



SHORT GENOME REPORT

Open Access



# High-quality permanent draft genome sequence of *Bradyrhizobium* sp. strain WSM1743 - an effective microsymbiont of an *Indigofera* sp. growing in Australia

Leila Eshraghi<sup>1,2</sup>, Sofie E. De Meyer<sup>1</sup>, Rui Tian<sup>1</sup>, Rekha Seshadri<sup>3</sup>, Natalia Ivanova<sup>3</sup>, Amrita Pati<sup>3</sup>, Victor Markowitz<sup>4</sup>, Tanja Woyke<sup>3</sup>, Nikos C. Kyrpides<sup>3,5</sup>, Ravi Tiwari<sup>1</sup>, Ron Yates<sup>6</sup>, John Howieson<sup>1</sup> and Wayne Reeve<sup>1\*</sup>

## Abstract

*Bradyrhizobium* sp. strain WSM1743 is an aerobic, motile, Gram-negative, non-spore-forming rod that can exist as a soil saprophyte or as a legume microsymbiont of an *Indigofera* sp. WSM1743 was isolated from a nodule recovered from the roots of an *Indigofera* sp. growing 20 km north of Carnarvon in Australia. It is slow growing, tolerates up to 1 % NaCl and is capable of growth at 37 °C. Here we describe the features of *Bradyrhizobium* sp. strain WSM1743, together with genome sequence information and its annotation. The 8,341,956 bp high-quality permanent draft genome is arranged into 163 scaffolds and 167 contigs, contains 7908 protein-coding genes and 75 RNA-only encoding genes and was sequenced as part of the Root Nodule Bacteria chapter of the Genomic Encyclopedia of Bacteria and Archaea project.

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

## Introduction

Rhizobia are soil-dwelling bacteria that have acquired the ability to establish associations with leguminous plants to symbiotically fix nitrogen. After infection of the plant, the rhizobia become established within root nodules and can fix atmospheric dinitrogen gas into ammonia using a reaction that is catalyzed by the nitrogenase enzyme [1]. The export of fixed nitrogen to the plant improves growth and productivity under N-limiting environmental conditions. The effective use of the symbiosis leads to sustainable cropping systems with a net positive impact on the environment [2]. In Australia, the majority of productive legumes and their rhizobia in agricultural systems have been deliberately, or accidentally, introduced since European settlement [3]. However, recently, there has been an interest in the diversity of Australian native legumes and their microsymbionts [4].

The northwest of Western Australia is an ideal landscape to discover rhizobia nodulating indigenous legume flora [4] and is an area low in introduced legumes and inoculants. In 1996, an extensive survey was conducted of the area revealing a range of indigenous legume genera including a number of *Indigofera* spp. [4]. In Australia, this species has been found at dispersed locations in the Northern Territory, Queensland and Western Australia on dark brown clay loams and frequently on lands under cultivation [5]. The Australian *Indigofera* spp., based on their habitat, can be placed into three categories; i) shrubs, including *I. brevidens*, *I. australis*, *I. adesmiiifolia* and some members of the *I. pratensis* group, which occur mainly on better soil types in the east coast, ii) perennial herbs, such as *I. baileyi*, *I. efoliata*, *I. triflora*, *I. georgei*, *I. rugosa* and members of the *I. triflora* and *I. pratensis* groups, which occur in the more arid, or seasonally dry, parts of Australia, iii) annual herbs, uncommon amongst the endemic species, including the two annual species, *I. haplophylla* and *I.*

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

<sup>1</sup>Centre for Rhizobium Studies, Murdoch University, Murdoch, Western Australia

Full list of author information is available at the end of the article



*ammobia*, which occur in the monsoon tropics and the Tanami and Great Sandy Deserts, respectively [6]. The native species with wide extra-Australian distributions (particularly *I. colutea*, *I. hirsuta*, *I. linnaei* and *I. linifolia*) occur in a variety of habitats, mostly towards the northern parts of Australia. It is likely that these taxa now inhabit a greater range than they did before European settlement, and the Australian populations of these species may have been augmented by the introduction of seed from non-Australian sources [6].

