

5-2015

High-quality permanent draft genome sequence of *Bradyrhizobium* sp Th.b2, a microsymbiont of *Amphicarpaea bracteata* collected in Johnson City, New York

Rui Tian

Matthew Parker

Binghamton University--SUNY

Rekha Seshadri

T. B.K. Reddy

Victor Markowitz

See next page for additional authors

Follow this and additional works at: https://orb.binghamton.edu/bio_fac

 Part of the [Biology Commons](#)

Recommended Citation

Tian, Rui; Parker, Matthew; Seshadri, Rekha; Reddy, T. B.K.; Markowitz, Victor; Ivanova, Natalia; Pati, Amrita; Woyke, Tanja; Baeshen, Mohammed; Baeshen, Nabih; Kyrpides, Nikos; and Reeve, Wayne, "High-quality permanent draft genome sequence of *Bradyrhizobium* sp Th.b2, a microsymbiont of *Amphicarpaea bracteata* collected in Johnson City, New York" (2015). *Biological Sciences Faculty Scholarship*. 10.

https://orb.binghamton.edu/bio_fac/10

This Article is brought to you for free and open access by the Biological Sciences at The Open Repository @ Binghamton (The ORB). It has been accepted for inclusion in Biological Sciences Faculty Scholarship by an authorized administrator of The Open Repository @ Binghamton (The ORB). For more information, please contact ORB@binghamton.edu.

Authors

Rui Tian, Matthew Parker, Rekha Seshadri, T. B.K. Reddy, Victor Markowitz, Natalia Ivanova, Amrita Pati, Tanja Woyke, Mohammed Baeshen, Nabih Baeshen, Nikos Kyrpides, and Wayne Reeve



SHORT GENOME REPORT

Open Access

High-quality permanent draft genome sequence of *Bradyrhizobium* sp. Th.b2, a microsymbiont of *Amphicarpaea bracteata* collected in Johnson City, New York

Rui Tian¹, Matthew Parker², Rekha Seshadri³, TBK Reddy³, Victor Markowitz⁴, Natalia Ivanova³, Amrita Pati³, Tanja Woyke³, Mohammed N Baeshen^{6,7}, Nabih A Baeshen^{5,6}, Nikos Kyrpides^{3,5} and Wayne Reeve^{1*}

Abstract

Bradyrhizobium sp. Th.b2 is an aerobic, motile, Gram-negative, non-spore-forming rod that was isolated from an effective nitrogen-fixing root nodule of *Amphicarpaea bracteata* collected in Johnson City, New York. Here we describe the features of *Bradyrhizobium* sp. Th.b2, together with high-quality permanent draft genome sequence information and annotation. The 10,118,060 high-quality draft genome is arranged in 266 scaffolds of 274 contigs, contains 9,809 protein-coding genes and 108 RNA-only encoding genes. This rhizobial genome was sequenced as part of the DOE Joint Genome Institute 2010 Genomic Encyclopedia for Bacteria and Archaea-Root Nodule Bacteria (GEBA-RNB) project.

Keywords: Root-nodule bacteria, Nitrogen fixation, Symbiosis, *Alphaproteobacteria*, GEBA-RNB

Introduction

Strain Th.b2 is a representative of a widely distributed *Bradyrhizobium* lineage used by several common legumes indigenous to forested habitats in eastern North America. Strain Th.b2 was sampled in 1991 from a population of the annual legume *Amphicarpaea bracteata* in Johnson City, NY. Surveys of other *A. bracteata* populations in the eastern United States based on 20 isozyme markers found that strains similar or identical to Th.b2 were present in 19 of 24 sites across six states (IL, IN, WI, MI, NY, PA [1]). Based on both isozyme data and rRNA sequencing, isolates that were similar or identical to Th.b2 were also detected in nodule samples from two common herbaceous perennial legumes, *Apios americana* and *Hylodesmum glutinosum*, that often occur in woodland habitats together with *Amphicarpaea bracteata* [2]. A multilocus sequence analysis found strains in North Carolina populations of *A. bracteata* that were similar or identical to Th.b2 [3], and also detected a highly similar strain on another herbaceous

perennial legume, *Desmodium paniculatum*, that is widely distributed across eastern North America [4].

