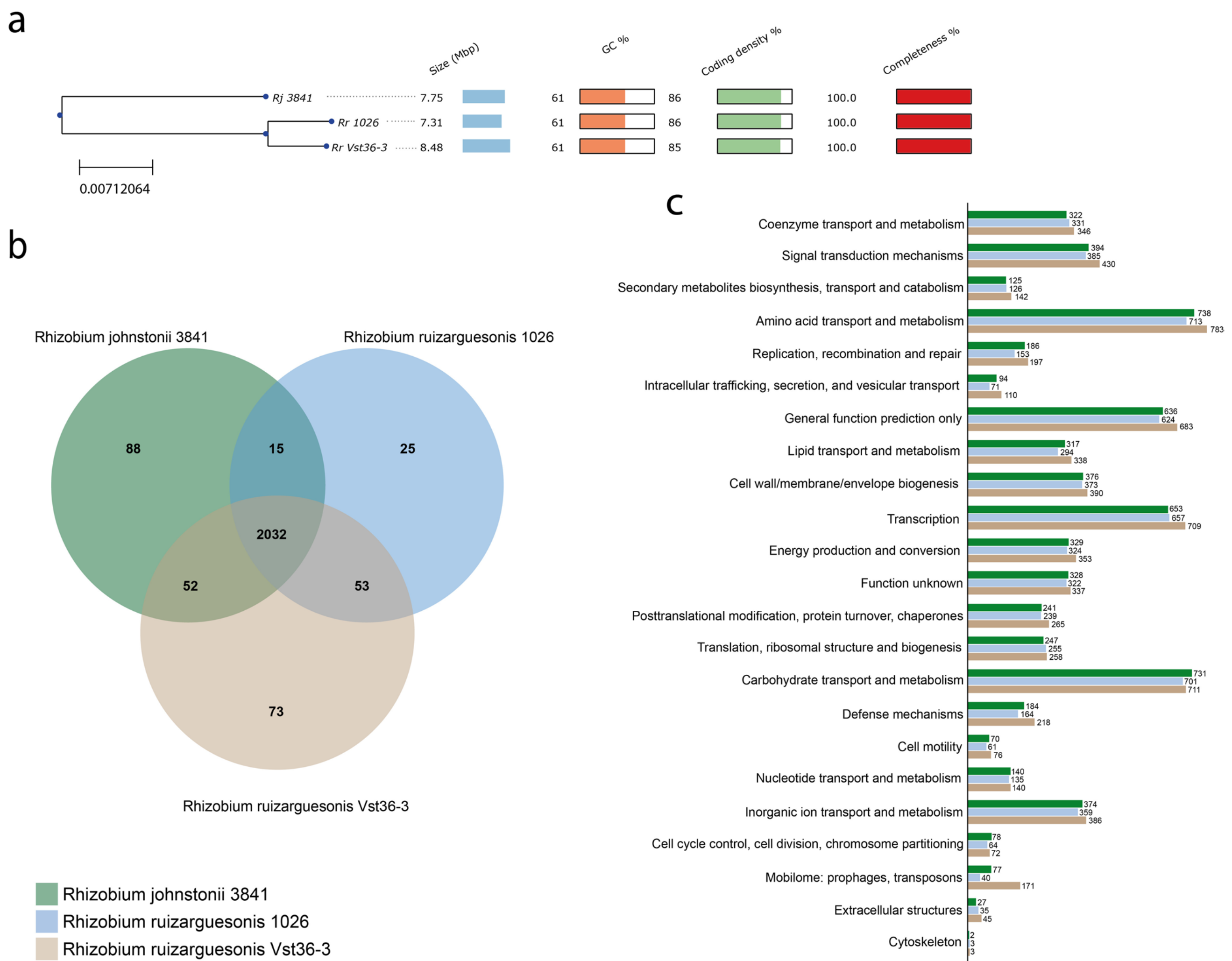


Fig. 7 Phylogenetic relationship and assembly statistics (**a**), comparison of common and specific COGs (**b**) and number of proteins from different COGs (**c**) for *R. johnstonii* 3841, *R. ruizarguesonis* RCAM 1026 and *R. ruizarguesonis* Vst36-3 strains

