



OPEN ACCESS

EDITED BY

Robert Beiko,
Dalhousie University, Canada

REVIEWED BY

Sheo Shankar Pandey,
Ministry of Science and Technology, India
Chandni Sidhu,
Max Planck Society, Germany
Nadia Riera,
Pasteur Institute of Montevideo, Uruguay
Frank Guzman,
Federal University of Rio Grande do Sul, Brazil

*CORRESPONDENCE

Donald Lawrence Smith

✉ Donald.smith@mcgill.ca

RECEIVED 09 October 2023

ACCEPTED 30 January 2024

PUBLISHED 13 February 2024

CITATION

Ilangumaran G, Subramanian S and Smith DL (2024) Complete genome sequences of *Rhizobium* sp. strain SL42 and *Hydrogenophaga* sp. strain SL48, microsymbionts of *Amphicarpaea bracteata*. *Front. Microbiomes* 3:1309947. doi: 10.3389/frmbi.2024.1309947

COPYRIGHT

© 2024 Ilangumaran, Subramanian and Smith. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Complete genome sequences of *Rhizobium* sp. strain SL42 and *Hydrogenophaga* sp. strain SL48, microsymbionts of *Amphicarpaea bracteata*

Gayathri Ilangumaran, Sowmyalakshmi Subramanian and Donald Lawrence Smith*

Department of Plant Science, McGill University, Sainte-Anne-de-Bellevue, QC, Canada

This study comprehensively analyzed two distinct rhizobacterial strains, *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48, through whole genome *de novo* sequencing. Isolated from root nodules of *Amphicarpaea bracteata*, a native legume related to soybean, they were selected to explore beneficial rhizobacteria from native plant relatives. Utilizing Illumina and Nanopore sequencers and MaSuRCA assembly, their complete genetic information was elucidated. *Rhizobium* sp. SL42 has a 4.06 Mbp circular chromosome and two plasmids with 60% GC content, while *Hydrogenophaga* sp. SL48 exhibits a 5.43 Mbp circular chromosome with 65% GC content. Genetic analysis identified them as new species, supported by ANI values (77.72% for SL42 and 83.39% for SL48) below the threshold. The genomic analysis unraveled a plethora of genes encoding diverse metabolic functions, secretion systems for substance transport, quorum sensing for coordination, and biosynthetic gene clusters suggesting the production of bioactive compounds. These functional properties contribute to plant growth stimulation, reflecting the symbiotic relationship of rhizobacteria with plants, potentially involving nitrogen fixation and growth-promoting compounds. This research contributes valuable knowledge about plant-microbe interactions and plant growth promotion by these two strains of rhizobacteria.

KEYWORDS

whole genome sequencing, Illumina, nanopore, gene functions, biosynthetic gene clusters

1 Introduction

Plants belonging to the Leguminosae family engage in symbiotic relationships with nitrogen-fixing bacteria, commonly referred to as rhizobia, which reside within their root nodules. These nodules are known not only to harbor nitrogen-fixing symbionts but also other bacteria that potentially function as plant growth-promoting rhizobacteria (PGPR) (Bai et al., 2002). *Amphicarpaea bracteata* (L.) Fernald, a wild legume native to North America, particularly Canada and the lower 48 states of the USA, is one such plant of interest. It predominantly thrives in woody, shaded areas of wetlands but can also be found in similar non-wetland habitats in some regions. This herbaceous perennial species, which grows into a vine, annually produces flowers, pods, and seeds (USDA, 2017). Its seeds and roots are edible, and indigenous communities have historically used it for both nutritional and medicinal purposes (Moerman, 1998).

Importantly, *A. bracteata* shares a close evolutionary relationship with the cultivated soybean (*Glycine max* (L.) Merrill) and hosts symbiotic bacteria from the *Bradyrhizobium* genus, albeit with different genotypes (Sterner and Parker, 1999). In a prior study, we isolated culturable members of the *A. bracteata* nodule phytomicrobiome and examined their potential benefits for soybean plants (Ilangumaran et al., 2021). Two isolates, *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48, emerged as promising candidates by enhancing salt tolerance and promoting soybean growth and maturity under controlled conditions (Ilangumaran et al., 2021, 2022). The taxonomy, morphological and physiological characteristics of these two strains are given in Table 1. Members of the Genus *Rhizobium* are well-known for their ability to form nodules and fix atmospheric nitrogen, underscoring the significance of the *Rhizobium* sp. SL42 isolate. Additionally, *Rhizobium* sp. SL42 was also capable of forming small nodules with soybean plants (data not shown). The isolation of a *Rhizobium* species from *A. bracteata* nodules is an intriguing discovery since the dominant symbiont is *Bradyrhizobium* (Sterner and Parker, 1999). Furthermore, the genus *Hydrogenophaga* is composed of bacteria that utilize hydrogen as an energy source, oxidizing it through the enzyme hydrogenase (Contzen et al., 2000). Our previous studies (cited above) have revealed associations between *Hydrogenophaga* species and plant roots, a novel finding. However, the molecular functions related to their roles in promoting plant growth and protection remain unexplored.

The primary objective of this study is to sequence the genomes of the isolates *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48 using high-throughput next-generation sequencing technology. We aim to comprehensively analyze the whole genome sequences using available platforms to characterize the features of the genome that are relevant to the plant growth-promoting characteristics of these bacteria.

The significance of whole genome sequencing of beneficial rhizobacteria are substantial for advancing our understanding of microbial-mediated plant growth promotion. This research elucidates the genomic underpinnings that enable these bacteria to enhance plant development, offering tangible prospects for

agricultural innovation. By mapping the genomes of these microorganisms, we gain insights into the natural mechanisms that can be harnessed to amplify crop health and yield. This knowledge not only advances our fundamental understanding of symbiotic relationships in agriculture but also holds promise for the development of regenerative approaches. Recognizing these genetic attributes could lead to the adoption of rhizobacteria in agricultural systems, contributing to a more sustainable and robust agroecosystems. Such biotechnological applications of beneficial

TABLE 1 Taxonomic classification and general features of *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48.

| Property | <i>Rhizobium</i> sp. SL42 | <i>Hydrogenophaga</i> sp. SL48 |
|-----------------------------|--|--|
| Classification: Domain | Bacteria | Bacteria |
| Phylum | Proteobacteria | Proteobacteria |
| Class | Alphaproteobacteria | Betaproteobacteria |
| Order | Rhizobiales | Burkholderiales |
| Family | Rhizobiaceae | Comamonadaceae |
| Genus | <i>Rhizobium</i> | <i>Hydrogenophaga</i> |
| Species | unidentified | unidentified |
| Gram stain | Negative | Negative |
| Cell shape | Rod | Rod |
| Motility | Motile | Motile |
| Temperature range | Mesophile | Mesophile |
| Optimum temperature | 25-30 °C (min. temp. 4 °C) | 25-30 °C (min. temp. 4 °C) |
| pH range; Optimum | 7.0 | 7.0 |
| Carbon source | Mannitol | Mannitol |
| Habitat | Soil, root nodule on host | Soil, root nodule on host |
| Salinity | Up to 250 mM NaCl | Up to 100 mM NaCl |
| Oxygen requirement | Aerobic | Aerobic |
| Biotic relationship | Free-living/symbiont | Free-living/symbiont |
| Pathogenicity | Non-pathogenic | Non-pathogenic |
| Biosafety level/ Risk Group | 1 | 1 |
| Isolation | Root nodule of <i>Amphicarpaea bracteata</i> | Root nodule of <i>Amphicarpaea bracteata</i> |
| Geographic location | Sainte-Anne-de-Bellevue, Canada | Sainte-Anne-de-Bellevue, Canada |
| Latitude | 45.404 °N | 45.404 °N |
| Longitude | 73.934 °W | 73.934 °W |
| Altitude | 50 m | 50 m |
| Sample collection | July 2017 | July 2017 |

rhizobacteria stand to revolutionize agricultural practices, paving the way towards an environmentally friendly and resilient agricultural future.