Based on these field surveys, the *Bradyrhizobium* lineage represented by strain Th.b2 appears to be relatively host-specific to legumes in these four genera (*Amphicarpaea*, *Apios*, *Desmodium*, *Hylodesmum*), because widespread sampling of sympatric legumes in eleven other genera have not detected this group [3,5,6]. However, inoculation experiments are needed to understand whether the Th.b2 lineage lacks the ability to nodulate these other genera, or alternatively, may simply be a poor competitor for nodulation in the presence of other bacterial strains that are their preferred symbionts.

It should also be noted that the eastern North American symbionts of *Amphicarpaea*, *Apios*, *Desmodium* and *Hylodesmum* are not phylogenetically homogeneous at housekeeping loci. Horizontal transfer of the symbiosis island (SI) region of the *Bradyrhizobium* chromosome [7] from a member of the Th.b2 clade to a distantly related *Bradyrhizobium* lineage has apparently enabled the recipient to gain the ability to interact with some of the normal legume hosts of the Th.b2 clade [3].

* Correspondence: W.Reeve@murdoch.edu.au

¹Centre for Rhizobium Studies, Murdoch University, Murdoch, Australia
Full list of author information is available at the end of the article

Bacteria that are closely related to Th.b2 have also been found in Japan associated with an Asian species of *Amphicarpaea* (*A. edgeworthii*) [6]. Surprisingly, strain Th.b2 lacks the ability to form nodules on *A. edgeworthii*, although Japanese strains from *A. edgeworthii* are effective nitrogen-fixing symbionts for the American legume *A. bracteata* [8,9]. These differences appear to be related to variation between related East Asian and North American strains in the synthesis of rhizobitoxine [8].

Here we provide an analysis of the high-quality permanent draft genome sequence of *Bradyrhizobium* sp. Th.b2, one of the rhizobial genomes sequenced as part of the DOE Joint Genome Institute 2010 Genomic Encyclopedia for Bacteria and Archaea-Root Nodule Bacteria (GEBA-RNB) project proposal [10], whose properties may provide useful insights about the evolution of symbiotic specificity and its relationship to SI region horizontal transfer in *Bradyrhizobium*.

Organism information

Classification and features

Bradyrhizobium sp. Th.b2 is a motile, non-sporulating, non-encapsulated, Gram-negative strain in the order *Rhizobiales* of the class *Alphaproteobacteria*. The rod shaped form has dimensions of approximately 0.5 μm in width and 1.5-2.0 μm in length (Figure 1 Left and Center). It is relatively slow growing, forming colonies after 6–7 days when grown on half strength Lupin Agar ($\frac{1}{2}$ LA) [11], tryptone-yeast extract agar (TY) [12] or a modified yeast-mannitol agar (YMA) [13] at 28°C. Colonies on $\frac{1}{2}$ LA are opaque, slightly domed and moderately mucoid with smooth margins (Figure 1 Right).

Figure 2 shows the phylogenetic relationship of *Bradyrhizobium* sp. Th.b2 in a 16S rRNA gene sequence based tree. This strain is phylogenetically most closely related to the type strains *Bradyrhizobium icense* LMTR 13^T and *Bradyrhizobium paxllaeri* LMTR 21^T, with a 16S rRNA gene sequence identity of 99.77% to the

corresponding gene sequence of each type strain based on alignment using the EzTaxon-e server [14,15].

Minimum Information about the Genome Sequence (MIGS) is provided in Table 1 and Additional file 1: Table S1.

Symbiotaxonomy

Strain Th.b2 was isolated in 1991 from a population of the annual legume *Amphicarpaea bracteata* in Johnson City, NY. Isolates that were similar or identical to Th.b2 were also detected in nodule samples from two common herbaceous perennial legumes, *Apios americana* and *Hylodesmum glutinosum*, that often occur in woodland habitats together with *Amphicarpaea bracteata* [2]. Th.b2 lacks the ability to form nodules on the Asian species *Amphicarpaea. edgeworthii*, which is associated with a strain closely related to Th.b2 from Japan [6,8].

