

Lawrence Berkeley National Laboratory

LBL Publications

Title

High-quality permanent draft genome sequence of *Ensifer* sp. PC2, isolated from a nitrogen-fixing root nodule of the legume tree (Khejri) native to the Thar Desert of India

Permalink

<https://escholarship.org/uc/item/6pb4703k>

Journal

Environmental Microbiome, 11(1)

ISSN

2524-6372

Authors

Gehlot, Hukam Singh

Ardley, Julie

Tak, Nisha

et al.

Publication Date

2016

DOI

10.1186/s40793-016-0157-7

Peer reviewed

SHORT GENOME REPORT

Open Access



High-quality permanent draft genome sequence of *Ensifer* sp. PC2, isolated from a nitrogen-fixing root nodule of the legume tree (Khejri) native to the Thar Desert of India

Hukam Singh Gehlot¹, Julie Ardley², Nisha Tak¹, Rui Tian², Neetu Poonar¹, Raju R. Meghwal¹, Sonam Rathi¹, Ravi Tiwari², Wan Adnawani², Rekha Seshadri³, T. B. K. Reddy³, Amrita Pati³, Tanja Woyke³, Manoj Pillay⁴, Victor Markowitz⁴, Mohammed N. Baeshen⁵, Ahmed M. Al-Hejin⁵, Natalia Ivanova³, Nikos Kyrpides^{3,5} and Wayne Reeve^{2*}

Abstract

Ensifer sp. PC2 is an aerobic, motile, Gram-negative, non-spore-forming rod that was isolated from a nitrogen-fixing nodule of the tree legume *P. cineraria* (L.) Druce (Khejri), which is a keystone species that grows in arid and semi-arid regions of the Indian Thar desert. Strain PC2 exists as a dominant saprophyte in alkaline soils of Western Rajasthan. It is fast growing, well-adapted to arid conditions and is able to form an effective symbiosis with several annual crop legumes as well as species of mimosoid trees and shrubs. Here we describe the features of *Ensifer* sp. PC2, together with genome sequence information and its annotation. The 8,458,965 bp high-quality permanent draft genome is arranged into 171 scaffolds of 171 contigs containing 8,344 protein-coding genes and 139 RNA-only encoding genes, and is 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.

Keywords: Root-nodule bacteria, Nitrogen fixation, Symbiosis, *Ensifer*, *Prosopis*

Introduction

The genus *Prosopis* (family *Leguminosae*, sub-family *Mimosoideae* [1]) comprises about 44 species that are widely distributed in the world's semi-arid regions, mostly in North and South America with a few species found in Africa and south west Asia [2–4]. Several species have been widely introduced throughout the world over the last 200 years [5]. *Prosopis* may have evolved from *P. africana* (Guill. & Perr.) Taub., in which various character traits and small genome size (392–490 Mbp) indicate that it is a primitive species [2]. According to Burkart [2], *Prosopis* is an old genus that diverged early into several principal lineages, with some of these

lineages producing more recent episodes of speciation. This is supported by a recent molecular dating analysis that places the divergence of the New World *Prosopis* Sections during the Oligocene (33.9 to 23.03 Mya) [6], which is remarkably ancient considering that the sub-family *Mimosoideae* originated between 42–50 Mya [7]. Section *Prosopis* consists of three species, *Prosopis cineraria* (L.) Druce, *P. farcta* (Banks et Sol.) Eig. and *P. koelziana* Burkart, which are native to North Africa and Asia [6].

P. cineraria is endemic to arid and semi-arid regions of the Indian Thar Desert and is designated as the state tree of Rajasthan [8]. It symbolizes the sacred mythological “Kalpa Vriksh” (wish tree) of the desert and is historically important, as it has been worshiped since ancient times by many rural communities in these arid regions. *P. cineraria* is a multipurpose tree used as food,

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

²Centre for Studies, Murdoch University, Murdoch, Western Australia, Australia

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

fodder, shelter and medicine by the local inhabitants. It is an important component of agro forestry, agrisilvicultural and silvopastoral systems in the alkaline soil of the Thar Desert. The tree is extremely drought and salt tolerant, having a deep root system (>100 metres) that helps in acquiring nutrients and moisture from deeper soil layers. It produces green pods that are rich in nutrients and antioxidants and eaten as a vegetable in the hot summer [9]. *P. cineraria* is a good candidate for rehabilitation of dry, marginal or degraded lands of low fertility and/or high salinity. It plays a vital role as a soil binder in the stabilization of sand dunes and enriches poor desert soil by fixing atmospheric nitrogen in association with its rhizobial microsymbionts [10–13].