Since there is a paucity of information regarding microsymbionts of *Indigofera*, a collection of root nodules was therefore obtained from the most prevalent *Indigofera* spp. present in northwest Australia and the microsymbionts from these nodules were then isolated. One microsymbiont, *Bradyrhizobium* sp. strain WSM1743, was isolated from a nodule recovered from the roots from an indigenous *Indigofera* sp. growing in red-brown sandy loam 40 m above sea level. The plant was located in natural bush land, approximately 20 km Northeast of the town Carnarvon in Western Australia [4]. The collection area has a warm semi-arid climate with a long-term mean seasonal rainfall of 226 mm.

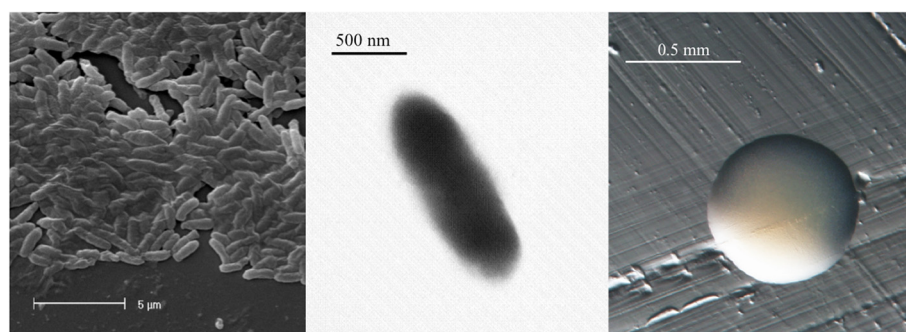
Strain WSM1743 was identified as a *Bradyrhizobium* sp. based on 16S rRNA typing [4]. Most *Bradyrhizobium* spp., including WSM1743, cannot grow on sucrose or lactose, which may indicate the lack of a disaccharide uptake system [7]. However, WSM1743, unlike other *Bradyrhizobium* spp., is able to grow at 37 °C and this ability could be a specific adaptation to the high soil temperatures experienced in the northwest of Western Australia [4]. Here we present a summary classification and a set of general features for this microsymbiont together with a description of its genome sequence and annotation done as part of the DOE Joint Genome Institute 2010 Genomic Encyclopedia for Bacteria and Archaea-Root Nodule Bacteria project [8] (Additional file 2).