Genome sequencing information

Genome project history

This organism was selected for sequencing on the basis of its environmental and agricultural relevance to issues in global carbon cycling, alternative energy production, and biogeochemical importance, and is part of the Genomic Encyclopedia of *Bacteria* and *Archaea*, Root Nodulating Bacteria (GEBA-RNB) project at the U.S. Department of Energy, Joint Genome Institute (JGI). The genome project is deposited in the Genomes On-Line Database [16] and a high-quality permanent draft genome sequence in IMG [17]. Sequencing, finishing and annotation were performed by the JGI using state of the art sequencing technology [18]. A summary of the project information is shown in Table 2.

Growth conditions and genomic DNA preparation

Bradyrhizobium sp. Th.b2 was cultured to mid logarithmic phase in 60 ml of TY rich media on a gyratory shaker at 28°C [19]. DNA was isolated from the cells using a CTAB (Cetyl trimethyl ammonium bromide) bacterial genomic DNA isolation method [20].

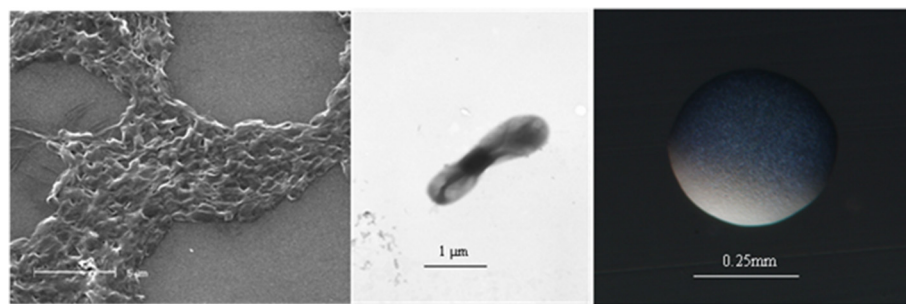


Figure 1 Images of *Bradyrhizobium* sp. Th.b2 using scanning (Left) and transmission (Center) electron microscopy as well as light microscopy to visualize colony morphology on solid media (Right).