Prosopis is a promiscuous genus, being nodulated by a wide range of taxonomically diverse rhizobia. Mesquite (Torr.) in the Sonoran Desert, California is nodulated by diverse strains of fast- and slow-growing rhizobia [14]. *Mesorhizobium chacoense* CECT 5336^T is a microsymbiont of *Prosopis alba* Griseb. growing in the Chaco Arido region in Argentina [15], whereas in Spain is nodulated by strains of *Ensifer medicae*, *E. meliloti* and *Rhizobium giardinii* [16]. In Africa, the introduced *Prosopis* species *P. chilensis* (Molina) Stuntz, *P. cineraria*, *P. juliflora* (Sw.) DC. and *P. pallida* (Willd.) Kunth are reported to nodulate with strains of *Ensifer arboris*, *E. kostiense*, *E. saheli* and *E. teranga* [17, 18] and *P. juliflora* also forms effective symbioses with strains of *Mesorhizobium plurifarium* [19] and *Rhizobium etli* [20]. Nodulation of *P. cineraria* growing in its native range was first described by Basak and Goyal [10]. Recently, *P. cineraria* and other native legumes growing in the alkaline soils of the Thar desert have been reported to nodulate with a dominant novel group of *Ensifer* strains (PC2, TW10, TP13, RA9, TV3 and TF7) that are closely related to African and Australian *Ensifer* strains on the basis of

16S rRNA sequence similarity, but form a distinct, well-separated cluster [21, 22].

The indigenous rhizobia of wild tree legumes growing in such arid and harsh environments have superior tolerance to abiotic factors such as salt stress, elevated temperatures and drought and can be used as inoculants for wild as well as crop legumes cultivated in reclaimed desert lands [10]. Because of its ability to nodulate the keystone species *P. cineraria* as well as crop legumes such as *Vigna radiata* (L.) R.Wilczek and *V. unguiculata* (L.) Walp. [21], strain PC2 has therefore been selected as part of the DOE Joint Genome Institute 2010 *Genomic Encyclopedia for Bacteria and Archaea-Root Nodule Bacteria* (GEBA-RNB) sequencing project [23]. Here we present a summary classification and a set of general features for *Ensifer* sp. strain PC2, together with a description of its genome sequence and annotation.

Organism information

Classification and features

Ensifer sp. PC2 is a motile, Gram-negative strain in the order *Rhizobiales* of the class *Alphaproteobacteria*. The rod shaped form (Fig. 1 Left and Center) has dimensions of approximately 0.3–0.5 μm in width and 1.25–1.5 μm in length. It is fast growing, forming colonies within 3–4 days when grown on half strength Lupin Agar [24], tryptone-yeast extract agar (TY) [25] or a modified yeast-mannitol agar (YMA) [26] at 28 °C. Colonies on $\frac{1}{2}$ LA are white, opaque, slightly domed and slightly mucoid with smooth margins (Fig. 1 Right).

Figure 2 shows the phylogenetic relationship of *Ensifer* sp. PC2 in a 16S rRNA sequence based tree. This strain is the most similar to *Ensifer saheli* LMG 7837^T based on the 16S rRNA gene alignment, with sequence identities of 99.41 % over 1,366 bp, as determined using the

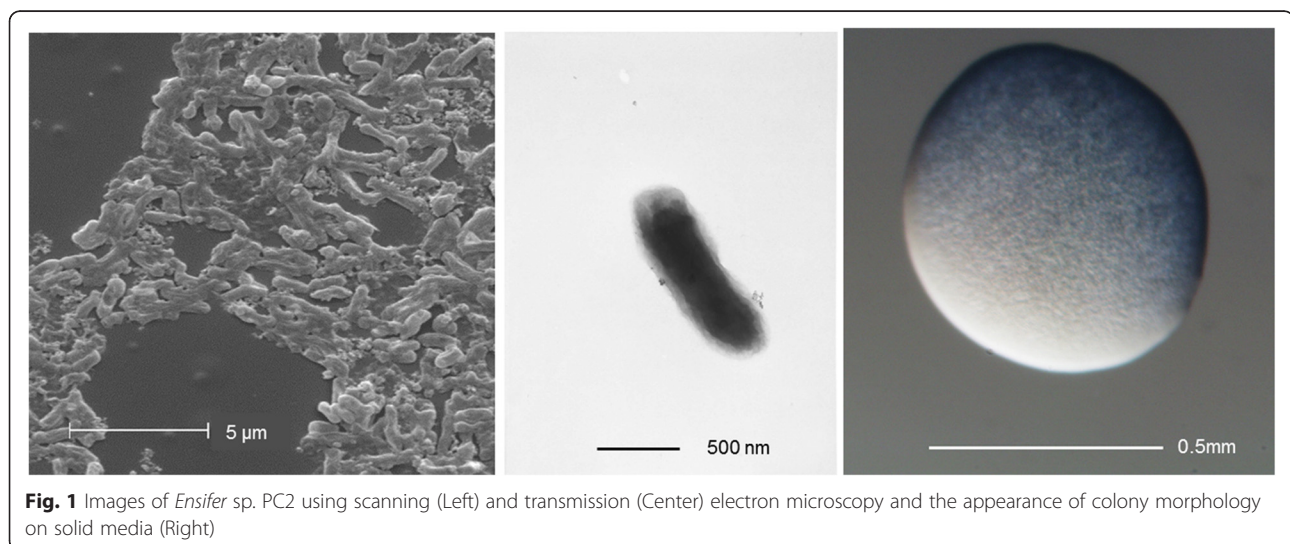


Fig. 1 Images of *Ensifer* sp. PC2 using scanning (Left) and transmission (Center) electron microscopy and the appearance of colony morphology on solid media (Right)

