



# Complete Genome Sequences of Eight *Rhizobium* Symbionts Associated with Common Bean (*Phaseolus vulgaris*)

Rosa Isela Santamaría, Patricia Bustos, Olga María Pérez-Carrascal,\* Fabiola Miranda-Sánchez,\* Pablo Vinuesa, Irma Martínez-Flores, Soledad Juárez, Luis Lozano, Esperanza Martínez-Romero, Miguel Ángel Cevallos, David Romero, Guillermo Dávila,\* Ernesto Ormeño-Orrillo,\* Víctor González

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

**ABSTRACT** We present here the high-quality complete genome sequences of eight strains of *Rhizobium*-nodulating *Phaseolus vulgaris*. Comparative analyses showed that some of them belonged to different genomic and evolutionary lineages with common symbiotic properties. Two novel symbiotic plasmids (pSyms) with *P. vulgaris* specificity are reported here.

There is ample diversity of *Rhizobium* species able to form nodules and fix nitrogen in symbiosis with common bean (1). Although they are commonly referred to as *Rhizobium etli* and *Rhizobium phaseoli*, they may correspond to distinct genomic lineages (2–4). To identify them, we obtained their complete genome sequences. *Rhizobium* isolates were from common bean nodules from different places in Mexico and elsewhere.

Genome sequences were generated with genomic DNA purified with the QIAamp DNA minikit (Qiagen). Illumina MiSeq 2 × 250 and 454 (MOgene LC, St. Louis, MO, USA) technologies were used for DNA sequencing of libraries with inserts of about 200 bp, 2 kb, and 3 kb in length. Paired-end reads were assembled using Newbler 2.5.3 (Roche), Velvet 1.1.06 (5), SSPACE-Basic 2.0 (6), and Consed version 23 (7), with deep coverage between 81 and 529×. Assemblies were guided partially with reference to the complete genomes of *Rhizobium etli* CFN42 and *Rhizobium phaseoli* CIAT652 and with NUCMER alignments (8). Additional long PacBio sequence reads were generated to verify the assembly of the Bra5 and IE4771 genomes. Open reading frames (ORFs) were predicted with Glimmer 3 (9); ORF frame corrections and manual annotations were done in Artemis 12.0 (10) and compared with GenBank (11), InterPro (12), and ISfinder (<http://www-is.biotoul.fr>) databases.

Table 1 summarizes the genomic features and the GenBank accession numbers of each genome. The lengths of the genomes range from 6.4 to 7.3 Mb, with an overall G+C content of 61%. There are between 3 and 5 plasmids per strain, with sizes ranging from 166 to 1,144 kb. The symbiotic plasmids, or pSyms, are 393 to 530 kb in length, with 59% average G+C content.

Average nucleotide identity by MUMmer (ANIm) and a phylogenetic tree, inferred from concatenated nucleotide regions shared among these genomes and others reported in GenBank (13), distinguished Bra5 as belonging to *R. phaseoli* and NCX12 as *R. etli*. Strains Kim5, IE4803, and IE4771 had an ANIm of higher than 95% among them but less than 95% with *R. etli* and *R. phaseoli*. CIAT894, TAL182, and NXC14 were unrelated by ANIm and phylogenetically divergent from other *Rhizobium* species. Therefore, they might represent other genomic lineages and species within the *Rhizobium* genus. The pSyms of IE4803, Bra5, Kim5, TAL182, and CIAT894 are quite similar (ANIm, about 97 to 98%), while strains NXC12 and NXC14 have common pSyms (98%

Received 19 May 2017 Accepted 25 May 2017 Published 27 July 2017

**Citation** Santamaría RI, Bustos P, Pérez-Carrascal OM, Miranda-Sánchez F, Vinuesa P, Martínez-Flores I, Juárez S, Lozano L, Martínez-Romero E, Cevallos MÁ, Romero D, Dávila G, Ormeño-Orrillo E, González V. 2017. Complete genome sequences of eight *Rhizobium* symbionts associated with common bean (*Phaseolus vulgaris*). Genome Announc 5: e00645-17. <https://doi.org/10.1128/genomeA.00645-17>.

**Copyright** © 2017 Santamaría et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Víctor González, [vgonzal@ccg.unam.mx](mailto:vgonzal@ccg.unam.mx).

\* Present address: Olga María Pérez Carrascal, Département de Sciences Biologiques, Université de Montréal, Montréal, Québec, Canada; Fabiola Miranda-Sánchez, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA; Guillermo Dávila, Laboratorio Internacional de Investigación sobre el Genoma Humano, Universidad Nacional Autónoma de México, Juriquilla, Querétaro, Mexico; Ernesto Ormeño-Orrillo, Laboratorio de Ecología Microbiana y Biotecnología, Universidad Nacional Agraria La Molina, Lima, Perú.