2 Materials and methods

2.1 Bacterial culture and whole genome sequencing

The isolates *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48 were initially grown on yeast extract mannitol (YEM) agar plates at 25°C from glycerol stock cultures. Single well-developed colonies were cultivated in YEM broth and incubated at 25°C, 150 rpm, until they reached an OD_{600nm} of 1.0. Following incubation, the cells were collected by centrifugation at 5000 × g for 10 minutes and then resuspended in 5 mL of broth.

Genomic DNA extraction was performed from 2 mL of cells using the DNeasy Ultraclean Microbial Kit (QIAGEN, Venlo, Netherlands). The isolated DNA, with a concentration of 15 µg in 25 µl Tris buffer, was subsequently dried using a vacuum-free evaporator (Centrifan PE, KD Scientific, Holliston, USA) and stored in DNASTable (Biomatrica, San Diego, USA) for *de novo* whole genome sequencing at Genotypic Technology Ltd., Bangalore, India.

The samples underwent quality control, showing optimal yield and concentration (Supplementary Table S1). DNA integrity was evaluated by agarose gel electrophoresis (Supplementary Figure S1). For quality control and Sanger sequencing, the concentration and purity of the genomic DNA were assessed using the Nanodrop Spectrophotometer 2000 (Thermo Fisher Scientific, Waltham, USA) and the Qubit dsDNA HS assay kit (Thermo Fisher Scientific, Waltham, USA). PCR amplification of 16S rRNA was conducted using 30–50 ng of the genomic DNA as the template with Takara ExTaq (Takara Bio, Shiga, Japan) and the resulted PCR product was purified and subjected to Sanger sequencing (Sanger et al., 1977) (Supplementary Table S1).

For Illumina sequencing, libraries were constructed using the Nextera XT DNA Library Preparation protocol (Illumina, San Diego, USA). Illumina-compatible sequencing libraries were quantified using the Qubit fluorometer and their fragment size distribution analyzed on Agilent TapeStation. The libraries were then sequenced on an Illumina HiSeq X Ten sequencer (Illumina) using 150 bp paired-end chemistry. In preparation for Nanopore sequencing, the Native barcoding kit (EXP-NBD114) from Oxford Nanopore Technology (ONT, Oxford, GB) was used. Native barcode ligation was performed, followed by adapter ligation. The sequencing library was eluted for sequencing on GridION X5 (ONT) using SpotON flow cell R9.4 (FLO-MIN106) in a 48-hour sequencing protocol. Nanopore raw reads (in ‘fast5’ format) were base-called (in ‘fastq5’ format) (Wick et al., 2019) using Fast base calling configuration and de-multiplexed using Guppy v2.3.4.

Quality control, trimming, and initial analyses were conducted using Commander, an NGS analysis tool developed by Genotypic Technology, Bangalore, India. Trimalore, a standard tool (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), was

utilized to remove low-quality reads from Illumina raw data, while Nanopore raw reads underwent processing with Porechop (<https://github.com/rrwick/Porechop>). Only high-quality reads were retained from both platforms. The quality threshold value for Illumina data filtering was Q30 and Nanopore data filtering was Q7.

Hybrid assembly was performed using MaSuRCA v3.3.7 hybrid assembler (Zimin et al., 2013). Gene prediction was carried out using PROKKA (Seemann, 2014), and the predicted proteins were annotated against the UniProt protein database using the DIAMOND BLASTp program for gene ontology and annotation (Buchfink et al., 2015). Pathway analysis utilized the Kyoto Encyclopedia of Genes and Genomes (KEGG) orthology (KO) database of molecular functions (Kanehisa et al., 2016) to assign KO identifiers to genes and proteins and generate pathway maps. The assembled genome was also used for Simple Sequence Repeats (SSR) prediction using the MISA (Micro Satellite identification tool) software (Beier et al., 2017) (Supplementary Figure S2). Additionally, the genome was independently annotated by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (Tatusova et al., 2016).

2.2 Genome analysis

2.2.1 Phylogenetic analysis using IQ-TREE

To analyze the phylogenetic relationships of two bacterial strains based on their 16S rRNA sequences, we first retrieved related sequences through NCBI’s BLAST search (Zhang et al., 2000). Highly similar sequences were optimized for Megablast (Morgulis et al., 2008), with an expectation threshold of 0.05, and a target of 100 sequences, deliberately excluding those from uncultured clones. The sequences were then aligned using Multiple Sequence Comparison by Log-Expectation – MUSCLE (Edgar, 2004) integrated into EBI sequence analysis tools (Madeira et al., 2022), resulting in a multiple sequence alignment (MSA) file in Phylip interleaved format. Subsequently, IQ-TREE (Nguyen et al., 2014) was employed for constructing maximum likelihood phylogenetic trees using its standard settings. The number of bootstrap alignments and maximum iterations were set to 1000, with a correlation coefficient of 0.99, to ensure robust tree construction. It conducted both the Ultrafast Bootstrap analysis (Hoang et al., 2017) and the SH-aLRT branch test (Guindon et al., 2010). The input MSA files for *Rhizobium* had 22 sequences with 1360 nucleotide sites, and *Hydrogenophaga* had 20 sequences with 1409 nucleotide sites. IQ-TREE included ModelFinder (Kalyaanamoorthy et al., 2017) analysis, which identified the HKY+F+I model as the best fit according to Bayesian Information Criterion (BIC). The resulting maximum-likelihood trees were generated in Newick format. For visualization and interpretation, iTOL v5 (Letunić and Bork, 2021) was used, where the trees were displayed and annotated, rooted at their outgroup taxon, and bootstrap values were presented. Statistical measures such as Log-likelihood, AIC, and BIC scores, as well as the identification of near-zero internal branches, provided insights into the robustness and reliability of the phylogenetic trees.

2.2.2 Genome comparison visualization

The circular representation of the chromosome for *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48 was generated using CGView, a tool that facilitates visualizing genomic structures and the arrangement of open reading frames (ORFs). Additionally, the Blast ring image generator (BRIG) circular map was employed to facilitate a detailed comparison between the genomes of *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48 with those of other closely related species in the *Rhizobium* and *Hydrogenophaga* genera available in the NCBI Genome database (Alikhan et al., 2011).

2.2.3 Genome data-mining

The genomes of *Rhizobium* and *Hydrogenophaga* were compared with those in the NCBI database using the M1CR0B1AL1Z3R web server – <https://microbializer.tau.ac.il/> (Avram et al., 2019). This server uses algorithms like Prodigal for ORF extraction, MMSEQS2 for homolog detection, and MCL for clustering. Orthologous groups were analyzed using MAFFT for amino acid sequence alignment and RAxML for species tree inference based on core-proteome alignment. The pan-genome analysis included GC content, ORF number distribution, and orthologous group size distribution. Only complete, fully assembled genomes were used to accurately identify orthologs and study horizontal gene transfer and bacterial evolution.

2.2.4 Identifying biosynthetic gene clusters

To search the genome sequences for secondary metabolite biosynthetic gene clusters, *fasta* files of *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48 genome sequence were queried for antiSMASH 5.0 bacterial sequence analysis. The genomes were analyzed using KnownClusterBlast, ClusterBlast, SubClusterBlast, MIBiG cluster comparison, ActiveSiteFinder, RREFinder, Cluster 5 Pfam analysis, Pfam-based GO term annotation and TIGRFam analysis to identify as many gene clusters as possible (Blin et al., 2019).