## Organism information

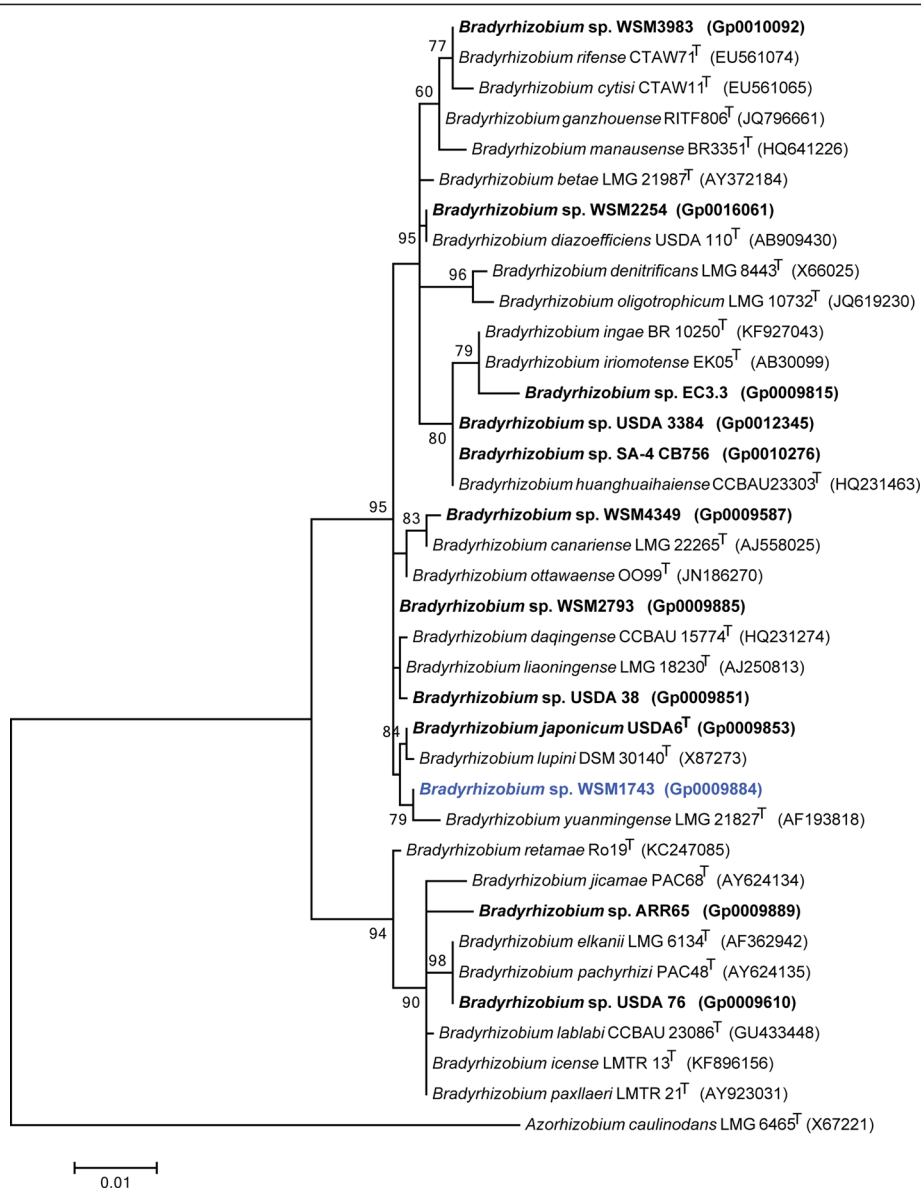
### Classification and features

*Bradyrhizobium* sp. strain WSM1743 is a motile, Gram-negative non-spore-forming rod (Fig. 1 Left and Center) in the order *Rhizobiales* of the class *Alphaproteobacteria*. It is slow growing, forming colonies within 7–10 days when grown on half strength Lupin Agar [9] at 28 °C. Colonies on ½LA are white-opaque, slightly domed and moderately mucoid with smooth margins (Fig. 1 Right). This strain was isolated together with 7 other bacteria from native *Indigofera* plants and physiologically characterised. Strain WSM1743 was identified as slow growing with poor growth in 1 % NaCl, no growth at 2 to 3 % NaCl and average growth in pH5 to 9 [4]. *Bradyrhizobium* type strains have a slow generation time (9 to 18 h) and fail to grow in media containing 2 % NaCl [10], which indicates that WSM1743 belongs to this genus. The maximal growth temperature for most *Bradyrhizobium* strains is 33 to 35 °C, with many strains failing to grow above 34 °C [10]. However, WSM1743 was able to grow at 37 °C on ½LA medium [4], and therefore extending the temperature range for *Bradyrhizobium*.

Figure 2 shows the phylogenetic relationship of *Bradyrhizobium* sp. strain WSM1743 in a 16S rRNA gene sequence based tree. This strain is phylogenetically the most related to the RNB type strains *B. japonicum* USDA 6<sup>T</sup>, *B. lupini* DSM30140<sup>T</sup> and *B. yuanmingense* LMG 21827<sup>T</sup> with sequence identities to the WSM1743 16S rRNA gene sequence of 99.78 %, 99.71 % and 99.63 %, respectively, as determined using the EzTaxon-server [11]. *B. japonicum* USDA6<sup>T</sup> was originally isolated in Japan from *Glycine max* root nodules and is able to nodulate and fix nitrogen effectively with several other *Glycine* species and *Macroptillium atropurpureum* [12]. *B. lupini* DSM30140<sup>T</sup> is a microsymbiont of *Lupinus luteus* and *Lupinus angustifolius* [13]. *B. yuanmingense* B071<sup>T</sup> was isolated from *Lespedeza cuneata* root nodules from China but is also able to nodulate and fix