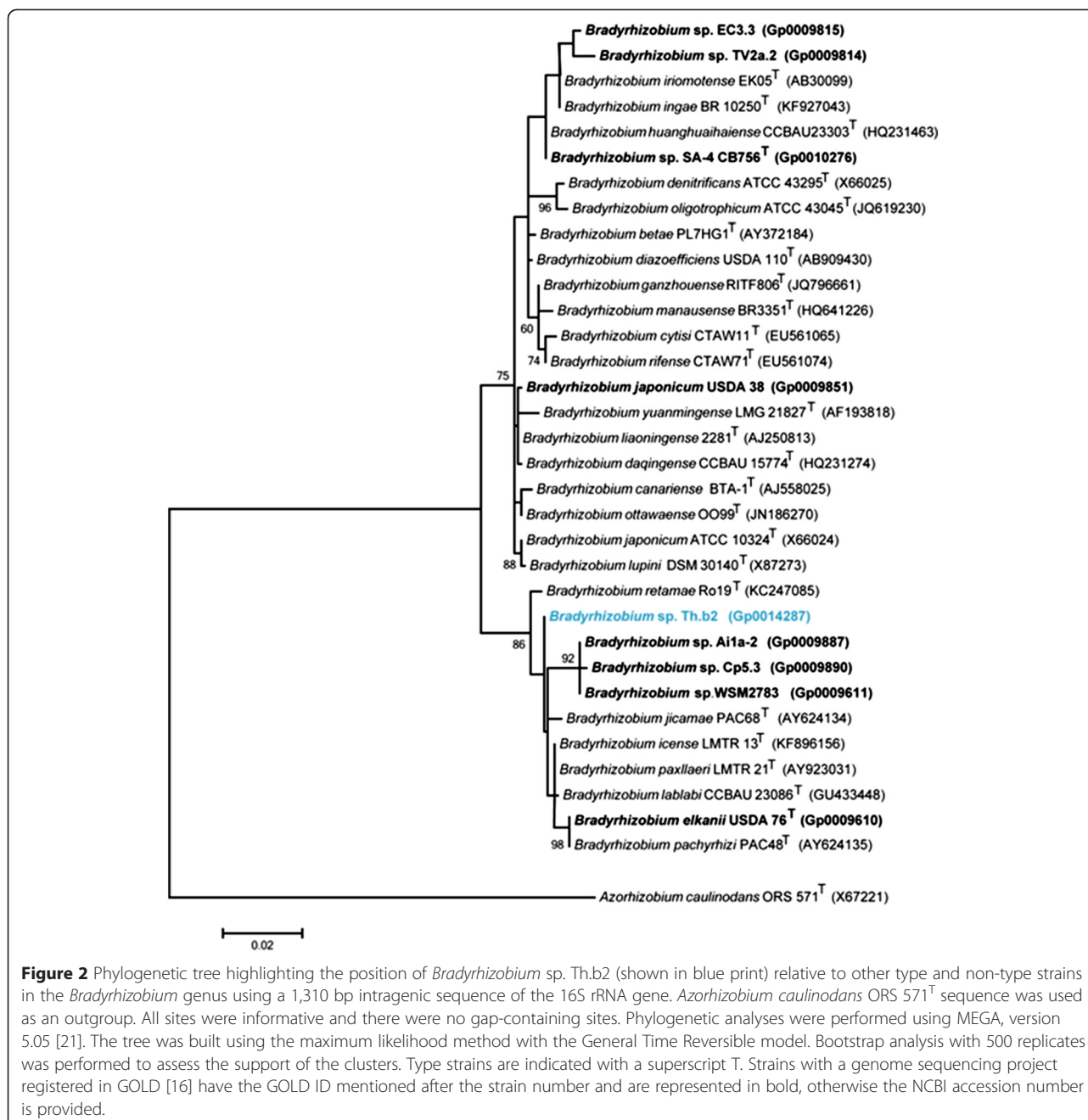


Figure 2 Phylogenetic tree highlighting the position of *Bradyrhizobium* sp. Th.b2 (shown in blue print) relative to other type and non-type strains in the *Bradyrhizobium* genus using a 1,310 bp intragenic sequence of the 16S rRNA gene. *Azorhizobium caulinodans* ORS 571^T sequence was used as an outgroup. All sites were informative and there were no gap-containing sites. Phylogenetic analyses were performed using MEGA, version 5.05 [21]. The tree was built using the maximum likelihood method with the General Time Reversible model. Bootstrap analysis with 500 replicates was performed to assess the support of the clusters. Type strains are indicated with a superscript T. Strains with a genome sequencing project registered in GOLD [16] have the GOLD ID mentioned after the strain number and are represented in bold, otherwise the NCBI accession number is provided.

Genome sequencing and assembly

The draft genome of *Bradyrhizobium* sp. th.b2 was generated at the DOE Joint Genome Institute (JGI) using the Illumina technology [22]. An Illumina standard shotgun library was constructed and sequenced using the Illumina HiSeq 2000 platform which generated 20,348,156 reads totaling 3,052.2 Mbp. All general aspects of library construction and sequencing were performed at the JGI and details can be found on the JGI website [23]. All raw Illumina sequence data was passed through DUK, a filtering program developed at JGI,

which removes known Illumina sequencing and library preparation artifacts (Mingkun L, Copeland A, Han J, Unpublished). Following steps were then performed for assembly: (1) filtered Illumina reads were assembled using Velvet (version 1.1.04) [24], (2) 1–3 Kbp simulated paired end reads were created from Velvet contigs using wgsim [25], (3) Illumina reads were assembled with simulated read pairs using Allpaths-LG (version r42328) [26]. Parameters for assembly steps were: 1) Velvet (velveth: 63 –shortPaired and velvetg: –very clean yes –exportFiltered yes –min contig lgth 500 –scaffolding no –cov cutoff 10) 2) wgsim (–e 0 –l

Table 1 Classification and general features of *Bradyrhizobium* sp. Th.b2 in accordance with the MIGS recommendations [27] published by the Genome Standards Consortium [28]

