



Universidad Nacional Autónoma de México

Facultad de Ciencias

TESIS

Predicción de islas genómicas, análisis filogenéticos y genómicos de una cepa de *Bradyrhizobium* aislada de los nódulos de *Lysiloma* sp. en Cuernavaca, Morelos

QUE PARA OBTENER EL TÍTULO DE

B I Ó L O G A

PRESENTA

Karen Lizbeth Claro Mendoza

Tutora de tesis

Dra. María Esperanza Martínez Romero



Ciudad Universitaria, Cd. Mx., 2022



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

*A mi madre y padre, quienes con sus cuidados y apoyo han sembrado
mis logros académicos presentes y futuros.*

*A mis hermanos, Lucy y Miguel, quienes me apoyan y son mis
compañeros de vida.*

*A Adán, quien con su bella presencia en mi vida, me ha motivado a
construir la mejor versión de mí.*

*A mis amigos de la universidad, Caro, Erandhi, Leonardo y Nadia;
a quienes admiro y cuya amistad es invaluable para mí.*

AGRADECIMIENTOS ACADÉMICOS

Agradezco a la Dra. María Esperanza Martínez Romero, mi tutora, por su valiosa asesoría para la elaboración de esta tesis y por brindarme la oportunidad de trabajar en su grupo de investigación.

Agradezco a la M. en C. Diana Hernández Oaxaca por sus valiosos comentarios y enseñanzas que me fueron de gran ayuda para la realización de este trabajo.

Agradezco al Taller de Ciencia para Jóvenes, 2015 (proyecto PAPIME-PE100915) que se llevó a cabo en el Centro de Geociencias, Campus Juriquilla-UNAM a cargo del Dr. Juan Martín Gómez González; por impulsarme de manera decisiva en el camino de la ciencia.

ÍNDICE

RESUMEN	4
1. INTRODUCCIÓN	4
2. OBJETIVOS.....	6
2.1 Objetivo general.....	6
2.2 Objetivos particulares	7
3. MATERIALES Y MÉTODOS.....	7
3.1 Aislamiento de bacterias de los nódulos, condiciones de crecimiento y hábitat	7
3.2 Extracción de ADN, PCR, patrones ERIC y secuenciación.....	8
3.3 Ensayos de nodulación en <i>Phaseolus vulgaris</i> , <i>Leucaena leucocephala</i> y <i>Lysiloma</i> sp.....	9
3.4 Ensamblaje y anotación del genoma	10
3.5 Cálculo de la identidad de nucleótidos promedio (ANI)	10
3.6 Análisis filogenéticos.....	11
3.7 Predicción de las islas genómicas.....	12
3.7.1 Identificación de sistemas conjugativos que podrían movilizar a las islas simbióticas.....	12
4. RESULTADOS.....	13
4.1 Aislamiento de cepas bacterianas de los nódulos de <i>Lysiloma</i> sp. y huellas ERIC-PCR ...	13
4.2 Secuenciación del genoma de <i>Bradyrhizobium</i> sp. CCGB01	17
4.3 Análisis filogenéticos de <i>Bradyrhizobium</i> sp. CCGB01 y cálculo de la identidad de nucleótidos promedio (ANI)	18
4.4 Ensayos de nodulación	26
4.5 Predicción de islas genómicas identificación de genes simbióticos de <i>Bradyrhizobium</i> sp. CCGB01	28
4.5.1 Identificación de sistemas conjugativos que podrían movilizar a las islas simbióticas.....	33
5. DISCUSIÓN.....	35
5. 1 Clasificación taxonómica de <i>Bradyrhizobium</i> sp. CCGB01.....	35
5.2 Incongruencia entre las filogenias de los genes <i>glnII</i> y <i>recA</i> , las filogenias de <i>nodC</i> y <i>nifH</i> y el análisis filogenómico	35
5.3 Evaluación de la capacidad de nodulación de <i>Bradyrhizobium</i> sp. CCGB01	36
5.4 La mayoría de los genes simbióticos de <i>Bradyrhizobium</i> sp. CCGB01 se encuentran en una isla genómica	37
5.4.1 Identificación de sistemas conjugativos que podrían movilizar a las islas simbióticas.....	41
6. CONCLUSIONES	43
7. PERSPECTIVAS	44
8. REFERENCIAS BIBLIOGRÁFICAS	45
9. MATERIAL SUPLEMENTARIO	53

RESUMEN

1. INTRODUCCIÓN

Los inoculantes microbianos son un método alternativo para incrementar la productividad de los cultivos y reducir el uso de pesticidas y fertilizantes químicos (Wang, et al., 2018). Entre los organismos que componen los inoculantes se pueden encontrar los rizobios. El término *rizobio* se refiere a las bacterias de la clase Proteobacteria que fijan nitrógeno atmosférico y que pueden establecer relaciones simbióticas con plantas leguminosas (Ormeño-Orrillo, 2020); entre los rizobios se encuentran las bacterias del género *Bradyrhizobium*.

Para el género *Bradyrhizobium* se han reportado 74 especies de las cuales 63 poseen nomenclaturas y publicaciones avaladas hasta la fecha (LPSN, Junio, 2021). Para designar las especies no se utiliza únicamente las secuencias 16S rARN debido a que estas secuencias se encuentran muy conservadas en *Bradyrhizobium*, sino que comúnmente se designan con base en análisis de secuencias multilocus (MLSA) (porcentaje de identidad mayor a 97%) de codificadores de proteínas *housekeeping* de copia única y en el método de identidad de nucleótidos promedio (ANI) de todo el genoma (ANI mayor a 95-96%) (Martens et al., 2018; Varghese et al., 2015; Durán et al., 2014; Richter y Rosselló-Móra, 2009). Las bacterias de este género nodulan el mayor rango de leguminosas (Parker, 2015); forman nódulos con plantas de la subfamilia Caesalpinioideae como *Dimorphandra wilsonii* Rizz, *Erythrophleum fordii* y *Leucaena leucocephala* (Fonseca et al., 2012; Yao et al., 2015; Wang et al., 1999), y también son simbioses de leguminosas de la subfamilia Papalionoideae como *Phaseolus lunatus* y *Erytrina brucei* (Ormeño-Orrillo et al., 2015; López-López et al., 2013; Aserse et al., 2017). Se ha sugerido que en *Bradyrhizobium* se encuentran los ancestros de los simbioses de las leguminosas (Martinez-Romero, 2009; Parker, 2015).

Los genes simbióticos que permiten fijar nitrógeno atmosférico y formar nódulos con plantas leguminosas de los géneros *Bradyrhizobium* y *Mesorhizobium* se encuentran ocasionalmente en plásmidos y con mayor frecuencia en islas genómicas en el cromosoma (Zahran, 2017). Generalmente se denominan islas

genómicas (GIs) a los grupos de genes consecutivos que se adquirieron por transferencia genética horizontal (HT) (Bertelli et al., 2019). Las GIs se pueden clasificar de acuerdo a los tipos de genes que contienen: simbióticos, patogénicos, metabólicos y de resistencia (Juhás et al., 2009).

Las GIs presentan diversas características que permiten su predicción. El tamaño de las GIs es de 10 a 200 kb aproximadamente y generalmente contienen elementos genéticos móviles como integrones, elementos de inserción de secuencias o transposones (Juhás et al., 2009; Dobrindt et al., 2004; Boyd et al., 2009); elementos integrativos, elementos conjugativos y profagos (Langille et al., 2010). Las GIs suelen estar flanqueadas por genes tARN y repeticiones directas perfectas de 16 a 20 pb; presentan diferentes contenidos de GC y frecuencias de este dinucleótido, uso de codones o firma de k-meros respecto al cromosoma bacteriano, y en ellas prevalecen proteínas hipotéticas (Dobrindt et al., 2004). También, las GIs se pueden identificar cuando la secuencia de genes que flanquean a las GIs no se encuentra interrumpida en las especies, aislados o cepas cercanas al organismo que posee la GI (Langelli et al., 2010).

Algunas GIs pueden ser movilizables y clasificarse como elementos genéticos móviles (MGEs). Evidencia experimental y análisis filogenéticos han permitido observar que diversos genes simbióticos en los rizobios son transmitidos horizontalmente mediante MGEs (Wardell et al., 2022). Por otro lado, algunas GIs pierden su movilidad después de ser adquiridas; de ésta forma, dan lugar a las denominadas GIs degeneradas o a otros tipos de MGEs (Bellanger et al., 2014)

Los MGEs se clasifican en elementos conjugativos e integrativos (ICEs), elementos integrativos y movilizables (IMEs), y los elementos movilizables en *cis* (CIMEs) (Bellanger et al., 2014). Los elementos conjugativos e integrativos (ICEs) son elementos móviles que se integran en los cromosomas de las células hospedadoras capaces de escindirse a sí mismos y transmitirse horizontalmente a otras células por conjugación (Bellanger et al., 2014; Bothelo y Schulenburg, 2021). Para su transmisión, los ICEs requieren genes que codifican para proteínas del sistema de secreción tipo 4 (T4SS), una proteína acopladora (T4CP), una endonucleasa llamada relaxasa y factores accesorios del complejo relaxosoma

(RAFs) (Ramsay y Firth, 2017). Una integrasa lleva a cabo la recombinación de repeticiones directas del sitio de anclaje *attB* del cromosoma del hospedero y los sitios *attI* o *attP* del ICE. Para su transferencia, el ICE es escindido en una sola hebra por integrasas específicas y proteínas accesorias, en el sitio de origen de transferencia (*oriT*) y la relaxasa se encarga de replicar y circular la hebra. Posteriormente, la copia es cubierta por proteínas de unión a ADN de un sola hebra (SSb) formando el complejo proteico denominado relaxosoma, que es guiado por una proteína acopladora a otro complejo proteico de transferencia conjugativa tipo IV para ser depositada en la célula receptora (Delavat et al, 2017; McLeod et al., 2006; Llosa et al. 2002; Lee et al., 2010; Ramsay y Firth, 2017). Sin embargo, existen otros elementos que no codifican toda esta maquinaria de conjugación pero que son capaces de transmitirse, tal es el caso de los elementos integrativos y movilizables (IMEs) y los elementos movilizados en *cis* (CIMES).

Los IMEs se caracterizan, al igual que los plásmidos movilizables, por ser elementos capaces de aprovechar la maquinaria de ICEs, transposones conjugativos o plásmidos conjugativos (Delavat et al, 2017; Ramsay y Firth, 2017). Por otro lado, las GIs movilizables pueden mediar la transferencia de ADN cromosomal adyacente en *cis*; los elementos movilizados de esta forma se denominan CIMES y se caracterizan por no presentar proteínas de conjugación ni recombinación, pero sí por estar flanqueados por sitios *att* (Bellanger et al., 2014; Pavlovic et al., 2004).

En este trabajo, el objetivo fue evaluar la capacidad de formar nódulos de bacterias del género *Bradyrhizobium* aisladas de nódulos de *Lysiloma* sp. en Cuernavaca, Morelos, determinar su identidad taxonómica, realizar la predicción y análisis de sus islas simbióticas e identificar si éstas podrían ser movilizables.

2. OBJETIVOS

2.1 Objetivo general

Evaluar la capacidad de nodulación de bacterias del género *Bradyrhizobium* aisladas de nódulos de *Lysiloma* sp. en Cuernavaca, Morelos, determinar su identidad taxonómica, realizar la predicción y análisis de sus islas simbióticas e

identificar si éstas podrían ser movilizables.

2.2 Objetivos particulares

- Aislar bacterias del género *Bradyrhizobium* y agrupar a las diferentes cepas mediante las secuencias consenso intergénicas repetitivas de enterobacterias (ERIC).
- Realizar análisis filogenéticos con el gen 16S rARN de las cepas con las diferentes huellas ERIC-PCR.
- Realizar ensayos de nodulación en *Leucaena leucocephala*, *Phaseolus vulgaris* y *Lysiloma* sp. con los diferentes genotipos.
- Obtener, ensamblar y anotar el genoma de la cepa con el patrón ERIC-PCR más representativo
- Realizar análisis filogenéticos para determinar la taxonomía de la cepa con el patrón ERIC-PCR más representativo
- Predecir las islas genómicas del genoma de la cepa con el patrón ERIC-PCR más frecuente.
- Identificar si las islas simbióticas podrían ser movilizables con base en la predicción bioinformática de los sistemas conjugativos de secreción tipo IV (T4SS) del genoma y con base en los sistemas de conjugación descritos en la literatura.

3. MATERIALES Y MÉTODOS

3.1 Aislamiento de bacterias de los nódulos, condiciones de crecimiento y hábitat

En el laboratorio se germinaron plántulas de *Lysiloma* sp. a 30°C como se describe previamente (Martínez-Romero y Rosenblueth, 1990) y se inocularon con una suspensión de suelo de la barranca del Tecolote, Cuernavaca, Morelos, México (18° 95' 50" N; 99° 27' 66" O, 1852 m.s.n.m) en donde se colectaron las semillas de *Lysiloma* sp.

Los nódulos se desinfectaron con alcohol al 70% por 1 minuto, después se

les agregó una solución de cloralex al 25% y se agitó por 3 minutos y finalmente se realizaron lavados con agua tridestilada (Rogel et al., 2001). El contenido de los nódulos y las colonias purificadas se cultivaron a 30 °C en medio YM (Vincent, 1970) sólido con cicloheximida 10 µg/mL para inhibir el crecimiento de hongos. Para la extracción de ADN, las bacterias se cultivaron en medio YM líquido y se incubaron a 30 °C. En el caso del ADN para la cepa de la que se secuenció el genoma, las bacterias también se cultivaron en medio líquido PY (Toledo et al., 2003). Los aislados fueron almacenados en glicerol a - 70 °C en el Centro de Ciencias Genómicas, Cuernavaca, Morelos, México.

3.2 Extracción de ADN, PCR, patrones ERIC y secuenciación

El ADN se extrajo con el kit High Pure FFPE DNA Isolation (Roche) siguiendo las instrucciones del fabricante. El ADN se visualizó en geles de agarosa mediante electroforesis y se cuantificó con el espectrómetro de nucleótidos NanoDrop ND1000 (Thermo FisherScientific, USA).

Para agrupar a las diferentes cepas del género *Bradyrhizobium*, se obtuvieron los patrones ERIC-PCR según el protocolo de Versalovic et al. (1991); este método fue validado por de Brujin (1992) para diversos géneros de la familia *Rhizobiaceae*, entre ellos *Bradyrhizobium*.

Los genes 16S rARN se amplificaron usando los oligonucleótidos fD1 y rD1 (Weisburg et al., 1991). Los productos de PCR del gen 16S rRNA se purificaron con el kit Thermo Scientific GeneJET PCR Purification siguiendo las instrucciones del fabricante sin usar el buffer de elución; en su lugar se utilizó agua purificada con el sistema Mili-Q®. Los perfiles ERIC fueron visualizados mediante electroforesis en geles de agarosa así como los productos de amplificación del gen 16S rARN.

La secuenciación de los genes 16S rARN se realizó mediante el método Sanger con los oligonucleótidos 518F (5'-CCAGCAGCCGCGGTAATACG-3') y 800R (5'-TACCAGGGTATCTAATCC-3') (Ghyselinck et al., 2013). La secuenciación del genoma de *Bradyrhizobium* sp. *CCGB01* se llevó a cabo a través de la plataforma Nova Seq6000 de Illumina con lecturas pareadas de 150 pb. Tanto la secuenciación

de los genes 16S rARN como la secuenciación del genoma fueron realizados por Macrogen Inc. (Seúl, Corea del Sur). También, se llevaron a cabo secuenciaciones de lecturas largas del genoma del mismo aislado, mediante la plataforma ONT MinION en la Unidad Universitaria de Secuenciación Masiva y Bioinformática del Instituto de Biotecnología de la UNAM.