**Fig. 1** Images of *Bradyrhizobium* sp. strain WSM1743 using scanning (Left) and transmission (Center) electron microscopy as well as light microscopy to visualize the colony morphology on a solid media (Right)



**Fig. 2** Phylogenetic tree highlighting the position of *Bradyrhizobium* sp. strain WSM1743 (shown in blue print) relative to other type and non-type strains in the *Bradyrhizobium* genus using a 1,251 bp internal region of the 16S rRNA gene. *Azorhizobium caulinodans* LMG 6465<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 [17]. 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 [18] are shown in bold and have the GOLD ID mentioned after the strain number, otherwise the NCBI accession number is provided

nitrogen effectively with *Vigna unguiculata* and *Glycyrhiza uralensis* [14]. Additionally, a recent report showed that *B. yuanmingense* and *B. japonicum* are the preferred microsymbionts of *Vigna unguiculata* and *Vigna radiate* in the subtropical region of China [15].

Minimum Information about the Genome Sequence (MIGS) [16] of WSM1743 is provided in Table 1 and Additional file 1: Table S1.

### Symbiotaxonomy

*Bradyrhizobium* sp. strain WSM1743 was isolated from *Indigofera* sp. nodules collected at site 32, north of Carnarvon, Western Australia [4]. The site of collection contained several Australian native legumes, with a soil pH of 7.5. Symbiotic interactions of *Bradyrhizobium* sp. strain WSM1743 were assessed on three annual, one biennial and nine perennial exotic legume species that

**Table 1** Classification and general features of *Bradyrhizobium* sp. strain WSM1743 in accordance with the MIGS recommendations [15] published by the Genome Standards Consortium [19]

MIGS ID	Property	Term	Evidence code
	Current classification	Domain <i>Bacteria</i>	TAS [20]
		Phylum <i>Proteobacteria</i>	TAS [21, 22]
		Class <i>Alphaproteobacteria</i>	TAS [21]
		Order <i>Rhizobiales</i>	TAS [23]
		Family <i>Bradyrhizobiaceae</i>	TAS [24]
		Genus <i>Bradyrhizobium</i>	TAS [10]
		Species sp.	IDA
	Gram stain	Negative	TAS [12]
	Cell shape	Rod	IDA
	Motility	Motile	IDA
	Sporulation	Non-sporulating	IDA
	Temperature range	Mesophile	IDA
	Optimum temperature	28 °C	IDA
	pH range; Optimum	5 - 9	TAS [4]
	Carbon source	Glutamate, L-arabinose	TAS [4]
MIGS-6	Habitat	Soil, root nodule, on host	IDA
MIGS-6.3	Salinity	Not reported	
MIGS-22	Oxygen requirement	Aerobic	IDA
MIGS-15	Biotic relationship	Free living, symbiotic	IDA
MIGS-14	Pathogenicity	Non-pathogenic	NAS
MIGS-4	Geographic location	20 km north of Carnarvon	TAS [4]
MIGS-5	Nodule collection date	July 1996	TAS [4]
MIGS-4.1	Latitude	-24.770	IDA
MIGS-4.2	Longitude	113.702	IDA
MIGS-4.3	Depth	Up to 1 m	IDA
MIGS-4.4	Altitude	11 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). These evidence codes are from the Gene Ontology project [25]

have agricultural use, or potential use, in southern Australia. WSM1743 consistently nodulated with the exotic legume species, *Macroptilium atropurpureum* and *Phaseolus vulgaris*, inconsistently with *Ononis natrix* and did not form nodules with *Argyrolobium uniflorum*, *Chamaecytisus proliferus*, *Sutherlandia microphylla*, *Hedysarum coronarium*, *Medicago sativa*, *Ornithopus sativus*, *O. compressus*, *Trifolium burchellianum*, *T. polymorphum* and

*T. uniflorum*. Strain WSM1743 was able to consistently nodulate the Australian native legumes, *Acacia saligna*, *Kenmedia prorepens* and *K. coccinea*, but could not nodulate *Swainsona pterostylis*, *S. formosa* and *S. maccullochiana* [4]. Additionally it was noted that the isolate could not nodulate *Indigofera brevidens*, an indigenous *Indigofera* found in the same location as the host of WSM1743 [4].