3 Results

3.1 Genome properties

The samples underwent quality control, showing optimal yield and concentration (Supplementary Table S2). DNA integrity was evaluated by agarose gel electrophoresis (Supplementary Figure S2). Based on the *16S rRNA* gene sequence analysis, strain SL42 was identified as *Rhizobium* with 98% identity to *Rhizobium ipomoeae* strain NFB1 and is closely related to the taxon *R. ipomoeae* shin9-1T (TaxID: 1210932) and the type strain was isolated from a water convolvulus field (Sheu et al., 2016). Strain SL48 was identified as *Hydrogenophaga* with 99% identity to *Hydrogenophaga taeniospiralis* CCUG 15921T strain NBRC 102512 and it was the most closely related type strain (TaxID: 1281780). *De novo* sequencing hits for SL42 were for *Rhizobium* with 98% identity and SL48 for *Hydrogenophaga* with 93% identity. Further analysis was performed using reference genomes available on NCBI,

Rhizobium flava YW14 (TaxID: 1335061) for SL42 and *Hydrogenaphaga pseudoflava* DSM 1034 (TaxID: 47421) for SL48. The average nucleotide identity – ANI (Yoon et al., 2017) values were compared with the reference genomes of the type strains of closely related species, and both strains exhibited values (77.72% for SL42 and 83.39% for SL48) well below the threshold level for designation as novel species.

The Illumina-compatible sequencing library had an average fragment size of 580 bp (Supplementary Figure S3), and pre-processing retained over 2 million paired-end reads for both samples (Supplementary Tables S3–S12). MaSuRCA v3.3.7 was used for hybrid assembly of Illumina and Nanopore reads, with sufficient sequencing coverage. The genome of *Rhizobium* sp. SL42 consists of a 4.06 Mbp circular chromosome and two circular plasmids of 750237 bp and 351829 bp with a GC content of 60%. The genome of *Hydrogenophaga* sp. SL48 consists of one 5.43 Mbp circular chromosome with a GC content of 65% (Table 2). Whole genome sequences were submitted to NCBI genome database (Table 3).

3.2 Insights from the genome sequencing

Gene prediction and annotation were conducted for both *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48, revealing comprehensive gene ontology and protein predictions (Supplementary Figures S4, S5). Our genomic analysis uncovered a diverse array of predicted genes associated with various biological functions. These gene clusters encompassed flagella, chemotaxis, homoserine lactone, and multidrug resistance, along with regulatory and transport proteins. Notably, *Rhizobium* sp. SL42 possessed genes encoding Type I and Type IV secretion systems, while *Hydrogenophaga* sp. SL48 carried genes for Type II and Type

TABLE 2 Assembly statistics for whole genome sequencing and annotation summary of *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48.

| Assembly statistics | <i>Rhizobium</i> sp. SL42 | <i>Hydrogenophaga</i> sp. SL48 |
|----------------------|---------------------------|--------------------------------|
| Contigs Generated | 3 | 1 |
| Contig Chromosome 1 | 4063937 bp | 5433040 |
| Contig Plasmid 1 | 351829 bp | |
| Contig Plasmid 2 | 750237 bp | |
| Total Contigs Length | 5166003 | 5433040 |
| N50 value | 4063937 | 5433040 |
| GC mol% | 60 | 65 |
| Coverage Illumina | 166.96 | 181.35 |
| Coverage Nanopore | 127.73 | 115.18 |
| Annotation | <i>Rhizobium</i> | <i>Hydrogenophaga</i> |
| CDS (genes) | 4727 | 5077 |
| Proteins (Annotated) | 4642 | 4937 |

TABLE 3 Whole genome sequencing project information NCBI Genome database.

| | <i>Rhizobium</i> sp. SL42 | <i>Hydrogenophaga</i> sp. SL48 |
|--------------------|---|--|
| Name | <i>Rhizobium</i> sp. strain: SL42 Genome | <i>Hydrogenophaga</i> sp. strain: SL48 Genome |
| Accession number | CP063397; CP063398; CP063399 | CP063400 |
| BioProject | PRJNA669345 | PRJNA669344 |
| BioSample | SAMN16451206 | SAMN16451201 |
| Locus Tag | IM739 | IM738 |
| Tax ID | 1210932 | 1904254 |
| Genome size | 4.06 Mbp | 5.43 Mbp |
| Assembly method | MaSuRCA 3.3.7 | MaSuRCA 3.3.7 |
| Assembly name | MGM_Rhim_1 | MGM_Hyga_1 |
| Reference Title | Genome sequence of <i>Rhizobium</i> sp. strain SL42 | Genome sequence of <i>Hydrogenophaga</i> sp. strain SL48 |
| Reference authors: | Ilangumaran, G., Subramanian, S., and Smith, D. | Ilangumaran, G., Subramanian, S., and Smith, D. |

IV secretion systems, as well as hydrogenase enzymes. Both strains also exhibited genes related to essential cellular processes, including photosystem I, nodulation, nitrogen fixation, heat shock, cold shock proteins, hypoxic response, iron chelation, and carotenoid synthesis. Genomic features contributing to these biological functions and related plant-microbe interactions is presented in Table 4.

The KEGG pathway mapping further elucidated the functions associated with the predicted proteins in these bacteria, encompassing bacterial motility proteins, secretion system proteins, bacterial chemotaxis, flagellar assembly, peptidoglycan biosynthesis, and quorum sensing. Intriguingly, each strain possessed unique proteins that set them apart. In *Rhizobium* sp. SL42, proteins related to photosynthesis, carbon fixation, and carotenoid biosynthesis pathways were identified, while in *Hydrogenophaga* sp. SL48, proteins involved in the biosynthesis of the vancomycin group of antibiotics were present (Supplementary Tables S13, S14).

Phylogenetic analyses of the 16S rRNA gene and house-keeping genes *gyrB*, *recA*, and *rpoD* indicated that *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48 exhibit distinct evolutionary characteristics (Supplementary Figures S6, S7). Using 16S rRNA gene sequences of related strains, comprehensive phylogenetic analyses in IQ-TREE enabled a detailed understanding of the evolutionary relationships between them and suggested potential novelty of *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48 within their respective genera, as indicated by distinct branching patterns (Figures 1, 2).

TABLE 4 Genes related to PGPR attributes in the genome of *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48.