3.3 Ensayos de nodulación en *Phaseolus vulgaris*, *Leucaena leucocephala* y *Lysiloma* sp.

Se realizaron ensayos de nodulación y se midió la reducción de acetileno de las cepas seleccionadas en *Lysiloma* sp., *Phaseolus vulgaris* y *Leucaena leucocephala* como describe previamente Rogel et al. (2001). Las semillas de *P. vulgaris* se desinfectaron con etanol al 70% por 5 minutos, con una solución de cloralex al 25% por 15 minutos y posteriormente se realizaron 10 enjuagues con agua tridestilada. Las semillas de *Leucaena leucocephala* se trataron con ácido sulfúrico y posteriormente repitieron los mismos pasos de desinfección usados para *P. vulgaris*. Se germinaron semillas de *L. leucocephala* y *P. vulgaris* en agar H₂O a 30 °C en condiciones de oscuridad. Después, las semillas germinadas se transfirieron a matraces de 250 mL con medio sólido Fahraeus (0.7% de agar) (Fahraeus, 1957). Las cepas representativas de cada perfil ERIC se suspendieron en agua estéril y se ajustaron a una densidad óptica (OD₆₀₀) de 0.3. Posteriormente, se inocularon 300 µL a las plántulas por triplicado, las cuales crecieron en condiciones de invernadero a temperatura constante de 30° C con un ciclo 12 horas luz, 12 de oscuridad por 24 días. Se usó como control positivo la cepa *Rhizobium leucaenae* CFN299 y plantas inoculadas con agua tridestilada como control negativo. Se llevó a cabo un conteo de los nódulos en *P. vulgaris* cada tercer día por 19 días y en *L. leucocephala* por 38 días. No se logró llevar a cabo el conteo de nódulos en *Lysiloma* sp. debido a las condiciones imperantes derivadas de la emergencia sanitaria del COVID-19. Para *P. vulgaris* y *Lysiloma* sp. se midió la reducción de acetileno en un cromatógrafo de gases (Varian 3300 Gas Chromatograph) con tres plántulas por cada tratamiento y por grupo control. Los datos se analizaron con una ANOVA de una vía seguida de

una prueba de TukeyHSD para comparaciones múltiples con el programa RStudio v.2022.02.3.

3.4 Ensamblaje y anotación del genoma

La calidad de las lecturas de Illumina se revisó con el programa FastQC v. 0.11.8 y la limpieza con Trim Galore v. 0.6.6. El genoma se ensambló *de novo* con los programas Velvet v. 1.2.10 y SPAdes v. 3.12.0. La mejor longitud de los k-meros se calculó mediante el programa VelvetOptimiser v. 2.2.5. Los ensamblajes de Velvet y SPAdes se compararon con el programa QUAST v.4.1. para el ensamblaje se formaron andamiajes con el programa SSPace v.3.0.

La predicción génica y su anotación funcional se llevaron a cabo con los programas RASTtk v. 1.3.0. (Brettin et al., 2015) y PROKKA (v.1.12). Estas anotaciones fueron evaluadas mediante la anotación de los genes simbióticos con BLASTX (Altschul et al., 1980) con la base de datos del *National Center for Biotechnology Information* (NCBI) que incluye a todas las secuencias de proteínas no redundantes de todos los organismos. Para lograr ensamblar completamente el genoma se decidió llevar a cabo una segunda secuenciación de lecturas largas mediante la plataforma ONT MinION. El ensamblaje se realizó con el programa Unicycler v. 0.4.8 (Wick et al., 2017) que primero ensambla las lecturas cortas con el programa SPAdes en un gráfico de ensamblaje preciso y conectado para después usar las lecturas largas para encontrar las mejores rutas a través del gráfico y así lograr cerrar el genoma. La predicción génica y anotación del genoma se realizó con RASTtk. La secuencia en crudo se encuentra disponible bajo el acceso BioProject PRJNA842432.

3.5 Cálculo de la identidad de nucleótidos promedio (ANI)

Para el cálculo de la identidad de nucleótidos promedio (ANI) se descargaron las últimas versiones de los genomas de las cepas tipo o de las cepas representativas de cada especie del género *Bradyrhizobium* así como genomas del clado al que pertenece *B. japonicum* USDA 6^T (superclado *japonicum*) reportados hasta el mes de junio de 2021. Los genomas se obtuvieron de la base de datos del sitio *National*

Center for Biotechnology Information (NCBI). Se utilizó el método de ANIm con el programa PYANI v.0.2.7 (Pritchard et al., 2016) para calcular el ANI entre los genomas y los porcentajes de identidad se graficaron en un mapa de calor.

3.6 Análisis filogenéticos

En el caso de los análisis filogenéticos, se utilizaron las secuencias de los genes 16S rARN, *glnII*, *recA*, *nodC* y *nifH* y también se obtuvo un árbol filogenómico. Las secuencias parciales 16S rARN se editaron con el programa Bioedit v.7.0.5 (Hall, 1999) y su alineamiento se realizó tomando en cuenta la estructura secundaria con el programa Infernal v.1.1 (Nawrocki y Eddy, 2013) mediante la plataforma en línea *Ribosomal Database Project* v. 11.5 (Cole et al., 2014). Como secuencia externa se utilizó la secuencia parcial 16S rARN de *Bosea thiooxidans* DSM9653^T.

Los alineamientos de las secuencias de los genes *recA*, *glnII*, *nodC* y *nifH* se realizaron con base en los alineamientos de las proteínas mediante el servidor en línea RevTrans v.2.0 (Wernersson y Pedersen, 2003); los alineamientos de las proteínas con el programa MUSCLE v.3.8.31 (Edgar, 2004) y los alineamientos de nucleótidos con el programa MAFFT v.7.309 (Hofacker et al, 2002). Para que las secuencias fueran de la misma longitud, se cortaron los alineamientos poco informativos manualmente con el programa MEGA X (Kumar et al., 2018). Las secuencias de los genes *recA* y *glnII* se concatenaron con el mismo programa. En todos los casos anteriores, los modelos de sustitución de nucleótidos se calcularon con el programa ModelTest-NG v.0.1.1 (Darriba et al., 2017) y los árboles se infirieron con IQ-Tree v. 1.6.11 (Nguyen et al., 2015) configurando el programa al mejor modelo de acuerdo con el criterio de información Akaike (AIC) y con mil pseudo réplicas de ultrafast bootstrap (UFBoot). El árbol de genes concatenados también se infirió con el programa MrBayes v.3.2.7a; las probabilidades posteriores bayesianas se calcularon a partir de dos simulaciones del método Metropolis-Hastings de Monte Carlo en cadena de Markov (MCMC) independientes y convergentes ejecutadas durante 3×10^6 generaciones. En todos los análisis filogenéticos anteriores se calcularon los porcentajes de identidad entre las secuencias a partir de los alineamientos con el programa en línea *Sequence Identity*

And Similarity (SIAS) de la Universidad Complutense de Madrid. Los árboles se visualizaron y editaron con el programa en línea iTOL v.6.5.6 (Letunic y Bork, 2021).

Para obtener el árbol filogenómico, se utilizó el programa OrthoFinder v.2.2.6 (Emms y Kelly, 2019), el cual se configuró para máxima verosimilitud con FastTree v.1.10 (Price et al., 2010). Se utilizaron genomas de cepas de bacterias del género *Nitrobacter* como grupo externo.

3.7 Predicción de las islas genómicas

Para la predicción de islas genómicas (GIs) con el genoma completo de *B. sp. CCGB01* se utilizó el programa en línea Island Viewer 4 (Bertelli et al., 2017). Este programa predice GIs con base en la composición de la secuencia y en comparaciones genómicas (Bertelli et al., 2017) y emplea cuatro programas: IslandPath-DIMOB, SIGI-HMM, Islander e IslandPick. Para la selección de islas en este trabajo, se tomó en cuenta la integración de los resultados de los programas IslandPath-DIMOB, SIGI-HMM e IslandPick, ya que los resultados del programa Islander no se encontraban disponibles.

Los MGEs suelen contener genes con funciones desconocidas, por lo que requieren una anotación más exhaustiva (Cury et al, 2020). Por esto último, se llevó a cabo una re-anotación de las islas simbióticas con el programa eggNOG-mapper v. v.2.1.6 (Cantalapiedra et al., 2021), el cual infiere las funciones de los genes con base en los grupos de genes ortólogos (OGs) de la bases de datos de eggNOG que cuentan con filogenias previamente computarizadas. Las anotaciones de RASTtk y eggNOG-mapper se integraron usando un script de bash propio ([Repositorio que contiene el script](#)).

3.7.1 Identificación de sistemas conjugativos que podrían movilizar a las islas simbióticas

Para identificar los sistemas conjugativos de tipo 4 (T4SS), se utilizaron tres programas: el módulo CONJscan del programa MacSyFinder, ICEfinder v.1.0 y oriTfinder v.1.1 (Abby et al, 2016; Abby et al, 2014; Liu et al., 2019; Li et al., 2018).

Los tres programas comparan perfiles de proteínas de modelos ocultos de Markov. CONJscan usa información sobre la organización y presencia de los componentes de los sistemas conjugativos para identificarlos (Cury et al., 2020) y el programa ICEfinder identifica elementos conjugativos e integrativos (ICEs) y elementos integrativos y movilizables (IMEs). Por otro lado, el servidor en línea oriTfinder identifica el sitio de origen de transferencia *oriT* para elementos genéticos móviles (MGEs); adicionalmente también identifica proteínas involucradas en el T4SS.

Las predicciones de las islas genómicas y de los sistemas conjugativos se visualizaron en figuras elaboradas con los programas DNA plotter y Dna Features v.3.1.0 (Carver et al., 2009; Zulkower y Rosser, 2020).

4. RESULTADOS

4.1 Aislamiento de cepas bacterianas de los nódulos de *Lysiloma* sp. y huellas ERIC-PCR

En total se obtuvieron 69 aislados bacterianos de 6 plantas de *Lysiloma* sp. De estos aislados, se analizaron solo 25. Para diferenciar entre las diferentes cepas, se utilizó el método ERIC-PCR. Se observaron 10 patrones ERIC diferentes, el patrón CCGB01 fue el que se presentó con mayor frecuencia (8/25, Tabla 1, Figura 1).

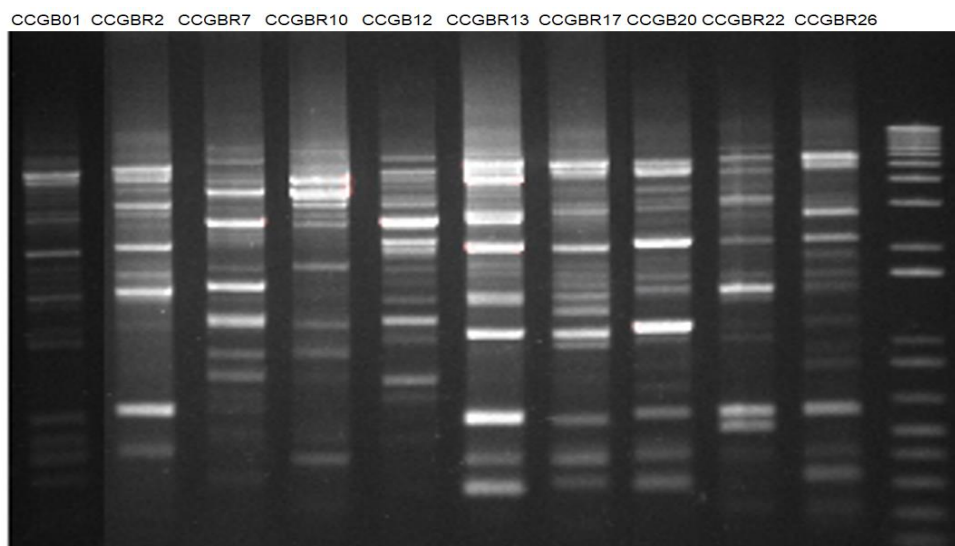


Figura 1. Huellas ERIC-PCR de los nódulos de *Lysiloma* sp. Para obtener una de las cepas más representativas de los nódulos de *Lysiloma* sp. se obtuvieron los patrones ERIC-PCR de los

primeros 25 aislados de 5 diferentes plántulas de *Lysiloma* sp. Los patrones se visualizaron en geles de agarosa mediante electroforesis. Se observaron 10 patrones diferentes (CCGB01, CCGBR2, CCGBR7, CCGBR10, CCGB12, CCGBR13, CCGBR17, CCGB20, CCGBR22 y CCGBR26). En la última hilera se muestra el marcador molecular Plus DNA Ladder de 1kb.

Tabla 1. Frecuencia de cada huella ERIC-PCR. Para obtener una de las cepas más representativas de los nódulos de *Lysiloma* sp. se obtuvieron los patrones ERIC-PCR de los primeros 25 aislados obtenidos. Se observaron 10 patrones diferentes, el patrón CCGB01 fue el más frecuente (8/25) y se presentó en 3 de 5 plántulas analizadas de *Lysiloma* sp.

Cepas representantes de cada huella ERIC-PCR	Frecuencia de las huellas ERIC-PCR	No. de plántulas de <i>Lysiloma</i> sp. en las que se encontraron cada huella ERIC-PCR
CCGB01	8	3 (plantas 1, 3 y 5)
CCGBR2	2	2 (plantas 1 y 4)
CCGBR7	4	1 (planta 1)
CCGBR10	2	1 (planta 1)
CCGB12	2	1 (planta 2)
CCGBR13	1	1 (planta 2)
CCGBR17	2	2 (plantas 2 y 3)
CCGB20	2	2 (plantas 3 y 5)
CCGBR22	1	1 (planta 4)
CCGBR26	1	1 (planta 5)
Total de huellas=10	Total de aislados= 25	Total de plántulas = 5

Se obtuvo un árbol filogenético de máxima verosimilitud de los genes 16S rARN para las diez cepas que presentaron los 10 diferentes patrones ERIC y para todas las especies del género *Bradyrhizobium* descritas hasta la fecha en la que se hizo este trabajo (71 especies) (Tabla S1). El modelo de sustitución de nucleótidos bajo el que se construyó el árbol fue HKY+I+G con el programa IQ-Tree v.1.6.11. Se observa que las diez cepas se agrupan (Figura 2), aunque la cepa CCGBR10 se separa y se agrupa con *B. liaoningense* 2281^T. El porcentaje de identidad entre las 10 cepas de este trabajo se encuentra entre el 99.24 y el 100%. En adelante, a la cepa con el patrón ERIC más frecuente CCGB01 se le refiere como *Bradyrhizobium* sp. *CCGB01*.

El mayor porcentaje de identidad del gen 16S rRNA la cepa *B. sp. CCGB01* fue de 99.91% con la secuencia de *B. shewense* ERR11^T. El árbol filogenético concuerda con los últimos publicados (Rejili et al., 2020; Ferrarz et al., 2020).

Las secuencias parciales de los genes *16S rARN* se encontrarán depositadas en la base de datos de GenBank con los siguientes números de acceso: *B. sp. CCGB01*- OP185277; *CCGB12*- OP185278; *CCGB20*- OP185279; *CCGBR10*- OP185280; *CCGBR13*- OP185281; *CCGBR17*- OP185282; *CCGBR2*- OP185283; *CCGBR22*- OP185284; *CCGBR26*- OP185285; *CCGBR7*- OP185286 (Tabla S1).