## 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, The Root Nodulating Bacteria chapter (GEBA-RNB) project at the U.S. Department of Energy, Joint Genome Institute [8]. The genome project is deposited in the Genomes On-Line Database [18] and the high-quality permanent draft genome sequence in IMG [26]. Sequencing, finishing and annotation were performed by the JGI using state of the art sequencing technology [27]. A summary of the project information is shown in Table 2.

### Growth conditions and genomic DNA preparation

*Bradyrhizobium* sp. strain WSM1743 was cultured to mid logarithmic phase in 60 ml of TY rich media on a gyratory shaker at 28 °C [29]. DNA was isolated from the cells using a Cetyl trimethyl ammonium bromide bacterial genomic DNA isolation method [30].

### Genome sequencing and assembly

The draft genome of *Bradyrhizobium* sp. strain WSM1743 was generated at the DOE Joint Genome

**Table 2** Genome sequencing project information for *Bradyrhizobium* sp. strain WSM1743

MIGS ID	Property	Term
MIGS-31	Finishing quality	High-quality permanent draft
MIGS-28	Libraries used	Illumina Std PE
MIGS-29	Sequencing platforms	Illumina Hiseq 2000
MIGS-31.2	Fold coverage	440x Illumina
MIGS-30	Assemblers	Velvet 1.1.04; ALLPATHS-LG V. r39750
MIGS-32	Gene calling method	Prodigal 1.4
	Locus Tag	YU9
	Genebank ID	AXAZ00000000
	Genbank date of release	December 12, 2013
	GOLD ID	Gp0009884 [28]
	BIOPROJECT	PRJNA162991
MIGS-13	Source material identifier	WSM1743
	Project relevance	Symbiotic N <sub>2</sub> fixation, agriculture

Institute [31]. An Illumina standard shotgun library was constructed and sequenced using the Illumina HiSeq 2000 platform, which generated 14,683,452 reads totaling 2.2 Gbp. All general aspects of library construction and sequencing performed at the JGI can be found at the JGI web site [31]. 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). Artifact filtered sequence data was then screened and trimmed according to the k-mers present in the dataset (Mingkun L, Copeland A, Han J. unpublished). High-depth k-mers, presumably derived from MDA amplification bias, cause problems in the assembly, especially if the k-mer depth varies in orders of magnitude for different regions of the genome. Reads with high k-mer coverage (>30x average k-mer depth) were normalized to an average depth of 30x. Reads with an average kmer depth of less than 2x were removed. Following steps were then performed for assembly: (1) normalized Illumina reads were assembled using Velvet version 1.1.04 [32] (2) 1–3 Kbp simulated paired end reads were created from Velvet contigs using wgsim [33] (3) normalized Illumina reads were assembled with simulated read pairs using Allpaths-LG (version r39750) [34]. Parameters for assembly steps were: 1) Velvet optimizing parameters (--v -s 51 -e 71 -i 2 -t 1 -f "-shortPaired -fastq \$FASTQ" -o "-ins\_length 250 -min\_contig\_lgth 500) 2) wgsim version 0.3.0 (-e 0-1 76-2 76 -r 0 -R 0 -X 0) 3) Allpaths-LG (Prepare AllpathsInputs: 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 167 contigs in 163 scaffolds. The total size of the genome is 8.3 Mbp and the final assembly is based on 1238 Mbp of Illumina data, which provides an average of 440x coverage.

#### Genome annotation

Genes were identified using Prodigal [35], as part of the DOE-JGI genome annotation pipeline [36, 37]. 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 [38] was used to find tRNA genes, whereas ribosomal RNA genes were found by searches against models of the ribosomal RNA genes built from SILVA [39]. 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 [40]. Additional gene prediction analysis and manual functional annotation was performed within the Integrated Microbial Genomes-Expert Review system [41] developed by the Joint Genome Institute, Walnut Creek, CA, USA.