| <i>Rhizobium</i> sp. SL42 | | |
|---|---|------------------|
| Gene | Function | # Genes encoding |
| <i>bcr</i> | Bicyclomycin resistance | 3 |
| <i>bdlA</i> ; <i>bigR</i> | Biofilm | 6 |
| <i>cspA</i> ; <i>cspE</i> ; <i>cspG</i> | cold shock protein, cold shock-like protein | 7 |
| <i>entS</i> ; <i>fepC</i> ; <i>fepD</i> ; <i>fepG</i> | Enterobactin | 7 |
| <i>hspQ</i> | heat shock protein | 1 |
| <i>rhtB</i> | homoserine lactone efflux protein | 7 |
| <i>hrp1</i> | hypoxic response protein | 1 |
| <i>yfeA</i> ; <i>hemH</i> | Iron chelation | 2 |
| <i>lptA</i> ; <i>lptB</i> ; <i>lptG</i> ; <i>lapA</i> ; <i>lapB</i> | lipopolysaccharide assembly and export proteins | 8 |
| <i>mdtA</i> ; <i>mdtB</i> ; <i>mdtC</i> ; <i>mdtN</i> ; <i>mdtK</i> ; <i>mdtE</i> | Multidrug resistance | 24 |
| <i>mrpA</i> ; <i>mrpB</i> ; <i>mrpC</i> ; <i>mrpD</i> ; <i>mrpE</i> ; <i>mrpG</i> ; <i>mrpF</i> | Na(+)/H(+) antiporter subunit | 7 |
| <i>fixK</i> | nitrogen fixation regulation | 4 |
| <i>nodM</i> , <i>noIR</i> | Nodulation | 1 |
| <i>pleC</i> | Non-motile and phage-resistance protein | 3 |
| <i>envZ</i> | Osmolarity sensor protein | 1 |
| <i>hemF</i> | Oxygen-dependent coproporphyrinogen-III oxidase | 1 |
| <i>ycf3</i> ; <i>regA</i> | Photosynthesis | 2 |
| <i>crtI</i> ; <i>crtB</i> ; <i>carA2</i> | Phytoene | 3 |
| | Putative signaling | 22 |
| <i>fpvA</i> ; <i>fhuA</i> ; <i>ftsY</i> ; <i>chvE</i> ; <i>cheD</i> ; <i>fhuE</i> ; | Receptor | 12 |
| <i>aroK</i> ; <i>aroA</i> ; <i>aroE</i> ; <i>quiA</i> | Shikimate pathway | 5 |
| <i>chaA</i> | Sodium-potassium/proton antiporter | 1 |
| <i>potA</i> ; <i>potB</i> ; <i>potD</i> | Spermidine/putrescine | 25 |
| <i>gerE</i> | Spore germination protein | 1 |
| <i>soj</i> | Sporulation initiation inhibitor | 2 |
| <i>prsD</i> ; <i>prsE</i> ; | Type I secretion system | 15 |
| <i>virB4</i> ; <i>virB9</i> ; <i>virB10</i> , <i>virB11</i> | Type IV secretion system | 4 |
| <i>clcB</i> | Voltage-gated ClC-type chloride channel | 1 |

(Continued)

TABLE 4 Continued

| <i>Hydrogenophaga</i> sp. SL48 | | |
|---|--|------------------|
| Gene | Function | # Genes encoding |
| | Acid shock protein | 1 |
| arpC | Antibiotic efflux pump outer membrane protein | 1 |
| bcr | Bicyclomycin resistance | 2 |
| icaR | Biofilm operon regulator | 1 |
| ble | Bleomycin resistance | 1 |
| kfoC | Chondroitin synthase | 2 |
| cspA; cspG | cold shock protein, cold shock-like protein | 2 |
| fas6 | Cytokinin riboside 5'-monophosphate phosphoribohydrolase | 1 |
| entS | Enterobactin | 1 |
| fbpC | Fe(3+) ions import ATP-binding protein | 1 |
| hslR | heat shock protein | 1 |
| rhtB | homoserine lactone efflux protein | 5 |
| hypF; hypB; hypD | Hydrogenase maturation factor | 3 |
| hrp1 | hypoxic response protein | 1 |
| hemH; sirB | Iron chelation | 2 |
| lptA; lptB; lptC; lptG; lptF; lapA; lapB | lipopolysaccharide assembly and export proteins | 16 |
| mdtB; mdtN; mdtE mdtA; mdtC; mexR; mexA; mdtD; mdtH; mdtG | Multidrug resistance | 13 |
| gerN; mrpA; mrpD; mrpE; mrpG; mrpF; mnhC1 | Na(+)/H(+) antiporter subunit | 7 |
| hoxF; hoxU; hoxY; hoxH | NAD-reducing hydrogenase HoxS subunit | 4 |
| fixK | nitrogen fixation regulation | 1 |
| nifH; nifD; nifK | Nitrogenase iron protein | 5 |
| nifW | Nitrogenase-stabilizing/protective protein | 1 |
| nodD | Nodulation protein | 4 |
| envZ | Osmolarity sensor protein | 2 |
| osmY | Osmotically-inducible protein Y | 3 |
| hemF | Oxygen-dependent coproporphyrinogen-III oxidase | 1 |
| | Periplasmic [NiFeSe] hydrogenase subunit | 2 |

(Continued)

TABLE 4 Continued

| <i>Hydrogenophaga</i> sp. SL48 | | |
|---|--|------------------|
| Gene | Function | # Genes encoding |
| kcsA | pH-gated potassium channel | 1 |
| regA | Photosynthesis | 1 |
| crtB | Phytoene | 1 |
| | Putative signaling | 1 |
| cheD; cirA; aer; ftsY; fhuA; chvE; fucA; fhuE; | Receptor | 8 |
| cbbS1; cbbL; rlp2 | Ribulose biphosphate carboxylase | 3 |
| rubA; hrb | Rubredoxin | 2 |
| aroL; aroA; aroE; quiA; ydiB | Shikimate pathway | 6 |
| chaA | Sodium-potassium/proton antiporter | 1 |
| potA; potB; potD; | Spermidine/putrescine | 12 |
| spsA | Spore coat polysaccharide biosynthesis protein | 1 |
| srkA | stress response kinase A | 1 |
| iaaM | Tryptophan 2-monooxygenase | 1 |
| xpsD; gspE; gspF; epsE; epsF; hxcR; xcpQ; xcpV; xcpT; pulD; | Type II secretion system | 13 |
| virB1; virB4; virB8; virB10, virB1; ptlf | Type IV secretion system | 7 |

Circular genome mapping of *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48 identified unique regions, ORFs, and specific genetic elements exclusive to these strains. These genomic comparisons with other strains within their respective genera highlighted their genetic distinctiveness (Figures 3, 4). M1CR0B1AL1Z3R web server was utilized to analyze genomic features and identify orthologous genes in *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48, comparing them with complete genomes from the NCBI database. The analysis a comprehensive pan-genome analysis covering GC content, ORF distribution, and orthologous group sizes. We focused exclusively on complete and fully assembled genomes to ensure accurate ortholog identification, crucial for studying horizontal gene transfer and bacterial evolution (Supplementary Data Sheet 1, 2).

3.3 Finding secondary metabolites using Anti-SMASH

The AntiSMASH analysis revealed the presence of biosynthetic gene clusters encoding secondary metabolites in both *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48 strains (Table 5). Some of these clusters were unique to each strain. In *Rhizobium* sp. SL42, several

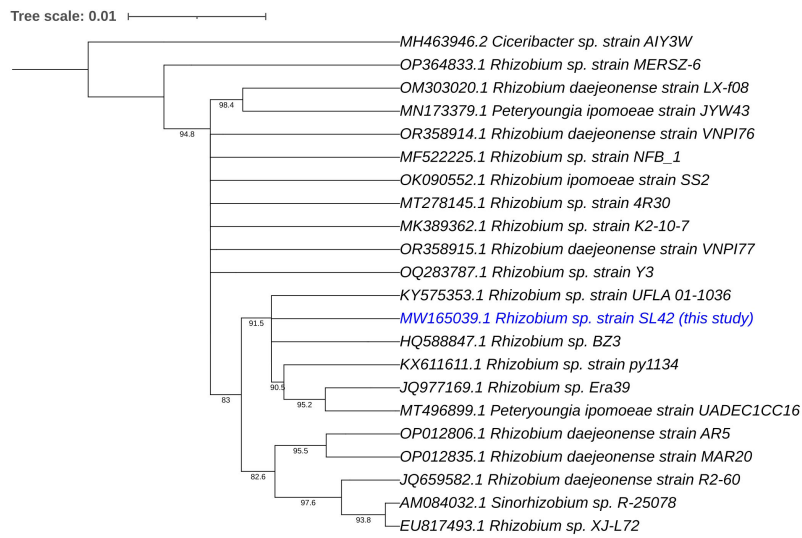


FIGURE 1
Phylogenetic analysis of *Rhizobium* sp. SL42 (in blue text) based on 16S rRNA gene sequences of related strains. The evolutionary history was inferred using the Maximum Likelihood (ML) method. The bootstrap consensus tree, constructed from 1000 replicate alignments, illustrates the inferred evolutionary history of the taxa analyzed. The percentages displayed next to the branches represent the proportion of these replicate trees in which the associated taxa clustered together, as observed in the bootstrap test. This ML tree was generated employing the Ultrafast Bootstrap analysis and the SH-aLRT branch test. The phylogenetic analyses were conducted using IQ-TREE and tree was generated using iTOL.