Discussion

Taxonomy analysis of a new rhizobial isolate Vst36-3 validated the assignment of this strain to *R. ruizarguesonis* with average nucleotide identity (ANI) 98.02% compared to typical *R. ruizarguesonis* UPM1133 T. The nodulation test in pea and vetch plants inoculated with the rhizobial strain *R. ruizarguesonis* Vst36-3 revealed significant differences between this strain and the typical rhizobial strains *R. johnstonii* 3841, *R. ruizarguesonis* RCAM 1026, all of which belong to the complex of species of *R. leguminosarum* (*Rlc*). We observed that the new rhizobial strain *R. ruizarguesonis* Vst36-3 was able to form effective nodules in vetch, while it forms ineffective nodules in pea. At the same time, according to the literature data, the majority of isolates of *R. ruizarguesonis* are equally effective in establishing symbioses with both host genera [6].

The distinctive characteristic of the *R. ruizarguesonis* species may be their ability to process hydrogen generated during nitrogen fixation, which provides the potential to establish a more effective symbiosis [6]. However, the activity of rhizobial hydrogenase in root nodules is affected by the host legume. As an example, lentil plants do not allow high symbiotic expression of hydrogenase from *R. leguminosarum* bv. *viciae* hup^{+} (Hydrogen Uptake positive, HUP^{+}) strains, while hydrogenase is induced at a high level in bacteroids of hosts such as *Vicia* and *Pisum* [28].

Furthermore, crop species were observed to interact with symbionts in a more selective way than wild legumes [29]. As an example, *P. sativum* had a more limited range in its symbionts compared to any of the wild legumes [29]. The study of microheterogenicity among common symbionts of a group of related legume species showed that the majority of strains isolated from vetch formed a nitrogen-fixing