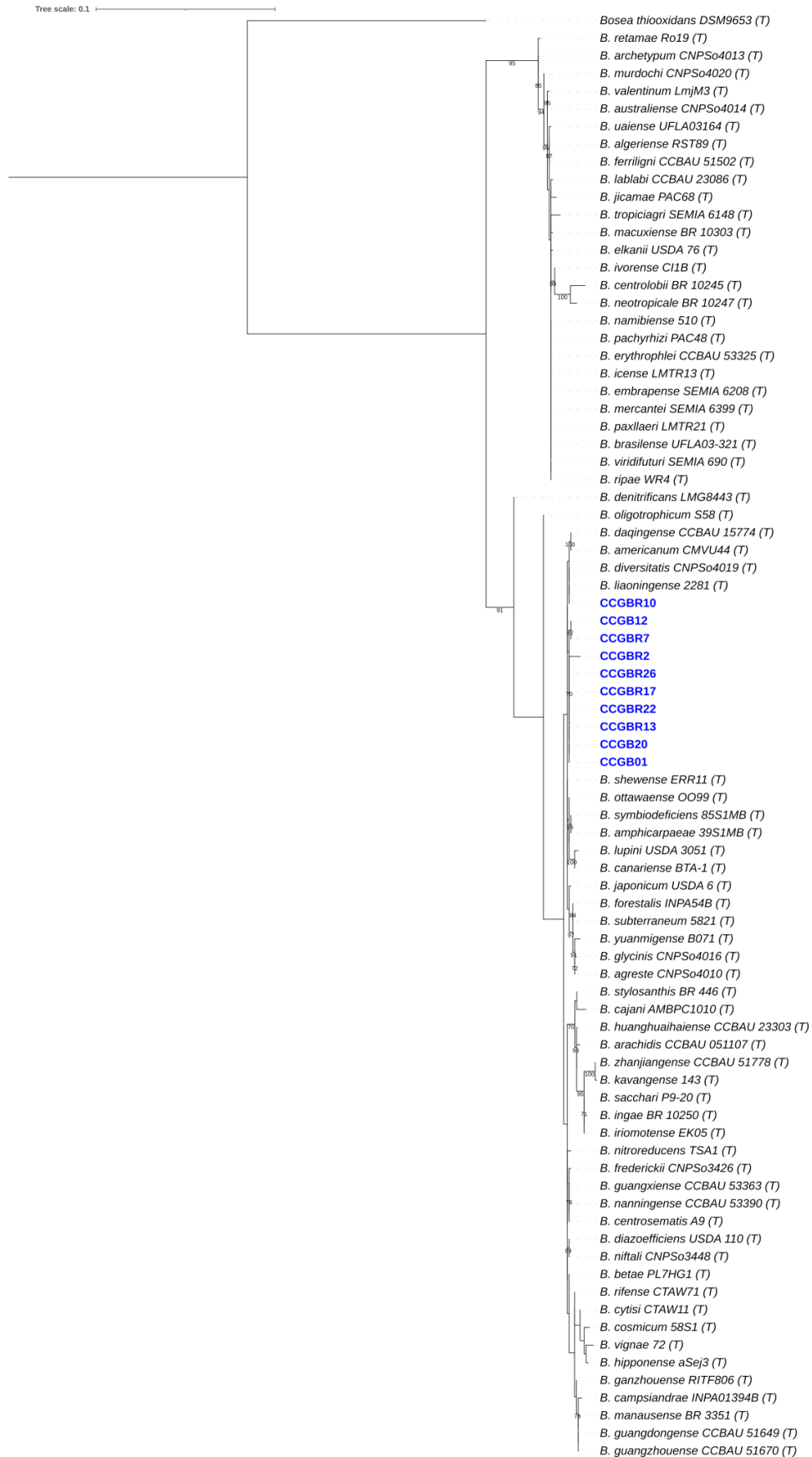


Figura 2. Árbol filogenético de máxima verosimilitud de secuencias de genes parciales 16S rARN de *Bradyrhizobium* sp. CCGB01 y cepas tipo del género con un modelo de sustitución HKY+I+G. La longitud del alineamiento fue de 1190 pb. Se muestran los valores de UFBoot (≥ 70) de 1000 pseudo réplicas; el soporte representa directamente la probabilidad de que el clado sea confiable y a partir de valores ≥ 95 los clados comienzan a ser confiables (Minh et al., 2013). En azul se resaltan las cepas analizadas en este trabajo. La secuencia de *Bosea thiooxidans* DSM9653^T se utilizó como grupo externo. La barra muestra las sustituciones esperadas por sitio; (T)= cepa tipo.

4.2 Secuenciación del genoma de *Bradyrhizobium* sp. *CCGB01*

Se secuenció el genoma de la cepa *CCGB01* (*B. sp. CCGB01*) que pertenece al grupo que presentó el patrón ERIC-PCR más abundante de las 25 cepas analizadas. El tamaño del genoma completo es de 9,317,328 pb y está comprendido en un solo cromosoma. El porcentaje G+C es de aproximadamente 63.8. Además, el genoma posee 9,336 secuencias codificantes, 54 tARNs y 3 rARNs. Estos valores se asemejan a los valores presentados por las cepas tipo *B. japonicum* USDA 6^T y *B. diazoefficiens* USDA 110^T que cuentan con sus genomas completos y bien caracterizados (Tabla 2).

Se obtuvo un análisis de calidad del genoma de *B. sp. CCGB01* anotado con RASTtk mediante las herramientas EvalG con una re-implementación del algoritmo checkM (Parks et al., 2014) y el programa EvalCon a través de la plataforma en línea PATRIC v. 3.6.9 (Davis et al., 2020). De acuerdo con las medidas que este análisis reporta (% de consistencia gruesa = 99; % consistencia dura = 95.3; % de integridad = 100 y % de contaminación = 3.8), se puede establecer que el genoma de *B. sp. CCGB01* posee una buena calidad. Para consultar el significado de cada medida, ver la Tabla S2 en el Material Suplementario.

Tabla 2. Comparación de propiedades del genoma de *Bradyrhizobium* sp. *CCGB01* con los genomas de *B. japonicum* USDA 6^T y *B. japonicum* USDA 110^T. Los datos se obtuvieron usando las herramientas en línea del programa PATRIC v. 3.6.9 (Davis et al., 2020).

Característica genómica	Nombre de la cepa		
	<i>B. sp. CCGB01</i>	<i>B. japonicum</i> USDA 6 ^T	<i>B. diazoefficiens</i> USDA 110 ^T
Tamaño del genoma (pb)	9,317,328	9,207,384	9,105,828
CDS(total)	9,336	9,447	8,489
CDS (codificantes)	9,336	8,829	8,220
rARNs	3	6	3
tARNs	54	51	52
Pseudo genes (total)	NA	NA	213
Regiones repetidas	63	128	148
Contenido GC	63.785	63.665	64.06

Islas simbióticas	Sí	Sí	Sí
Plásmidos	Ninguno	Ninguno	Ninguno

4.3 Análisis filogenéticos de *Bradyrhizobium* sp. *CCGB01* y cálculo de la identidad de nucleótidos promedio (ANI)

Para los análisis filogenéticos, se seleccionaron los genes *housekeeping recA* y *glnII* ya que ambos se encuentran disponibles para todas las especies del género *Bradyrhizobium*. Además, estos dos genes se han utilizado en análisis de MLSA como soporte para la identificación de especies nuevas en el género (Aráujo et al., 2017; Peix et al., 2015; Vinuesa et al., 2005). La filogenia individual de *recA* ubica a *B. sp. CCGB01* en un clado bien soportado (ultrafast bootstrap ≥ 95) con *B. liaoningense* LMG 18230^T (Figura S1). La filogenia del gen *glnII* agrupa a *B. sp. CCGB01* con *B. canariense* BTA-1^T; el clado presenta un valor de ultrafast bootstrap mayor a 95% (Figura S2). Por otro lado, tanto la filogenia bayesiana y de máxima verosimilitud de genes concatenados *glnII* y *recA* presentaron valores de ultrafast bootstrap y de probabilidad posterior bayesiana bajos para la clasificación de *B. sp. CCGB01*. Ambos árboles indican que *B. iriomotense* EK05^T y *B. ingae* BR 1250^T comparte ancestro con la cepa *B. sp. CCGB01*. La probabilidad posterior del nodo que ubica en esta posición a *B. sp. CCGB01* es de 0.71 y el valor ultrafast bootstrap es de 38 (no se muestra en la filogenia) (Figuras 4 y S3). *B. liaoningense* LMG 18230^T presenta porcentaje de identidad más alto (95.79%) a los genes concatenados *glnII* y *recA* de *B. sp. CCGB01*.

Un árbol filogenómico de especies de máxima verosimilitud se infirió de acuerdo a un alineamiento concatenado de proteínas de 1131 ortogrupos para *B. sp. CCGB01*, cepas tipo del género *Bradyrhizobium*, y cepas del clado al que pertenece *B. japonicum* USDA 6^T, clado al que mostró pertenecer *B. sp. CCGB01* según análisis filogenómicos preliminares que se realizaron con el mismo programa OrthoFinder v 2.2.6. En este árbol, la cepa *B. sp. CCGB01* se ubica en la base de dicho clado y está soportado por un alto valor local según la prueba Shimodaira-Hasegawa (Figura 5).

Se realizó un árbol filogenético de los genes *nodC* para *B. sp. CCGB01* y cepas que representan las 12 simbiovariedades que se han descrito para el género *Bradyrhizobium* (Rejili et al., 2020). La secuencia del gen *nodC* de la cepa que se analiza en este trabajo se agrupó en un clado bien soportado con *B. iriomotense* EK05^T, *B. ingae* BR 1250^T, *B. manausense* BR 3351^T, *B. stylosanthis* BR 446^T y *B.americanum* CMVU44^T (ultrafast bootstrap \geq 95%); la simbiovariedad *phaseolarum* es la más cercana a *B. sp. CCGB01*, la cual está comprendida por *B.americanum* CMVU44^T y *B. stylosanthis* BR 446^T (Figura 6).

También se realizó un árbol de máxima verosimilitud para el gen *nifH* de las cepas tipo y otras cepas del superclado *japonicum*. No obstante, es importante aclarar que el alineamiento se hizo con base en 255 nucleótidos, debido a que en el momento de realizar este trabajo, se cuenta únicamente con secuencias de *nifH* de *B. ingae* de esta longitud, y la relación filogenética de esta especie con *B. sp. CCGB01* es de interés para este estudio. Al igual que el gen *nodC*, las secuencias *nifH* más cercanas a las de *B. sp. CCGB01* fueron las secuencias de *B. ingae* BR 10250^T (92.94% de identidad) y *B. iriomotense* EK05^T (90.98%). La filogenia del gen *nifH* concuerda con la filogenia del gen *nodC* en cuanto a la composición del clado que contiene a *B. sp. CCGB01* así como muestra un valor alto de ultrafast bootstrap (ultrafast bootstrap > 95%); sin embargo, en este árbol la secuencia de *B. ingae* BR 10250^T se ubica como ancestro de *B. sp. CCGB01* (Figura 7) mientras que en la filogenia del gen *nodC*, *B. sp. CCGB01* se encuentra como ancestro de *B. ingae* BR 10250^T (Figura 6). El árbol de los genes concatenados *glnII* y *recA*, el árbol filogenómico ni los árboles de los genes simbióticos (*nodC* y *nifH*) concuerdan en cuanto a la agrupación de *B. sp. CCGB01*.

El porcentaje de identidad ANI, con base en el alineamiento de genomas, más alto que presentó *B. sp. CCGB01* fue de 91.2% con el genoma de *B. japonicum* J5. Los porcentajes de identidad ANI con cepas no pertenecientes al clado *japonicum* fueron menores al 90% (Tabla 3). En el mapa de calor de los porcentajes de identidad ANI, se puede observar cómo la agrupa el clado al que pertenece *B. japonicum* USDA 6^T (del superclado *japonicum*) y cómo la cepa analizada en este trabajo se encuentra en la base de este (Figura 3).

Los números de acceso de los genomas y secuencias utilizadas en este trabajo se encuentran en la Tabla S1.

Tabla 3. Porcentajes de identidad de promedio de nucleótidos (ANI) de comparaciones de las cepas más cercanas a *B. sp. CCGB01* con base en el análisis filogenómico. El porcentaje de ANI más alto con la cepa de estudio es el de *B. japonicum* cepa J5 (91.234%).

Nombre de la cepa	ANI (%)	Nombre de la cepa	ANI (%)
<i>B. japonicum</i> J5	91.234	<i>B. japonicum</i> SZCCT0280	91.054
<i>B. liaoningense</i> SZCCT0399	91.171	<i>B. japonicum</i> USDA 6 ^T	91.048
<i>B. sp</i> MOS001	91.109	<i>B. japonicum</i> E109	91.047
<i>B. japonicum</i> CCBAU 15354	91.096	<i>B. japonicum</i> USDA 38	91.045
<i>B. japonicum</i> CCBAU 15517	91.089	<i>B. japonicum</i> CCBAU 25435	91.040
<i>B. japonicum</i> SZCCT0402	91.088	<i>B. liaoningense</i> SZCCT0293	91.036
<i>B. japonicum</i> CCBAU 83623	91.087	<i>B. japonicum</i> Is-34	91.028
<i>B. japonicum</i> SZCCT0395	91.076	<i>B. japonicum</i> UBMA 197	91.009
<i>B. japonicum</i> USDA 123	91.062	<i>B. japonicum</i> FN1	90.968
<i>B. japonicum</i> SEMIA 5079	91.062	<i>B. japonicum</i> NBRC 14783	90.873
<i>B. japonicum</i> CCBAU 15618	91.058	<i>B. niftali</i> CNPSo 3448 ^T	89.711

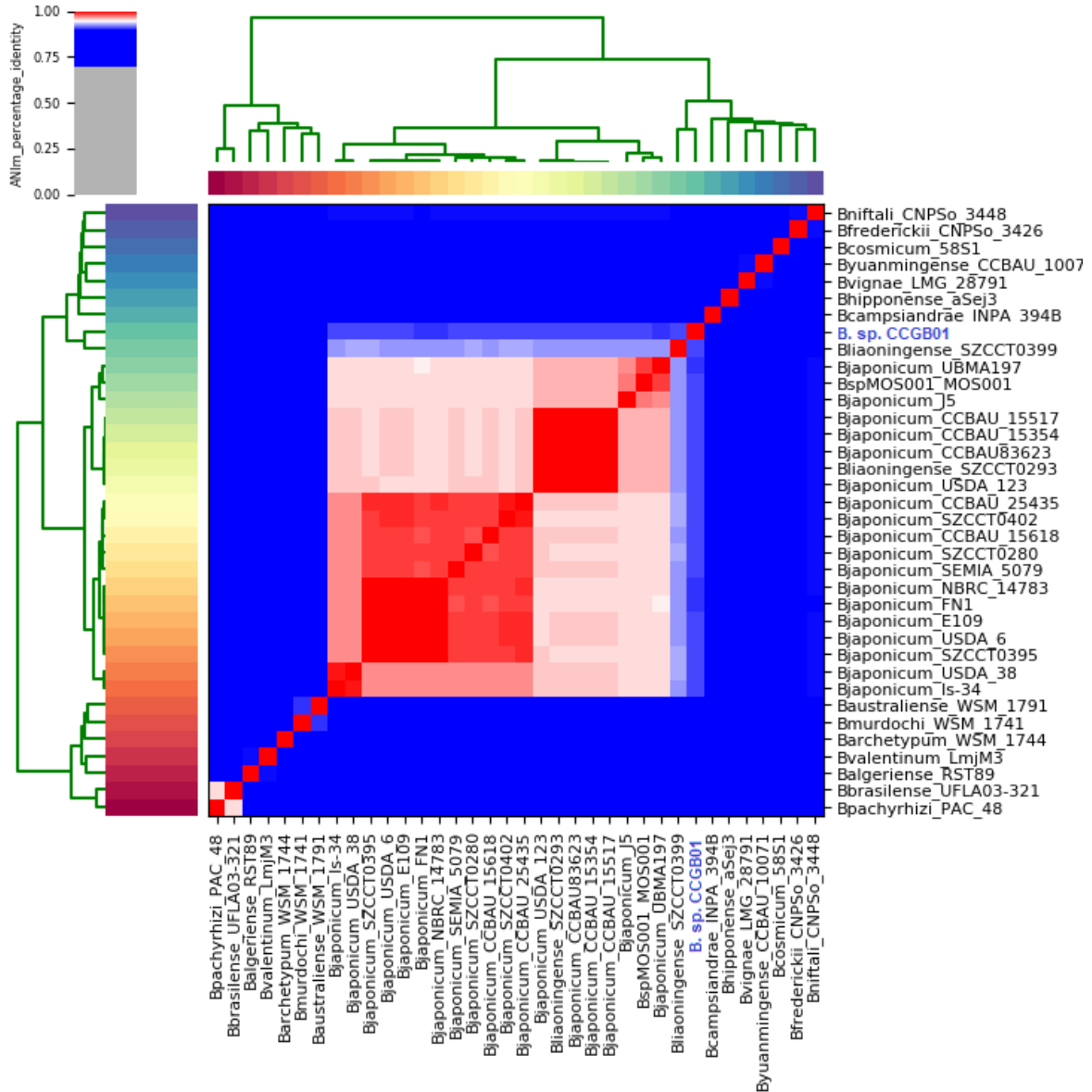


Figura 3. Mapa de calor de los porcentajes de identidad de la identidad de nucleótidos promedio (ANI) para *Bradyrhizobium* sp. CCGB01. Se observa que la cepa *B. sp. CCGB01* se encuentra en la base del clado que contiene a *Bradyrhizobium japonicum* USDA 6T