regions were identified, including TfuA-related, Terpene, Hserlactone, and Bacteriocin. *Hydrogenophaga* sp. SL48 displayed diverse secondary metabolite coding regions, including Arylpolyene, Terpene, T1PKS, NRPS-like, NRPS, Bacteriocin, Siderophore, and Betalactone. Some of these regions exhibited similarities to known clusters, such as Xanthomonadin I, Desferrioxamine E, and Mycosubtilin (Supplementary Figures S8, S9).

4 Discussion

4.1 Phylogenetic analysis and genome comparisons

The phylogenetic analyses based on the 16S rRNA gene sequences of related strains facilitated a deeper understanding of

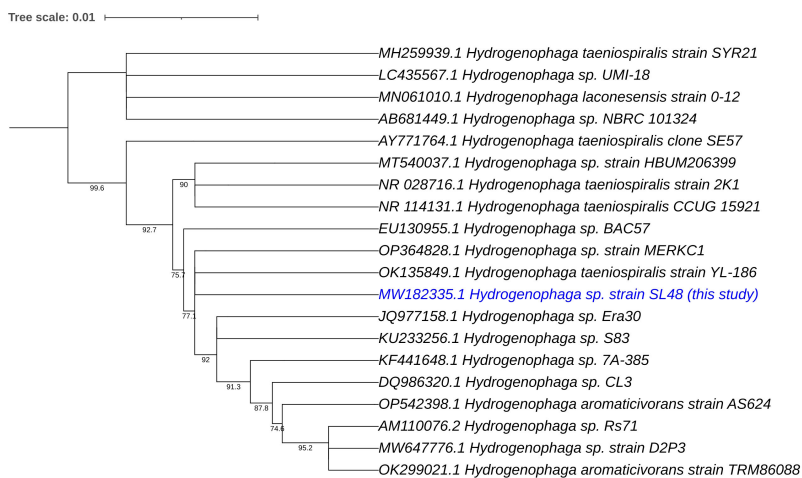


FIGURE 2
Phylogenetic analysis of *Hydrogenophaga* sp. SL48 (in blue text) based on 16S rRNA gene sequences of related strains. The evolutionary history was inferred using the Maximum Likelihood (ML) method. The bootstrap consensus tree, constructed from 1000 replicate alignments, illustrates the inferred evolutionary history of the taxa analyzed. The percentages displayed next to the branches represent the proportion of these replicate trees in which the associated taxa clustered together, as observed in the bootstrap test. This ML tree was generated employing the Ultrafast Bootstrap analysis and the SH-aLRT branch test. The phylogenetic analyses were conducted using IQ-TREE and tree was generated using iTOL.

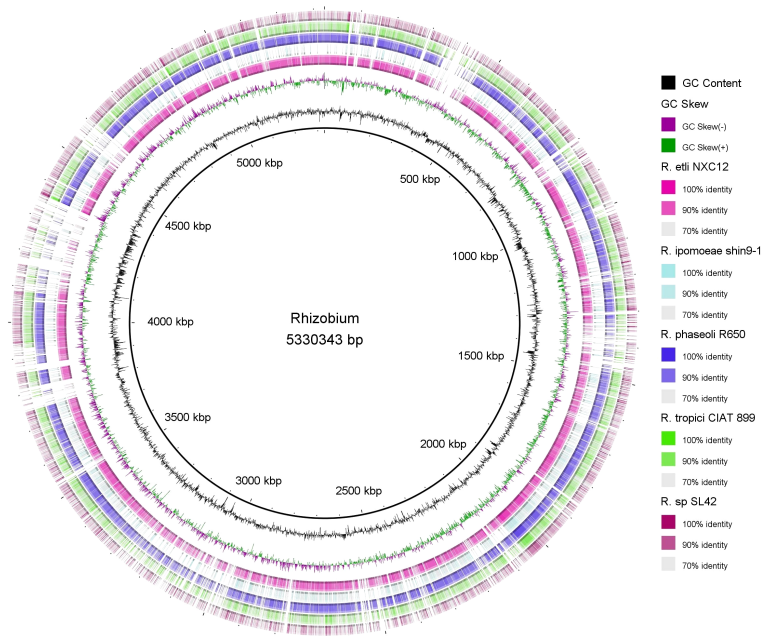


FIGURE 3
Map of the circular genome of *Rhizobium* sp. SL42, created with CGview, which accentuates the strain’s unique regions and open reading frames (ORFs), and includes BRIG visualizations for comparative genomics. This map facilitates a comparison with other strains within the same genera, highlighting the unique genetic elements that are specific to *Rhizobium* sp. SL42.

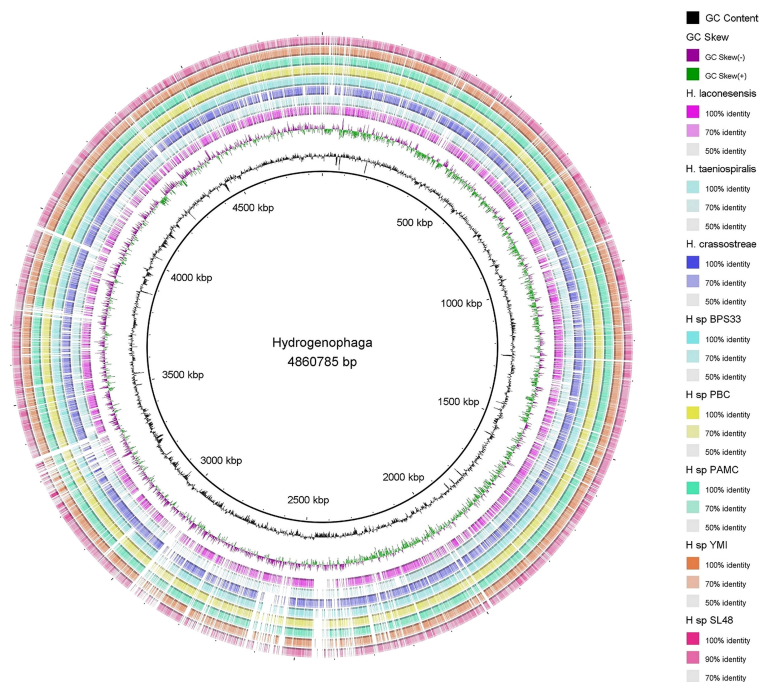


FIGURE 4
Map of the circular genome of *Hydrogenophaga* sp. SL48, created with CGview, which accentuates the strain’s unique regions and open reading frames (ORFs), and includes BRIG visualizations for comparative genomics. This map facilitates a comparison with other strains within the same genera, highlighting the unique genetic elements that are specific to *Hydrogenophaga* sp. SL48.