MIGS ID	Property	Term	Evidence code
	Classification	Domain <i>Bacteria</i>	TAS [29]
		Phylum <i>Proteobacteria</i>	TAS [30,31]
		Class <i>Alphaproteobacteria</i>	TAS [31,32]
		Order <i>Rhizobiales</i>	TAS [33]
		Family <i>Bradyrhizobiaceae</i>	TAS [34]
		Genus <i>Bradyrhizobium</i>	TAS [35]
		Species <i>Bradyrhizobium</i> sp.	IDA
	Gram stain	Negative	IDA
	Cell shape	Rod	IDA
	Motility	Motile	IDA
	Sporulation	Non-sporulating	NAS
	Temperature range	Mesophile	NAS
	Optimum temperature	28°C	NAS
	pH range; Optimum	Unknown	NAS
	Carbon source	Varied	NAS
	Energy source	Chemoorganotroph	NAS
MIGS-6	Habitat	Soil, root nodule, host	TAS [1]
MIGS-6.3	Salinity	Non-halophile	NAS
MIGS-22	Oxygen requirement	Aerobic	NAS
MIGS-15	Biotic relationship	Free living, symbiotic	TAS [1]
MIGS-14	Pathogenicity	Non-pathogenic	NAS
	Biosafety level	1	TAS [36]
	Isolation	Root nodule of <i>Amphicarpaea bracteata</i>	TAS [1]
MIGS-4	Geographic location	Johnson City, New York	TAS [1]
MIGS-5	Sample collection date	1991	IDA
MIGS-4.1	Latitude	42.107	IDA
MIGS-4.2	Longitude	-75.9691	IDA
MIGS-4.3	Depth	5 cm	IDA
MIGS-4.4	Altitude	255 m	IDA

Evidence codes – IDA: Inferred from Direct Assay; TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). Evidence codes are from the Gene Ontology project [37,38].

100 -2 100 -r 0 -R 0 -X 0) 3) Allpaths-LG (PrepareAllpathsInputs: PHRED 64 = 1 PLOIDY = 1 FRAG_COVERAGE = 125 JUMP_COVERAGE = 25 LONG_JUMP_COV = 50, RunAllpathsLG: THREADS = 8 RUN = std_shredpairs TARGETS = standard VAPI_WARN_ONLY = True OVERWRITE = True). The final draft assembly contained 274 contigs in 266 scaffolds. The total size of the genome is 10.1 Mbp and the final assembly is based on 1,216.8 Mbp of Illumina data, which provides an average 120.4x coverage of the genome.

Genome annotation

Genes were identified using Prodigal [39], as part of the DOE-JGI genome annotation pipeline [40,41] The predicted

CDSs were translated and used to search the National Center for Biotechnology Information (NCBI) non-redundant database, UniProt, TIGRFam, Pfam, KEGG, COG, and InterPro databases. The tRNAScanSE tool [42] was used to find tRNA genes, whereas ribosomal RNA genes were found by searches against models of the ribosomal RNA genes built from SILVA [43]. Other non-coding RNAs such as the RNA components of the protein secretion complex and the RNase P were identified by searching the genome for the corresponding Rfam profiles using INFERNAL [44]. Additional gene prediction analysis and manual functional annotation was performed within the Integrated Microbial Genomes-Expert Review (IMG-ER) system [45]

Table 2 Project information

MIGS ID	Property	Term
MIGS-31	Finishing quality	High-quality permanent draft
MIGS-28	Libraries used	Illumina Standard PE
MIGS-29	Sequencing platforms	Illumina HiSeq2000
MIGS-31.2	Fold coverage	Illumina, 120.4x
MIGS-30	Assemblers	Velvet version 1.1.04; Allpaths-LG version r42328
MIGS-32	Gene calling method	Prodigal 1.4
	Locus Tag	K359
	GenBank ID	AUGA000000000
	GenBank Date of Release	June 13, 2014
	GOLD ID	Gp0014287 [46]
	BIOPROJECT	195826
MIGS-13	Source Material Identifier	Th.b2
	Project relevance	Symbiotic N ₂ fixation, agriculture

developed by the Joint Genome Institute, Walnut Creek, CA, USA.

Genome properties

The genome is 10,118,060 nucleotides with 63.25% GC content (Table 3) and comprised of 266 scaffolds. From a total of 9,919 genes, 9,809 were protein encoding and 108 RNA only encoding genes. The majority of genes (70.75%) were assigned a putative function whilst the remaining genes were annotated as hypothetical. The distribution of genes into COGs functional categories is presented in Table 4.