#### Genome properties

The genome is 8,341,956 nucleotides with 63.37 % GC content (Table 3) and comprised of 163 scaffolds of 167 contigs. From a total of 7983 genes, 7908 were protein encoding and 75 RNA only encoding genes. The majority of genes (71.51 %) were assigned a putative function whilst the remaining genes were annotated as hypothetical. The distribution of genes into COG functional categories is presented in Table 4.

#### Conclusion

*Bradyrhizobium* sp. WSM1743 belongs to a group of Alpha-rhizobia microsymbionts from native Australian legumes and was isolated from a nodule of an *Indigofera* species growing 20 km north of Carnarvon in northwestern Australia. Phylogenetic analysis revealed that WSM1743 is most closely related to *B. japonicum* USDA 6<sup>T</sup>, which was obtained from *Glycine max* root nodules from Japan and is able to nodulate and fix nitrogen effectively with several other *Glycine* species and *Macroptilium atropurpureum* [12]. Strain WSM1743 has been shown to nodulate with *Macroptilium atropurpureum* and endemic Australian legumes including *Acacia saligna*, *Kennedia prorepens* and *K. coccinea* [4].

**Table 3** Genome statistics for *Bradyrhizobium* sp. strain WSM1743

Attribute	Value	% of total
Genome size (bp)	8,341,956	100.00
DNA coding (bp)	6,951,810	83.34
DNA G + C (bp)	5,286,166	63.37
DNA scaffolds	163	
Total genes	7,983	100.00
Protein-coding genes	7,908	99.06
RNA genes	75	0.94
Pseudo genes	12	0.15
Genes in internal clusters	465	5.82
Genes with function prediction	5,709	71.51
Genes assigned to COGs	4,824	60.43
Genes with Pfam domains	6,012	75.31
Genes with signal peptides	840	10.52
Genes with transmembrane proteins	1,784	22.35
CRISPR repeats	1	

**Table 4** Number of protein coding genes of *Bradyrhizobium* sp. strain WSM1743 associated with the general COG functional categories

Code	Value	% age	Description
J	190	3.51	Translation, ribosomal structure and biogenesis
A	0	0.00	RNA processing and modification
K	369	6.82	Transcription
L	155	2.86	Replication, recombination and repair
B	2	0.04	Chromatin structure and dynamics
D	28	0.52	Cell cycle control, Cell division, chromosome partitioning
V	82	1.51	Defense mechanisms
T	213	3.93	Signal transduction mechanisms
M	253	4.67	Cell wall/membrane/envelope biogenesis
N	94	1.74	Cell motility
U	112	2.07	Intracellular trafficking, secretion, and vesicular transport
O	180	3.33	Posttranslational modification, protein turnover, chaperones
C	357	6.60	Energy production and conversion
G	403	7.45	Carbohydrate transport and metabolism
E	642	11.86	Amino acid transport and metabolism
F	88	1.63	Nucleotide transport and metabolism
H	207	3.82	Coenzyme transport and metabolism
I	333	6.15	Lipid transport and metabolism
P	285	5.27	Inorganic ion transport and metabolism
Q	242	4.47	Secondary metabolite biosynthesis, transport and catabolism
R	659	12.17	General function prediction only
S	519	9.59	Function unknown
-	3,159	39.57	Not in COGS

The total is based on the total number of protein coding genes in the genome

A comparison of the genome of WSM1743 to that of USDA 6<sup>T</sup> [42] reveals that WSM1743 has a lower GC content, gene count, coding base count %, rRNAcount, COG % and transmembrane %. In contrast, the paralogs % is much higher for WSM1743 than for USDA 6<sup>T</sup> (81.75 % versus 42.54 %). These two genomes are included within a group of 54 *Bradyrhizobium* genomes that have been deposited into the IMG database [26]. Within this group, strains known to symbiotically fix nitrogen all contain the nitrogenase-RXN MetaCyc pathway that is characterized by the multiprotein nitrogenase complex. The genome of *Bradyrhizobium* sp. WSM1743, in conjunction with the other *Bradyrhizobium* genomes, will be important for on-going comparative and functional analyses of the plant microbe interactions required for the successful establishment of native Australian legume symbioses.

