

# UC Merced

## UC Merced Previously Published Works

### Title

High-quality permanent draft genome sequence of Bradyrhizobium sp. Tv2a.2, a microsymbiont of Tachigali versicolor discovered in Barro Colorado Island of Panama

### Permalink

<https://escholarship.org/uc/item/09w510qp>

### Journal

Environmental Microbiome, 10(1)

### ISSN

1944-3277

### Authors

Tian, Rui

Parker, Matthew

Seshadri, Rekha

et al.

### Publication Date

2015

### DOI

10.1186/s40793-015-0006-0

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



SHORT GENOME REPORT

Open Access

# High-quality permanent draft genome sequence of *Bradyrhizobium* sp. Tv2a.2, a microsymbiont of *Tachigali versicolor* discovered in Barro Colorado Island of Panama

Rui Tian<sup>1</sup>, Matthew Parker<sup>2</sup>, Rekha Seshadri<sup>3</sup>, TBK Reddy<sup>3</sup>, Victor Markowitz<sup>4</sup>, Natalia Ivanova<sup>3</sup>, Amrita Pati<sup>3</sup>, Tanja Woyke<sup>3</sup>, Mohammed N Baeshen<sup>6,7</sup>, Nabih A Baeshen<sup>5</sup>, Nikos Kyrpides<sup>3,5</sup> and Wayne Reeve<sup>1\*</sup>

## Abstract

*Bradyrhizobium* sp. Tv2a.2 is an aerobic, motile, Gram-negative, non-spore-forming rod that was isolated from an effective nitrogen-fixing root nodule of *Tachigali versicolor* collected in Barro Colorado Island of Panama. Here we describe the features of *Bradyrhizobium* sp. Tv2a.2, together with high-quality permanent draft genome sequence information and annotation. The 8,496,279 bp high-quality draft genome is arranged in 87 scaffolds of 87 contigs, contains 8,109 protein-coding genes and 72 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

Legumes engage in nitrogen-fixation symbioses with bacterial partners from at least 13 genera of *Proteobacteria* [1-4]. Despite the high extent of phylogenetic diversity of root nodule bacteria, the very broad distribution of one particular genus (*Bradyrhizobium*) across host legume clades suggests that bacteria in this genus may have been the first legume symbionts [5]. *Bradyrhizobium* interacts with the widest diversity of legume clades (at least 24 of ca. 33 nodule-forming legume tribes; [6]) and is associated with nodulating groups that represent early branching lineages [7] in all three legume subfamilies [8,9]. Analysis of basal *Bradyrhizobium* lineages that are associated with early-diverging legume groups may thus shed light on the origins of this symbiosis.

Here we report the genome sequence of one such organism, *Bradyrhizobium* strain Tv2a.2. Strain Tv2a.2 was sampled in 1997 from the tree *Tachigali versicolor* on Barro Colorado Island, Panama, a biological preserve with an old-growth moist tropical forest [10]. *Tachigali*

is one of just a handful of nodule-forming genera in the legume Subfamily Caesalpinioideae [11], which is comprised of the earliest branching lineages in the legume family [7]. *Tachigali versicolor* is a large canopy tree with an unusual monocarpic life history, in which trees grow for decades without flowering. They produce just a single crop of seeds, and then die [12].

Strain Tv2a.2 is a typical representative of the nodule symbionts that are associated with *Tachigali* in this tropical forest habitat [13], and appears to represent a unique early-diverging lineage of *Bradyrhizobium*. Phylogenetic analyses have placed Tv2a.2 somewhere near the early split in the genus between two large superclades represented by *B. diazoefficiens* USDA 110 and *B. elkanii* USDA 76. However, its exact position near the base of the *Bradyrhizobium* tree varies to some extent in different analyses, depending on the loci, the strains included, and the method of tree analysis [5,13]. For example, a Bayesian analysis of 16S rRNA sequences from the type strains of 21 *Bradyrhizobium* species and strain ORS278 placed Tv2a.2 as the earliest diverging *Bradyrhizobium* lineage [14].

Here we provide an analysis of the complete genome sequence of Tv2a.2, one of the rhizobial genomes sequenced

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

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

as part of the DOE Joint Genome Institute 2010 Genomic Encyclopedia for Bacteria and Archaea-Root Nodule Bacteria (GEBA-RNB) project proposal [15], whose properties should help to clarify early events in the diversification of the genus *Bradyrhizobium* as a whole.