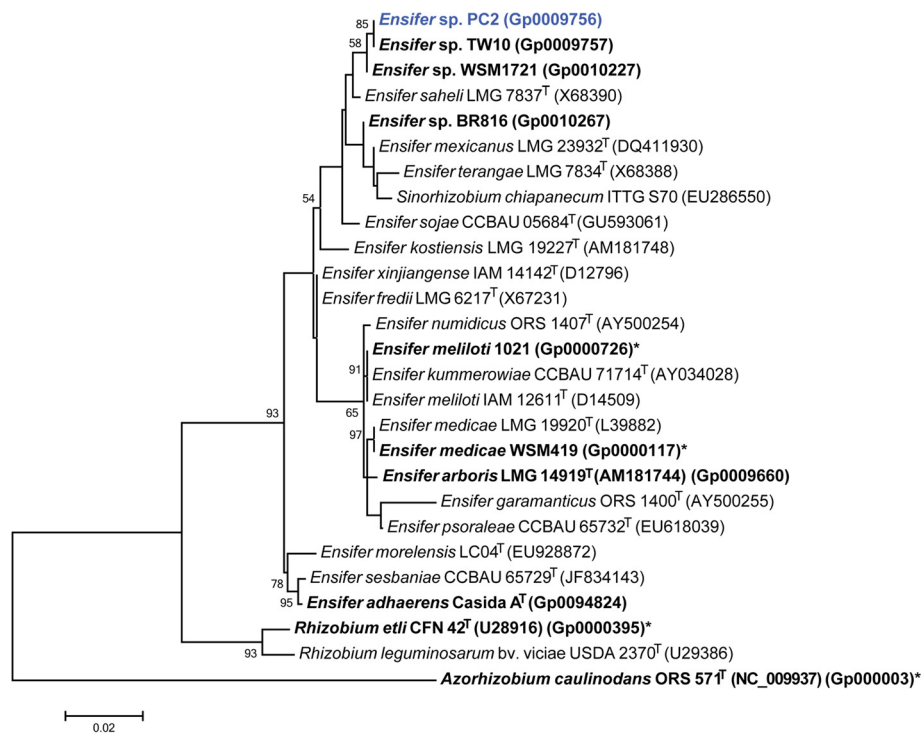


Fig. 2 Phylogenetic tree showing the relationship of *Ensifer* sp. PC2 (shown in bold blue print) to *Ensifer* spp. and other root nodule bacteria species in the order *Rhizobiales*, based on aligned sequences of the 16S rRNA gene (1,283 bp internal region). (The species name “*Sinorhizobium chiapanecum*” has not been validly published.) *Azorhizobium caulinodans* ORS 571^T was used as an outgroup. All sites were informative and there were no gap-containing sites. Phylogenetic analyses were performed using MEGA, version 6 [44]. The tree was built using the Maximum-Likelihood method with the General Time Reversible model [45]. Bootstrap analysis [46] 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 [29] are in bold font and the GOLD ID is provided after the GenBank accession number, where this is available. Finished genomes are indicated with an asterisk.

EzTaxon-e database, which contains the sequences of validly published type strains [27]. The PC2 16S rRNA gene sequence has 100 % sequence identity with that of another Indian Thar Desert rhizobial strain, *Ensifer* sp. TW10, isolated from a nodule of the perennial legume *Tephrosia wallichii* [22]. Minimum Information about the Genome Sequence for PC2 is provided in Table 1 and Additional file 1: Table S1.

Symbiotaxonomy

Ensifer sp. strain PC2 is able to nodulate and fix nitrogen with both mimosoid and papilionoid legume hosts. It is interesting to note that sp. PC2 is able to nodulate and fix nitrogen with *Acacia saligna* (Labill.) Wendl., a promiscuous legume tree that mainly nodulates with species of in its native southwestern Australia range [28]. PC2 also effectively nodulates the Central American mimosoid tree *Leucaena leucocephala* (Lam.) de Wit. PC2 appears to be a relatively promiscuous strain that has potential to be used as an inoculant for crop legumes species such as *Vigna radiata* (L.) Wilczek and *V. unguiculata* (L.) Walp.. The symbiotic characteristics of sp. strain PC2 on a range of selected hosts are summarised in Additional file 2: Table S2.

