

# Draft genome sequence of multidrug-resistant *Bradyrhizobium* sp. BL isolated from a sewage treatment plant in China

Peng Bao (✉ [pbao@iue.ac.cn](mailto:pbao@iue.ac.cn))

Institut of urban environment <https://orcid.org/0000-0002-8070-0589>

Guo-Xiang Li

institut of urban environment

Yu-Qin He

institut of urnam environment

Yi Dai

Ningbo beilun water affairs

---

## Short report

**Keywords:** multidrug-resistant, *Bradyrhizobium*, municipal wastewater treatment plant, ecological versatility

**Posted Date:** November 27th, 2019

**DOI:** <https://doi.org/10.21203/rs.2.17832/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

The genus *Bradyrhizobium* is considered to be widespread and abundant group of symbiotic bacteria in many plant-soil ecosystems. However, the ecological versatility of this phylogenetic group remains highly understudied in man-made ecosystems, mainly due to the lack of pure cultures and genomic data. To further expand our understanding of this genus for human health, we analyzed the high quality draft genome of *Bradyrhizobium* strain BL, isolated from a municipal wastewater treatment plant in Ningbo, China. The *Bradyrhizobium* sp. BL draft genome has a total size of 7,718,431 bp with an overall G + C content of 46.43%. From a total of 7236 predicted sequences, 7176 and 60 are protein and RNA coding sequences, respectively. Moreover, 63.51% of the predicted genes were assigned into to Clusters of Orthologous Groups (COG) functional categories. The *Bradyrhizobium* sp. BL genome contains various defense mechanisms against antibiotics that up to predicted 60 antibiotic resistance coding genes. The *Bradyrhizobium* sp. BL genome contains 237 termed virulence factors coding genes which show its potential pathogenicity. This study provides important insights into the genomic diversity of the genus *Bradyrhizobium* and provides a foundation for future comparative genomic studies that will generate a better understanding of the antibiotic resistance process.

## Introduction

Antibiotics are hailed as the greatest harvest of the medical profession in the 20th century. They have been widely used in medical, agricultural and animal husbandry since they were discovered. However, long-term excessive use of antibiotics or abuse of antibiotics will make a huge difference to ourselves and our living environment. The threat also leads to the widespread spread of antibiotic-resistant bacteria [1]. At present, a large number of studies have found that resistant microorganisms are widely spread in water, soil, air and other media and in animals and plants [2, 3].

The sewage treatment plant is a gathering place for urban sewage and wastewater. The antibiotic pollutants discharged from various pollution sources are collected into urban sewage treatment plants through different channels [4, 5], some of the resistant bacteria in the sewage are removed by the treatment technology, and the other part of the resistant bacteria are discharged into the environment through the water outlet of the sewage treatment plant, causing secondary impact on the urban environment such as natural water. Therefore, understanding and studying the resistant microorganisms in the effluent of sewage treatment plants is particularly important for the biological risk assessment and control of wastewater treatment and reuse [4].

In this experiment, we carried out water sampling research from the effluent of Beilun Yandong Wastewater Treatment Plant in Ningbo, and isolated and identified a kind of resistance microorganisms of sulfamethoxazole, trimethoprim and erythromycin. We performed the whole genome sequencing of this resistant microorganism isolated from the sewage treatment plant and determined that the resistant microorganism belongs to *Bradyrhizobium*. The dataset has been submitted to the European Nucleotide Archive (ENA) and is reported here.

# Materials And Methods

Isolation of the bacterial strain.

Sewage sample was diluted 30 times and pipetted 100  $\mu$ L diluent in YEP-Agar medium (Yeast extract 5.0 g/L, peptone 10.0 g/L, NaCl 5.0 g/L, and agar 15.0 g/L) for 72 hours at 30°C. The growing bacterial colonies were sub-cultured. Here, we selected and isolated sulfamethoxazole (8.0 mg/L), trimethoprim (4.0 mg/L) and erythromycin (4.0 mg/L) resistant bacteria. After three generation purifications, a purified isolate BL was obtained.