**TABLE 1** General features of *Rhizobium* complete genomes

Species	Strain	Origin	Genome length (bp)	No. of chromosomes, no. of plasmids	G+C content (%)	GenBank accession no.	Reference
<i>R. etli</i>	NXC12	Huautla, Mexico	6,756,853	1, 5	61	CP020906 to CP020911	14
<i>R. phaseoli</i>	Bra5	Brazil	6,665,454	1, 4	61	CP020896 to CP020900	1
<i>Rhizobium</i> spp.	IE4803	Puebla, Mexico	6,997,434	1, 4	61	CP007641 to CP007645	15
	Kim5	Idaho, USA	6,817,255	1, 4	61	CP021124 to CP021128	16
	CIAT894	Colombia	6,657,947	1, 5	61	CP020947 to CP020952	1
	TAL182	Hawaii, USA	6,402,377	1, 5	61	CP021024 to CP021029	1
	IE4771	Puebla, Mexico	7,057,405	1, 5	61	CP006986 to CP006991	15
	NXC14	Huautla, Mexico	6,686,318	1, 3	61	CP021030 to CP021033	14

ANI<sub>m</sub>) but are distinct from the IE4771 pSym that is unique in this collection. Extensive genomic comparisons within the *Rhizobium* genus point out the ecological convergence toward nodulation of a common host due to the presence of similar and divergent pS<sub>ym</sub>s.

**Accession number(s).** Table 1 shows the GenBank accession numbers of the eight strains.

### ACKNOWLEDGMENTS

We appreciate the technical assistance of José Espiritu and Víctor del Moral.

This work was supported by grants from PAPIIT-UNAM (IN 208414-3) and CONACyT (CB131499) to V.G.

### REFERENCES

- Piñero D, Martínez E, Selander RK. 1988. Genetic diversity and relationships among isolates of *Rhizobium leguminosarum* biovar phaseoli. *Appl Environ Microbiol* 54:2825–2832.
- Acosta JL, Eguiarte LE, Santamaría RI, Bustos P, Vinuesa P, Martínez-Romero E, Dávila G, González V. 2011. Genomic lineages of *Rhizobium etli* revealed by the extent of nucleotide polymorphisms and low recombination. *BMC Evol Biol* 11:305. <https://doi.org/10.1186/1471-2148-11-305>.
- López-Guerrero MG, Ormeño-Orrillo E, Velázquez E, Rogel MA, Acosta JL, González V, Martínez J, Martínez-Romero E. 2012. *Rhizobium etli* taxonomy revised with novel genomic data and analyses. *Syst Appl Microbiol* 35:353–358. <https://doi.org/10.1016/j.syapm.2012.06.009>.
- Ribeiro RA, Ormeño-Orrillo E, Dall'Agnol RF, Graham PH, Martínez-Romero E, Hungria M. 2013. Novel *Rhizobium* lineages isolated from root nodules of the common bean (*Phaseolus vulgaris* L.) in Andean and Mesoamerican areas. *Res Microbiol* 164:740–748. <https://doi.org/10.1016/j.resmic.2013.05.002>.
- Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <https://doi.org/10.1101/gr.074492.107>.
- Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27:578–579. <https://doi.org/10.1093/bioinformatics/btq683>.
- Gordon D, Abajian C, Green P. 1998. Consed: a graphical tool for sequence finishing. *Genome Res* 8:195–202. <https://doi.org/10.1101/gr.8.3.195>.
- Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. *Genome Biol* 5:R12. <https://doi.org/10.1186/gb-2004-5-2-r12>.
- Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with Glimmer. *Nucleic Acids Res* 27:4636–4641. <https://doi.org/10.1093/nar/27.23.4636>.
- Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. 2000. Artemis: sequence visualization and annotation. *Bioinformatics* 16:944–945. <https://doi.org/10.1093/bioinformatics/16.10.944>.
- Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2009. GenBank. *Nucleic Acids Res* 37:D26–D31.
- Apweiler R, Attwood TK, Bairoch A, Bateman A, Birney E, Biswas M, Bucher P, Cerutti L, Corpet F, Croning MD, Durbin R, Falquet L, Fleischmann W, Gouzy J, Hermjakob H, Hulo N, Jonassen I, Kahn D, Kanapin A, Karavidopoulou Y, Lopez R, Marx B, Mulder NJ, Oinn TM, Pagni M, Servant F, Sigrist CJ, Zdobnov EM. 2001. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Res* 29:37–40. <https://doi.org/10.1093/nar/29.1.37>.
- Pérez Carrascal OM, VanInsberghe D, Juárez S, Polz MF, Vinuesa P, González V. 2016. Population genomics of the symbiotic plasmids of sympatric nitrogen-fixing *Rhizobium* species associated with *Phaseolus vulgaris*. *Environ Microbiol* 18:2660–2676. <https://doi.org/10.1111/1462-2920.13415>.
- Miranda-Sánchez F, Rivera J, Vinuesa P. 2016. Diversity patterns of *Rhizobiaceae* communities inhabiting soils, root surfaces and nodules reveal a strong selection of rhizobial partners by legumes. *Environ Microbiol* 18:2375–2391. <https://doi.org/10.1111/1462-2920.13061>.
- Silva C, Vinuesa P, Eguiarte LE, Martínez-Romero E, Souza V. 2003. *Rhizobium etli* and *Rhizobium gallicum* nodulate common bean (*Phaseolus vulgaris*) in a traditionally managed milpa plot in Mexico: population genetics and biogeographic implications. *Appl Environ Microbiol* 69:884–893. <https://doi.org/10.1128/AEM.69.2.884-893.2003>.
- Josephson KL, Pepper IL. 1984. Competitiveness and effectiveness of strains of *Rhizobium phaseoli* isolated from the Sonoran Desert. *Soil Biol Biochem* 16:651–655. [https://doi.org/10.1016/0038-0717\(84\)90086-5](https://doi.org/10.1016/0038-0717(84)90086-5).