Table 3 Genome statistics for *Bradyrhizobium* sp. Th.b2

Attribute	Value	% of Total
Genome size (bp)	10,118,060	100.00
DNA coding (bp)	8,412,367	83.14
DNA G + C (bp)	6,399,174	63.25
DNA scaffolds	266	100
Total genes	9,917	100.00
Protein coding genes	9,809	98.91
RNA genes	108	1.09
Pseudo genes	0	0.00
Genes in internal clusters	713	7.19
Genes with function prediction	7,016	70.75
Genes assigned to COGs	5,576	56.23
Genes with Pfam domains	71.85	72.45
Genes with signal peptides	978	9.86
Genes coding transmembrane helices	2,166	21.84
CRISPR repeats	0	0.00

Table 4 Number of genes associated with the general COG functional categories

Code	Value	% of total (6,228)	COG category
J	199	3.20	Translation, ribosomal structure and biogenesis
A	0	0.00	RNA processing and modification
K	520	8.35	Transcription
L	197	3.16	Replication, recombination and repair
B	3	0.05	Chromatin structure and dynamics
D	30	0.48	Cell cycle control, cell division, chromosome partitioning
V	103	1.65	Defense mechanisms
T	248	3.98	Signal transduction mechanisms
M	290	4.66	Cell wall/membrane/envelope biogenesis
N	72	1.16	Cell motility
U	118	1.89	Intracellular trafficking, secretion, and vesicular transport
O	200	3.21	Posttranslational modification, protein turnover, chaperones
C	432	6.94	Energy production and conversion
G	382	6.13	Carbohydrate transport and metabolism
E	702	11.27	Amino acid transport and metabolism
F	81	1.30	Nucleotide transport and metabolism
H	208	3.34	Coenzyme transport and metabolism
I	391	6.28	Lipid transport and metabolism
P	338	5.43	Inorganic ion transport and metabolism
Q	301	4.83	Secondary metabolite biosynthesis, transport and catabolism
R	799	12.83	General function prediction only
S	614	9.86	Function unknown
-	4,341	43.77	Not in COGS

Conclusions

Bradyrhizobium sp. Th.b2 was isolated from a root nodule of *Amphicarpaea bracteata* collected from Johnson City, New York. Little is currently known of the symbiotic associations of its host *Amphicarpaea bracteata*. This strain belongs to a member of a widely distributed *Bradyrhizobium* lineage, isolated from diverse legume hosts in North, Central and South America and South Africa. Phylogenetically, Th.b2 is separated from the most closely related species *Bradyrhizobium icense* LMTR 13^T and *Bradyrhizobium paxllaeri* LMTR 21^T, both isolated from root nodules of *Phaseolus lunatus* (Lima bean) in Peru [47]. Th.b2 may therefore be a novel species of *Bradyrhizobium*. A total of 25 *Bradyrhizobium* genomes have now been sequenced as part of the GEBA-RNB project [10]. Of these 25 strains, Th.b2 has the second largest genome size (10.1 Mbp), gene count (9,917) and COG % and the lowest coding base count % (83.17). The genome attributes of *Bradyrhizobium* sp.

Th.b2, in conjunction with other *Bradyrhizobium* genomes from GEBA-RNB project, will be important for the understanding of the biogeography of *Bradyrhizobium* spp. interactions required for the successful establishment of effective symbioses with their diverse hosts.

Additional file

Additional file 1: Table S1. Associated MIGS record.

Abbreviations

GEBA-RNB: Genomic Encyclopedia for Bacteria and Archaea-Root Nodule Bacteria; JGI: Joint Genome Institute; ½LA: half strength Lupin Agar; TY: Tryptone Yeast; YMA: Yeast Mannitol Agar; CTAB: Cetyl Trimethyl Ammonium Bromide.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

MP supplied the strain and background information for this project and the DNA to the JGI, TR performed all imaging, TR and WR drafted the paper, MNB and NAB provided financial support and all other authors were involved in sequencing the genome and/or editing the final paper. All authors read and approved the final manuscript.

Acknowledgements

This work was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Berkeley National Laboratory under contract No. DE-AC02-05CH11231. We thank Gordon Thompson (Murdoch University) for the preparation of SEM and TEM photos. We would also like to thank the Center of Nanotechnology at King Abdulaziz University for their support.