## Additional files

**Additional file 1: Table S1.** Associated MIGS record for WSM1743. (PDF 74 kb)

**Additional file 2: Annotation Summary, GenBank Accession Summary, Strain ID Summary, Plant Name Summary, Scientific Name Summary and Reference Search Summary.** (DOC 78 kb)

## Abbreviations

GEBA-RNB: Genomic encyclopedia of bacteria and archaea – root nodule bacteria; JGI: Joint genome institute; TY: Trypton yeast; CTAB: Cetyl trimethyl ammonium bromide; WSM: Western Australian soil microbiology; IMG-ER: Integrated microbial genomes-expert review; NCBI: National centre for biotechnology information; N<sub>2</sub>: Dinitrogen.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

RY and JH supplied the strain and background information for this project, TR supplied DNA to JGI and performed all imaging, LE and SDM drafted the paper, RT provided financial support and all other authors were involved in sequencing the genome and editing the final manuscript. 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, Lawrence Livermore National Laboratory under Contract No. DE-AC52-07NA27344, and Los Alamos National Laboratory under contract No. DE-AC02-06NA25396.

## Author details

<sup>1</sup>Centre for Rhizobium Studies, Murdoch University, Murdoch, Western Australia. <sup>2</sup>Centre for Phytophthora Science and Management (CPSM), Murdoch University, Murdoch, Western Australia. <sup>3</sup>DOE Joint Genome Institute, Walnut Creek, CA, USA. <sup>4</sup>Biological Data Management and Technology Center, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. <sup>5</sup>Department of Biological Sciences, King Abdulaziz, Jeddah, Saudi Arabia. <sup>6</sup>Department of Agriculture and Food, South Perth, Western Australia.

Received: 9 December 2014 Accepted: 8 October 2015

Published online: 26 October 2015

## References

1. Terpolilli JJ, Hood GA, Poole PS. What determines the efficiency of N<sub>2</sub>-fixing *Rhizobium*-Legume symbioses. *Adv Microb Physiol.* 2012;60:325–89.
2. Bohlool BB, Ladha JK, Garrity DP, George T. Biological nitrogen fixation for sustainable agriculture - A perspective. *Plant Soil.* 1992;141:1–11.
3. Donald CM. Temperate pasture species. Canberra: Australian National University Press; 1970.
4. Yates RJ, Howieson JG, Nandasena KG, O'Hara GW. Root-nodule bacteria from indigenous legumes in the north-west of Western Australia and their interaction with exotic legumes. *Soil Biol Biochem.* 2004;36:1319–29.
5. Wilson PG, Rowe R. A revision of the Indigofereae (Fabaceae) in Australia. 2. Indigofera species with trifoliate and alternately pinnate leaves. *Telopea.* 2008;12:293–307.
6. Wilson PG, Rowe R. A revision of the Indigofereae (Fabaceae) in Australia. 1. Indigofera and the simple or unifoliate species of Indigofera. *Telopea.* 2004;10:651–82.
7. Glenn AR, Reeve WG, Tiwari RP, Dilworth MJ. Acid tolerance in root nodule bacteria. *Novartis Found Symp.* 1999;221:112–26. discussion 126–130.
8. Reeve W, Ardley J, Tian R, Eshragi L, Yoon J, Ngamwisetkun P, et al. A genomic encyclopedia of the root nodule bacteria: Assessing genetic diversity through a systematic biogeographic survey. *Stand Genomic Sci* In press.
9. Howieson JG, Ewing MA, D'antuono MF. Selection for acid tolerance in *Rhizobium meliloti*. *Plant Soil.* 1988;105:179–88.

