Data in Brief

Genome sequence of *Bradyrhizobium* sp. LMTR 3, a diazotrophic symbiont of Lima bean (*Phaseolus lunatus*)

Ernesto Ormeño-Orrillo a,⁎, Luis Rey b, David Durán b, Carlos A. Canchaya c, Doris Zúñiga-Dávila a, Juan Imperial b,d, Esperanza Martínez-Romero e, Tomás Ruiz-Argüeso b

a Laboratorio de Ecología Microbiana y Biotecnología, Departamento de Biología, Facultad de Ciencias, Universidad Nacional Agraria La Molina, Lima, Peru

b Departamento de Biotecnología y Biología Vegetal, ETSI Agronómica, Alimentaria y de Biosistemas, Universidad Politécnica de Madrid, and Centro de Biotecnología y Genómica de Plantas (UPM-INIA), Spain

c Departamento de Bioquímica, Genética e Immunología, Universidad de Vigo, Vigo 36310, Spain

d CSIC, Madrid, Spain

e Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico

**A B S T R A C T**

*Bradyrhizobium* sp. LMTR 3 is a representative strain of one of the geno(species) of diazotrophic symbionts associated with Lima bean (*Phaseolus lunatus*) in Peru. Its 7.83 Mb genome was sequenced using the Illumina technology and found to encode a complete set of genes required for nodulation and nitrogen fixation, and additional genes putatively involved in root colonization. Its draft genome sequence and annotation have been deposited at GenBank under the accession number MAXC00000000.

1. Direct link to deposited data

   https://www.ncbi.nlm.nih.gov/nuccore/MAXC00000000

2. Introduction

Diazotrophic bacteria known as rhizobia engage in mutualistic symbioses with leguminous plants providing fixed nitrogen to their symbionts. The *Bradyrhizobium* genus includes rhizobia able to associate with several important leguminous crops like soybean, cowpea, peanut and Lima bean. *Bradyrhizobium* sp. LMTR 3 is a symbiont of Lima bean (*Phaseolus lunatus*) and was isolated from a root nodule collected in an agricultural field in La Molina (Lima, Peru) [1]. In comparison to symbionts obtained from other bean *Phaseolus* species, little genomic information is available for *P. lunatus* rhizobia. Here we present the genome sequence and annotation of strain LMTR 3 which is a representative of one of the four geno(species) of *Bradyrhizobium* symbionts associated with Lima bean in Peru [2].

3. Experimental design, materials and methods

3.1. Strain growth and DNA isolation

Strain LMTR 3 was grown in yeast extract mannitol medium [3] for 5 days at 30 °C. Total DNA was extracted from pelletized cells using the DNA Isolation Kit for Cells and Tissues (Roche) according to the manufacturer’s instructions. Integrity of the extracted DNA was evaluated by agarose gel electrophoresis while quantity and quality were determined using a Nanodrop 2000 spectrophotometer.

⁎ Corresponding author.

E-mail address: eormeno@lamolina.edu.pe (E. Ormeño-Orrillo).
3.2. Sequencing and assembly

Genome sequencing was performed using the Illumina MiSeq platform. 300 bp paired-end reads were generated from a library constructed using 500 bp DNA fragments with the Illumina TruSeq DNA nano kit following manufacturers’ instructions. Reads were quality-trimmed using the options SLIDINGWINDOW:4:15 and MINLEN:50 of Trimmomatic [4] and de novo assembled with SPAdes [5]. Genome completeness was evaluated with the program BUSCO [6].

3.3. Bioinformatics

The NCBI Prokaryotic Genome Annotation Pipeline (PGAP) [7] was used for gene prediction and automatic annotation. Additionally, the sequence was sent to the Rapid Annotations using Subsystems Technology server (RAST) [8] for complementary functional annotation. Average nucleotide identity (ANI) values were calculated with JSpecies [9] using MUMmer for alignment. The GGDC webservice [10] was used to obtain digital DNA-DNA hybridization (DDH) values. A neighbor-joining dendrogram depicting phylogenomic relationships was constructed using ANI values expressed as fractions with the PAST software [11]. All programs and webservers were used with default parameters.

4. Data description

4.1. Genome features

The final assembly had an average coverage of ~100× and was composed of 55 contigs ranging in size from 358 to 1,005,583 bp with a N50 size of 423,838 bp. Genome size was estimated at 7.83 Mbp and the average G+C content was 62.27%. Completeness was predicted at 100% by the program BUSCO [6]. NCBI’s PGAP predicted 7137 CDS, 48 tRNAs and 3 rRNAs genes. Functions could be assigned to 67% of the CDS. The distribution of CDS genes into RAST functional categories is shown in Fig. 1. The genome seemed to be devoid of extrachromosomal replicons due to the absence of genes coding for plasmid replication proteins.

4.2. Phylogenomics

LMTR 3 showed < 70% digital DDH and < 94% ANI values to all related type strains, indicating that it belongs to a not yet described
species \[9,10\]. \textit{B. retamae} Ro19\textsuperscript{T} was the most related strain to LMTR 3 with which it shares 55.2% DDH and 92.4% ANI. Phylogenomic relationships of strain LMTR 3 with related and reference strains are shown in Fig. 2.