Author details

¹Centre for Rhizobium Studies, Murdoch University, Murdoch, Australia. ²Binghamton University, State University of New York, New York, USA. ³DOE Joint Genome Institute, Walnut Creek, California, USA. ⁴Biological Data Management and Technology Center, Lawrence Berkeley National Laboratory, Berkeley, California, USA. ⁵Department of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia. ⁶Center of Nanotechnology, King Abdulaziz University, Jeddah, Saudi Arabia. ⁷Department of Biological Sciences, Faculty of Science, Jeddah University, Jeddah, Saudi Arabia.

Received: 13 February 2015 Accepted: 16 April 2015

Published online: 16 May 2015

References

- Parker MA, Spoerke JM. Geographic structure of lineage associations in a plant-bacterial mutualism. *J Evolution Biol.* 1998;11:549–62.
- Parker MA. Relationships of bradyrhizobia from the legumes *Apios americana* and *Desmodium glutinosum*. *Appl Environ Microbiol.* 1999;65:4914–20.
- Parker MA. Legumes select symbiosis island sequence variants in *Bradyrhizobium*. *Mol Ecol.* 2012;21:1769–78.
- Isely D. The *Desmodium paniculatum* (L.) DC. (*Fabaceae*) complex revisited. *Sida.* 1983;10:142–58.
- Parker MA, Kennedy DA. Diversity and relationships of bradyrhizobia from legumes native to eastern North America. *Can J Microbiol.* 2006;52:1148–57.
- Parker MA. The spread of *Bradyrhizobium* lineages across host legume clades: from Abarema to Zygia. *Microb Ecol.* 2015;69:630–40.
- Kaneko T, Nakamura Y, Sato S, Minamisawa K, Uchiyama T, Sasamoto S, et al. Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110 (supplement). *DNA Res.* 2002;9:225–56.
- Parker MA, Peters NK. Rhizobitoxine production and symbiotic compatibility of *Bradyrhizobium* from Asian and North American lineages of *Amphicarpaea*. *Can J Microbiol.* 2001;47:889–94.
- Parker MA. rRNA and *dnaK* relationships of *Bradyrhizobium* sp. nodule bacteria from four papilionoid legume trees in Costa Rica. *Syst Appl Microbiol.* 2004;27:334–42.
- Reeve W, Ardley J, Tian R, Eshragi L, Yoon JW, Ngamwisetkun P, et al. A genomic encyclopedia of the root nodule bacteria: assessing genetic diversity through a systematic biogeographic survey. *Stand Genomic Sci.* 2015;10:14.
- Howieson JG, Ewing MA, D'antuono MF. Selection for acid tolerance in *Rhizobium meliloti*. *Plant Soil.* 1988;105:179–88.
- Beringer JE. R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol.* 1974;84:188–98.
- Vincent JM. *A Manual for the Practical Study of Root-nodulating Bacteria*. International Biological Programme. UK: Blackwell Scientific Publications, Oxford; 1970.
- Kim O-S, Cho Y-J, Lee K, Yoon S-H, Kim M, Na H, et al. Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int J Syst Evol Microbiol.* 2012;62:716–21.
- EZTaxon [http://eztaxon-e.ezbiocloud.net/]
- Pagani I, Liolios K, Jansson J, Chen IM, Smirnova T, Nosrat B, et al. The Genomes OnLine Database (GOLD) v. 4: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res.* 2012;40:D571–9.
- Markowitz VM, Chen I-MA, Palaniappan K, Chu K, Szeto E, Pillay M, et al. IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Res.* 2014;42:D560–7.
- Mavromatis K, Land ML, Brettin TS, Quest DJ, Copeland A, Clum A, et al. The fast changing landscape of sequencing technologies and their impact on microbial genome assemblies and annotation. *PLoS One.* 2012;7, e48837.
- Reeve WG, Tiwari RP, Worsley PS, Dilworth MJ, Glenn AR, Howieson JG. Constructs for insertional mutagenesis, transcriptional signal localization and gene regulation studies in root nodule and other bacteria. *Microbiology.* 1999;145:1307–16.
- Protocols and sample preparation information [http://jgi.doe.gov/collaborate-with-jgi/pmo-overview/protocols-sample-preparation-information/]
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol.* 2011;28:2731–9.
- Bennett S. Solexa Ltd. *Pharmacogenomics.* 2004;5:433–8.
- JGI: Joint genome institute [http://www.jgi.doe.gov]
- Zerbino D, Birney E. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* 2008;18:821–9.
- Reads simulator wgsim [https://github.com/lh3/wgsim]
- Gnerre S, MacCallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, et al. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *P Natl A Sci.* 2011;108:1513–8.
- Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, et al. Towards a richer description of our complete collection of genomes and metagenomes "Minimum Information about a Genome Sequence" (MIGS) specification. *Nature Biotechnol.* 2008;26:541–7.
- Field D, Amaral-Zettler L, Cochrane G, Cole JR, Dawyndt P, Garrity GM, et al. The Genomic standards Consortium. *PLoS Biol.* 2011;9, e1001088.
- Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *P Natl A Sci USA.* 1990;87:4576–9.
- Garrity GM, Bell JA, Lilburn T. Phylum XIV. *Proteobacteria* phyl. nov. In: Garrity GM, Brenner DJ, Kreig NR, Staley JT, editors. *Bergey's Manual of Systematic Bacteriology*. Volume 2. Second ed. New York: Springer - Verlag; 2005. p. 1.
- Validation of publication of new names and new combinations previously effectively published outside the IJSEM. List no. 106. *Int J Syst Evol Microbiol.* 2005;55:2235–38.
- Garrity GM, Bell JA, Lilburn T. Class I. *Alphaproteobacteria* class. In: Garrity GM, Brenner DJ, Kreig NR, Staley JT, editors. *Bergey's Manual of Systematic Bacteriology*. Volume 2. Second ed. New York: Springer - Verlag; 2005. p. 1.
- Kuykendall LD. Order VI. *Rhizobiales* ord. nov. In: Garrity GM, Brenner DJ, Kreig NR, Staley JT, editors. *Bergey's Manual of Systematic Bacteriology*. Second ed. New York: Springer - Verlag; 2005. p. 324.