TABLE 5 AntiSMASH results of secondary metabolite coding regions of *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48.

| <i>Rhizobium</i> sp. SL42 | | <i>Hydrogenophaga</i> sp. SL48 | |
|---------------------------|--------------|--------------------------------|------------------------|
| Region | Type | Region | Type |
| Region 1.1 | TfuA-related | Region 1 | Arylpolyene |
| Region 1.2 | Terpene | Region 2 | Terpene |
| Region 1.3 | Hserlactone | Region 3 | T1PKS, NRPS-like, NRPS |
| Region 1.4 | Bacteriocin | Region 4 | Bacteriocin |
| Region 2.1 | NRPS, T1PKS | Region 5 | Siderophore |
| Region 2.2 | TfuA-related | Region 6 | Betalactone |
| Region 3.1 | Hserlactone | | |

the evolutionary dynamics of these bacterial strains and strongly suggested that *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48 might represent novel species within their respective genera. The distinct branching patterns and genetic divergence from known species in their taxonomic groups provide compelling evidence of their uniqueness. The circular genome mapping of these strains revealed unique genomic regions, open reading frames (ORFs), and distinctive elements exclusive to *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48. This demonstrates their genetic distinctiveness and supports the hypothesis that they may represent novel species. The genomic comparisons with other strains within their respective genera highlight these differences.

The use of the M1CR0B1AL1Z3R web server for the computation of key genomic features and identification of orthologous genes is valuable for understanding the genetic relationships of these strains. The identification of orthologous genes can provide insights into the evolutionary history and relatedness of these strains to other known species. The pan genome analysis, which includes aspects such as GC content, ORF distribution, and orthologous group sizes, is essential for gaining a comprehensive understanding of the genetic diversity within the studied bacterial genera. It can help uncover the potential influence of horizontal gene transfer (HGT) events on bacterial evolution, shedding light on how genetic diversity is shaped in these organisms.

4.2 Biological significance of genomic features

The genomic analysis has revealed a diverse array of genes associated with various biological functions in both *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48. These features play a crucial role in plant-microbe interactions and plant growth enhancement. The presence of Type I and Type IV secretion systems in *Rhizobium* sp. SL42 is significant because these systems are known to be involved in the communication between rhizobia and leguminous plants (Nelson and Sadowsky, 2015). They play a crucial role in the establishment of symbiotic relationships, leading to effective

nodulation and nitrogen fixation. The Type IV secretion system, in particular, is implicated in the transfer of symbiotic signals and genetic material between the bacterium and the plant (Black et al., 2012). This system is essential for the successful formation of nodules on plant roots and the subsequent fixation of atmospheric nitrogen, making it an integral part of the nitrogen fixation process.

Similarly, *Hydrogenophaga* sp. SL48 carries genes for Type II and Type IV secretion systems, which may contribute to interactions with plant roots and environmental adaptation. The presence of hydrogenase enzymes in *Hydrogenophaga* sp. SL48 is also noteworthy. Hydrogenase enzymes are involved in the metabolism of hydrogen gas, which can be a source of electrons for nitrogenase enzymes responsible for nitrogen fixation (Mishra et al., 2019). In nitrogen-fixing bacteria, hydrogenase enzymes play a role in maintaining the redox balance necessary for efficient nitrogen fixation. Therefore, the presence of hydrogenase enzymes in *Hydrogenophaga* sp. SL48 suggests its potential involvement in nitrogen fixation processes.

Additionally, the presence of genes related to photosynthesis, carbon fixation, nitrogen fixation, and carotenoid biosynthesis in *Rhizobium* sp. SL42, and the genes for the biosynthesis of antibiotics in *Hydrogenophaga* sp. SL48, highlight their potential to produce growth-promoting compounds. These genomic features are of significant biological importance as they can contribute to the production of compounds that stimulate plant growth and protect plants from pathogens (Levy et al., 2018).

4.3 Ecological and agricultural applications

Plant growth promoting rhizobacteria (PGPR) produce bioactive substances that improve plant growth and alleviate stress. Understanding the behavior of PGPR when inoculated onto plants is important for their application in agriculture. Some of these compounds are also essential for plant root colonization (Bloemberg and Lugtenberg, 2001). The whole genome sequencing analysis revealed the genes harbored in the genomes of *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48 that might play key roles in plant-microbe interactions. PGPR are known to produce auxins, gibberellins, cytokinins and ethylene and manipulate phytohormone balance in plants. PGPR stimulate root proliferation by excretion of indole-3-acetic acid (IAA) into the rhizosphere, thus enhancing uptake of water and nutrients (Sukumar et al., 2013). Several PGPR also secrete cytokinins that have been detected in cell free medium (Garcia de Salamone et al., 2001). Genes encoding IAA and cytokinin biosynthesis (*iaaM* and *fas6*) were present in *Hydrogenophaga* sp. SL48. Volatile organic compounds (VOCs) produced by bacteria help in plant development and stress responses (Bailly and Weisskopf, 2012). Polyamines play important physiological and protective roles in plants. *Bacillus megaterium* BOFC15 secretes spermidine, a polyamine leading to enhanced cellular polyamine levels in Arabidopsis. Inoculation with the bacterium resulted in an increase in biomass, changed root architecture and elevated photosynthetic capacity. The plants also exhibited higher drought

tolerance and abscisic acid content under water deficit (osmotic stress) (Zhou et al., 2016a). Both strains possess multiples genes that encode for spermidine/putrescine compounds (potA, osmY, envZ). These genomic findings have implications for potential ecological and agricultural applications. Understanding the genomic features of these strains can lead to the development of more effective PGPR inoculants for agricultural use.

Genes encoding the production of secondary metabolites found using Anti-SMASH showed that the PGPR produces antibiotics such as thiopeptides, polyketides and bacteriocins that suppress pathogens. These findings underscore the potential of these strains to produce bioactive compounds, further contributing to their functional properties in plant growth stimulation and ecosystem interactions reported in our previous studies (Ilangumaran et al., 2022). Bacterial surface factors like flagellins and o-antigen of lipopolysaccharides induce systemic resistance (ISR) whereas, analogs of salicylic acid, jasmonic acid and ethylene elicit systemic acquired resistance (SAR) in plants (Ping and Boland, 2004; Lugtenberg and Kamilova, 2009; Pieterse et al., 2014). A bacteriocin, thuricin 17, isolated from the soybean endosymbiont *Bacillus thuriengenesis* NEB 17, when applied as foliar spray or root drench stimulated the growth of soybean and corn (Subramanian et al., 2016). Siderophores are iron chelators produced by some microorganisms and enhance plant growth under iron-depleted conditions where they are used as the method for accessing scarce iron and also act as biocontrol agents by reducing the availability of iron for pathogens (Saha et al., 2016). The identified biosynthetic gene clusters for antibiotics, polyketides, and bacteriocins might suppress plant pathogens, thus contributing to plant health and growth.

4.4 Rhizosphere colonization

Genes involved in the pathways of cell motility, chemotaxis, lipopolysaccharide synthesis and biofilm formation suggested that they might play important roles in rhizosphere colonization of *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48. Plant beneficial bacteria present in the rhizosphere are in proximity to roots and many are known to form biofilms, which aid in the successful colonization of root surfaces and adjacent soil particles and thwart pathogenic bacteria. Biofilms are structured communities of bacterial cells living adherent to a surface embedded in an extracellular polysaccharide matrix. Biofilms of beneficial bacteria play a crucial role in plant growth promoting effects, as they aid in the colonization of root surfaces and the inhibition of pathogenic bacteria (Ramey et al., 2004). Plant roots exude signal compounds that regulate plant-bacteria interactions and trigger chemotaxis in bacteria, towards the rhizosphere (Fan et al., 2012). For example, flavonoids secreted by roots determine the legume-rhizobia symbiotic associations while malate and citrate are found to interact with *Bacillus* and *Pseudomonas* strains (Badri and Vivanco, 2009). There are genes related to the metabolism of these compounds in *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48, suggesting that they possibly take part in plant-microbe interactions.