## Organism information

### Classification and features

*Bradyrhizobium* sp. Tv2a.2 is a motile, non-sporulating, non-encapsulated, Gram-negative strain in the order *Rhizobiales* of the class *Alphaproteobacteria*. The rod shaped form (Figure 1 Left, Center) has dimensions of approximately 0.5  $\mu\text{m}$  in width and 1.5-2.0  $\mu\text{m}$  in length. It is relatively slow growing, forming colonies after 6–7 days when grown on half strength Lupin Agar ( $\frac{1}{2}$ LA) [16], tryptone-yeast extract agar (TY) [17] or a modified yeast-mannitol agar (YMA) [18] 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. Tv2a.2 in a 16S rRNA gene sequence based tree. This strain is phylogenetically the most related to *Bradyrhizobium* sp. EC3.3 based on a 16S rRNA gene sequence identity of 99.31% as determined using BLAST analysis [19]. Tv2a.2 is also related to the type strains *Bradyrhizobium ingae* BR 10250<sup>T</sup> and *Bradyrhizobium iriomotense* EK05<sup>T</sup> with 16S rRNA gene sequence identities of 99.16 % and 99.08%, respectively, based on results from the EzTaxon-e server [20,21].

Minimum Information about the Genome Sequence (MIGS) of Tv2a.2 is provided in Table 1 and Additional file 1: Table S1.

### Symbiotaxonomy

*Bradyrhizobium* strain Tv2a.2 was isolated from nodules of *Tachigali versicolor* found in a tropical forest on Barro

Colorado Island, Panama [10]. Due to the highly erratic pattern of seed production from this host, no seeds of this legume were available to authenticate the symbiotic proficiency of strain Tv2a.2. Nodulation and nitrogen fixation was therefore tested on two promiscuous legumes (*Vigna unguiculata*, *Macroptilium atropurpureum*) and revealed that nodules could only develop on *M. atropurpureum*. Acetylene reduction assays also showed that these nodules lacked nitrogenase activity [13]. A further indication that Tv2a.2 may be relatively host-specific is the fact that extensive sampling of other legume hosts in Panama (and elsewhere in the Neotropics) have never recovered strains belonging to the Tv2a.2 lineage from any legume taxa other than *T. versicolor* [9].

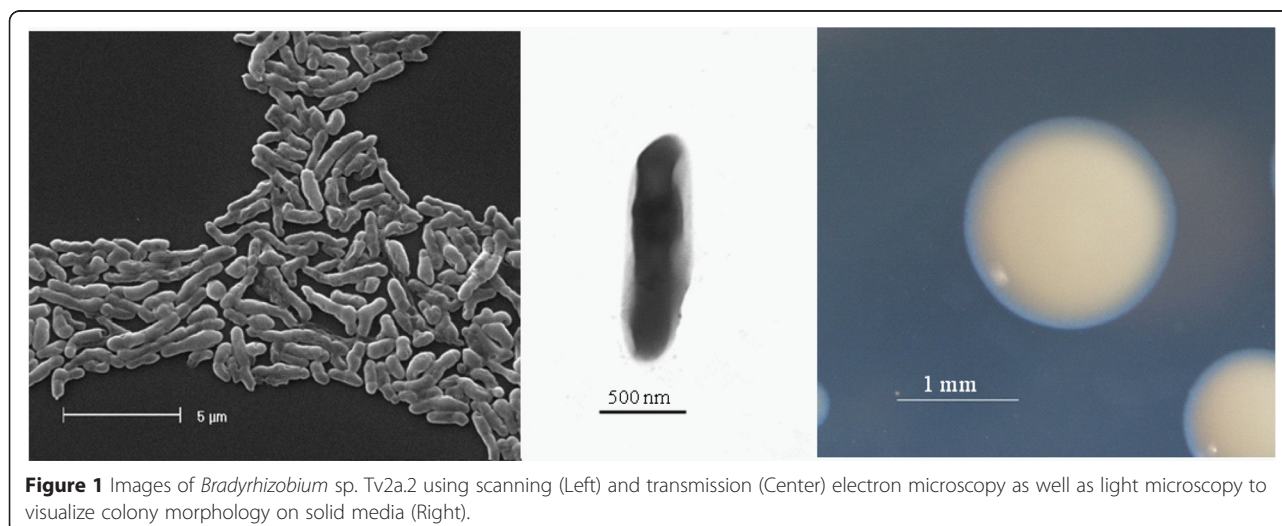
## Genome sequencing and annotation information