34. Garrity GM, Bell JA, Lilburn T. Family VII. *Bradyrhizobiaceae* fam. nov. In Bergey's Manual of Systematic Bacteriology. Volume 2. Second edition. Edited by Brenn DJ. New York: Springer - Verlag; 2005: 438
35. Jordan DC. Transfer of *Rhizobium japonicum* Buchanan 1980 to *Bradyrhizobium* gen. nov., a genus of slow-growing, root nodule bacteria from leguminous plants. *Int J Syst Bacteriol.* 1982;32:136–9.
36. Biological agents: technical rules for biological agents. TRBA:466.
37. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. *The Gene Ontology Consortium Nat Genet.* 2000;25:25–9.
38. Guide to GO evidence codes [<http://www.geneontology.org/GO.evidence.shtml>]
39. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics.* 2010;11:119.
40. Mavromatis K, Ivanova NN, Chen IM, Szeto E, Markowitz VM, Kyrpides NC. The DOE-JGI standard operating procedure for the annotations of microbial genomes. *Stand Genomic Sci.* 2009;1:63–7.
41. Chen IM, Markowitz VM, Chu K, Anderson I, Mavromatis K, Kyrpides NC, et al. Improving microbial genome annotations in an integrated database context. *PLoS One.* 2013;8, e54859.
42. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 1997;25:955–64.
43. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, et al. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* 2007;35:7188–96.
44. Infernal: inference of RNA alignments [<http://infernal.janelia.org/>]
45. Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC. IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics.* 2009;25:2271–8.
46. GOLD ID for *Bradyrhizobium* sp. Th.b2 [<https://gold.jgi-psf.org/projects?id=14287>]
47. Duran D, Rey L, Mayo J, Zuniga-Davila D, Imperial J, Ruiz-Argueso T, et al. *Bradyrhizobium paxllaeri* sp. nov. and *Bradyrhizobium icense* sp. nov., nitrogen-fixing rhizobial symbionts of Lima bean (*Phaseolus lunatus* L.) in Peru. *Int J Syst Evol Microbiol.* 2014;64:2072–8.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