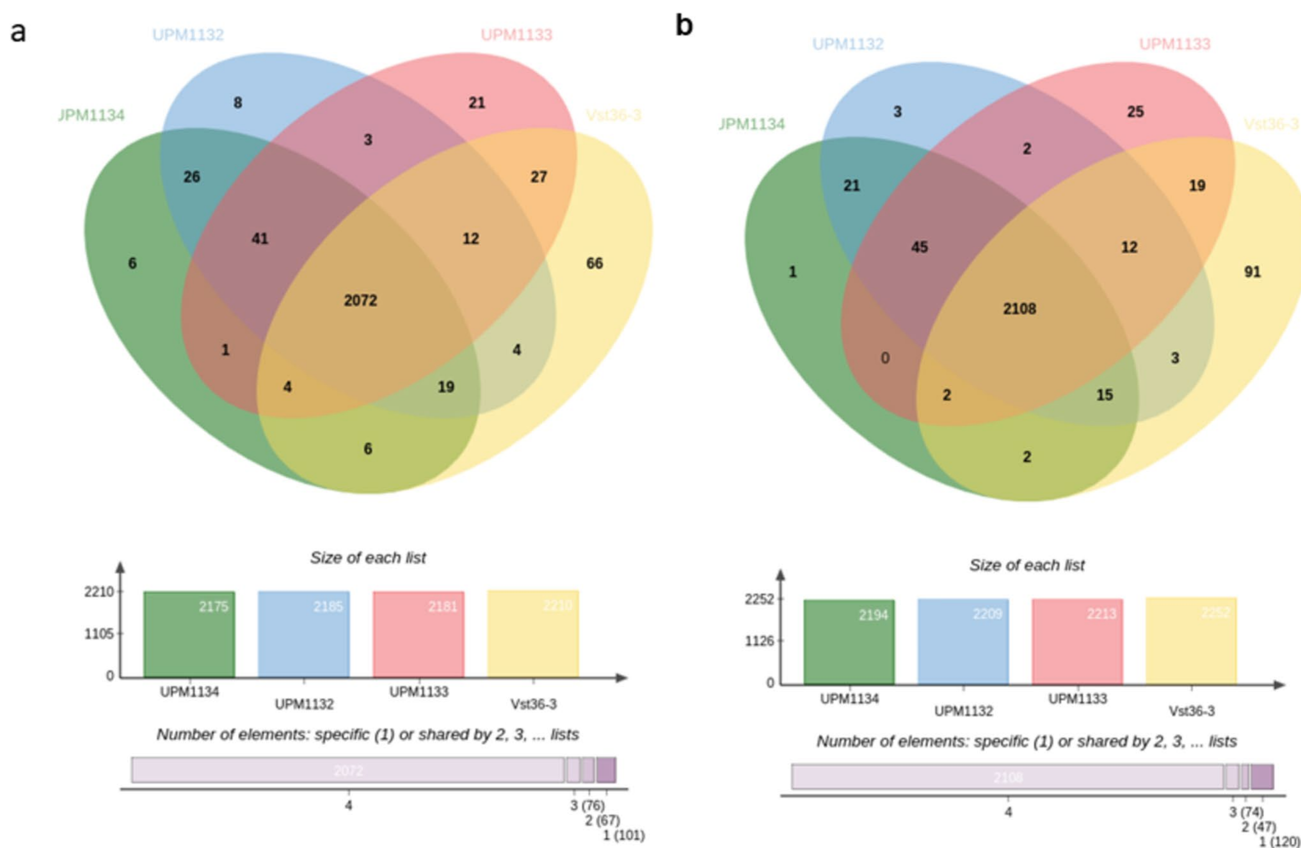


Fig. 8 Comparison of common and specific COGs (**a**) and number of proteins from different KEGGs (**b**) for *R. ruizarguesonis* Vst36-3 and *R. ruizarguesonis* UPM1132, UPM1133, and UPM1134 strains

symbiosis (Fix^+) in *P. sativum* plants with only a few cases of inefficiency (Fix^-).

The diversity of responses to different rhizobia in host plants during bacteroid formation and nitrogen fixation, indicates that there is a host specificity in these processes. However, a similar pattern of regulation of more than 100 upregulated genes has been demonstrated between pea (*P. sativum*) and vetch (*V. cracca*) bacteroids isolated from 28-day-old nodules inoculated with *R. leguminosarum* bv. *viciae*. In addition, many of these genes are not found in other rhizobia, which highlights the plasticity of the rhizobial responses to their respective hosts [30].

Therefore, despite *Rlc* rhizobial strains can establish effective symbiotic relationships with members of the genera *Pisum, Vicia, Lathyrus, Lens,* and *Vavilovia*, which belong to the Fabeae tribe, there are certain genotypes within these strains that are more suited to specific plant hosts [31]. There are many different reasons for strain inefficiency (structurally different Nod factors, immunity reactions, various sets of NCR peptides and microRNAs). Investigation of the causes of ineffective symbiosis can help identify the factors that contribute to the specificity of the interaction between rhizobia and legumes. Our analysis highlights the plasticity of the *R. ruizarguesonis* Vst36-3 responses compared to other strains. Generally, the components necessary for quorum sensing, some metabolic processes, and hydrogenase activity regulation, T3SS and T4SS systems were not found in typical *R. ruizarguesonis* UPM1133 T strain, that makes the Vst36-3 strain promising to study. This indicates the need for further comparative analysis using a wider range of *R. ruizarguesonis* genomes. Analysis of genome sequencing data indicates that *R. ruizarguesonis* Vst36-3 possesses the largest genome and the highest number of predicted proteins compared to typical rhizobial strains R. *johnstonii* 3841 and *R. ruizarguesonis* RCAM 1026. It makes it more attractive for searching and investigating new factors of host specificity.

Conclusions

Whole genome sequencing and de novo genome assembly were performed for a new rhizobial isolate from *Vicia sativa* root nodules. However, the nodulation test showed that *R. ruizarguesonis* Vst36-3 may demonstrate various specificity in relation to plant hosts from the Fabeae tribe, such as pea

and vetch. It is in contrast to typical strains of the *Rhizobium leguminosarum* species complex, which formed effective nodules as in pea as in vetch. Phylogenetic analysis allowed us to identify this strain as *Rhizobium ruizarguesonis* Vst36-3. Analysis of the Sym plasmid containing the *nod* and *nif* genes revealed that *R. ruizarguesonis* Vst36-3 has a complete suite of essential genes for the development of symbiosis. Investigating the causes of ineffective symbiosis can help identify factors that contribute to the specificity of the interaction between rhizobia and legumes.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00284-025-04265-3.