### Genome project history

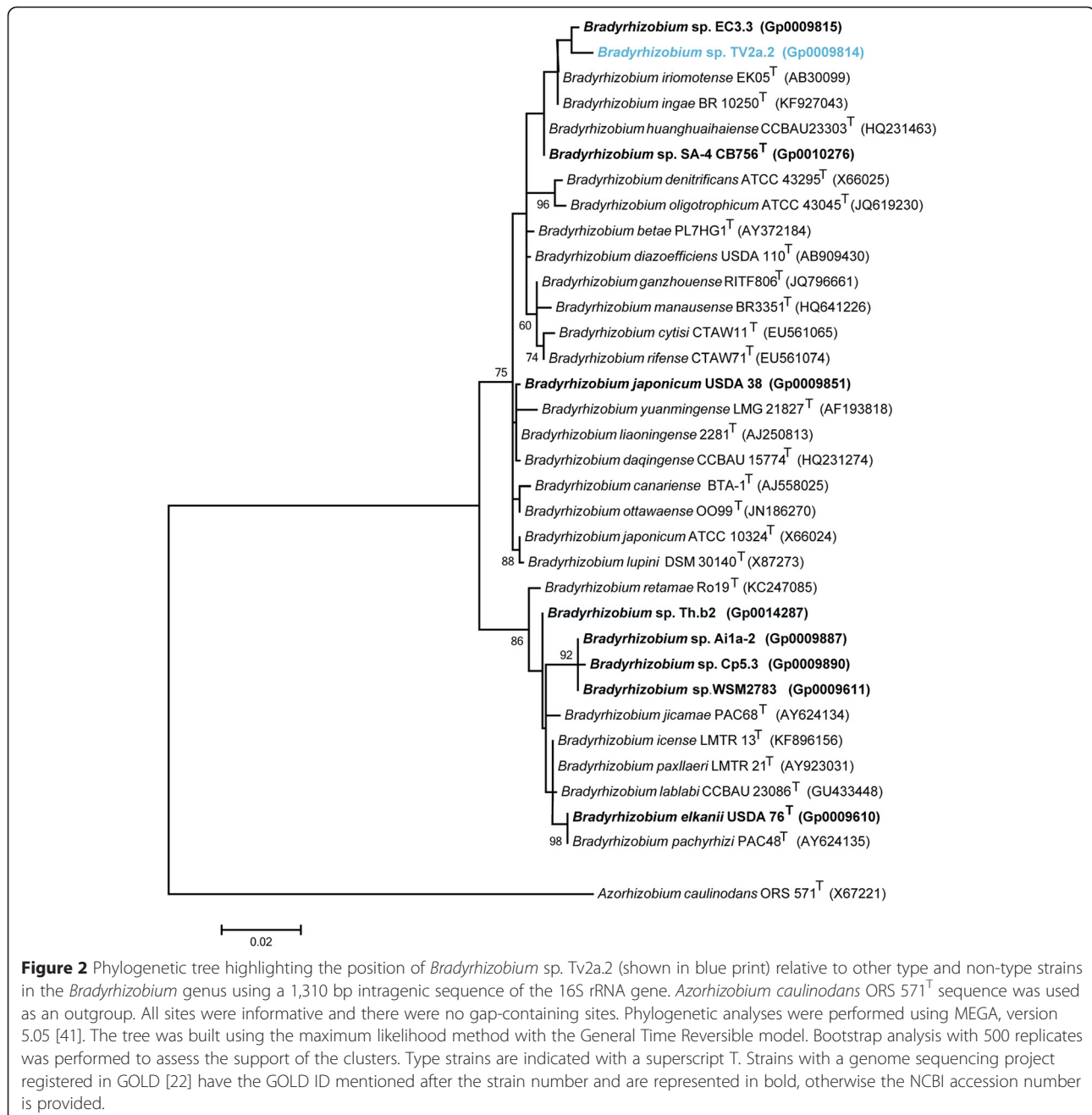
This organism was selected for sequencing on the basis of its environmental and agricultural relevance to issues in global carbon cycling, alternative energy production, and biogeochemical importance, and is part of the Genomic Encyclopedia of Bacteria and Archaea, Root Nodulating Bacteria (GEBA-RNB) project at the U.S. Department of Energy, Joint Genome Institute (JGI). The genome project is deposited in the Genomes OnLine Database [22] and a high-quality permanent draft genome sequence in IMG [23]. Sequencing, finishing and annotation were performed by the JGI using state of the art sequencing technology [24]. A summary of the project information is shown in Table 2.