Genome sequencing and assembly.

We got water samples containing resistant microorganisms from a sewage treatment plant outlet, then extracted strains and genomic DNA from water samples. Genomic DNA was isolated by using a commercial kit (Thermo Fisher). The quality of isolated DNA was checked using a Qubit fluorimeter (Thermo Fisher). Genomic DNA sequencing were performed using Illumina Hiseq™ to obtain raw image data, which were then converted to the original sequencing reads by CASAVA Base Calling analysis. The raw data quality value and other information were statistically analyzed, and the quality of the sampled data was visually evaluated using FastQC, then the data was processed by Trimmomatic to obtain clean data [6]. For the clean data that have been obtained, the SPAdes was used to splice [7], and the contigs complement GAP obtained by splicing were used by GapFiller [8]. We used PrInSeS-G for sequence correction, correcting clipping errors during splicing and insertion of small fragments [9].

Gene prediction.

Prokka is a collection of genetic element prediction tools that call Prodigal to predict coding genes, Aragorn predicts tRNA, RNAmmer predicts rRNA, and Infernal predicts miscRNA [10]. The predicted gene elements are eventually aggregated and completed the preliminary note. Meanwhile, we used RepeatModerler to perform the repeated sequence prediction of the assembly results, and then used RepeatMasker to find the position and frequency of each type of repeat sequence on the genome segment.

Genome annotation and protein classification.

We used NCBI Blast to compare gene protein sequences with multiple databases to obtain functional annotation information [11]. These databases are VFDB, CARD, PHI-base and other virulence factors and drug resistance databases as well as CDD, KOG, COG, NR, NT, PFAM, Swissprot, TrEMBL, KEGG. At the same time, we used HMMER3 to compare the gene protein sequence with the CAZy database to obtain functional annotation information [12]. In addition, we used TMHMM for transmembrane protein prediction analysis [13], SignalP for signal peptide prediction analysis [14], LipoP for lipoprotein prediction analysis [15], ProtCamp for protein subcellular localization analysis, IslandPath-DIOMB for genomic island prediction analysis, and PhiSpy for prophage prediction analysis.

Comparative genomics and phylogenetic analysis.

We used NCBI Blast to compare the predicted 16S rRNA gene sequence with the NCBI database [11], obtained the information of the homologous strain, and constructed the phylogenetic tree. The phylogenetic tree was built using neighbor-joining method [16] and the evolutionary distances were computed using the Maximum Composite Likelihood method [17] available on MEGA7 phylogenetic suite [18]. The PGAP was used to perform pan-genome analysis and homologous gene cluster analysis based on the homologous strain gene information [19], and a phylogenetic tree was constructed based on the homologous gene.

## Results And Discussion

### Genome sequencing information Genome project history.

The genus *Bradyrhizobium* within the Alphaproteobacteria is traditionally associated with legumes and was proposed by Jordan [20] for all slow-growing strains from the genus *Rhizobium*. *Bradyrhizobium* species are symbiotic bacteria inducing the formation of nitrogen-fixing nodules on the roots of plants [21], and could be resistance to different antibiotics [22]. Because *Bradyrhizobium* sp. BL was isolated from wastewater sewage treatment plant, its genome was expected to yield insights that microbiological characteristics of antibiotic resistance in wastewater treatment. The draft genome sequence of *Bradyrhizobium* sp. BL was completed on 27/07/2013. The high-quality draft genome is available in the European Nucleotide Archive (ENA) under study accession number PRJEB34501. Phylogenetic analysis of strain BL built by 16sr DNA sequences, compared with representative *Bradyrhizobium* species strains showed that it was a novel specie of the genus *Bradyrhizobium* (Figure 1). This result confirmed by pan-genome based phylogenetic analysis of *Bradyrhizobium* sp. BL, based on the homologous strains gene information (Figure 2).