During colonization, PGPR assimilate substances released by the roots and in turn, produce bioactive compounds that promote plant growth or ameliorate stress (Xie et al., 2014). Recent advances in high-throughput strategies have led to detailed investigations of plant-microbe interactions and the differential effects of root exudates on mechanisms of rhizobacteria that are crucial to the beneficial effects observed. The differentially expressed genes or proteins were mainly those involved in nutrient utilization and transport, chemotaxis, secretion, quorum sensing, extracellular matrix, synthesis of volatile compounds, and antibiotic production (Fan et al., 2012; Beauregard et al., 2013; Kierul et al., 2015; Mwita et al., 2016; Zhou et al., 2016b). The presence of genes related to the metabolism of root-exuded compounds, such as flavonoids and organic acids, in these strains indicates their ability to respond to plant signals and interact with the rhizosphere. This is essential for establishing beneficial relationships with plants and enhancing nutrient uptake.

Understanding the dynamic function of bacterial cells and regulatory networks related to enzyme metabolism, transport and utilization of nutrients, signal transduction proteins and root colonization pattern is important in determining their potential applications in agriculture (Kierul et al., 2015; Zhang et al., 2015; Mwita et al., 2016; Zhou et al., 2016b). The genomic features in *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48 hold great potential for enhancing their ability to establish beneficial relationships with soybeans. These features, including nutrient acquisition, hormone production, and bioactive compound production, are of clear biological significance in the context of plant-microbe interactions and growth enhancement. This suggests promising applications for these strains in agriculture to improve crop productivity and promote sustainable farming practices.

5 Summary

Genomic analysis of *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48 revealed the presence of genes encoding various metabolic functions and compound synthesis. These strains were selected for sequencing as part of the isolation and characterization of beneficial rhizobacteria from native relatives of cultivated plants, chosen based on their ecological and agricultural importance in plant-microbe interactions, stress tolerance, and growth promotion. The sequencing aimed to determine their plant growth promoting and genetic characteristics. The findings highlight the functional diversity of these bacterial strains and suggest their potential roles in ecological interactions and adaptation strategies. The presence of essential genes associated with plant-microbe interactions, along with unique proteins, presents promising avenues for exploring their contributions to plant growth stimulation and antibiotic production. The study enriches our understanding of microbial diversity and potential biotechnological applications.

The combined results of biosynthetic gene cluster analyses, phylogenetic assessments, and genomic mapping provide compelling evidence for the novelty of *Rhizobium* sp. SL42 and *Hydrogenophaga* sp. SL48 within their respective genera. Unique biosynthetic gene clusters, distinct phylogenetic positions, and

exclusive genomic regions further support their potential classification as entirely new species. These findings underscore the significance of these bacterial strains in terms of microbial diversity and their potential contributions to plant growth stimulation and ecological interactions. The discovery of novel bacterial species enhances our understanding of microbial ecology and opens promising avenues for utilizing the beneficial properties of these bacteria in agriculture and the environment.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary Material](#).

Author contributions

GI: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft. SS: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. DS: Formal Analysis, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Publication of this article has been made possible by the Biomass Canada Cluster. The Biomass Canada Cluster is managed by BioFuelNet

References

- Alikhan, N. F., Petty, N. K., Zakour, N. L. B., and Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12, 402. doi: 10.1186/1471-2164-12-402
- Avram, O., Rapoport, D., Portugez, S., and Pupko, T. (2019). M1CR0B1A1L1Z3R-a user-friendly web server for the analysis of large-scale microbial genomics data. *Nucleic Acids Research* 47 (W1), W88–W92. doi: 10.1093/nar/gkz423
- Badri, D. V., and Vivanco, J. M. (2009). Regulation and function of root exudates. *Plant Cell Environ.* 32, 666–681. doi: 10.1111/j.1365-3040.2009.01926.x
- Bai, Y. M., D'Aoust, F., Smith, D. L., and Driscoll, B. T. (2002). Isolation of plant-growth-promoting *Bacillus* strains from soybean root nodules. *Can. J. Microbiol.* 48, 230–238. doi: 10.1139/W02-014
- Bailey, A., and Weiskopf, L. (2012). The modulating effect of bacterial volatiles on plant growth: current knowledge and future challenges. *Plant Signal. Behav.* 7, 79–85. doi: 10.4161/psb.7.1.18418
- Beauregard, P. B., Chai, Y. R., Vlamakis, H., Losick, R., and Kolter, R. (2013). *Bacillus subtilis* biofilm induction by plant polysaccharides. *P. Natl. Acad. Sci. U.S.A.* 110, E1621–E1630. doi: 10.1073/pnas.1218984110
- Beier, S., Thiel, T., Munch, T., Scholz, U., and Mascher, M. (2017). MISA-web: a web server for microsatellite prediction. *Bioinformatics* 33, 2583–2585. doi: 10.1093/bioinformatics/btx198
- Black, M., Moolhuijzen, P., Chapman, B., Barrero, R., Howieson, J., Hungria, M., et al. (2012). The genetics of symbiotic nitrogen fixation: comparative genomics of 14 rhizobia strains by resolution of protein clusters. *Genes* 3, 138–166. doi: 10.3390/genes3010138
- Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., et al. (2019). antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 47, W81–W87. doi: 10.1093/nar/gkz310
- Bloemberg, G. V., and Lugtenberg, B. J. J. (2001). Molecular basis of plant growth promotion and biocontrol by rhizobacteria. *Curr. Opin. Plant Biol.* 4, 343–350. doi: 10.1016/S1369-5266(00)00183-7
- Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176
- Contzen, M., Moore, E. R. B., Blumel, S., Stolz, A., and Kamper, P. (2000). *Hydrogenophaga intermedia* sp. nov., a 4-aminobenzenesulfonate degrading organism. *Syst. Appl. Microbiol.* 23, 487–493. doi: 10.1016/S0723-2020(00)80022-3
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Fan, B., Carvalhais, L. C., Becker, A., Fedoseyenko, D., von Wiren, N., and Borriss, R. (2012). Transcriptomic profiling of *Bacillus amyloliquefaciens* FZB42 in response to maize root exudates. *BMC Microbiol.* 12, 116. doi: 10.1186/1471-2180-12-116

Canada and is funded through the Canadian Agricultural Partnership's, AgriScience Program, Agriculture and Agri-Food, Canada.

Acknowledgments

The authors thank members of the Smith lab for their help and support during the due course of this project. Genomics service provided by Genotypic, India for whole genome sequencing is highly appreciated.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frmbi.2024.1309947/full#supplementary-material>