### Growth conditions and genomic DNA preparation

*Bradyrhizobium* sp. Tv2a.2 was cultured to mid logarithmic phase in 60 ml of TY rich media on a gyratory shaker at 28°C [25]. DNA was isolated from the cells



**Figure 1** Images of *Bradyrhizobium* sp. Tv2a.2 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. Tv2a.2* (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<sup>T</sup> 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 [41]. 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 [22] have the GOLD ID mentioned after the strain number and are represented in bold, otherwise the NCBI accession number is provided.

using a CTAB (Cetyl trimethyl ammonium bromide) bacterial genomic DNA isolation method [26].

#### Genome sequencing and assembly

The draft genome of *Bradyrhizobium sp. Tv2a.2* was generated at the DOE Joint Genome Institute (JGI) using the Illumina technology [27]. An Illumina standard shotgun library was constructed and sequenced using the Illumina HiSeq 2000 platform which generated 8,336,316 reads totaling 1250.45 Mbp. All general aspects of library

construction and sequencing were performed at the JGI and details can be found on the JGI website [28]. 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) [29], (2) 1–3 Kbp simulated paired end reads were created from Velvet contigs using wgsim [30], (3) Illumina reads

**Table 1 Classification and general features of *Bradyrhizobium* sp. Tv2a.2 in accordance with the MIGS recommendations [42] published by the Genome Standards Consortium [43]**

MIGS ID	Property	Term	Evidence code
	Classification	Domain <i>Bacteria</i>	TAS [44]
		Phylum <i>Proteobacteria</i>	TAS [45,46]
		Class <i>Alphaproteobacteria</i>	TAS [46,47]
		Order <i>Rhizobiales</i>	TAS [48]
		Family <i>Bradyrhizobiaceae</i>	TAS [49]
		Genus <i>Bradyrhizobium</i>	TAS [50]
		Species <i>Bradyrhizobium</i> sp.	IDA
	Gram stain	Negative	IDA
	Cell shape	Rod	IDA
	Motility	Motile	IDA
	Sporulation	Non-sporulating	NAS
	Temperature range	Unknown	NAS
	Optimum temperature	28°C	NAS
	pH range; Optimum	Unknown	
	Carbon source	Varied	NAS
	Energy source	Chemoorganotroph	NAS
MIGS-6	Habitat	Soil, root nodule, host	TAS [10]
MIGS-6.3	Salinity	Non-halophile	NAS
MIGS-22	Oxygen requirement	Aerobic	NAS
MIGS-15	Biotic relationship	Free living, symbiotic	TAS [10]
MIGS-14	Pathogenicity	Non-pathogenic	NAS
	Biosafety level	1	TAS [51]
	Isolation	Root nodule of <i>Tachigali versicolor</i>	TAS [10]
MIGS-4	Geographic location	Barro Colorado Island, Panama	TAS [10]
MIGS-5	Sample collection	1997	IDA
MIGS-4.1	Latitude	9.1663	IDA
MIGS-4.2	Longitude	- 79.8248	IDA
MIGS-4.3	Depth	5 cm	IDA
MIGS-4.4	Altitude	28 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 [52,53].

were assembled with simulated read pairs using Allpaths-LG (version r39750) [31]. Parameters for the assembly steps were 1) velvet: --v -s 51 -e 71 -i 2 -t 1 -f “-short-Paired -fastq \$FASTQ” -o “-ins\_length 250 -min\_contig\_lgth 500” for Velvet and 2) wgsim: -e 0-1 76-2 76 -r 0 -R 0 -X 0. The final draft assembly contained 87 contigs in 87 scaffolds. The total size of the genome is 8.5 Mb with an average of 109.04x coverage of the genome.

**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	109.04x
MIGS-30	Assemblers	Velvet version 1.1.04; Allpaths-LG version r39750
MIGS-32	Gene calling method	Prodigal 1.4
	Locus Tag	A3AI
	GenBank ID	AXAI000000000
	GenBank Date of Release	September 30, 2013
	GOLD ID	Gp0009814 [54]
	BIOPROJECT	165315
MIGS-13	Source Material Identifier	Tv2a.2
	Project relevance	Symbiotic N <sub>2</sub> fixation, agriculture

### Genome annotation

Genes were identified using Prodigal [32], as part of the DOE-JGI genome annotation pipeline [33,34]. The predicted CDSs were translated and used to search the National Center for Biotechnology Information non-redundant database, UniProt, TIGRFam, Pfam, KEGG, COG, and InterPro databases. The tRNAScanSE tool [35] was used to find tRNA genes, whereas ribosomal RNA genes were found by searches against models of the ribosomal RNA genes built from SILVA [36]. Other non-coding RNAs such as the RNA components of the protein