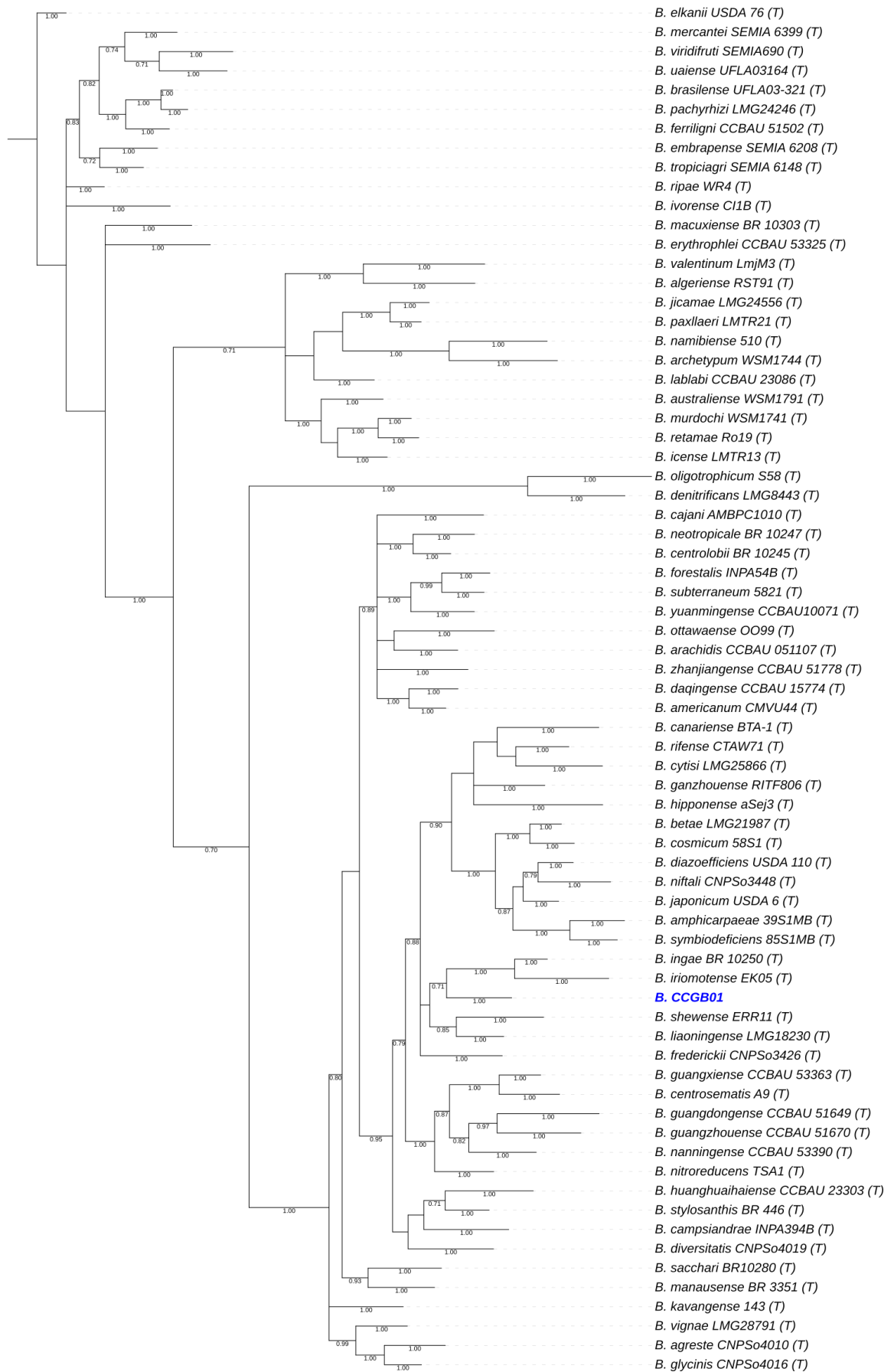


Figura 4. Árbol filogenético bayesiano de los genes concatenados *glnII* y *recA* de *Bradyrhizobium* sp. CCGB01 y cepas tipo del género sin enraizar con partición por gen. Modelo de sustitución GTR+I+G. La longitud de los alineamientos fueron de: *glnII*: 504 pb; *recA*: 377 pb. La barra muestra las sustituciones esperadas por sitio. Se muestran las probabilidades posteriores bayesianas (≥ 0.7); (T)= cepa tipo.

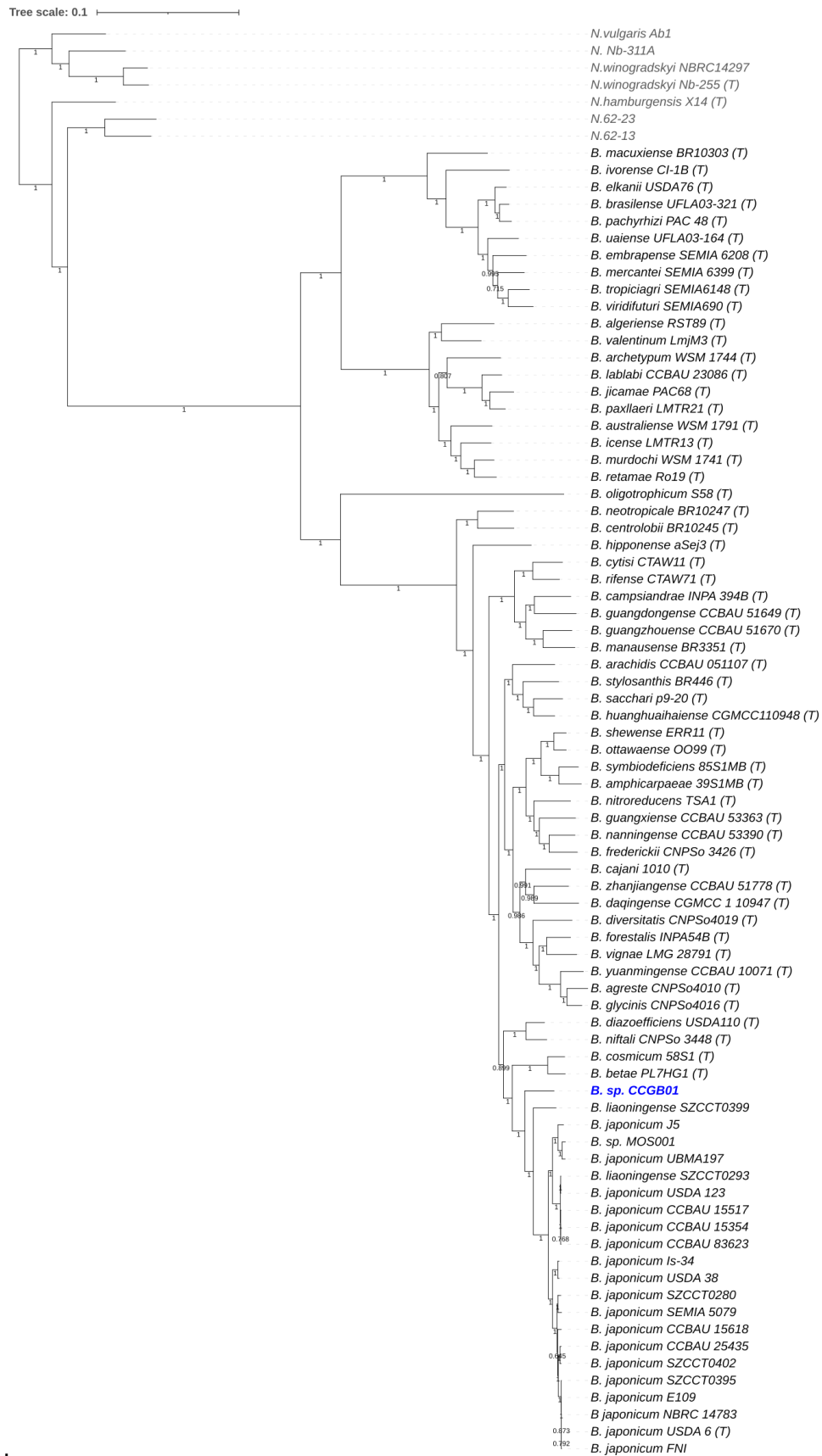


Figura 5. Árbol filogenómico de máxima verosimilitud aproximada inferido del alineamiento concatenado de proteínas de 1131 ortogrupos para *B. sp. CCGB01*, cepas tipo del género y 20 cepas del clado al que pertenece USDA 6^T con un modelo de sustitución JTT+ CAT. Al menos 96,4% de las especies en cualquiera de los ortogrupos utilizados en la filogenia tienen representados genes de copia única. Las cepas de *Nitrobacter* (en gris) se utilizaron como grupo externo. Se muestran los valores de soporte locales calculados con la prueba Shimodaira-Hasegawa; (T)= cepa tipo. Se muestra en azul la cepa *B. sp. CCGB01*.

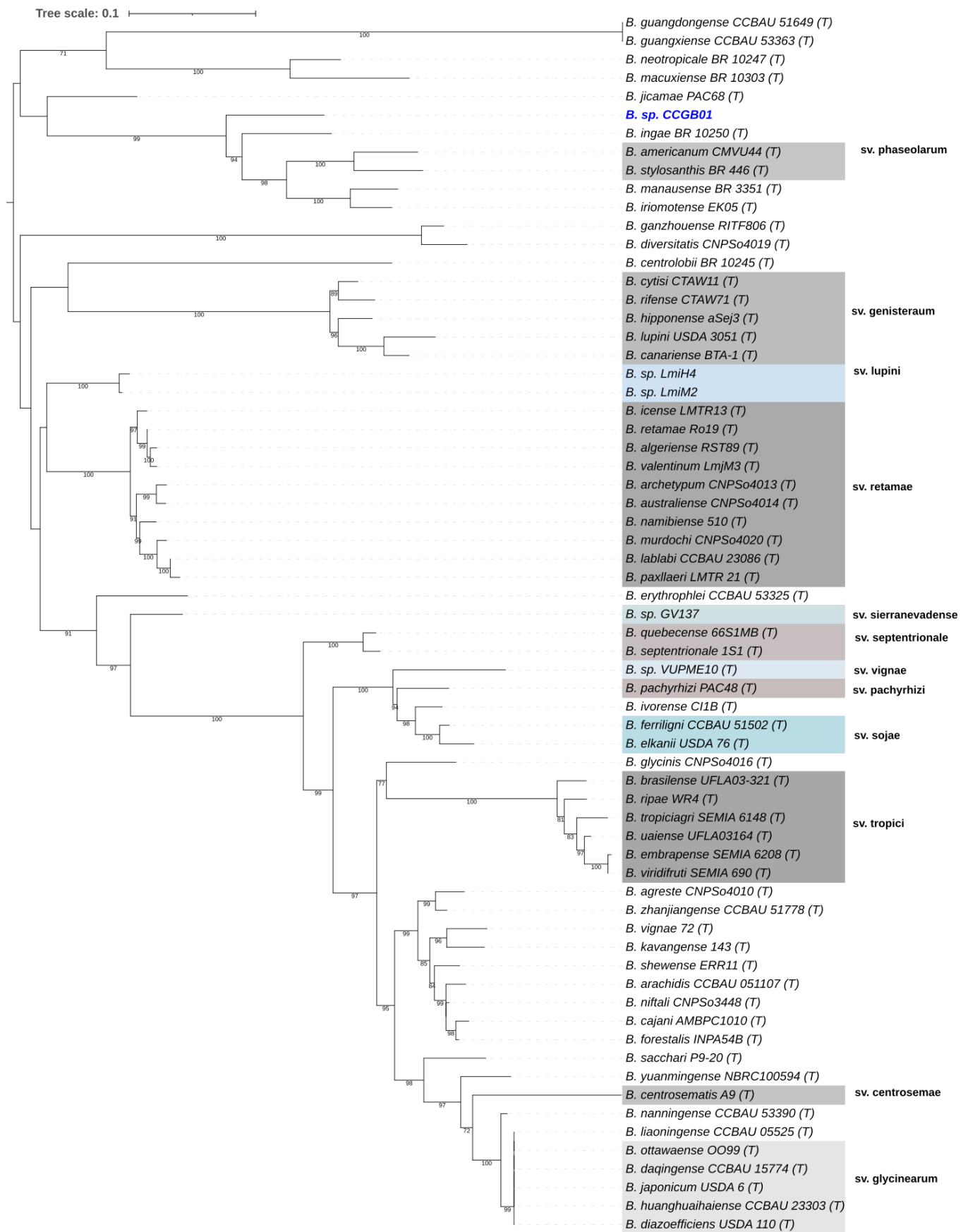


Figura 6. Árbol filogenético de máxima verosimilitud del gen *nodC* para *Bradyrhizobium* sp. CCGB01 y cepas tipo del género sin enraizar con un modelo de sustitución GTR+I+G. La longitud del alineamiento fue de 386 pb. Se muestran los valores de UFBoot (≥ 70) de 1000 pseudo réplicas; el soporte representa directamente la probabilidad de que el clado sea confiable y a partir de valores ≥ 95 los clados comienzan a ser confiables (Minh et al., 2013). En los recuadros coloreados en escala de grises y azules, se resaltan las diferentes simbiovariedades descritas para *Bradyrhizobium*; (T)= cepa tipo

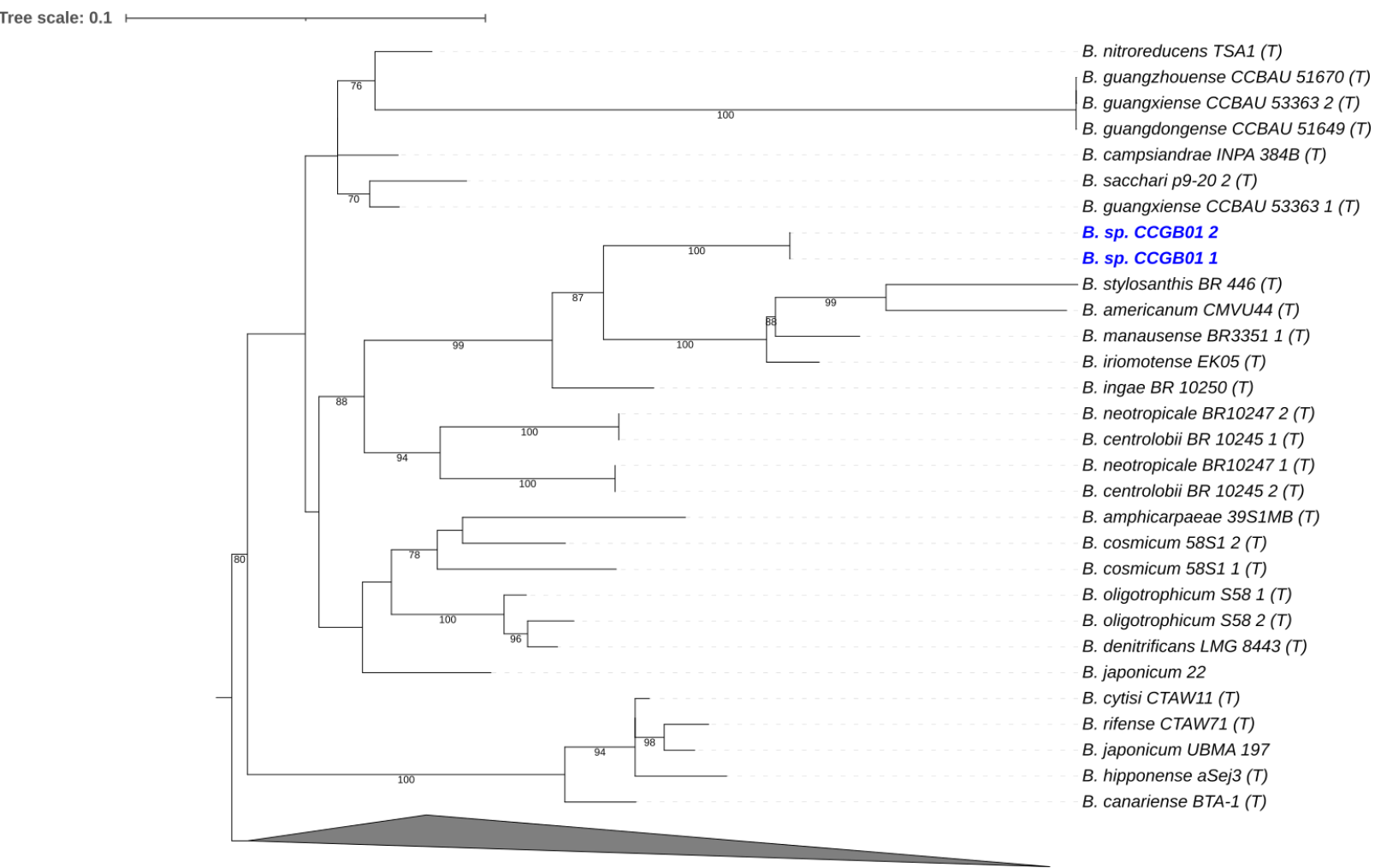


Figura 7. Árbol filogenético de máxima verosimilitud de los genes *nifH* para *Bradyrhizobium* sp. CCGB01 y cepas tipo del género sin enraizar con un modelo de sustitución TrN+I+G4 La longitud del alineamiento fue de 255 pb. Se muestran los valores de UFBoot (≥ 70) de 1000 pseudo réplicas; el soporte representa directamente la probabilidad de que el clado sea confiable y a partir de valores ≥ 95 los clados comienzan a ser confiables (Minh et al., 2013). El clado colapsado contiene otras especies tipo del género *Bradyrhizobium*, estas se despliegan en la Figura S4 del Material Suplementario; (T)= cepa tipo

4.4 Ensayos de nodulación

Para medir la capacidad de nodulación de las cepas con las 10 huellas que se identificaron mediante el método ERIC-PCR, se realizaron ensayos de nodulación en *Phaseolus vulgaris*, *Leucaena leucocephala* y *Lysiloma* sp. Además, se llevó a cabo la cinética de nodulación en *P. vulgaris* y se midió la reducción de acetileno en *P. vulgaris* y *Lysiloma* sp.