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 project at the U.S. Department of Energy, Joint Genome Institute. The genome project is deposited in the Genomes OnLine Database [29] and a high-quality permanent draft genome sequence is deposited in IMG [30]. Sequencing, finishing and annotation were performed by the JGI [31]. A summary of the project information is shown in Table 2.

Growth conditions and genomic DNA preparation

Ensifer sp. PC2 was streaked onto TY solid medium [25, 32] and grown at 28 °C for three days to obtain well grown, well separated colonies, then a single colony was selected and used to inoculate 5 ml TY broth medium. The culture was grown for 48 h on a gyratory shaker (200 rpm) at 28 °C. Subsequently 1 ml was used to inoculate 60 ml TY broth medium and grown on a

Table 1 Classification and general features of *Ensifer* sp. PC2 in accordance with the MIGS recommendations [47] published by the Genome Standards Consortium [48]

MIGS ID	Property	Term	Evidence code ^a
	Current classification	Domain Bacteria	TAS [49]
		Phylum <i>Proteobacteria</i>	TAS [50, 51]
		Class <i>Alphaproteobacteria</i>	TAS [52, 53]
		Order <i>Rhizobiales</i>	TAS [54]
		Family <i>Rhizobiaceae</i>	TAS [55]
		Genus <i>Ensifer</i>	TAS [56–58]
		Species <i>Ensifer</i> sp.	IDA
		Strain: PC2	TAS [21]
		Gram stain	Negative
	Cell shape	Rod	IDA
	Motility	Motile	IDA
	Sporulation	Non-sporulating	NAS
	Temperature range	10–40 °C	IDA
	Optimum temperature	28 °C	IDA
	pH range; Optimum	5–9.5; 6.5–8	IDA
	Carbon source	Mannitol, tryptone, yeast extract	TAS [21]
MIGS-6	Habitat	Soil; root nodule on host (<i>Prosopis</i> (L.) Druce)	TAS [21]
MIGS-6.3	Salinity	0.89–2.0 % (w/v)	NAS
MIGS-22	Oxygen requirement	Aerobic	TAS [21]
MIGS-15	Biotic relationship	Free living, symbiotic	TAS [21]
MIGS-14	Pathogenicity	Biosafety level 1	TAS [59]
MIGS-4	Geographic location	Jodhpur, Indian Thar Desert	TAS [21]
MIGS-5	Sample collection	October, 2009	TAS [21]
MIGS-4.1	Latitude	26.27061	TAS [21]
MIGS-4.2	Longitude	73.021177	TAS [21]
MIGS-4.3	Depth	0–10 cm	NAS
MIGS-4.4	Altitude	234 m	TAS [21]

^aEvidence 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 [60], [<http://geneontology.org/page/guide-go-evidence-codes>]

gyratory shaker (200 rpm) at 28 °C until OD_{600nm} 0.6 was reached. DNA was isolated from 60 ml of cells using a CTAB bacterial genomic DNA isolation method [<http://jgi.doe.gov/collaborate-with-jgi/pmo-overview/protocols-sample-preparation-information/>]. Final concentration of the DNA was 0.5 mg ml⁻¹.

Genome sequencing and assembly

The draft genome of sp. PC2 was generated at the JGI using the Pacific Biosciences (PacBio) technology. A PacBio SMRTbell™ library was constructed and sequenced on the PacBio RS platform, which generated 403,200 filtered subreads totaling 1.1 Gbp. All general aspects of library construction and sequencing performed at the JGI can be found on the JGI website [<http://jgi.doe.gov/>].

The raw reads were assembled using HGAP (version: 2.0.12.0.1) [33]. The final draft assembly contained 171 contigs in 171 scaffolds, totalling 8.5 Mbp in size. The input read coverage was 181.5x.

Genome annotation

Genes were identified using Prodigal [34] as part of the DOE-JGI genome annotation pipeline [35, 36]. The predicted CDSs were translated and used to search the National Center for Biotechnology Information nonredundant database, UniProt, TIGRFam, Pfam, KEGG, COG, and InterPro databases. The tRNAScanSE tool [37] was used to find tRNA genes, whereas ribosomal RNA genes were found by searches against models of the ribosomal RNA genes built from SILVA [38]. Other non-coding RNAs such as the RNA components of the protein secretion complex

Table 2 Genome sequencing project information for *Ensifer* sp. PC2

MIGS ID	Property	Term
MIGS-31	Finishing quality	High-quality draft
MIGS-28	Libraries used	Pacbio SMRTbell™ library
MIGS-29	Sequencing platforms	Pacific Biosciences RS
MIGS-31.2	Fold coverage	181.5x
MIGS-30	Assemblers	HGAP (version: 2.0.12.0.1)
MIGS-32	Gene calling methods	Prodigal 1.4
	Locus Tag	B077 [http://www.ncbi.nlm.nih.gov/bioproject/?term=B077]
	GenBank ID	LATE00000000
	GenBank Date of Release	Apr 20 2015
MIGS-13	GOLD ID	Gp0009756 [https://gold.jgi-psf.org/project?id=9756]
	BIOPROJECT ID	PRNJA169749
	Source Material Identifier	PC2, WSM4384
	Project relevance	Symbiotic N ₂ fixation, agriculture