**Table 3 Genome statistics for *Bradyrhizobium* sp. Tv2a.2**

Attribute	Value	% of total
Genome size (bp)	8,496,279	100.00
DNA coding (bp)	7,163,193	84.31
DNA G + C (bp)	5,284,500	62.20
DNA scaffolds	87	100
Total genes	8,181	100.00
Protein coding genes	8,109	99.12
RNA genes	72	0.88
Pseudo genes	7	0.09
Genes in internal clusters	665	8.13
Genes with function prediction	5,967	72.94
Genes assigned to COGs	4,871	59.54
Genes with Pfam domains	6,080	74.32
Genes with signal peptides	866	10.59
Genes with transmembrane helices	1,836	22.44
CRISPR repeats	0	0.00



secretion complex and the RNase P were identified by searching the genome for the corresponding Rfam profiles using INFERNAL [37]. Additional gene prediction analysis and manual functional annotation was performed within the Integrated Microbial Genomes-Expert Review (IMG-ER) system [38] developed by the Joint Genome Institute, Walnut Creek, CA, USA.

### Genome properties

The genome is 8,496,279 nucleotides with 62.20% GC content (Table 3) and comprised of 87 scaffolds. From a total of 8,181 genes, 8,109 were protein encoding and 72 RNA only encoding genes. The majority of genes (72.94%) 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 4 Number of genes associated with the general COG functional categories**

Code	Value	% of total (5,458)	COG category
J	184	3.37	Translation, ribosomal structure and biogenesis
A	0	0.00	RNA processing and modification
K	378	6.93	Transcription
L	142	2.60	Replication, recombination and repair
B	2	0.04	Chromatin structure and dynamics
D	29	0.53	Cell cycle control, cell division, chromosome partitioning
V	99	1.81	Defense mechanisms
T	222	4.07	Signal transduction mechanisms
M	256	4.69	Cell wall/membrane/envelope biogenesis
N	62	1.14	Cell motility
U	105	1.92	Intracellular trafficking, secretion, and vesicular transport
O	195	3.57	Posttranslational modification, protein turnover, chaperones
C	426	7.81	Energy production and conversion
G	343	6.28	Carbohydrate transport and metabolism
E	625	11.45	Amino acid transport and metabolism
F	82	1.50	Nucleotide transport and metabolism
H	204	3.74	Coenzyme transport and metabolism
I	346	6.34	Lipid transport and metabolism
P	281	5.15	Inorganic ion transport and metabolism
Q	245	4.49	Secondary metabolite biosynthesis, transport and catabolism
R	711	13.03	General function prediction only
S	521	9.55	Function unknown
-	3,310	40.46	Not in COGS

### Conclusions

*Bradyrhizobium* sp. Tv2a.2 was collected in 1997 from a nodule of the tree *Tachigali versicolor* on Barro Colorado Island, Panama. Based on 16S rRNA gene analyses, Tv2a.2 is phylogenetically the most closely related to *Bradyrhizobium* sp. EC3.3 (a strain isolated from a nodule of *Erythrina costaricensis* collected from Barro Colorado Island, Panama) and to the type strains *Bradyrhizobium ingae* BR 10250<sup>T</sup> and *Bradyrhizobium iriomotense* EK05<sup>T</sup> isolated from Inga laurina (Sw.) Willd. growing in the Cerrado Amazon region, State of Roraima, Brazil [39] and from *Entada koshunensis*, a legume available in Okinawa, Japan [40], respectively. Strain Tv2a.2 is one of 25 *Bradyrhizobium* genomes that were sequenced within the GEBA-RNB project [15]; of these, the Tv2a.2 genome has the fifth lowest genome size (8.5 Mbp), gene count (8,181) and Pfam percentage (74.32%) amongst these strains. The specific genome attributes of *Bradyrhizobium* sp. Tv2a.2 compared to the other *Bradyrhizobium* genomes will be important to understand the interactions required for the successful establishment of an effective symbiosis with the host *Tachigali versicolor*.