### *Genome properties.*

The *Bradyrhizobium* sp. BL draft genome consists of 61,051,394 reads and has a total size of 7,718,431 bp with an overall G + C content of 46.43% (Table 1). From a total of 7236 predicted sequences, 7176 and 60 are proteins and RNA coding sequences, respectively. Moreover, 63.51% of the predicted genes were assigned into to Clusters of Orthologous Groups (COG) functional categories (Table 2).

### *Insights from the genome sequence.*

*Bradyrhizobium* sp. BL shares a core genome including 2061 orthologous protein clusters with 15 representative *Bradyrhizobium* species strains (Figure 3). Moreover, *Bradyrhizobium* sp. BL features 550 unique genes (Figure 3). The *Bradyrhizobium* sp. BL genome contains various defense mechanisms against antibiotics that up to predicted 60 antibiotic resistance coding genes. *Bradyrhizobium* sp. BL encodes multidrug resistance proteins (MdtA, MdtB, MdtC), multidrug efflux pump subunit (AcrB), multidrug efflux pump subunit (AcrB), and efflux pump membrane transporter (BepE, BepG), which keep it

alive under stress of sulfamethoxazole, trimethoprim and erythromycin (supplementary information table 1 and 2). Moreover, *Bradyrhizobium* sp. BL genome contains 237 termed virulence factors coding genes which show its potential pathogenicity (supplementary information table 3 and 4).

## Conclusions

In this study, the genome of *Bradyrhizobium* sp. BL was analyzed, demonstrating that *Bradyrhizobium* sp. BL shares a significant amount of genomic features with other representatives of the genus. The results regarding growth conditions and antibiotic resistance of *Bradyrhizobium* sp. BL clearly suggest a niche differentiation between different species of *Bradyrhizobium* genus. The *Bradyrhizobium* sp. BL genome will facilitate a better understanding of the metabolic versatility of the genus *Bradyrhizobium* and will be useful for future comparative studies, especially those with a focus on antibiotic resistance.

## Abbreviations

Mdt

multidrug resistance proteins; Acr:multidrug efflux pump subunit; Bep:efflux pump membrane transporter.

## Declarations

### Additional file

supplementary information table 1: Summary table of notes on drug resistance functional proteins.

supplementary information table 2: Comment on The Comprehensive Antibiotic Resistance Database protein table. supplementary information table 3: Summary table of notes on Virulence Factors.

supplementary information table 4: Comment on Virulence Factors of Pathogenic Bacteria data base.

### Abbreviations

Mdt: multidrug resistance proteins; Acr: multidrug efflux pump subunit; Bep: efflux pump membrane transporter.

### Acknowledgements

Not available.

### Authors' contributions

Guo-Xiang Li did the bacterium isolation and identification. Guo-Xiang Li and Yu-Qin He did the draft of manuscript. Guo-Xiang Li, Yu-Qin He and Yi Dai did the assembly and annotation. Peng Bao supervised the study. All authors discussed, revised and approved the final manuscript.

### Funding

This research was financially supported by Ningbo Science and Technology People-Benefit Project (No. 2017C50009), Ningbo Leading Talent Program.

### Competing interests

The authors declare that they have no competing interests. The data used to support the findings of this study are included within the article and the supplementary information file.

### Author details

<sup>1</sup> Key Lab of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen, 361021, P. R. China;

<sup>2</sup> Ningbo Urban Environment Observation and Station, Chinese Academy of Sciences, Ningbo, 315800, P. R. China;

<sup>3</sup> University of Chinese Academy of Sciences, Beijing, 100049, P. R. China;

<sup>4</sup> Ningbo Beilun water affairs limited.

## References

[1] Martínez, J.L., "Environmental pollution by antibiotics and by antibiotic resistance determinants," *Environmental Pollution*, vol. 157, no. 11, pp. 2893–2902, 2009.

[2] Andersson, D.I., "Persistence of antibiotic resistant bacteria," *Current opinion in microbiology*, vol. 6, no. 5, pp. 452–456, 2003.

[3] Pruden A, Pei R, Storteboom H, Carlson KH, "Antibiotic resistance genes as emerging contaminants: studies in Northern Colorado," *Environmental Science & Technology*, vol. 40, no. 23, pp. 7445–7450, 2006.