Table 3 Genome statistics for *Ensifer* sp. PC2

Attribute	Value	% of Total
Genome size (bp)	8,458,965	100.00
DNA coding (bp)	7,124,539	84.22
DNA G + C (bp)	5,187,131	61.32
DNA scaffolds	171	100.00
Total genes	8483	100.00
Protein coding genes	8344	98.36
RNA genes	139	1.64
Pseudo genes	0	-
Genes in internal clusters	513	6.05
Genes with function prediction	6290	74.15
Genes assigned to COGs	5205	61.36
Genes assigned Pfam domains	6533	77.01
Genes with signal peptides	645	7.60
Genes with transmembrane helices	1733	20.43
CRISPR repeats	1	-

and the RNase P were identified by searching the genome for the corresponding Rfam profiles using INFERNAL [39]. Additional gene prediction analysis and manual functional annotation was performed within the Integrated Microbial Genomes platform [40] developed by the Joint Genome Institute, Walnut Creek, CA, USA [41].

Genome properties

The genome is 8,458,965 nucleotides with 61.32 % GC content (Table 3) and comprised of 171 scaffolds of 171 contigs. From a total of 8,483 genes, 8,344 were protein encoding and 139 RNA only encoding genes. The majority of protein-coding genes (76.34 %) were assigned a putative function whilst the remaining genes were annotated as

Table 4 Number of genes of sp. PC2 associated with general COG functional categories

Code	Value	%age of total (5,205)	Description
J	236	4.01	Translation, ribosomal structure and biogenesis
A	0	0.00	RNA processing and modification
K	514	8.74	Transcription
L	172	2.92	Replication, recombination and repair
B	2	0.03	Chromatin structure and dynamics
D	47	0.80	Cell cycle control, cell division, chromosome partitioning
Y	0	0.00	Nuclear structure
V	115	1.95	Defense mechanisms
T	271	4.61	Signal transduction mechanisms
M	331	5.63	Cell wall/membrane/envelope biogenesis
N	101	1.72	Cell motility
Z	0	0.00	Cytoskeleton
W	44	0.75	Extracellular structures
U	132	2.24	Intracellular trafficking, secretion, and vesicular transport
O	213	3.62	Posttranslational modification, protein turnover, chaperones
C	351	5.97	Energy production and conversion
G	548	9.31	Carbohydrate transport and metabolism
E	598	10.16	Amino acid transport and metabolism
F	116	1.97	Nucleotide transport and metabolism
H	277	4.71	Coenzyme transport and metabolism
I	227	3.86	Lipid transport and metabolism
P	309	5.25	Inorganic ion transport and metabolism
Q	171	2.91	Secondary metabolite biosynthesis, transport and catabolism
R	593	10.08	General function prediction only
S	395	6.71	Function unknown
X	120	2.04	Mobilome: prophages, transposons
-	3,278	38.64	Not in COGS

hypothetical. The distribution of genes into COGs functional categories is presented in Table 4.

Insights from the genome sequence

With a genome totaling 8.5 Mbp in size, *Ensifer* sp. PC2 is approximately 25 % larger than the average *Ensifer* genome in GenBank. Although PC2 shares 100 % 16S rRNA sequence identity and 99.17 Average Nucleotide Identity with *Ensifer* sp. TW10, also isolated from a Thar Desert woody legume, the genome of TW10 has a smaller size of 6.8 Mbp. PC2 contains over 1,000 genes that are not found in TW10, including two plasmid replication initiator proteins and a suite of genes (*vir/trb*) involved in conjugative transfer. From this it is assumed that the PC2 genome is multipartite and contains at least one conjugative plasmid. In PC2, 38.64 % of genes have not been assigned to a COG functional category, whereas in TW10, only 31.55 % have not been assigned to a COG functional category. Compared with TW10, PC2 has a much higher number of genes assigned to the mobilome category (54 and 120 genes, respectively) and to extracellular structures (29 and 44 genes, respectively).

Conclusion

Based on the 16S rRNA gene alignment, *Ensifer* sp. PC2 is most closely related to *Ensifer* sp. TW10 and *Ensifer* sp. WSM1721, two strains isolated from perennial legumes growing in arid climates and alkaline soils in India and Australia, respectively [21, 42]. *Ensifer fredii* strains isolated from Chinese soybean were also superdominant in sampling sites with alkaline-saline soils [43], which suggests that the biogeographic distribution of several *Ensifer* spp. is linked to their adaptation to alkaline soils. Further, this suggests that the symbiotic associations formed by promiscuous legumes, such as *Prosopis*, are likely to vary depending on which rhizobial genera are best adapted to the edaphic conditions in which the host is growing.