### Additional file

**Additional file 1: 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

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

Received: 13 February 2015 Accepted: 16 April 2015

Published online: 17 May 2015

## References

- Willems A. The taxonomy of rhizobia; an overview. *Plant Soil*. 2006;287:3–14.
- Gyaneshwar P, Hirsch AM, Moulin L, Chen WM, Elliott GN, Bontemps C, et al. Legume-nodulating betaproteobacteria: diversity, host range, and future prospects. *Mol Plant Microbe In*. 2011;24:1276–88.
- Ardley JK, Parker MA, De Meyer SE, Trengove RD, O'Hara GW, Reeve WG, et al. *Microvirga lupini* sp. nov., *Microvirga lotononidis* sp. nov. and *Microvirga zambiensis* sp. nov. are alphaproteobacterial root-nodule bacteria that specifically nodulate and fix nitrogen with geographically and taxonomically separate legume hosts. *Int J Syst Evol Microbiol*. 2012;62:2579–88.
- Maynaud G, Willems A, Soussou S, Vidal C, Maure L, Moulin L, et al. Molecular and phenotypic characterization of strains nodulating *Anthyllis vulneraria* in mine tailings, and proposal of *Aminobacter anthyllidis* sp. nov., the first definition of *Aminobacter* as legume-nodulating bacteria. *Syst Appl Microbiol*. 2012;35:65–72.
- Parker MA. The spread of *Bradyrhizobium* lineages across host legume clades: from *Abarema* to *Zygia*. *Microb Ecol*. 2014;69:630–640.
- Lewis GP. Legumes of the world. Royal Botanic Gardens, Kew: Richmond, UK; 2005.
- Doyle JJ. Phylogenetic perspectives on the origins of nodulation. *Mol Plant Microbe Int*. 2011;24:1289–95.
- Fonseca MB, Peix A, de Faria SM, Mateos PF, Rivera LP, Simoes-Araujo JL, et al. Nodulation in *Dimorphandra wilsonii* Rizz. (*Caesalpinioideae*), a threatened species native to the Brazilian Cerrado. *PLoS One*. 2012;7:e49520.
- Parker MA, Rousteau A. Mosaic origins of *Bradyrhizobium* legume symbionts on the Caribbean island of Guadeloupe. *Mol Phylogenet Evol*. 2014;77:110–5.
- Foster RB, Brokaw NVL. Structure and history of the vegetation on Barro Colorado Island. The ecology of a tropical forest: seasonal rhythms and long-term changes. Smithsonian Institution: Washington, DC; 1996.
- Sprent JI, Ardley JK, James EK. From North to South: a latitudinal look at legume nodulation processes. *South Afr J Bot*. 2013;89:31–41.
- Foster RB. *Tachigali versicolor* is a suicidal neotropical tree. *Nature*. 1977;268:24–26.
- Parker MA. Divergent *Bradyrhizobium* symbionts on *Tachigali versicolor* from Barro Colorado Island, Panama. *Syst Appl Microbiol*. 2000;23:585–90.
- Giraud E, Moulin L, Vallenet D, Barbe V, Cytryn E, Avarre JC, et al. Legumes symbioses: absence of Nod genes in photosynthetic bradyrhizobia. *Science*. 2007;316:1307–12.
- 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 the root-nodule bacteria. International biological programme. UK: Blackwell Scientific Publications, Oxford; 1970.
- NCBI BLAST [<http://blast.ncbi.nlm.nih.gov/Blast.cgi>]
- 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/>]
- 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.
- 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.
- 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.
- 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.
- 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.
- 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;35:1788–96.
- Infernal: inference of RNA alignments [<http://infernal.janelia.org/>]
- 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.
- da Silva K, De Meyer SE, Rouws LF, Farias EN, dos Santos MA, O'Hara G, et al. *Bradyrhizobium ingae* sp. nov., isolated from effective nodules of *Inga laurina* grown in Cerrado soil. *Int J Syst Evol Microbiol*. 2014;64:3395–401.
- Islam MS, Kawasaki H, Muramatsu Y, Nakagawa Y, Seki T. *Bradyrhizobium iriomotense* sp. nov., isolated from a tumor-like root of the legume *Entada koshunensis* from Iriomote Island in Japan. *Biosci Biotechnol Biochem*. 2008;72:1416–29.
- 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.
- 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.
- 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.
- 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
- 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.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Chery JM, et al. Gene ontology: tool for the unification of biology. *The Gene Ontology Consortium Nat Genet*. 2000;25:25–9.
- Guide to GO evidence codes [<http://www.geneontology.org/GO.evidence.shtml>]
- GOLD ID for *Bradyrhizobium* sp.Tv2a.2 [<https://gold.jgi-psf.org/projects?id=9814>]