[4] Lindsey, M., Meyer, M., Thurman, E., "Analysis of trace levels of sulfonamide and tetracycline antimicrobials in ground water and surface water using solid phase extraction and liquid chromatography/mass spectrometry," *Analytical Chemistry*, vol. 73, no. 19, pp. 4640–4646, 2001.

[5] Lindberg, R., Jarnheimer, P.-Å., Olsen, B., Johansson, M., Tysklind, M., "Determination of antibiotic substances in hospital sewage water using solid phase extraction and liquid chromatography/mass spectrometry and group analogue internal standards," *Chemosphere*, vol. 57, no. 10, pp. 1479–1488, 2004.

- [6] Bolger AM, Lohse M, Usadel B., "Trimmomatic: a flexible trimmer for Illumina sequence data," *Bioinformatics*, vol. 30, no. 15, pp. 2114–2120, 2014.
- [7] Bankevich A, Nurk S, Antipov D, et al., "SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing," *Journal of Computational Biology*, vol. 19, no. 5, pp. 455–477, 2012.
- [8] Boetzer M, Pirovano W, "Toward almost closed genomes with GapFiller," *Genome Biology*, vol. 13, no. 6, DOI. R56, 2012.
- [9] Massouras A, Hens K, Gubelmann C, et al., "Primer-initiated sequence synthesis to detect and assemble structural variants," *Nat Methods*, vol. 7, no. 7, pp. 485–486, 2010.
- [10] Seemann T., "Prokka: rapid prokaryotic genome annotation," *Bioinformatics*, vol. 30, no. 14, pp.2068–2069, 2014.
- [11] Altschul SF, Madden TL, Schäffer AA, et al., "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs," *Nucleic Acids Research*, vol. 25, no. 17, pp. 3389–3402, 1997.
- [12] Eddy SR., "A new generation of homology search tools based on probabilistic inference," *Genome Informatics*, vol. 23, no. 1, pp. 205–211, 2009.
- [13] Moller, S. Croning, M.D.R. Apweiler, R., "Evaluation of methods for the prediction of membrane spanning regions," *Bioinformatics*, vol. 17, no. 7, pp. 646–653, 2001.
- [14] Petersen TN, Brunak S, von Heijne G, Nielsen H., "SignalP 4.0: discriminating signal peptides from transmembrane regions," *Nat Methods*, vol. 8, no. 10, pp. 785–786, 2011.
- [15] Juncker AS, Willenbrock H, von Heijne G, Brunak S, Nielsen H, Krogh A., "Prediction of lipoprotein signal peptides in Gram-negative bacteria," *Protein Science: A Publication of the Protein Society*, vol. 12, no. 8, pp. 1652–1662, 2003.
- [16] Saitou N, Nei M., "The neighbor-joining method: a new method for reconstructing phylogenetic trees," *Molecular Biology and Evolution*, vol. 4, no. 4, pp. 406–425, 1987.
- [17] Tamura K, Nei M, Kumar S., "Prospects for inferring very large phylogenies by using the neighbor-joining method," *Proceedings of the National Academy of Sciences of The United States of America*, vol. 101, no. 30, pp. 11030–11035, 2004.
- [18] Kumar S, Stecher G, Tamura K., "MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets," *Molecular Biology and Evolution*, vol. 33, no. 7, pp. 1870–1874, 2016.
- [19] Zhao Y, Wu J, Yang J, Sun S, Xiao J, Yu J., "PGAP: pan-genomes analysis pipeline," *Bioinformatics*, vol. 28, no. 3, pp. 416–418, 2012.

[20] Jordan, D.C., "Transfer of *Rhizobium japonicum* Buchanan 1980 to *Bradyrhizobium* gen. nov., a genus of slow growing root nodule bacteria from leguminous plants," *International Journal of Systematic and Evolutionary Microbiology*, vol. 32, no. 1, pp. 136–139, 1982.