The ability of PC2 to fix nitrogen with both *P. cineraria* (L.) Druce and the crop legumes *Vigna radiata* (L.) R. Wilczek and *V. unguiculata* (L.) Walp. makes it a valuable inoculant strain for use in arid, alkaline regions such as the Thar desert. Analysis of the PC2 sequenced genome and comparison with the genomes of sequenced *Ensifer* spp. and other rhizobia will provide insights into the molecular basis of the patterns seen in rhizobial biogeographic distributions and associations with plant hosts and into the molecular determinants of rhizobial tolerance to arid and alkaline environments.

Additional files

Additional file 1: Table S1. Associated MIGS record for PC2 (DOCX 19 kb)

Additional file 2: Table S2. Nodulation and N₂ fixation properties of *Ensifer* sp. PC2 on selected legume hosts. (DOCX 17 kb)

Abbreviations

ANI: Average nucleotide identity; GEBA-RNB: Genomic encyclopedia for bacteria and archaea-root nodule bacteria; IMG: Integrated microbial genomes.

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 Thomson (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 and acknowledge King Abdulaziz University Vice President for Educational Affairs Prof. Alyoubi for his support. We sincerely acknowledge funding received from University Grant Commission, New Delhi, India for UGC-SAPII-CAS-I, UGC-BSR Research Start-Up- Grant (F.30-16/2014-BSR); Department of Biotechnology, Govt. of India (BT/PR11461/AGR/21/270/2008). We also thank the Crawford Fund Award-ATSE, Australia for funding HG and NT to conduct research at the CRS.

Authors' contribution

HG supplied the strain, DNA and background information for this project, TR supplied DNA to JGI and performed all imaging, JA and NT drafted the paper, MNB and AMA-H 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.

Competing interests

The authors declare that they have no competing interests.

Author details

¹BNF and Stress Biology Lab., Department of Botany, J.N. Vyas University, Jodhpur 342001, India. ²Centre for Studies, Murdoch University, Murdoch, Western Australia, Australia. 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.

Received: 4 September 2015 Accepted: 18 May 2016

Published online: 23 June 2016

References

- Lewis G, Schrire B, Mackinder B, Lock M. Legumes of the World. Richmond, Surrey: Royal Botanic Gardens, Kew; 2005.
- Burkart A. A monograph of the genus *Prosopis* (Leguminosae Subfam, Mimosoideae). J Arnold Arbor. 1976;57:219–49, 450–525.
- Felker P. Unusual physiological properties of the arid adapted tree legume *Prosopis* and their applications in developing countries. In: De la Barrera E, Smith WK, editors. Perspectives in Biophysical Plant Ecophysiology: A Tribute to Park S Nobel. Mexico City: Universidad Nacional Autónoma de México; 2009. p. 221–55.
- Sprent JI. Nodulation in Legumes. Richmond, Surrey: Royal Botanic Gardens, Kew; 2001.
- Pasiecznik NM, Harris PJC, Smith SJ: Identifying tropical *Prosopis* species: a field guide. Coventry: Henry Doubleday Research Association; 2003
- Catalano SA, Vilardi JC, Tosto D, Saidman BO. Molecular phylogeny and diversification history of *Prosopis* (Fabaceae: Mimosoideae). Biol J Linnean Soc. 2008;93:621–40.
- Lavin M, Herendeen PS, Wojciechowski MF. Evolutionary rates analysis of *Leguminosae* implicates a rapid diversification of lineages during the Tertiary. Syst Biol. 2005;54:575–94.
- Bhandari MM. Flora of the Indian Desert. Jodhpur: MPS Repros; 1990.