- Garcia de Salamone, I. E., Hynes, R. K., and Nelson, L. M. (2001). Cytokinin production by plant growth promoting rhizobacteria and selected mutants. *Can. J. Microbiol.* 47, 404–411. doi: 10.1139/w01-029
- Guindon, S., Dufayard, J., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PHYML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010
- Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2017). UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* 35, 518–522. doi: 10.1093/molbev/msx281
- Ilangumaran, G., Schwinghamer, T. D., and Smith, D. L. (2021). Rhizobacteria from root nodules of an indigenous legume enhance salinity stress tolerance in soybean. *Front. Sustain. Food Syst.* 4. doi: 10.3389/fsufs.2020.617978
- Ilangumaran, G., Subbramanian, S., and Smith, D. L. (2022). Soybean leaf proteomic profile influenced by rhizobacteria under optimal and salt stress conditions. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.809906
- Kalyanamamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A., and Jermini, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Met.* 14, 587–589. doi: 10.1038/nmeth.4285
- Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 44, D457–D462. doi: 10.1093/nar/gkv1070
- Kierul, K., Voigt, B., Albrecht, D., Chen, X. H., Carvalhais, L. C., and Borriss, R. (2015). Influence of root exudates on the extracellular proteome of the plant growth-promoting bacterium *Bacillus amyloliquefaciens* FZB42. *Microbiology* 161, 131–147. doi: 10.1099/mic.0.083576-0
- Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49, W293–W296. doi: 10.1093/nar/gkab301
- Levy, A., Conway, J. M., Dangl, J. L., and Woyke, T. (2018). Elucidating bacterial gene functions in the plant microbiome. *Cell Host Mic.* 24, 475–485. doi: 10.1016/j.chom.2018.09.005
- Lugtenberg, B., and Kamilova, F. (2009). Plant-growth-promoting rhizobacteria. *Annu. Rev. Microbiol.* 63, 541–556. doi: 10.1146/annurev.micro.62.081307.162918
- Madeira, F., Pearce, M., Tivey, A., Basutkar, P., Lee, J., Edbali, O., et al. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* 50, W276–W279. doi: 10.1093/nar/gkac240
- Mishra, A. K., Kaushik, M. S., and Tiwari, D. N. (2019). “Nitrogenase and Hydrogenase: Enzymes for Nitrogen Fixation and Hydrogen Production in Cyanobacteria,” in *Cyanobacteria*. Eds. A. K. Mishra, D. N. Tiwari and A. N. Rai (London, UK: Academic Press), 173–191. doi: 10.1016/B978-0-12-814667-5.00008-8
- Moerman, D. E. (1998). *Native American Ethnobotany* (Portland, Oregon, USA: Timber Press).
- Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T., Agarwala, R., and Schäffer, A. A. (2008). Database indexing for production MegaBLAST searches. *Bioinformatics* 24, 1757–1764. doi: 10.1093/bioinformatics/btn322
- Mwita, L., Chan, W. Y., Pretorius, T., Lyantagaye, S. L., Lapa, S. V., Avdeeva, L. V., et al. (2016). Gene expression regulation in the plant growth promoting *Bacillus atrophaeus* UCMB-5137 stimulated by maize root exudates. *Gene* 590, 18–28. doi: 10.1016/j.gene.2016.05.045
- Nelson, M. S., and Sadowsky, M. J. (2015). Secretion systems and signal exchange between nitrogen-fixing rhizobia and legumes. *Front. Plant Sci.* 6. doi: 10.3389/fpls.2015.00491
- Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. (2014). IQ-TREE: a fast and effective stochastic algorithm for estimating Maximum-Likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300
- Pieterse, C. M. J., Zamioudis, C., Berendsen, R. L., Weller, D. M., Van Wees, S. C. M., and Bakker, P. A. H. M. (2014). Induced systemic resistance by beneficial microbes. *Ann. Rev. Phytopathol.* 52, 347–375. doi: 10.1146/annurev-phyto-082712-102340
- Ping, L. Y., and Boland, W. (2004). Signals from the underground: bacterial volatiles promote growth in Arabidopsis. *Trends Plant Sci.* 9, 263–266. doi: 10.1016/j.tplants.2004.04.008
- Ramey, B. E., Koutsoudis, M., von Bodman, S. B., and Fuqua, C. (2004). Biofilm formation in plant-microbe associations. *Curr. Opin. Microbiol.* 7, 602–609. doi: 10.1016/j.mib.2004.10.014
- Saha, M., Sarkar, S., Sarkar, B., Sharma, B. K., Bhattacharjee, S., and Tribedi, P. (2016). Microbial siderophores and their potential applications: a review. *Environ. Sci. Pollut. Res. Int.* 23, 3984–3999. doi: 10.1007/s11356-015-4294-0
- Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* 74, 5463–5467. doi: 10.1073/pnas.74.12.5463
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Sheu, S. Y., Chen, Z. H., Young, C. C., and Chen, W. M. (2016). *Rhizobium ipomoeae* sp. nov., isolated from a water convolvulus field. *Int. J. Syst. Evol. Microbiol.* 66, 1633–1640. doi: 10.1099/ijsem.0.000875
- Sterner, J. P., and Parker, M. A. (1999). Diversity and relationships of *Bradyrhizobia* from *Amphicarpaea bracteata* based on partial nod and ribosomal sequences. *Syst. Appl. Microbiol.* 22, 387–392. doi: 10.1016/S0723-2020(99)80047-2
- Subbramanian, S., Souleimanov, A., and Smith, D. L. (2016). Proteomic studies on the effects of lipo-chitooligosaccharide and thuricin 17 under unstressed and salt stressed conditions in *Arabidopsis thaliana*. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.01314
- Sukumar, P., Legue, V., Vayssières, A., Martin, F., Tuskan, G. A., and Kalluri, U. C. (2013). Involvement of auxin pathways in modulating root architecture during beneficial plant-microorganism interactions. *Plant Cell Environ.* 36, 909–919. doi: 10.1111/pce.12036
- Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., et al. (2016). NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* 44, 6614–6624. doi: 10.1093/nar/gkw569
- USDA – United States Department of Agriculture Natural Resources Conservation Service. *PLANTS: Amphicarpaea bracteata (L.) Fernald, American hogpeanut*. Available online at: <https://plants.usda.gov/core/profile?symbol=AMBR2> (Accessed 12 Aug 2017).
- Wick, R. R., Judd, L. M., and Holt, K. E. (2019). Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biol.* 20, 129. doi: 10.1186/s13059-019-1727-y
- Xie, S. S., Wu, H. J., Zang, H. Y., Wu, L. M., Zhu, Q. Q., and Gao, X. W. (2014). Plant growth promotion by spermidine-producing *Bacillus subtilis* OKB105. *Mol. Plant-Microbe Interact.* 27, 655–663. doi: 10.1094/Mpmi-01-14-0010-R
- Yoon, S. H., Ha, S. M., Lim, J., Kwon, S., and Chun, J. (2017). A large-scale evaluation of algorithms to calculate average nucleotide identity. *Anton. Leeuw. Int. J. G.* 110, 1281–1286. doi: 10.1007/s10482-017-0844-4
- Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* 7, 203–214. doi: 10.1089/10665270050081478
- Zhang, N., Yang, D. Q., Wang, D. D., Miao, Y. Z., Shao, J. H., Zhou, X., et al. (2015). Whole transcriptomic analysis of the plant-beneficial rhizobacterium *Bacillus amyloliquefaciens* SQR9 during enhanced biofilm formation regulated by maize root exudates. *BMC Genomics* 16, 685. doi: 10.1186/s12864-015-1825-5
- Zhou, D. M., Huang, X. F., Chaparro, J. M., Badri, D. V., Manter, D. K., Vivanco, J. M., et al. (2016b). Root and bacterial secretions regulate the interaction between plants and PGPR leading to distinct plant growth promotion effects. *Plant Soil* 401, 259–272. doi: 10.1007/s11104-015-2743-7
- Zhou, C., Ma, Z., Zhu, L., Xiao, X., Xie, Y., Zhu, J., et al. (2016a). Rhizobacterial strain *Bacillus megaterium* BOFC15 induces cellular polyamine changes that improve plant growth and drought resistance. *Int. J. Mol. Sci.* 17, 976. doi: 10.3390/ijms17060976
- Zimin, A. V., Marçais, G., Puiu, D., Roberts, M., Salzberg, S. L., and Yorke, J. A. (2013). The MaSuRCA genome assembler. *Bioinformatics* 29, 2669–2677. doi: 10.1093/bioinformatics/btt476