10. Kuykendall LD. Genus I. *Bradyrhizobium*. In: Brenner DJ, Krieg NR, Staley JT, editors. *Bergey's Manual of Systematic Bacteriology*, vol. 2 part C. 2nd ed. New York: Springer; 2005. p. 438–43.
11. 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.
12. Jordan DC. NOTES: 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.
13. Eckhardt MM, Baldwin IL, Fred EB. Studies of the Root-Nodule Organism of *Lupinus*. *J Bacteriol*. 1931;21:273–85.
14. Yao ZY, Kan FL, Wang ET, Wei GH, Chen WX. Characterization of rhizobia that nodulate legume species of the genus *Lespedeza* and description of *Bradyrhizobium yuanmingense* sp nov. *Int J Syst Evol Microbiol*. 2002;52:2219–30.
15. Zhang YF, Wang ET, Tian CF, Wang FQ, Han LL, Chen WF, et al. *Bradyrhizobium elkani*, *Bradyrhizobium yuanmingense* and *Bradyrhizobium japonicum* are the main rhizobia associated with *Vigna unguiculata* and *Vigna radiata* in the subtropical region of China. *FEMS Microbiol Lett*. 2008;285:146–54.
16. 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. *Nat Biotechnol*. 2008;26:541–7.
17. 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.
18. Reddy TB, Thomas AD, Stamatis D, Bertsch J, Isbandi M, Jansson J, et al. The Genomes OnLine Database (GOLD) v. 5: a metadata management system based on a four level (meta)genome project classification. *Nucleic Acids Res*. 2015;43(Database issue):D1099–106. doi: 10.1093/nar/gku950.
19. Field D, Amaral-Zettler L, Cochrane G, Cole JR, Dawyndt P, Garrity GM, et al. The Genomic Standards Consortium. *PLoS Biol*. 2011;9:e1001088.
20. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. *Proc Natl Acad Sci U S A*. 1990;87:4576–9.
21. Garrity GM, Bell JA, Lilburn T. Class I. *Alphaproteobacteria* class. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. *Bergey's Manual of Systematic Bacteriology*. Secondth ed. New York: Springer - Verlag; 2005.
22. Validation of publication of new names and new combinations previously effectively published outside the IJSEM. *Int J Syst Evol Microbiol* 2005, 55:2235–2238
23. Kuykendall LD. Order VI. *Rhizobiales* ord. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. *Bergey's Manual of Systematic Bacteriology*. Secondth ed. New York: Springer - Verlag; 2005. p. 324.
24. Garrity GM, Bell JA, Liburn T. Family VII. *Bradyrhizobiaceae*. In: Brenner DJ, Krieg NR, Staley JT, editors. *Bergey's Manual of Systematic Bacteriology*. Volume 2 part C. New York: Springer; 2005. p. 438–75.
25. 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.
26. 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.
27. 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.
28. GOLD ID *Bradyrhizobium* sp. WSM1743 [https://gold.jgi-psf.org/project?id=9884]. Access date 23/10/2015.
29. Beringer JE. R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol*. 1974;84:188–98.
30. CTAB DNA extraction protocol [http://jgi.doe.gov/collaborate-with-jgi/pmo-overview/protocols-sample-preparation-information/]. Access date 23/10/2015.
31. JGI Website [http://www.jgi.doe.gov]. Access date 23/10/2015.
32. Zerbino D, Birney E. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res*. 2008;18:821–9.
33. wgsim [https://github.com/lh3/wgsim]. Access date 23/10/2015.
34. 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. *Proc Natl Acad Sci*. 2011;108:1513–8.
35. 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.
36. 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.
37. 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.
38. 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.
39. 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.
40. INFERNAL. Inference of RNA alignments [http://infernal.janelia.org]. Access date 23/10/2015.
41. 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.
42. Kaneko T, Maita H, Hirakawa H, Uchiike N, Minamisawa K, Watanabe A, et al. Complete genome sequence of the soybean symbiont *Bradyrhizobium japonicum* strain USDA6T. *Genes*. 2011;2:763–87.

**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