9. Rani B, Singh U, Sharma R, Gupta A, Dhawan NG, Sharama AK, Sharma S, Maheshwari RK. *Prosopis cineraria* (L.) Druce: A desert tree to brace livelihood in Rajasthan. *Asian J Pharmaceut Res Health Care*. 2013;5:58–64.
10. Basak MK, Goyal SK. Studies on tree legumes: nodulation pattern and characterization of the symbiont. *Ann Arid Zone*. 1975;14:367–70.
11. Felker P: Mesquite: an all-purpose leguminous arid land tree. In *New Agricultural Crops*. Edited by Ritchie GA. Golden, Colorado: American Association for the Advancement of Science, Westview Press; 1979: 89–132.
12. Hocking D. *Trees for Drylands*. Madison, Wisconsin: International Science Publisher; 1993.
13. Tewari JC, Pasiecznik NM, Harsh LN, Harris PJC. *Prosopis* species in the arid and semi-arid zones of India. In *Proceedings of the Prosopis Conference; Central Arid Zone Research Institute, Jodhpur, Rajasthan, India*. The *Prosopis* Society of India and the Henry Doubleday Research Association; 1993: 128 pp.
14. Thomas PM, Golly KF, Zyskind JW, Virginia RA. Variation of clonal, mesquite-associated rhizobial and bradyrhizobial populations from surface and deep soils by symbiotic gene region restriction fragment length polymorphism and plasmid profile analysis. *Appl Environ Microb*. 1994;60:1146–53.
15. Velázquez E, Igual JM, Willems A, Fernández MP, Muñoz E, Mateos PF, Abril A, Toro N, Normand P, Cervantes E, Gillis M, Martínez-Molina E. *Mesorhizobium chacoense* sp. nov., a novel species that nodulates *Prosopis alba* in the Chaco Arido region (Argentina). *Int J Syst Evol Microbiol*. 2001;51:1011–21.
16. Iglesias O, Rivas R, García-Fraile P, Abril A, Mateos PF, Martínez-Molina E, Velázquez E. Genetic characterization of fast-growing rhizobia able to nodulate *Prosopis alba* in North Spain. *FEMS Microbiol Lett*. 2007;277:210–6.
17. Räsänen L, Sprent J, Lindström K. Symbiotic properties of sinorhizobia isolated from *Acacia* and *Prosopis* nodules in Sudan and Senegal. *Plant Soil*. 2001;235:193–210.
18. Nick G, de Lajudie P, Eardly BD, Suomalainen S, Paulin L, Zhang X, Gillis M, Lindström K. *Sinorhizobium arboris* sp. nov. and *Sinorhizobium kostense* sp. nov., isolated from leguminous trees in Sudan and Kenya. *Int J Syst Bacteriol*. 1999;49: 1359–68.
19. de Lajudie P, Willems A, Nick G, Moreira F, Molouba F, Hoste B, Torck U, Neyra M, Collins MD, Lindström K, Dreyfus B, Gillis M. Characterization of tropical tree rhizobia and description of *Mesorhizobium plurifarum* sp. nov. *Int J Syst Bacteriol*. 1998;48:369–82.
20. Odee DW, Haukka K, McInroy SG, Sprent JI, Sutherland JM, Young JPW. Genetic and symbiotic characterization of rhizobia isolated from tree and herbaceous legumes grown in soils from ecologically diverse sites in Kenya. *Soil Biol Biochem*. 2002;34:801–11.
21. Gehlot HS, Panwar D, Tak N, Tak A, Sankhla IS, Poonar N, Parihar R, Shekhawat NS, Kumar M, Tiwari R, Ardley J, James EK, Sprent JI. Nodulation of legumes from the Thar desert of India and molecular characterization of their rhizobia. *Plant Soil*. 2012;357:227–43.
22. Tak N, Gehlot HS, Kaushik M, Choudhary S, Tiwari R, Tian R, Hill YJ, Bräu L, Goodwin L, Han J, Liolios K, Huntemann M, Palaniappan K, Pati A, Mavromatis K, Ivanova NN, Markowitz VM, Woyke T, Kyrpides NC, Reeve WG. Genome sequence of *Ensifer* sp. TW10; a *Tephrosia wallichii* (Biyani) microsymbiont native to the Indian Thar Desert. *Stand Genomic Sci*. 2013;9:304–14.
23. Reeve WG, Ardley JK, Tian R, Eshragi L, Yoon JW, Ngamwisetkun P, Sheshadri R, Ivanova NN, Kyrpides NC. A genomic encyclopedia of the root nodule bacteria: Assessing genetic diversity through a systematic biogeographic survey. *Stand Genomic Sci*. 2014;10:14.
24. Howieson JG, Ewing MA, D'antuono MF. Selection for acid tolerance in *Rhizobium meliloti*. *Plant Soil*. 1988;105:179–88.
25. Beringer JE. R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol*. 1974;84:188–98.
26. Vincent JM. *A manual for the practical study of the root-nodule bacteria*. International Biological Programme. Oxford, UK: Blackwell Scientific Publications; 1970.
27. Kim O-S, Cho Y-J, Lee K, Yoon S-H, Kim M, Na H, Park S-C, Jeon YS, Lee J-H, Yi H, Won S, Chun J. 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.
28. Marsudi NDS, Glenn AR, Dilworth MJ. Identification and characterization of fast- and slow-growing root nodule bacteria from South-Western Australian soils able to nodulate *Acacia saligna*. *Soil Biol Biochem*. 1999;31:1229–38.
29. Reddy TBK, Thomas AD, Stamatis D, Bertsch J, Isbandi M, Jansson J, Mallajosyula J, Pagani I, Lobos EA, Kyrpides NC. The Genomes OnLine Database (GOLD) v.5: a metadata management system based on a four level (meta)genome project classification. *Nucleic Acids Res*. 2014;43:D1099–106.
30. Markowitz VM, Chen I-MA, Palaniappan K, Chu K, Szeto E, Pillay M, Ratner A, Huang J, Woyke T, Huntemann M, Anderson I, Billis K, Varghese N, Mavromatis K, Pati A, Ivanova NN, Kyrpides NC. IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Res*. 2014;42:D560–7.
31. Mavromatis K, Land ML, Brettin TS, Quest DJ, Copeland A, Clum A, Goodwin L, Woyke T, Lapidus A, Klenk HP, Cottingham RW, Kyrpides NC. The fast changing landscape of sequencing technologies and their impact on microbial genome assemblies and annotation. *PLoS One*. 2012;7:e48837.
32. 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.
33. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Meth*. 2013;10:563–9.
34. 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.
35. 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.
36. Chen IMA, Markowitz VM, Chu K, Anderson I, Mavromatis K, Kyrpides NC, Ivanova NN. Improving microbial genome annotations in an integrated database context. *PLoS One*. 2013;8:e54859.
37. 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.
38. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glöckner FO. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res*. 2007;35:7188–96.
39. INFERNAL. Inference of RNA alignments [http://eddylab.org/infernal/]. Accessed 24 May 2016.
40. The Integrated Microbial Genomes (IMG) platform [http://img.jgi.doe.gov]. Accessed 24 May 2016.
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. 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.
43. Zhang YM, Li Y, Chen WF, Wang ET, Tian CF, Li QQ, Zhang YZ, Sui XH, Chen WX. Biodiversity and biogeography of rhizobia associated with soybean plants grown in the North China Plain. *Appl Environ Microb*. 2011;77:6331–42.
44. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol*. 2013;30(12):2725–9.
45. Nei M, Kumar S. *Molecular Evolution and Phylogenetics*. New York: Oxford University Press; 2000.
46. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*. 1985;39:783–91.
47. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen M, Angiuoli SV, Ashburner M, Axelrod N, Baldauf S, Ballard S, Boore JL, Cochrane G, Cole J, Dawyndt P, de Vos P, de Pamphilis C, Edwards R, Faruque N, Feldman R, Gilbert J, Gilna P, Glöckner FO, Goldstein P, Guralnick R, Haft D, Hancock D, 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.
48. Field D, Amaral-Zettler L, Cochrane G, Cole JR, Dawyndt P, Garrity GM, Gilbert J, Glöckner FO, Hirschman L, Karsch-Mizrachi I, Klenk H-P, Knight R, Kottmann R, Kyrpides N, Meyer F, San Gil I, Sansone S-A, Schriml LM, Sterk P, Tatusova T, Ussery DW, White O, Wooley J. The Genomic Standards Consortium. *PLoS Biology*. 2011;9:e1001088.
49. 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.
50. Validation of publication of new names and new combinations previously effectively published outside the IJSEM. *Int J Syst Evol Microbiol* 2005, 55: 2235–2238