En *P. vulgaris* todas las diferentes cepas lograron inducir la formación de nódulos y únicamente se observó la formación de nódulos inefectivos, con excepción de *B. sp. CCGB01*. La nodulación fue lenta en comparación con el grupo control. Únicamente las cepas CCGB12 y CCGBR2 (los nombres se corrigieron en todo el documento) indujeron ligeramente la formación de nódulos a los 10 días post-inoculación, mientras que el resto de los tratamientos indujeron la producción de nódulos a los 13 días post-inoculación (Figura 8). No se detectó la reducción de acetileno en el frijol, con excepción del caso de *B. sp. CCGB01*. La producción de etileno como producto de la reducción del acetileno que mostró *B. sp. CCGB01* fue variable y baja, ya que solo en una plántula de tres se detectaron 380 nmol C₂H₄/hg, lo cual resulta en un promedio de 126.66 nmol C₂H₄/hg.

En *L. leucocephala*, se llevó a cabo el registro de la formación de nódulos durante 38 días, pero no se observó la formación de nódulos salvo en las plantas inoculadas con la cepa control *R. leucaenae* CFN299. No se llevó a cabo la medición de reducción de acetileno en *L. leucocephala*.

En los ensayos con *Lysiloma* sp., todas las cepas nodularon y fueron positivas a la prueba de reducción de acetileno. El ANOVA mostró que al menos un promedio de etileno producido por alguna de las 10 cepas difería de algún otro ($f(9)=8.522$, $p=3.79e-05$). De acuerdo a la prueba de TukeyHSD para comparaciones múltiples, los promedios de la reducción de acetileno (etileno como producto de la reacción, nmol C₂H₄/hg) de las cepas CCGBR13 y CCGBR7 son significativamente menores a los promedios de las cepas CCGB12, CCGB20, CCGBR10, CCGBR22 y CCGBR26; además, el promedio de reducción de acetileno de la cepa CCGB12 es mayor de manera estadísticamente significativa al promedio de la cepa CCGBR2 ($p<0.01$) (Figura 9).([Repositorio con detalles de análisis estadístico](#)).

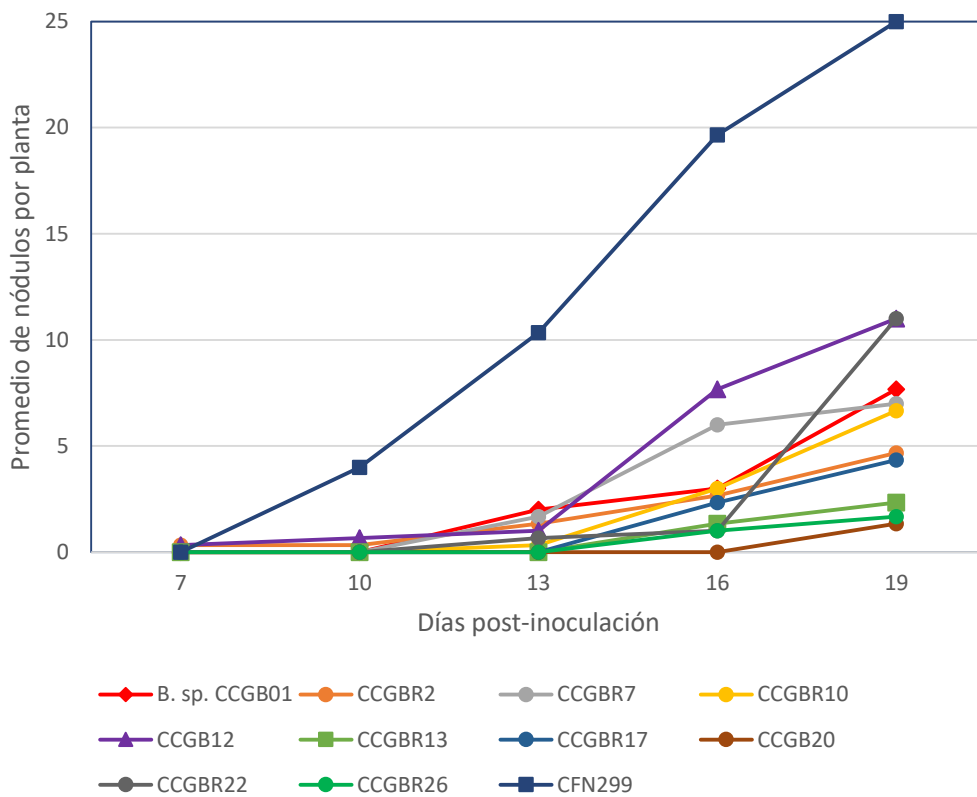


Figura 8. Cinética de nodulación de las cepas que representan las diferentes huellas genéticas identificadas con el método ERIC-PCR que se aislaron de *Lysiloma* sp en *Phaseolus vulgaris*. La cepa *Rhizobium leucaenae* CFN299 se utilizó como grupo control.

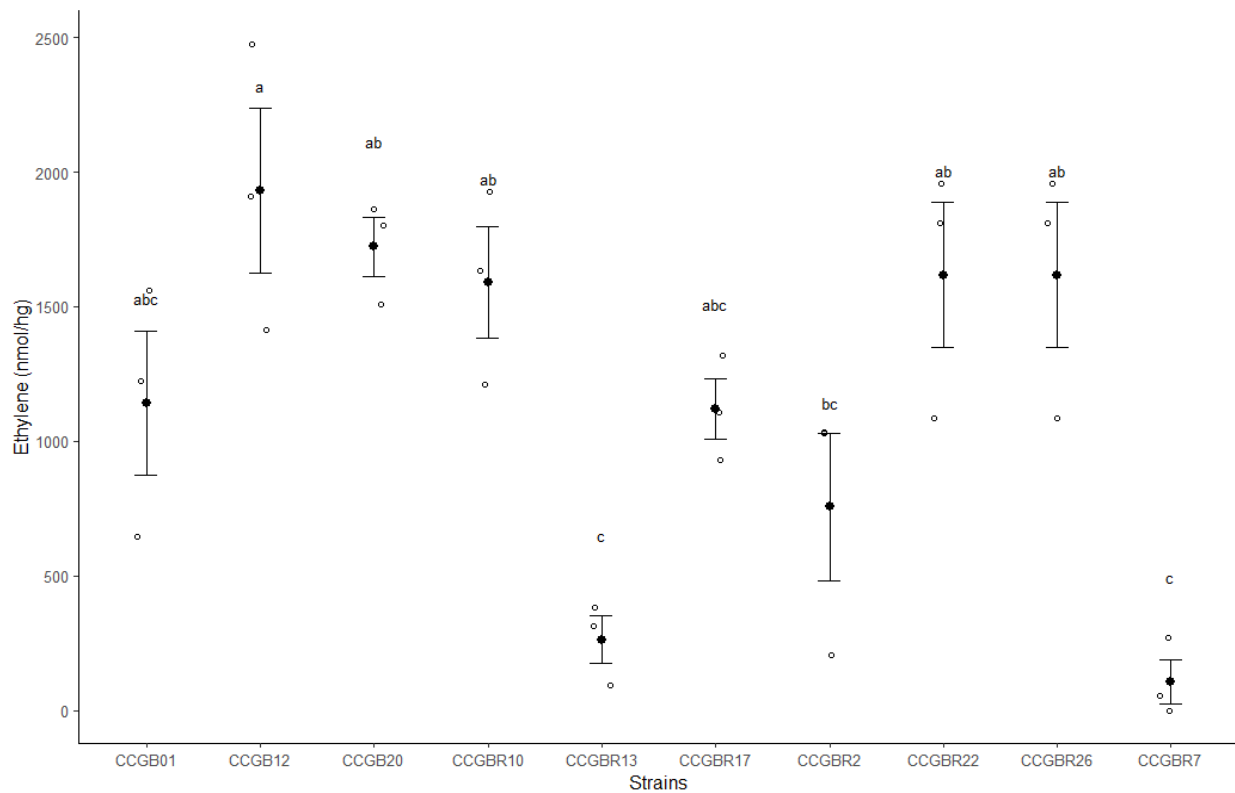


Figura 9. Comparación de nanomoles de etileno (C_2H_4) por hectogrammo (hg) como producto de la reducción del acetileno por las cepas que representan las diferentes huellas genéticas identificadas con el método ERIC-PCR que se aislaron de *Lysiloma* sp. El ensayo de nodulación se llevó a cabo en *Lysiloma* sp. Se muestran las barras de errores estándar. Las letras (a,b,c) indican las relaciones entre los grupos de acuerdo a la prueba estadística de TukeyHSD de comparaciones múltiples ($p < 0.01$).

4.5 Predicción de islas genómicas identificación de genes simbióticos de *Bradyrhizobium* sp. CCGB01

Para identificar las probables islas simbióticas de las predicciones de islas genómicas de Island Viewer 4 que se obtuvieron en este trabajo, se consideró que las islas presentaran al menos un gen involucrado en la nodulación. Se identificaron tres islas genómicas que cumplieron este criterio.

Se identificó una primera posible isla simbiótica (nombrada PSI1) de 421.144 kb con las coordenadas 1604779, 2025923 pb, que se ubica en una zona con menor contenido GC respecto al resto del cromosoma (Figura 10). En esta isla, el programa RASTtk anotó 231 proteínas hipotéticas, de las cuales 66 se identificaron con funciones conocidas con el programa eggNOG-mapper. Esta isla simbiótica contiene genes que codifican para genes de los clusters Nod (*nodABCDSJ*, *noel*, *noIY* y *noIO*), Fix (*fixABC* y *RX*) y Nif (*nifHNEXSTUBZQA* y *W*). Esta isla presenta dos copias de los genes *nifH* y *nodD* (Figura 11). Se encuentra flanqueada por tARNs. También contiene genes que suelen encontrarse en las islas genómicas que codifican para diversos elementos móviles, transposasas, integrasas de fagos y una probable integrasa/recombinasa; transportadores ABC, genes relacionados con las bombas de eflujo RND, reguladores de respuesta transcripcional de 2 componentes y un activador transcripcional de unión a acil-homoserina lactona de la familia LuxR y un gen que codifica para el factor RNA polimerasa σ^{54} . Además, contiene genes que codifican para las proteínas del T4SS, TrwB y TrbI (Tabla 4).

También se identificó otra isla genómica (PSI2), de 342.841 kb con las coordenadas 6974868, 7317709 pb. La isla PSI2 se encuentra en una región con menor contenido de GC respecto al cromosoma (Figura 10). En esta isla, el programa RASTtk anotó 278 proteínas hipotéticas, de las cuales 88 se identificaron con funciones conocidas con el programa eggNOG-mapper. Entre los genes que contiene esta isla se encuentra un gen que codifica para un regulador de respuesta transcripcional de 2 componentes de la familia NarL/FixJ, elementos móviles, proteínas hipotéticas, transportadores ABC, genes relacionados con las bombas de eflujo RND, genes relacionados con el eflujo de multi-fármacos y

genes relacionados al sistema *regBA* y al igual que PSI1, un regulador de respuesta transcripcional de dos componentes de la familia LuxR. PSI2 también se encuentra flanqueada por tARNs (Tabla 4).

Una tercera isla (PSI3), con las coordenadas 7899957,8128700 pb, presentó los genes simbióticos *nolL* y *nfeC*. Al igual que las islas anteriores, la isla se encuentra en una región con menor contenido GC respecto al cromosoma (Figura 10). El programa RASTtk anotó 76 proteínas hipotéticas de las cuales se lograron re-anotar 58 proteínas con el programa eggNOG-mapper. También, PSI3 contiene un gen que codifica para un regulador de respuesta transcripcional de dos componentes de la familia LuxR y un gen que codifica para una probable integrasa/recombinasa. PSI3 contiene genes que codifican para proteínas del T4SS: Trb(I,G,F,L,J,E,D,C,B,N),VirD4, VirD2 y TraF. Esta isla no se encuentra flanqueada por tRNAs (Tabla 4).

Fuera de las islas genómicas también se identificaron genes simbióticos anotados con RASTtk. Hay copias adicionales de los genes *nifB* y *nifU* (una copia de cada gen) así como un solo gen que codifica para *nodN*. Fuera de las islas, también se encuentra un cluster de los genes *fixGHIJKLMNOPQS* y un gen adicional *fixR* (entre las coordenadas 2671217, 2918793); además, dos genes anotados como reguladores transcripcionales de dos componentes de la familia NarL/FixJ otros dos genes *fixJ* y otro gen *fixK* que no se encuentran en el cluster anterior.

Tabla 4. Descripción general de las probables islas simbióticas de *B. sp. CCGB01*. Se incluye el tamaño aproximado de las islas genómicas, su ubicación, los genes relacionados con la nodulación, genes del sistema de conjugación tipo 4 (T4SS) y otros genes que suelen encontrarse en las islas genómicas; además se incluyen los tARNs que flanquean las islas.