[21] Tan, Z. Hurek, T. Vinuesa, P. Muller, P. Ladha, J.K. Reinhold-Hurek, B., "Specific detection of *Bradyrhizobium* and *Rhizobium* strains colonizing rice (*Oryza sativa*) roots by 16S-23S ribosomal DNA intergenic spacer-targeted PCR," *Applied and Environmental Microbiology*, vol. 67, no. 8, pp. 3655–3664, 2001.

[22] Date R, Hurse L., "Growth, competitiveness and effectiveness of spontaneous antibiotic-resistant strains of *Bradyrhizobium* for *Desmodium intortum* CV Greenleaf," *Soil Biology and Biochemistry*, vol. 24, no. 1, pp. 33–39, 1992.

## Tables

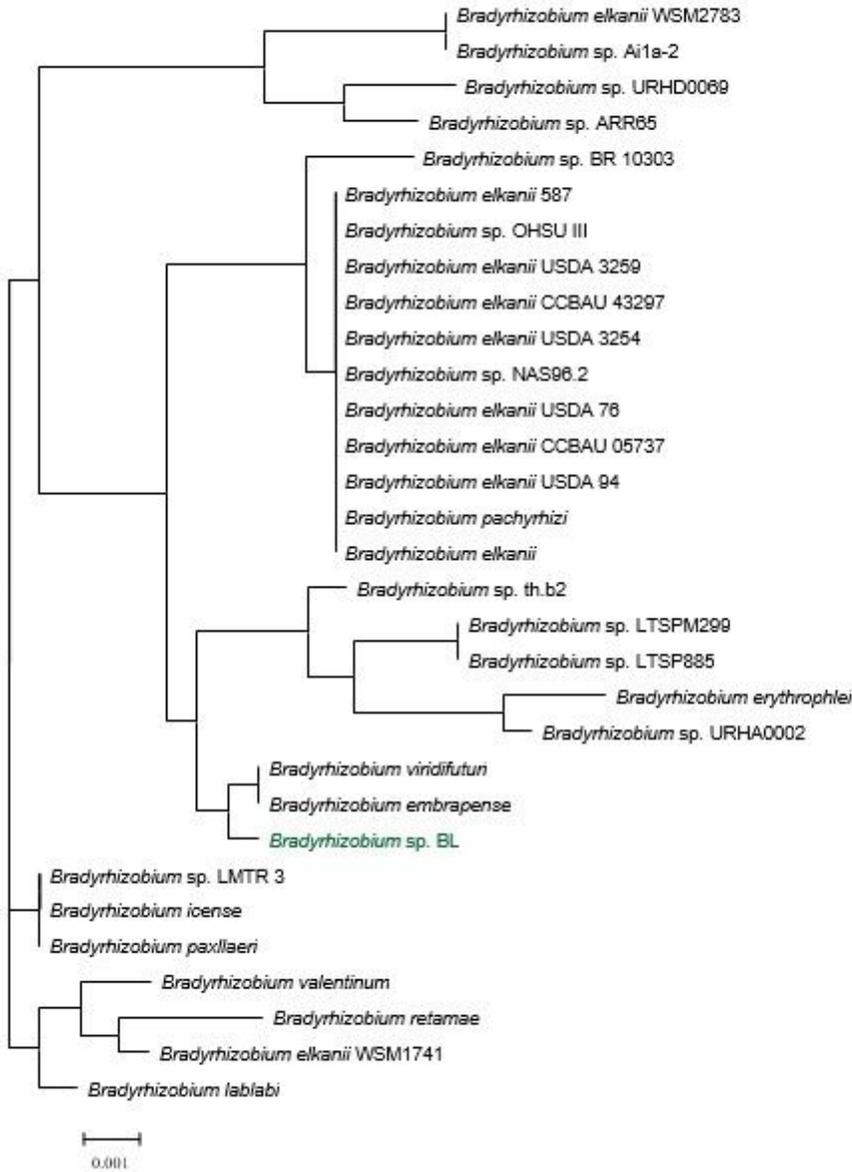
Table 1. *Bradyrhizobium* sp. BL genome characteristics.