4.3. Symbiosis island

Regions spanning 360 kb and 280 kb in two different contigs could be assigned as portions of a putative symbiosis island based on lower than genome average G + C content (59%) and on the presence of genes for nodulation and nitrogen fixation as well as for type III and IV secretion and hydrogen recycling which are typically found in rhizobial symbiotic compartments \[12\]. Genes or gene clusters encoding the proteins for Nod factor biosynthesis or transport included \textit{nodM}, \textit{nolL}, \textit{noel-noel-noeK} and \textit{nodD2-nodD1-nodABC-SUJ-nolo-nodZ-noe}. A complete set of genes for nitrogen fixation were found including those for nitrogenase structure (\textit{nifH, nifDK}) and maturation (\textit{nifZ}), FeMo-co biosynthesis (\textit{nifB, nifENX, nifQ, nifUS, nifV}), electron donation (\textit{fix-ABCX}), regulation (\textit{nifA}) and other functions (\textit{nifW, nifT}).

4.4. Other genes related to host colonization

Genes which may promote root colonization included those for biosynthesis of pili and adhesins, uptake systems and catabolic genes for carbohydrates and amino acids, biosynthesis of pili and adhesins, uptake systems and catabolic genes

4.4. Other genes related to host colonization

Genes which may promote root colonization included those for biosynthesis of pili and adhesins, uptake systems and catabolic genes for carbohydrates and amino acids, flagellar motility, and siderophore production and uptake. Furthermore, genes for synthesis of gibberellin and cytokinin phytohormones and for ethylene precursor cleavage could positively influence the interaction with its legume host.

Nucleotide sequence accession numbers

This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession MAXC00000000. The version described in this paper is version MAXC01000000.

Conflict of interest

The authors declare no conflicts of interest in this study.

Acknowledgements

This work was funded in whole or part by PNICP project 449-PNICP-BRI-2014. CC was funded by an “Isidro Parga Pondal” research fellowship from Xunta de Galicia (Spain). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

References

[1] E. Ormeño-Orrillo, P. Vinuesa, D. Zúñiga-Dávila, E. Martínez-Romero, Molecular diversity of native Bradyrhizobium isolated from Lima bean (\textit{Phaseolus lunatus L}) in Peru, Syst. Appl. Microbiol. 29 (2006) 253–262.
[2] E. Ormeño-Orrillo, D. Zúñiga-Dávila, E. Martínez-Romero, Biodiversity of nitrogen-fixing nodule bacteria associated with Lima bean (\textit{Phaseolus lunatus L}) in its domestication centers, in: A.S.F. Araújo, Á.C.A. Lopes, R.L.F. Gomes (Eds.), \textit{Phaseolus lunatus: Diversity, Growth and Production}, Nova Science Publishers, New York, 2015, pp. 91–102.
[3] J.M. Vincent, A Manual for the Practical Study of Root Nodule Bacteria, Blackwell Scientific Publications, Oxford, 1970.
[4] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics 30 (2014) 2114–2120.
[5] A. Bankevich, S. Nurk, D. Antipov, A.A. Gurevich, M. Dvorkin, A.S. Kulikov, V.M. Lesin, S.I. Nikolenko, S. Pham, A.D. Prijibelski, A.V. Pyshkin, A.V. Sirotkin, N. Vyahhi, G. Tesler, M.A. Alekseyev, P.A. Pevzner, SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing, J. Comput. Biol. 19 (2012) 455–477.
[6] F.A. Simão, R.M. Waterhouse, P. Ioannidis, E.V. Kriventseva, E.M. Zdobnov, BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs, Bioinformatics 31 (2015) 3219–3222.
[7] T. Tatusova, M. DiCuccio, A. Badretdin, V. Chetverin, E.P. Nawrocki, L. Laskovsk, A. Losmadze, K.D. Pruitt, M. Borodovsky, J. Ostell, NCBI prokaryotic genome annotation pipeline, Nuclieic Acids Res. 44 (2016) 6614–6624.
[8] R.K. Aziz, D. Bartels, A.A. Best, M. DeJongh, T. Ditz, R.A. Edwards, K. Edwards, C. Gerdes, E.M. Glass, M. Kubal, F. Meyer, G.J. Olsen, S. Oliner, A.L. Osterman, R.A. Overbeek, L.K. McNeil, D. Pauw, T. Papp, V. Prange, D.G. Pusch, C. Reich, R. Stevens, O. Vassieva, V. Vonstein, A. Wilke, O. Zharkikh. The RAST server: rapid annotations using subsystems technology, BMC Genomics 9 (2008) 75.
[9] M. Richter, R. Rossello-Mora, Shifting the genomic gold standard for the prokaryotic species definition, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 19126–19131.
[10] J. Meier-Kolthoff, A. Auch, H.-P. Kleik, M. Goker, Genome sequence-based species delimitation with confidence intervals and improved distance functions, BMC Bioinf. 14 (2013) 60.
[11] O. Hammer, D.A.T. Harper, P.D. Ryan, PAST: paleontological statistics software package for education and data analysis, Palaeontol. Electron. 4 (2001) 1–9.
[12] A. Siqueira, E. Ormeño-Orrillo, R. Souza, E. Rodrigues, L.G. Almeida, F. Barcellos, J.M. Vincent, A Manual for the Practical Study of Root Nodule Bacteria, Blackwell Scientific Publications, Oxford, 1970.
[13] O. Hammer, D.A.T. Harper, P.D. Ryan, PAST: paleontological statistics software package for education and data analysis, Palaeontol. Electron. 4 (2001) 1–9.
[14] A. Siqueira, E. Ormeño-Orrillo, R. Souza, E. Rodrigues, L.G. Almeida, F. Barcellos, J.M. Vincent, A Manual for the Practical Study of Root Nodule Bacteria, Blackwell Scientific Publications, Oxford, 1970.
[15] O. Hammer, D.A.T. Harper, P.D. Ryan, PAST: paleontological statistics software package for education and data analysis, Palaeontol. Electron. 4 (2001) 1–9.