51. Garrity GM, Bell JA, Phylum LT, XIV. *Proteobacteria* phyl. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. *Bergey's Manual of Systematic Bacteriology*, vol. 2. 2nd ed. Springer, New York: Part B; 2005. p. 1.
52. List of new names and new combinations previously effectively, but not validly, published. *International Journal of Systematic and Evolutionary Microbiology* 2006, 56:1–6.
53. Garrity GM, Bell JA, Lilburn T: Class I. *Alphaproteobacteria* class. nov. In *Bergey's Manual of Systematic Bacteriology*. Second edition. Edited by Garrity GM, Brenner DJ, Krieg NR, Staley JT: New York: Springer; 2005
54. Kuykendall LD: Order VI. *Rhizobiales* ord. nov. In *Bergey's Manual of Systematic Bacteriology*. Second edition. Edited by Garrity GM, Brenner DJ, Krieg NR, Staley JT: New York: Springer; 2005: 324
55. Kuykendall LD. Family I. *Rhizobizceae*. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. *Bergey's Manual of Systematic Bacteriology*. New York: Springer; 2005.
56. Kuykendall LD, Hashem FM, Wang ET: Genus VII. *Ensifer*. In *Bergey's Manual of Systematic Bacteriology*. Volume 2. Edited by Garrity GM, Brenner DJ, Krieg NR, Staley JT. New York: Springer; 2005: 358–361
57. Judicial Commission of the International Committee on Systematics of Prokaryotes. The genus name *Sinorhizobium* Chen et al. 1988 is a later synonym of *Ensifer* Casida 1982 and is not conserved over the latter genus name, and the species name '*Sinorhizobium adhaerens*' is not validly published. Opinion 84. *Int J Syst Evol Microbiol*. 2008;58:1973.
58. Casida LE. *Ensifer adhaerens* gen. nov., sp. nov.: a bacterial predator of bacteria in soil. *Int J Syst Bacteriol*. 1982;32:339–45.
59. Biological Agents: Technical rules for biological agents. [<http://www.baua.de/en/Topics-from-A-to-Z/BiologicalAgents/TRBA/TRBA.html>]. Accessed 24 May 2016.
60. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. *The Gene Ontology Consortium Nat Genet*. 2000;25:25–9.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