S. No	Name	Genome characteristics and resources
1	ENA Genome Accession Number	PRJEB34501
2	Sequence type	Illumina Hiseq
3	Total number of Reads	61,051,394
4	Read length	148.54
5	Overall coverage	>100×
6	Mapped reads	86.80%
7	Estimated genome size	7,718,431 bp
8	GC content	64.43%
9	Protein coding genes	7176
10	tRNA coding genes	57
11	rRNA coding genes	3

Table 2. Summary of the average number of genes and percentage of each genome representing each COG functional category associated with *Bradyrhizobium* sp. BL evaluated in this study.

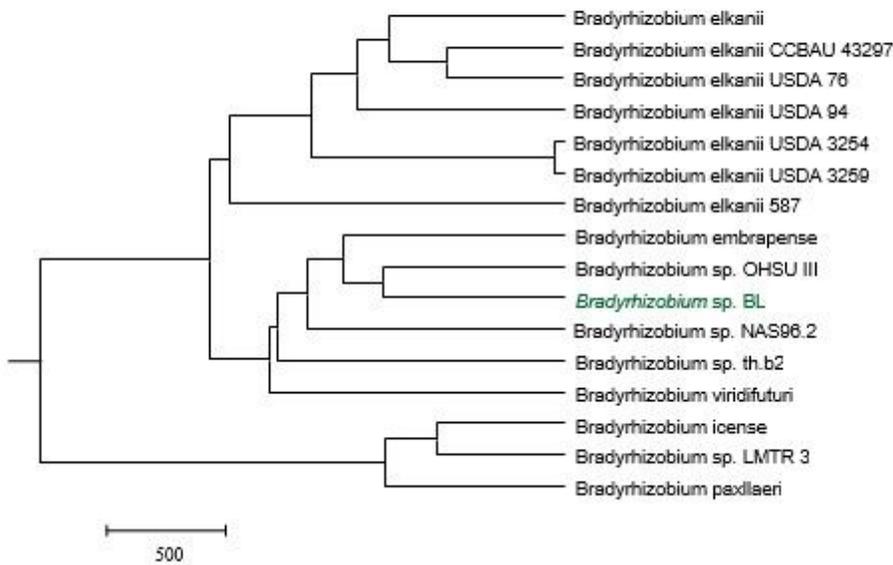
Code	Value	%age	Description
C	354	6.97	Energy production and conversion
D	30	0.58	Cell cycle control, cell division, chromosome partitioning
E	536	10.28	Amino acid transport and metabolism
F	64	1.23	Nucleotide transport and metabolism
G	336	6.45	Carbohydrate transport and metabolism
H	179	3.43	Coenzyme transport and metabolism
I	251	4.81	Lipid transport and metabolism
J	193	3.70	Translation, ribosomal structure and biogenesis
K	458	8.79	Transcription
L	127	2.44	Replication, recombination and repair
M	244	4.68	Cell wall/membrane/envelope biogenesis
N	27	0.52	Cell motility
O	202	3.87	Posttranslational modification, protein turnover, chaperones
P	305	5.85	Inorganic ion transport and metabolism