Acknowledgements The research was carried out using equipment of the Research Resource Centre "Molecular and Cell Technologies" at the Saint Petersburg State University (state assignment 125022803066-3) and the Core Centrum 'Genomic Technologies, Proteomics and Cell Biology' in ARRIAM.

Author Contributions A.V.D. wrote the manuscript, made bioinformatic analysis and genome assembly, prepared figures. E.A.S. purified DNA for genome sequencing, data analysis, prepared figures. A.M.D. made nodulation test. E.S.K. prepared the libraries. T.S.A. performed the experiments, prepared the libraries, genome sequencing. O.S. Y. measured the nitrogenase activity. O.N.K. performed the experiments with plants. O.P.O. performed the experiments with plants, data analysis. E.A.D. conceptualization, writing and editing, and supervision. All authors reviewed the manuscript.

Funding This work was supported financially by the Russian Science Foundation RSF 24-16-00180.

Data Availability The authors declare that all data supporting the findings of this study are available in this article and its Supplementary Information files.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The authors declare that they have no conflict of interest in the publication.

Ethical Approval and Consent to Participate Not applicable.

References

1. Sprent JI, Ardley J, James EK (2017) Biogeography of nodulated legumes and their nitrogen-fixing symbionts. New Phytol 215:40–56. https://doi.org/10.1111/nph.14474
2. Ellis THN, Smýkal P, Maxted N et al (2024) The taxonomic status of genera within the fabeae (Vicieae), with a special focus on pisum. Diversity 16:365. https://doi.org/10.3390/d16070365
3. Young JPW, Moeskjær S, Afonin A et al (2021) Defining the Rhizobium leguminosarum species complex. Genes (Basel) 12:111. https://doi.org/10.3390/genes12010111
4. Young JPW, Jorrin B, Moeskjær S, James EK (2023) Rhizobium brockwellii sp. nov., Rhizobium johnstonii sp. nov. and Rhizobium beringeri sp. nov., three genospecies within the Rhizobium leguminosarum species complex. Int J Syst Evol Microbiol 73:1–10. https://doi.org/10.1099/ijsem.0.005979
5. Saïdi S, Ramírez-Bahena M-H, Santillana N et al (2014) Rhizobium laguerreae sp. nov. nodulates Vicia faba on several continents. Int J Syst Evol Microbiol 64:242–247. https://doi.org/10.1099/ijs.0.052191-0
6. Jorrin B, Palacios JM, Peix Á, Imperial J (2020) Rhizobium ruizarguesonis sp. Nov., isolated from nodules of Pisum sativum L. Syst Appl Microbiol 43:126090. https://doi.org/10.1016/j.syapm.2020.126090
7. Jiao YS, Yan H, Ji ZJ et al (2015) Rhizobium sophorae sp. nov. and Rhizobium sophoriradicis sp. nov., nitrogen-fixing rhizobial symbionts of the medicinal legume Sophora flavescens. Int J Syst Evol Microbiol 65:497–503. https://doi.org/10.1099/ijs.0.068916-0
8. Rahi P, Giram P, Chaudhari D et al (2020) Rhizobium indicum sp. nov., isolated from root nodules of pea (Pisum sativum) cultivated in the Indian trans-Himalayas. Syst Appl Microbiol 43:126127. https://doi.org/10.1016/j.syapm.2020.126127
9. Jin Y, Liu H, Luo D et al (2016) DELLA proteins are common components of symbiotic rhizobial and mycorrhizal signalling pathways. Nat Commun 7:12433. https://doi.org/10.1038/ncomms12433
10. Young JPW, Crossman LC, Johnston AWB et al (2006) The genome of Rhizobium leguminosarum has recognizable core and accessory components. Genome Biol 7:R34. https://doi.org/10.1186/gb-2006-7-4-r34
11. Sánchez-Cañizares C, Jorrín B, Durán D et al (2018) Genomic diversity in the endosymbiotic bacterium Rhizobium leguminosarum. Genes (Basel) 9:60. https://doi.org/10.3390/genes9020060
12. Onischuk OP (2023) Biodiversity of the symbiotic systems formed by nodule bacteria Rhizobium leguminosarum with the leguminous plants of galegoid complex. Sel'skokhozyaistvennaya Biol 58:87–99. https://doi.org/10.15389/agrobiology.2023.1.87eng
13. Afonin A, Sulima A, Zhernakov A, Zhukov V (2017) Draft genome of the strain RCAM1026 Rhizobium leguminosarum bv. viciae. Genomics Data 11:85–86. https://doi.org/10.1016/j.gdata.2016.12.003
14. Ausubel FM, Brent R, Kingston RE et al (2004) Preface. Curr Protoc Mol Biol 66:1–1. https://doi.org/10.1002/0471142727.mbprefs66
15. Wick RR, Judd LM, Cerdeira LT et al (2021) Trycycler: consensus long-read assemblies for bacterial genomes. Genome Biol. https://doi.org/10.1186/s13059-021-02483-z
16. Wick RR, Judd LM, Holt KE (2023) Assembling the perfect bacterial genome using Oxford Nanopore and Illumina sequencing. PLoS Comput Biol. https://doi.org/10.1371/journal.pcbi.1010905
17. Kolmogorov M, Yuan J, Lin Y, Pevzner PA (2019) Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37:540–546. https://doi.org/10.1038/s41587-019-0072-8
18. Vaser R, Šikić M (2021) Time- and memory-efficient genome assembly with Raven. Nat Comput Sci 1:332–336. https://doi.org/10.1038/s43588-021-00073-4
19. Li H (2016) Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. Bioinformatics 32:2103–2110. https://doi.org/10.1093/bioinformatics/btw152
20. Wick RR, Holt KE (2022) Polypolish: short-read polishing of long-read bacterial genome assemblies. PLOS Comput Biol 18:e1009802. https://doi.org/10.1371/journal.pcbi.1009802
21. Tatusova T, Dicuccio M, Badretdin A et al (2016) NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res 44:6614–6624. https://doi.org/10.1093/nar/gkw569
22. Cumsille A, Durán RE, Rodríguez-Delherbe A et al (2023) GenoVi, an open-source automated circular genome visualizer for bacteria and archaea. PLoS Comput Biol. https://doi.org/10.1371/JOURNAL.PCBI.1010998