Características	PSI1	PSI2	PSI3
Coordenadas (pb)	1604779, 2025923	6974868, 7317709	7899957, 8128700
Tamaño (kb)	421.144	342.841	228.743
Genes simbióticos	Cluster <i>Nod</i> : (<i>nodABCDSEIJU,nolO, nolY</i> y <i>noel</i>) Cluster <i>Nif</i> : (<i>nifHNDKENXSTUBZQAWV</i>)	Cluster <i>Fix</i> : Regulador de respuesta transcripcional de 2 componentes de la familia NarL/FixJ	Cluster <i>Nod</i> : <i>nolL</i> y <i>nfeC</i>

	Cluster <i>Fix</i> (<i>fixABCR</i> y <i>X</i>)		
Genes del T4SS	<i>trbI</i> y <i>trwB</i>	NA	<i>trb(I,G,F,L,J,E,D,C,B,N)</i> , <i>virD4</i> , <i>virD2</i> y <i>traF</i>
Otros genes	<ul style="list-style-type: none"> - Elementos móviles, transposasas, integrasas de fagos, una probable integrasa/recombinasa - Transportadores ABC, genes relacionados con las bombas de eflujo RND - Reguladores de respuesta transcripcional de 2 componentes y activador transcripcional de unión a acil-homoserina lactona de la familia LuxR - Factor RNA polimerasa $\sigma 54$ 	<ul style="list-style-type: none"> - Un elemento móvil, una integrasa, integrasas de fago - Transportadores ABC, genes relacionados con las bombas de eflujo RND y otros genes anotados como genes relacionados con el eflujo de multi-fármacos - Regulador de respuesta transcripcional de dos componentes de la familia LuxR - genes <i>regBA</i> 	<ul style="list-style-type: none"> - una probable integrasa/recombinasa - Regulador de respuesta transcripcional de dos componentes de la familia LuxR
tRNAs	<p>Inicio: un tRNA-Met-CAT dirección forward .</p> <p>Final: un tRNA-Ile-TAT y un tRNA-Val-CAC dirección reversa</p>	<p>Inicio: dos tRNA-Pseudo-CTT en dirección reversa</p> <p>Final: un tRNA-Lys-CTT dirección reversa</p>	NA

* T4SS = Sistema de secreción tipo 4

* NA= No aplica

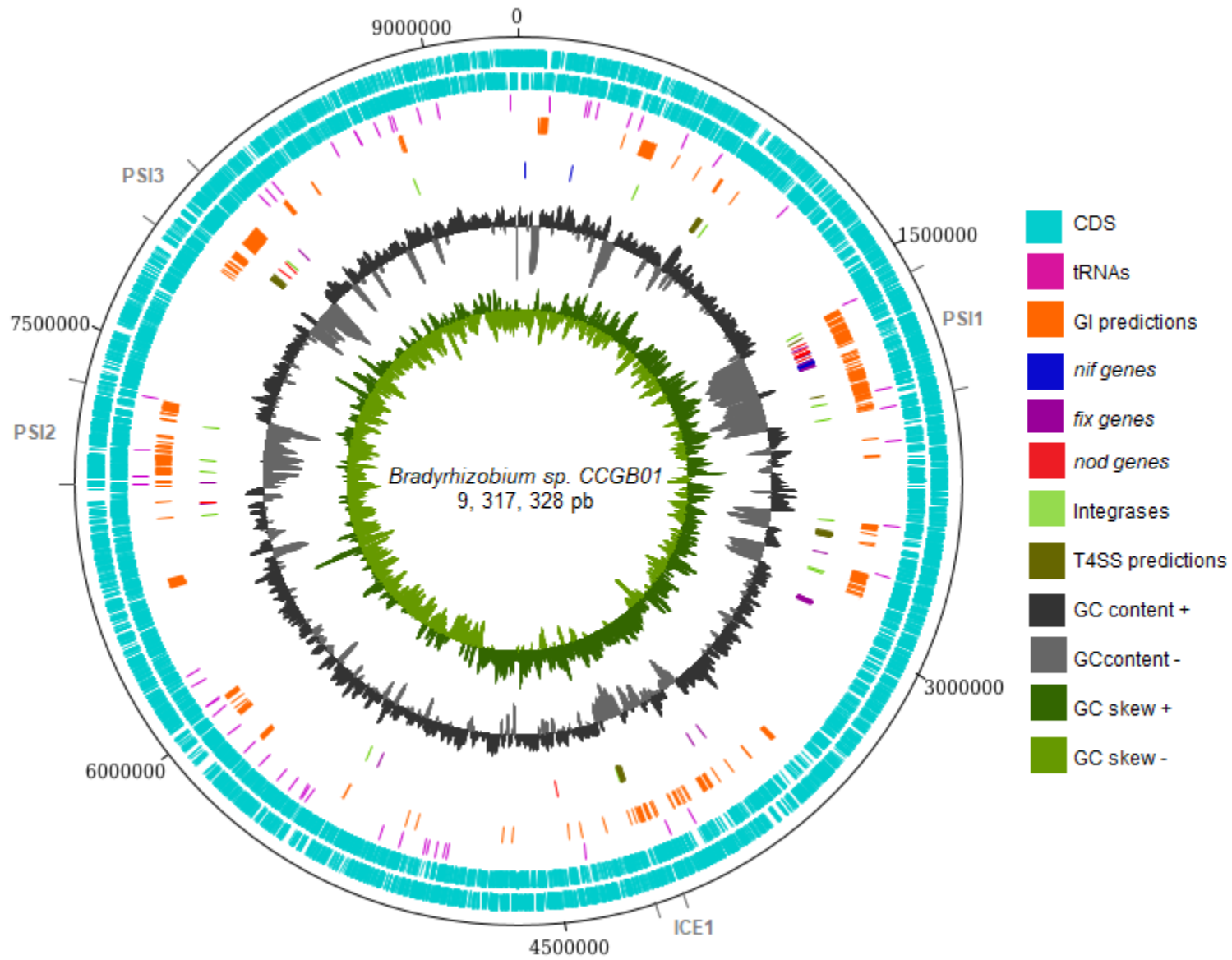


Figura 10. Mapa genómico circular de *Bradyrhizobium* sp. CCGB01. El círculo más externo muestra las coordenadas en pares de bases (pb), los siguientes círculos muestran las posiciones de los probables genes que codifican proteínas en sentido de las manecillas de reloj y en el sentido contrario. El quinto círculo muestra las predicciones de islas genómicas con el programa Island Viewer 4 (Bertelli et al, 2017) El sexto círculo muestra las posiciones de los probables genes que codifican para integrasas, genes *nod*, *nif* y *fix* y las predicciones de los sistemas conjugativos tipo 4 (T4SS). Los dos últimos círculos representan el contenido GC y el sesgo GC respectivamente. PSI= Probable isla simbiótica; ICE=Elemento Conjugativo e Integrativo. La visualización se realizó con DNA plotter (Carver et al., 2019).

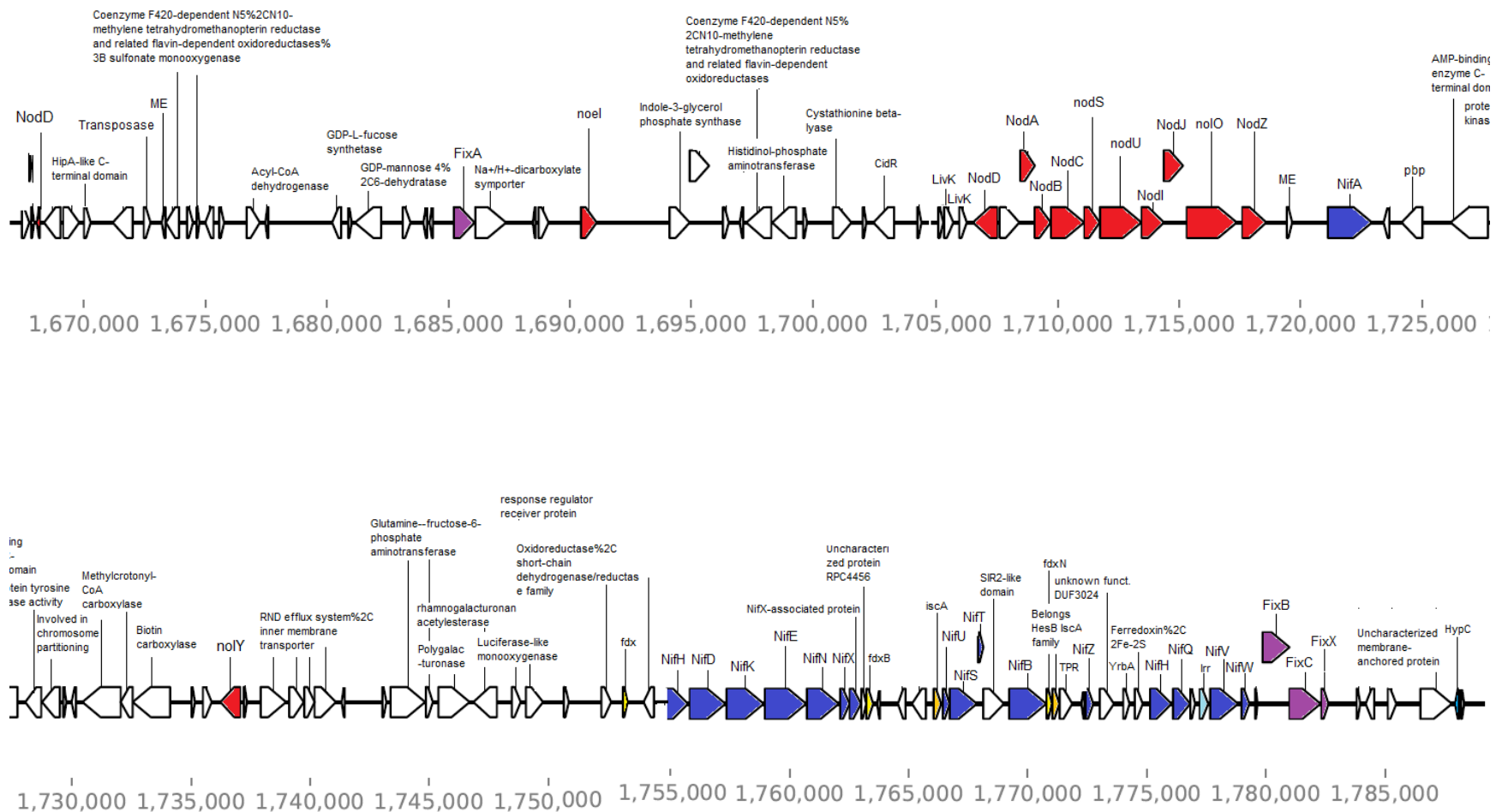


Figura 11. Representación de la región de la probable isla simbiótica PSI1 de *B. sp. CCGB01* en la que se encuentran los genes esenciales para la nodulación. En rojo se pueden observar los genes *nod*, en azul los genes *nif* y en morado los genes *fix*. Las cajas en blanco corresponden a proteínas hipotéticas.

4.5.1 Identificación de sistemas conjugativos que podrían movilizar a las islas simbióticas

Tanto el módulo CONJscan como con el programa oriTfinder detectaron cuatro regiones con componentes pertenecientes al T4SS. Ambos programas detectaron los componentes de los sistemas en las mismas regiones, sin embargo, a diferencia de oriTfinder, CONJscan logró identificar las proteínas relaxasas en cada una de las regiones, mientras que oriTfinder incluyó a las proteínas TraF y TrbN dentro de los sistemas T4SS identificados (Figura 12). El programa CONJscan no identificó T4SSs atípicos o incompletos y todos los sistemas que predijo el programa se clasificaron como T4SS de tipo T. El programa oriTfinder no identificó regiones de replicación (*oriT*).

Fuera de las cuatro regiones identificadas por CONJscan y oriTfinder, dentro de la posible isla simbiótica PSI1, se encuentran las proteínas TrwB y TrbI. Se logró identificar la proteína acopladora TrwB mediante oriTfinder y eggNOG-mapper.

El programa ICEfinder únicamente identificó un posible ICE (4094693, 4190549) con el sitio de inserción en las coordenadas 4096431, 4098638 pb, y con las secuencias directas repetidas *attL* (4094693, 4094707: aagatggccacaaca) y *attR* (4190535, 4190549: aagatggccacaaca).

- CONJscan and oriTfinder T4SS predictions
- CONJscan T4SS predictions
- oriTfinder T4SS predictions

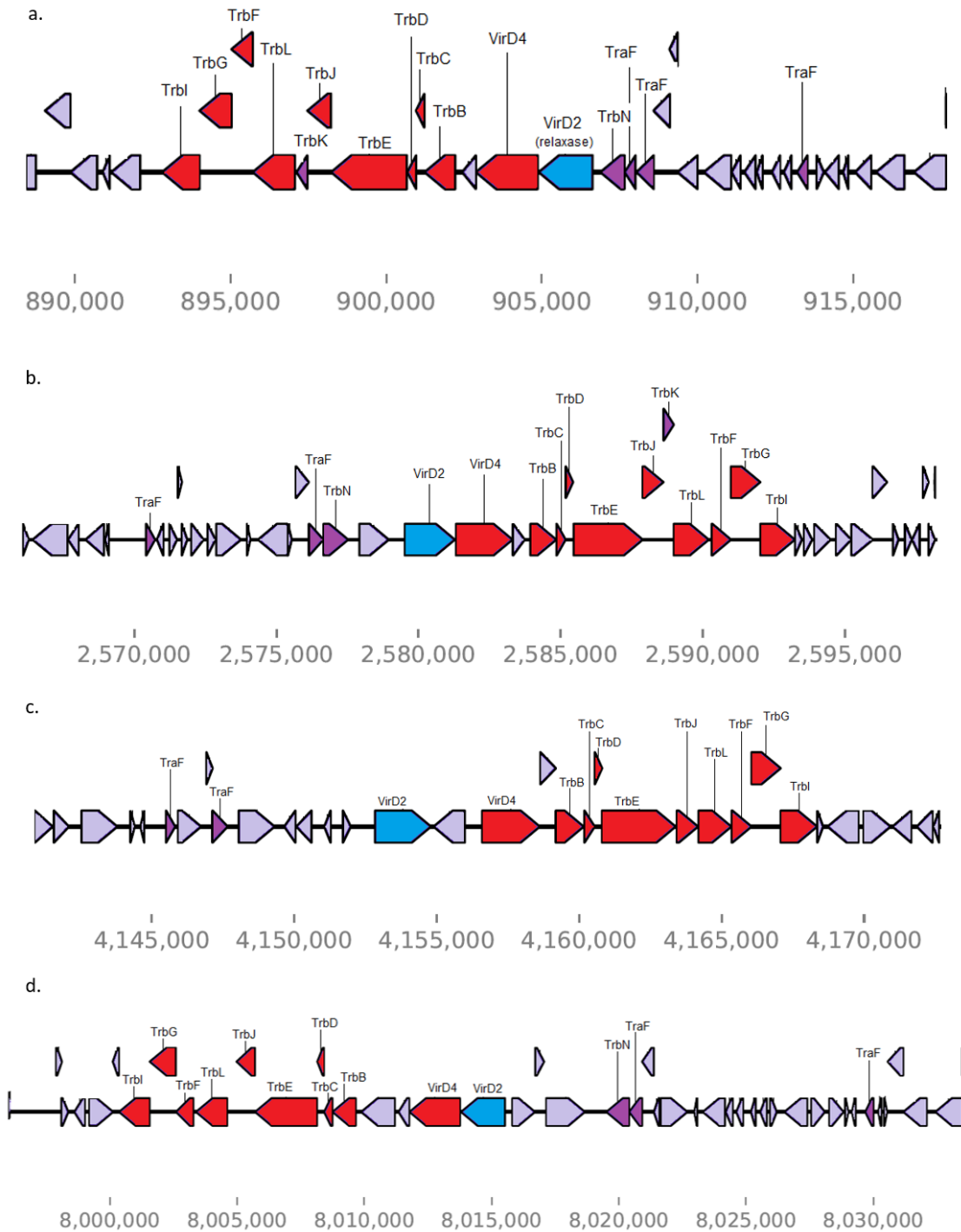


Figura 12. Representación de las predicciones de los T4SSs del genoma de *B. sp. CCGB01* con el módulo CONJscan y el programa oriTfinder. Tanto CONJscan como oriTfinder identificaron cuatro posibles T4SSs en las mismas regiones (a-d).

5. DISCUSIÓN

5.1 Clasificación taxonómica de *Bradyrhizobium* sp. CCGB01

B. sp. CCGB01 representa una nueva especie. De acuerdo con el análisis filogenómico, se realizó el cálculo de la identidad de nucleótidos promedio (ANI). El % ANI más alto fue de 91.2% con *B. japonicum* J5; esta cepa fue aislada de los nódulos de la soja (*Glycine max*) en Hokkaido, Japón. Para clasificar una cepa en una especie ya descrita, se ha propuesto un ANI mayor a 95-96% (Richter y Rosselló-Móra, 2009), lo cual permite clasificar a *B. sp. CCGB01* como una especie nueva. También se realizó un árbol filogenético bayesiano y otro de máxima verosimilitud de los genes *housekeeping* concatenados *glnII* y *recA*. Ambos árboles ubican a la cepa *B. sp. CCGB01* como ancestro de *B. iriomotense* EK05^T y *B. ingae* BR 1250^T; no obstante, el nodo que contiene a *B. sp. CCGB01* presenta una probabilidad posterior bayesiana menor a 0.9 y un valor bootstrap menor a 70% (Figura 4 y S3). Para mejorar esta filogenia se podrían incluir más genes *housekeeping*; sin embargo, con el análisis filogenómico es posible hacer comparaciones con genomas completos y no tomar únicamente en cuenta los genes *housekeeping*, ya que estos son pocos.