## Figures



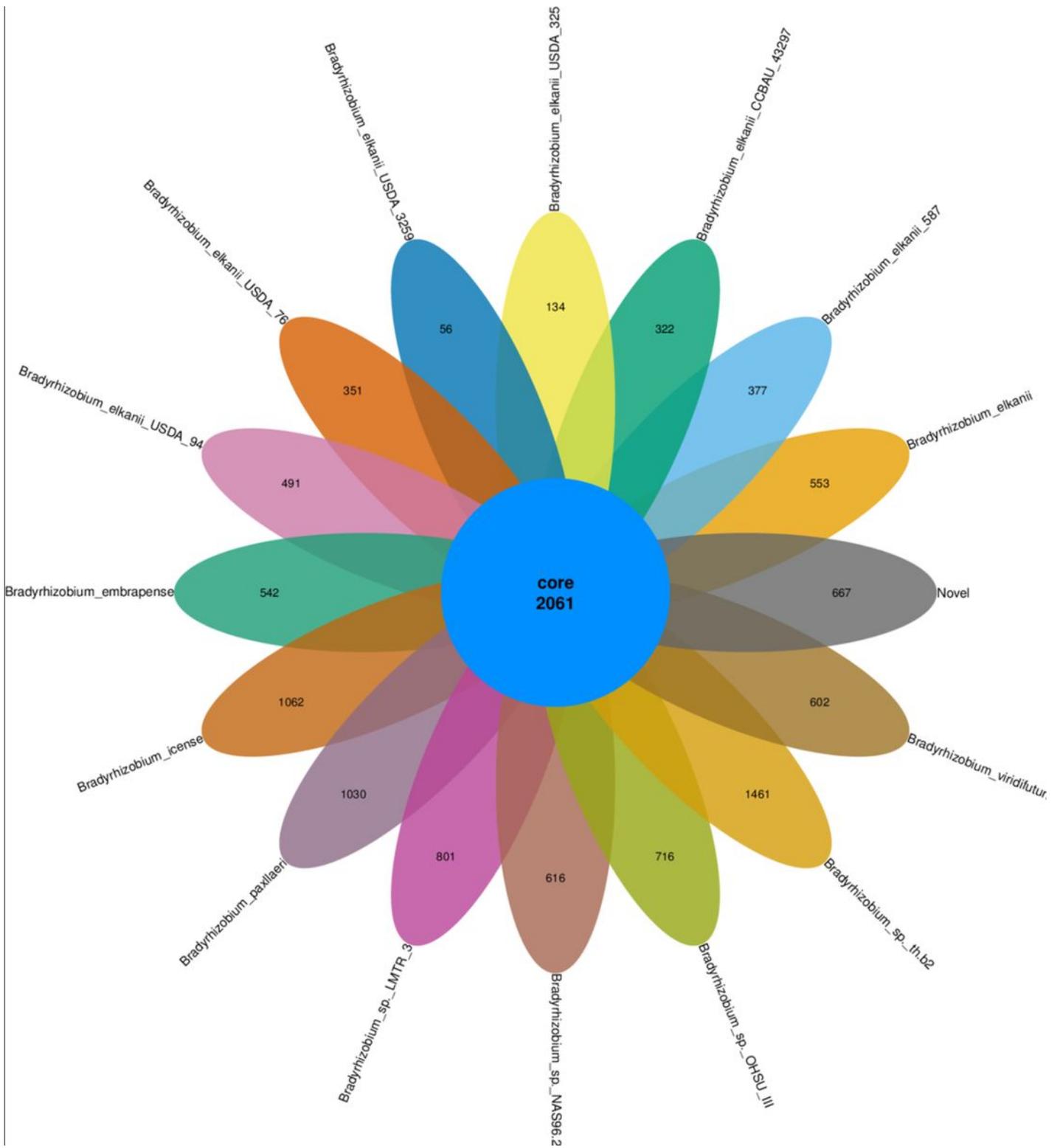
**Figure 1**

Phylogenetic analysis of *Bradyrhizobium* sp. BL, compared with representative *Bradyrhizobium* species strains. The phylogenetic tree was built using neighbor-joining method and the evolutionary distances were computed using the Maximum Composite Likelihood method available on MEGA7 phylogenetic suite. The scale bar indicates 0.001 substitutions per nucleotide position.



**Figure 2**

Pan-genome based phylogenetic analysis of *Bradyrhizobium sp. BL*, compared with representative *Bradyrhizobium* species strains. The phylogenetic tree was built using neighbor-joining method and the evolutionary distances were computed using the Maximum Composite Likelihood method available on MEGA7 phylogenetic suite. The scale bar indicates 500 substitutions per nucleotide position.



**Figure 3**

Flower plots of homologous genes of *Bradyrhizobium* sp. BL, compared with representative *Bradyrhizobium* species strains. Here, “novel” represents *Bradyrhizobium* sp. BL.