23. Marquis B, Pillonel T, Carrara A, Bertelli C (2024) zDB: bacterial comparative genomics made easy. mSystems. https://doi.org/10.1128/msystems.00473-24
24. Meier-Kolthoff JP, Göker M (2019) TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nat Commun. https://doi.org/10.1038/s41467-019-10210-3
25. Davis EO, Evans IJ, Johnston AWB (1988) Identification of nodX, a gene that allows Rhizobium leguminosarum biovar viciae strain TOM to nodulate Afghanistan peas. Mol Gen Genet MGG 212:531–535. https://doi.org/10.1007/BF00330860
26. Krehenbrink M, Downie JA (2008) Identification of protein secretion systems and novel secreted proteins in Rhizobium leguminosarum bv. viciae. BMC Genomics 9:55. https://doi.org/10.1186/1471-2164-9-55
27. Nelson MS, Sadowsky MJ (2015) Secretion systems and signal exchange between nitrogen-fixing rhizobia and legumes. Front Plant Sci 6:1–11. https://doi.org/10.3389/fpls.2015.00491
28. Brito B, Toffanin A, Prieto R-I et al (2008) Host-dependent expression of Rhizobium leguminosarum bv. viciae hydrogenase is controlled at transcriptional and post-transcriptional levels in legume nodules. Mol Plant Microbe Interact 21:597–604. https://doi.org/10.1094/MPMI-21-5-0597
29. Mutch LA, Young JPW (2004) Diversity and specificity of Rhizobium leguminosarum biovar viciae on wild and cultivated legumes. Mol Ecol 13:2435–2444. https://doi.org/10.1111/j.1365-294X.2004.02259.x
30. Karunakaran R, Ramachandran VK, Seaman JC et al (2009) Transcriptomic analysis of Rhizobium leguminosarum biovar viciae in symbiosis with host plants pisum sativum and vicia cracca. J Bacteriol 191:4002–4014. https://doi.org/10.1128/JB.00165-09
31. Jorrin B, Imperial J (2015) Population genomics analysis of legume host preference for specific rhizobial genotypes in the Rhizobium leguminosarum bv. viciae symbioses. Mol Plant Microbe Interact 28:310–318. https://doi.org/10.1094/MPMI-09-14-0296-FI

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.