5.2 Incongruencia entre las filogenias de los genes *glnII* y *recA*, las filogenias de *nodC* y *nifH* y el análisis filogenómico

Se realizó una filogenia de máxima verosimilitud del gen *nodC* para las cepas tipo del género *Bradyrhizobium* que representaran las 12 simbiovariedades que se han descrito hasta el momento (Rejili et al., 2020). *B. sp. CCGB01* se agrupó con *B. iriomotense* EK05^T, *B. ingae* BR 1250^T, *B. manausense* BR 3351^T, *B. stylosanthis* BR 446^T y *B. americanum* CMVU44^T en un clado bien soportado (Figura 6); el mayor porcentaje de identidad fue de 87.4% con *B. ingae* BR 10250^T. Esta especie fue aislada de los nódulos de la mimosa *Inga laurina* (Sw.) Willd, que pertenece a la misma tribu de *Lysiloma*, en una región con clima tropical Aw (clasificación Köppen) de Brasil; algunas especies del género se encuentran también en México (da Silva et al., 2014). Para clasificar una cepa en una simbiovariedad ya descrita, se ha propuesto una similitud de los genes *nodC* de

92.5% (Delamuta et al., 2014) por lo que, aunque *B. sp. CCGB01* es cercana a especies que pertenecen a la simbiovariedad *phaseolarum*, no se podría clasificar en una simbiovariedad con base en el gen *nodC*. Este árbol filogenético concuerda con los últimos publicados (Bromfield y Cloutier; Rejili et al., 2020; Ferrarz et al., 2020). También se realizó una filogenia del gen *nifH* (Figura 7), la cual concuerda con el último árbol publicado de este tipo para el género *Bradyrhizobium* (Bromfield y Cloutier, 2021); los resultados de esta filogenia respaldan los resultados que sugiere la filogenia del gen *nodC*.

Por otro lado, con el análisis filogenómico se infirió un árbol de especies de máxima verosimilitud en donde *B. sp. CCGB01* se ubica con un buen soporte en el clado que representa la cepa *B. japonicum* USDA 6^T (Figura 5). El árbol filogenómico es consistente con los últimos reportados para el género (Tao et al., 2021; Ormeño-Orrillo y Martínez-Romero, 2019; Avontuur et al., 2019). La discrepancia entre las filogenias de los genes *nodC* y *nifH* con el análisis filogenómico, indica que las propiedades simbióticas de *B. sp. CCGB01* pueden transferirse entre cepas y especies.

5.3 Evaluación de la capacidad de nodulación de *Bradyrhizobium sp. CCGB01*

Los ensayos de nodulación sugieren que la nodulación de *B. sp. CCGB01* está mejor adaptada en su hospedero ya que no se observó la formación de nódulos efectivos en *Leucaena leucocephala* y en *Phaseolus vulgaris* los nódulos mostraron una fijación de nitrógeno baja y variable. *B. sp. CCGB01* mostró una reducción variable y baja en *P. vulgaris* (126.66 nmol C₂H₄/hg en promedio; únicamente una planta mostró fijación de nitrógeno) mientras que en *Lysiloma sp.* el promedio fue de 1142.08 nmol C₂H₄/hg (Figura 9). El resto de las cepas que se delimitaron con base en el método ERIC-PCR no formaron nódulos efectivos ni mostraron fijación de nitrógeno ni en *P. vulgaris* ni en *L. leucocephala*. Se determinó que estas cepas pertenecen al género *Bradyrhizobium* de acuerdo a un árbol filogenético bayesiano de secuencias parciales del gen 16S rARN. Las secuencias de las 10 cepas con patrones ERIC diferentes, presentaron

porcentajes de identidad de 99.24 a 100% entre sí (Figura 2). El árbol filogenético concuerda con los últimos publicados para el género (Rejili et al., 2020; Ferrarz et al., 2020). Debido a que los análisis filogenéticos de los genes *nodC* y *nifH* mostraron que *B. ingae* BR 1250^T podría ser la cepa más cercana en cuanto a las características simbióticas, ensayos de nodulación en otros ejemplares de la tribu *Ingeae* ayudarían a esclarecer las capacidades de nodulación de *B. sp. CCGB01*.

5.4 La mayoría de los genes simbióticos de *Bradyrhizobium sp. CCGB01* se encuentran en una isla genómica

Se identificaron tres probables islas simbióticas que se denominaron PS1, PS2 y PS3. En la figura 10 se aprecia que las tres presentaron un contenido GC diferente, en este caso menor, respecto al resto del cromosoma, lo cual es una característica de los elementos que provienen de eventos de transferencia horizontal (Zhang et al., 2014; Juhas et al., 2009)(Figura 10).

La isla PSI1 que se identificó en el genoma de *B. sp. CCGB01* contiene genes que le permiten establecer relaciones simbióticas con su hospedero. Esta isla contiene genes *nod* (*ABCDIJSUZ*, *noIO*, *noIY* y *noel*), *nif* (*HDKENBXUSQTZA* y *W*) y *fix* (*ABCX* y *R*).

Los genes *nod* codifican para la síntesis y transporte de los factores de nodulación (NFs), los cuales son lipo-quitina-oligosacáridos necesarios para el establecimiento de la relación simbiótica con el hospedero. Los genes *nodABC* codifican para la síntesis del esqueleto de estas moléculas (Paudel et al., 2020; Long et al., 1996; Roche et al., 1996). En su isla simbiótica, *B. sp. CCGB01* contiene dos copias del gen *nodD*; el producto de este gen interactúa con los flavonoides que exudan las raíces del hospedero y desencadena una cascada de señalizaciones para la formación de los primordios nodulares (Perret et al., 2000). Diversas copias de los genes *nodD* podrían conferir mayor estabilidad del proceso de nodulación si hubiera una mutación en alguna de las copias y además, cada copia podría responder de diferente manera a los flavonoides de las plantas (Broughton et al., 2000). Los genes *nodI* y *J* están involucrados en el transporte de los NFs (Fernández-López et al., 1996). Los genes *nodSUZ*, *noIO* y *noel* se

encargan de decorar químicamente los NFs. El gen *noI/O* está involucrado en la 3-O(o 4-O)-carbamoilación del término no reductor de los NFs y *noel* en la 2-O-metilación del grupo fucosil; mientras que *nodS* codifica para una metiltransferasa involucrada en la N metilación, *nodU* para una enzima que lleva a cabo la 6-O-carbamoilación y *nodZ* para la enzima encargada de la fucosilación del residuo de N-acetilglucosamina reductor terminal (Jabbouri et al., 1998; López-Lara et al., 2001; Jabbouri et al., 1995; Quesada-Vincens et al., 1997). Estos genes son necesarios para establecer relaciones simbióticas solo con algunas leguminosas. *B. sp. CCGB01* también presenta el gen *noIY* el cual, de acuerdo con Dockendorff et al. (1994), es regulado por NodD1 y por isoflavonas. La delección mutacional de este gen no tiene efectos notables en la nodulación de la soja (*Glycine max*), caupí (*Vigna unguiculata*) ni en el siratro (*Macroptilium atropurpureum*); sin embargo, los efectos de su delección en el frijol mungo (*V. radiata*) fueron mayores.

La PSI1 de *B. sp. CCGB01* contiene dos copias del gen *nifH*. Dentro del género *Bradyrhizobium*, son pocas las cepas que contienen dos copias del gen *nifH*. Las cepas fotosintéticas de *Bradyrhizobium* se caracterizan por contener dos copias del gen *nifH*, así como *B. sacchari* p9-20^T y *B. guangxiense* CCBAU 53363^T que contienen un gen *nifH* en una isla simbiótica y otro en una isla de genes *nif*, característica de las cepas de *Bradyrhizobium* con estilo de vida libre (Tao et al., 2021). *nifHDK* son genes que codifican para la estructura catalítica de la enzima nitrogenasa de molibdeno mientras que los genes *nifENBXUSQ* codifican para la biosíntesis del cofactor FeMo-co, necesario para que la nitrogenasa lleve a cabo la lisis del triple enlace del N₂ (Addo y Dos Santos 2020; Rubio y Ludden, 2008). Los genes *nifZ* y *nifW* codifican para la maduración de la nitrogenasa y el gen *nifA* está involucrado en la regulación de los genes *nif* (Addo y Dos Santos, 2020; Rubio y Ludden, 2008; Jimenez-Vicente et al., 2019). Se desconoce la función del gen *nifT* (Paudel et al., 2020). Esta isla también contiene un gen *nifV*, el cual codifica para una enzima que sintetiza homocitrato, parte del cofactor Fe-Mo de la nitrogenasa (Nowen et al., 2017). La presencia del gen *nifV* es interesante ya que esta enzima no se encuentra en la mitad de los genomas del género *Bradyrhizobium* (Tao et al., 2021); en los casos en los que no existe el gen, el

homocitrato lo provee el hospedero. Tao et al. (2021), con base en análisis filogenéticos, apoyan la hipótesis que postula que el gen *nifV* estaba presente en el último ancestro común (LCA) de *Bradyrhizobium* que presentaban un estilo de vida libre, pero se perdieron en el superclado *B. japonicum*, cuando las relaciones simbióticas permitieron prescindir del gen. Fuera de las islas genómicas, *B. CCGB01* presenta una copia adicional del gen *nifB* y del gen *nifU*.

Los genes *fixABCX* codifican para un operón involucrado en el transporte de electrones que es necesario para la actividad de la nitrogenasa (Lindström y Mousavi, 2019). El gen *fixR* está involucrado en una de las dos cascadas de regulación de la transcripción de los genes *fix* y *nif* mediante el sistema de detección redox RegS-RegR (Dixon y Khan, 2004). *fixR*, junto con *nifA*, codifican para el activador transcripcional NifA (Bauer et al., 1998). El factor RNA polimerasa σ^{54} , que también se encuentra codificado en PSI1, interactúa con NifA (Lindström y Mousavi, 2019). Ni en las islas genómicas ni en el resto del genoma se encuentran anotados los genes *regS* o *regR*, sin embargo, en PSI2 se encuentra un sistema RegBA y otro sistema se encuentra antes de la PSI1, entre las coordenadas 9260309 y 9260206 pb. Se ha encontrado evidencia de similitud funcional de RegBA con RegSR (Emmerich et al., 2000).

La PSI3 posee dos genes involucrados en la simbiosis: *nolL* y *nfeC*. El gen *nolL* codifica para una enzima que agrega el grupo 4-O-acetil al residuo fucosílico de los NFs (Berck et al. 1999; Corvera et al. 1999). Por otro lado, el gen *nfeC*, que como su nombre indica, está involucrado en la eficiencia de la formación de los nódulos (Chun y Stacey, 1994).

Un gen anotado como regulador transcripcional de dos componentes de la familia NarL/FixJ se encuentra en la segunda isla genómica que se identificó (PSI2), otras dos copias anotadas con el mismo nombre se encuentran fuera de estas islas.

Fuera de las islas genómicas, se encuentra un cluster de los genes *fixGHIJKLMNOPQS* y un gen adicional *fixR*. FixJ y L son reguladores transcripcionales que responden al oxígeno, y trabajan con FixK, que forma parte de la superfamilia Crp-Fnr; ésta es una segunda vía de regulación de los genes *fix*

y *nif*. El operón *fixNOPQ* codifica para un complejo proteico citocromo oxidasa unido a la membrana que es inducido en condiciones microaerobias. Los genes *fixGHI* y *S*, están involucrados en la absorción y metabolismo del cobre, el cual forma parte del citocromo (Dixon y Kahn, 2004; Lindström y Mousavi, 2019; Fischer, 1994; Preisig et al., 1996).

Tanto la PSI1 como la PSI2 de *B. sp. CCGB01* presenta genes que codifican para bombas de eflujo resistentes a multifármacos (MRD) que se ha reportado que influyen en la competitividad de la nodulación y la fijación de nitrógeno. Estas bombas se pueden clasificar en cinco familias estructurales de proteínas (Álvarez-Ortega et al., 2013); de estas cinco familias, las islas genómicas de *B. sp. CCGB01* presentan genes relacionados a la superfamilia de las bombas de eflujo de resistencia/nodulación/división (RND) y a la superfamilia de casetes de unión a ATP (ABC). Se ha reportado que en *Sinorhizobium meliloti* y *Rhizobium etli*, proteínas de la superfamilia RND están involucradas en la competitividad de nodulación y en la fijación de nitrógeno (Eda et al., 2011; Gonzalez-Pasayo y Martinez-Romero, 2000; Lindemann et al., 2010). En el caso de las proteínas de la superfamilia ABC, se ha reportado que un transportador ABC de glicerol influye en la competitividad de la nodulación de *R. leguminosarum* (Ding et al., 2012).

PSI1 contiene un gen que codifica para la proteína TbrI de transferencia conjugativa; su presencia podría indicar que la isla es un elemento conjugativo e integrativo (ICE) o que es un elemento integrativo y movilizable (IME) (Bellanger et al., 2014).

Las tres islas cuentan con genes de la familia LuxR, los cuales están involucrados en el *quorum sensing*. Este mecanismo permite la regulación de genes en respuesta a la densidad poblacional (Krysciak et al., 2014). En el caso de los rizobios, se ha reportado que el *quorum sensing* se encuentra involucrado en la regulación de diferentes procesos como la transferencia de plásmidos e islas simbióticas, la síntesis de exopolisacáridos, motilidad, crecimiento, formación de biofilm, así como en la eficiencia de la nodulación (Calatrava-Morales et al., 2018; Lindemann et al., 2011).

Fuera de las islas genómicas se encuentra el gen *nodN*, cuya función no se ha descrito; sin embargo, se ha sugerido que en *Rhizobium meliloti* y *R. leguminosarum* bv. *viciae* comparte funciones con el gen *nodM*, el cual codifica para una glucosamina sintasa (Baev et al., 1992). También, fuera de las islas simbióticas se encuentran dos genes *fixJ*, cuyas funciones se describieron previamente.

De acuerdo a lo anterior, se puede concluir que en la isla simbiótica denominada PSI1 de *B. sp. CCGB01* se encuentran los genes esenciales para la síntesis de los NFs además de otros genes no esenciales que podrían aumentar su capacidad de nodulación y su afinidad por el hospedero, en este caso *Lysiloma* sp.

5.4.1 Identificación de sistemas conjugativos que podrían movilizar a las islas simbióticas

Se hizo una predicción bioinformática de los T4SSs del genoma de *B. sp. CCGB01* para identificar si las islas que contienen genes simbióticos podrían ser elementos genéticos móviles (MGEs). La PSI1, en la cual se encuentran todos los genes esenciales para la simbiosis, contiene genes que codifican para las proteínas TrwB y TrbI. La proteína TrwB es homóloga a la proteína acopladora VirD4, la cual se encarga de transferir el relaxosoma al canal de secreción (Llosa et al., 2002; Atmakuri et al., 2004). En el plásmido R388 se identificó la independencia estructural de TrwB del resto del T4SS (Larrea et al., 2013). Por otro lado, la proteína TrbI es homóloga a la proteína VirB10, la cual forma parte del canal del T4SS y es la proteína más grande de este canal (Jakubowski et al., 2009). Se ha observado que la ausencia de la proteína TrwE, otro homólogo de VirB10, no afecta la estabilidad del resto de los componentes del canal del T4SS, lo cual sugiere que se ensambla después de estos; además, la interacción entre las proteínas TrwB y VirB10 en diferentes sistemas se encuentra demostrada (Larrea et al., 2013). La presencia de estas proteínas en la PSI1 sugiere que podría tratarse de un ICE con un sistema de conjugación aún no descrito, de un IME, o de una GI movilizable degenerada.

La PSI1 podría ser movilizada por otros aparatos de conjugación codificados en otras regiones del genoma (Figuras 9 y 11), es decir que se clasificaría como un IME, ya que fuera de las islas genómicas, con el módulo CONJscan, se identificaron cuatro posibles T4SSs completos de tipo T. Se ha observado que los IMEs y plásmidos movilizables (Mob) pueden ser movilizados mediante diferentes estrategias. Los IMEs o Mobs pueden únicamente presentar el *oriT* y el gen que codifica para una relaxasa de plásmido movilizable y en ocasiones genes que codifican para factores accesorios del relaxosoma (RAFs) y así aprovechar el sistema de conjugación ajeno; además del *oriT* y la relaxasa, pueden contener una proteína acopladora (T4CP). Otras estrategias consisten en mimetizar el *oriT* para que el IME sea movilizado por relaxasas que no están codificadas en el MGE, es decir, son movilizadas *in trans*. En esta estrategia la relaxasa puede estar ausente y aprovechar relaxasas de dos tipos: una relaxasa de plásmido conjugativo o una relaxasa de tipo plásmido movilizable; esta clasificación se hizo con base en secuencia de aminoácidos y relaciones filogenéticas (Francia et al., 2004). Así mismo, en la movilización *in trans*, se puede encontrar un homólogo de la relaxasa de la que se aprovecha el MGE, por lo que en este último caso, se utilizan los RAFs codificados por el MGE y la relaxasa ajena a este. Otra estrategia consiste en que proteínas encargadas de iniciar el mecanismo de replicación del círculo rodante (RCR) actúen como relaxasas conjugativas (Ramsay y Firth, 2017). En estas estrategias, es necesario al menos un sitio *oriT*, por lo que además de identificar los T4SSs del genoma, se utilizó el programa *oriTfinder* para identificar los posibles orígenes de transcripción; sin embargo, no se identificó ninguno. Para hacer la búsqueda de los *oriT*, este programa utiliza la co-localización de las relaxasas homólogas, las cuales identifica con el programa HMMer y directamente hace búsquedas de similitud usando la base de datos *oriTDB-archived* (Li et al., 2018). Los sitios *oriT* se caracterizan por su diversidad y porque generalmente, en plásmidos conjugativos, se encuentran cercanos a los genes que codifican para el relaxosoma (Li et al., 2018; de la Cruz et al., 2010; Kiss et al., 2019)

Otra posibilidad es que PSI1 sea un elemento compuesto derivado de la acumulación en tándem (*'tandem accretion'*); este proceso sucede cuando un IME - o un CIME - se integra con un ICE y es movilizado por este en *cis* (Liu et al., 2019; Bellanger et al., 2014). Haskett et al. (2016), descubrieron un tipo de ICE tripartita en las bacterias del género *Mesorhizobium*, este ICE está conformado por partes separadas en el genoma. Recientemente, en *M. sp. AA22* se identificó el ICE, tripartita más grande hasta ahora (853, 775 bp) (Colombi et al., 2021). Este tipo peculiar de ICE, no es posible predecirlo con los programas bioinformáticos actuales. Otro caso particular en la detección de MGEs es la existencia de integrasas atípicas como la integrasa TcplM de *Clostridium perfringens*, que en la transferencia conjugativa del plásmido pCW3 toma la función de una relaxasa (Wisniewski et al., 2016). Debido a estos últimos casos particulares, también es posible que la PSI1 de *B. sp. CCGB01* sea una isla simbiótica movilizable cuyos MGEs no se puedan predecir aún con métodos bioinformáticos.

PSI2 no codifica para ninguna proteína del T4SS, pero sí codifica para una integrasa, por lo que PSI1, podría ser movilizada en *trans* o en *cis*. De esta manera sería posible que sea un IME o una GI movilizable degenerada.

Por último, a comparación de PSI1 y PSI2, dentro de la PSI3 se encuentra un cluster que codifica para proteínas del T4SS: *trb(I,G,F,L,J,E,D,C,B,N)*, *virD4*, *virD2* y *traF*. Este cluster se identificó con los programas CONJscan y oriTfinder; sin embargo, el programa ICEfinder no hizo predicciones de ningún ICE en esta región, sino únicamente un probable ICE que no contiene genes simbióticos y que por lo tanto no se discute en este trabajo. Es más probable que PSI3 sea un ICE cuyos sitios de inserción *att* y *oriT* aún no pueden ser predichos bioinformáticamente o bien, podría tratarse de una GI movilizable degenerada, que ya no es capaz de moverse o ser movilizadas.

6. CONCLUSIONES

De acuerdo con los análisis filogenéticos y de identidad de nucleótidos que se presentan en este trabajo, *B. sp. CCGB01* representa una nueva especie del

género *Bradyrhizobium* cercana a la cepa *B. japonicum* J5 (91.2% ANI con el genoma de *B. japonicum* J5).

Los genes que le confieren plasticidad genómica a *B. sp. CCGB01* para que logre la nodulación de su hospedero *Lysiloma* sp., se encuentran en su mayoría en una isla simbiótica, denominada PSI1. Además, en esta isla se encuentran genes que podrían aumentar la especificidad simbiótica de *B. sp. CCGB01* por *Lysiloma* sp. Los ensayos de nodulación de *B. sp. CCGB01* en *L. leucocephala*, *P. vulgaris* y en su hospedero, también sugieren que esta cepa es un simbiote específico de *Lysiloma* sp.

Las filogenias de los genes *nodC* y *nifH* no concuerdan con el análisis filogenómico ni con las filogenias de los genes *housekeeping* concatenados *glnII* y *recA*, lo cual también indica que PSI1 se puede transferir entre distintas bacterias. De acuerdo a las especies hasta ahora descritas, posiblemente *B. ingae* sea la especie más cercana a *B. sp. CCGB01* en cuanto a sus características simbióticas.

PSI1 podría ser una isla genómica movilizable, ya que de acuerdo a las predicciones bioinformáticas realizadas en este trabajo, el genoma de *B. sp. CCGB01* cuenta con cuatro sistemas de conjugación de secreción tipo IV y de acuerdo a la revisión en la literatura, podría contar con otros sistemas de conjugación que aún no son posibles de predecir de manera bioinformática o que aún no se han descrito.

7. PERSPECTIVAS

- Demostrar y evaluar la transferencia de la isla simbiótica PSI1 a cepas no simbióticas mediante experimentos de conjugación en el laboratorio.
- En caso de que PSI1 sea movilizable, definir qué tipo de MGE es de manera experimental e identificar los elementos requeridos para su movilización.
- Realizar el alineamiento múltiple y un análisis de motivos de las proteínas contenidas en todas las islas simbióticas del género *Bradyrhizobium* incluyendo las probables islas simbióticas de *B. sp. CCGB01*, así como un alineamiento de los T4SS.

8. REFERENCIAS BIBLIOGRÁFICAS

1. Abby S., Néron B., Ménager, H., Touchon, M. & Rocha, E. (2014). MacSyFinder: A Program to Mine Genomes for Molecular Systems with an Application to CRISPR-Cas Systems. In PLoS ONE, 9 (10), pp. e110726. doi:10.1371/journal.pone.0110726
2. Abby, S., Cury, J., Guglielmini, J., Néron, B., Touchon, M. & Rocha, E. (2016). Identification of protein secretion systems in bacterial genomes. In Scientific Reports, 6, pp. 23080. doi:10.1038/srep23080
3. Addo, M. & Dos Santos P. (2020). Distribution of Nitrogen Fixation Genes in Prokaryotes Containing Alternative Nitrogenases. ChemBioChem. doi:10.1002/cbic.20200002
4. Araújo, J., Flores-Félix, J. D., Igual, J. M., Peix, A., González-Andrés, F., Díaz-Alcántara, C. A., & Velázquez, E. (2017). Bradyrhizobium cajani sp. nov. isolated from nodules of Cajanus cajan. International journal of systematic and evolutionary microbiology, 67(7), 2236–2241. <https://doi.org/10.1099/ijsem.0.001932>
5. Aserse, A., Woyke, T., Kyrpides, N., Whitman, W., & Lindström, K. (2017). Draft genome sequences of Bradyrhizobium shewense sp. nov. ERR11T and Bradyrhizobium yuanmingense CCBAU 10071T. Standards in genomic sciences, 12, 74. <https://doi.org/10.1186/s40793-017-0283-x>
6. Atmakuri, K., Cascales, E. y Christie P. (2004). Energetic components VirD4, VirB11 and VirB4 mediate early DNA transfer reactions required for bacterial type IV secretion. Mol. Microbiol. 54, 1199–1211.
7. Aziz, R., Bartels, D., Best, A., DeJongh, M., et al. (2008) The RAST Server: Rapid Annotations using Subsystems Technology. BMC Genomics
8. Baev, N., Schultze, M., Barlier, I., Ha, D. C., Virelizier, H., Kondorosi, E., y Kondorosi, A. (1992). Rhizobium nodM and nodN genes are common nod genes: nodM encodes functions for efficiency of nod signal production and bacteroid maturation. Journal of bacteriology, 174(23), 7555–7565. <https://doi.org/10.1128/jb.174.23.7555-7565.1992>
9. Bauer, E., Kaspar, T., Fischer, H. M., & Hennecke, H. (1998). Expression of the fixR-nifA operon in Bradyrhizobium japonicum depends on a new response regulator, RegR. Journal of bacteriology, 180(15), 3853–3863. <https://doi.org/10.1128/JB.180.15.3853-3863.1998>
10. Bellanger, X., Payot, S., Leblond-Bourget, N. y Guedon, G. (2014) Conjugative and mobilizable genomic islands in bacteria: evolution and diversity. FEMS Microbiol Rev 38: 720-760
11. Berck, S., Perret, X., Quesada-Vincens, D., Promé, J. C., Broughton, W. J., & Jabbouri, S. (1999). NoIL of Rhizobium sp. strain NGR234 is required for O-acetyltransferase activity. J. Bacteriol, 181, 957-964
12. Bertelli, C., Laird, M. R., Williams, K. P., Simon Fraser University Research Computing Group, Lau, B. Y., Hoad, G., Winsor, G. L., & Brinkman, F. (2017). IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. Nucleic acids research, 45(W1), W30–W35. <https://doi.org/10.1093/nar/gkx343>
13. Bertelli, C., Tilley, K. E., & Brinkman, F. (2019). Microbial genomic island discovery, visualization and analysis. Briefings in bioinformatics, 20(5), 1685–1698. <https://doi.org/10.1093/bib/bby042>
14. Brettin, T., Davis, J., Disz, T., Edwards, R. & Gerdes S. (2015). RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep.
15. Broughton, W., Jabbouri S. & Perret X. (2000) Keys to symbiotic harmony. J Bacteriol., 182, 5641–5652.

16. Bromfield, E., & Cloutier, S. (2021). *Bradyrhizobium septentrionale* sp. nov. (sv. septentrionale) and *Bradyrhizobium quebecense* sp. nov. (sv. septentrionale) associated with legumes native to Canada possess rearranged symbiosis genes and numerous insertion sequences. *International journal of systematic and evolutionary microbiology*, 71(6), 004831. <https://doi.org/10.1099/ijsem.0.004831>
17. Boyd, E., Almagro-Moreno, S. & Parent, M. (2009). Genomic islands are dynamic, ancient integrative elements in bacterial evolution. *Trends in microbiology*, 17, 47-53. [10.1016/j.tim.2008.11.003](https://doi.org/10.1016/j.tim.2008.11.003).
18. Calatrava-Morales, N., McIntosh, M., & Soto, M. J. (2018). Regulation Mediated by N-Acyl Homoserine Lactone Quorum Sensing Signals in the Rhizobium-Legume Symbiosis. *Genes*, 9(5), 263. <https://doi.org/10.3390/genes9050263>
19. Cantalapiedra, C., Hernández-Plaza, A., Letunic, I., Bork, P. y Huerta-Cepas, J. (2021) eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Molecular Biology and Evolution*, (38)12, 5825–5829, <https://doi.org/10.1093/molbev/msab293>
20. Carver T, Thomson N, Bleasby A, Berriman M and Parkhill J (2009) DNAPlotter: circular and linear interactive genome visualization *Bioinformatics (Oxford, England)* 25;1;119-20 DOI: 10.1093/bioinformatics/btn578.
21. Chun, J., & Stacey, G. (1994). A *Bradyrhizobium japonicum* gene essential for nodulation competitiveness is differentially regulated from two promoters. *Molecular plant-microbe interactions : MPMI*, 7(2), 248–255. <https://doi.org/10.1094/mpmi-7-0248>
22. Colombi, E., Perry, B. J., Sullivan, J. T., Bekuma, A. A., Terpolilli, J. J., Ronson, C. W., & Ramsay, J. P. (2021). Comparative analysis of integrative and conjugative mobile genetic elements in the genus *Mesorhizobium*. *Microbial genomics*, 7(10), 000657. <https://doi.org/10.1099/mgen.0.000657>
23. Corvera, A., Promé, D., Promé, J. C., Martínez-Romero, E., and Romero, D. (1999). The *nolL* gene from *Rhizobium etli* determines nodulation efficiency by mediating the acetylation of the fucosyl residue in the nodulation factor. *Mol. Plant-Microbe Interact*, 13, 236-246
24. Cury, J., Abby, S., Doppelt-Azeroual, O., Néron, B. & Rocha, E. (2020). Identifying Conjugative Plasmids and Integrative Conjugative Elements with CONJscan. *Methods Mol Biol*. 2075:265-283. doi: 10.1007/978-1-4939-9877-7_19. PMID: 31584169.
25. da Silva, K., De Meyer, S. E., Rouws, L., Farias, E., Dos Santos, M., O'Hara, G., Ardley, J. K., Willems, A., Pitard, R. M., & Zilli, J. E. (2014). *Bradyrhizobium ingae* sp. nov., isolated from effective nodules of *Inga laurina* grown in Cerrado soil. *International journal of systematic and evolutionary microbiology*, 64(Pt 10), 3395–3401. <https://doi.org/10.1099/ijse.0.063727-0>
26. da Silva Filho, A., Raittz, R., Guizelini, D., De Pierri, C., Augusto, D., Dos Santos-Weiss, I., & Marchaukoski, J. (2018). Comparative Analysis of Genomic Island Prediction Tools. *Frontiers in genetics*, 9, 619. <https://doi.org/10.3389/fgene.2018.00619>
27. Davis J., Wattam A., Aziz R. et al. (2020). The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. *Nucleic Acids Res*. 8;48(D1):D606-D612. doi: 10.1093/nar/gkz943.
28. de Bruijn F. J. (1992). Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergeneric consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria. *Applied and environmental microbiology*, 58(7), 2180–2187. <https://doi.org/10.1128/aem.58.7.2180-2187.1992>

29. de la Cruz, F., Frost, L., Meyer, R., & Zechner, E. (2010) Conjugative DNA metabolism in Gram-negative bacteria. *FEMS Microbiol Rev* 34: 18–40.
30. Delavat, F., Miyazaki, R., Carraro, N., Pradervand, N. & Meer, J. (2017) The hidden life of integrative and conjugative elements, *FEMS Microbiology Reviews*, (41)4: 512–537, <https://doi.org/10.1093/femsre/fux008>
31. Dixon, R., & Kahn, D. (2004). Genetic regulation of biological nitrogen fixation. *Nature reviews. Microbiology*, 2(8), 621–631. <https://doi.org/10.1038/nrmicro954>
32. Dobrindt, U., Hochhut, B. & Hentschel, U. (2004) Genomic islands in pathogenic and environmental microorganisms. *Nat Rev Microbiol* 2(5):414-24. doi: 10.1038/nrmicro884. PMID: 15100694.
33. Dockendorff, T., Sharma, A., & Stacey, G. (1994). Identification and characterization of the nolYZ genes of *Bradyrhizobium japonicum*. *Molecular plant-microbe interactions : MPMI*, 7(2), 173–180. <https://doi.org/10.1094/mpmi-7-0173>
34. Durán, D., Rey, L., Mayo, J., Zúñiga-Dávila, D., Imperial, J., Ruiz-Argüeso, T., Martínez-Romero, E., & Ormeño-Orrillo, E. (2014). *Bradyrhizobium paxllaeri* sp. nov. and *Bradyrhizobium icense* sp. nov., nitrogen-fixing rhizobial symbionts of Lima bean (*Phaseolus lunatus* L.) in Peru. *International journal of systematic and evolutionary microbiology*, 64(Pt 6), 2072–2078. <https://doi.org/10.1099/ijs.0.060426-0>
35. Edgar, R. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5, 113. <https://doi.org/10.1186/1471-2105-5-113>
36. Eda, K. (2011). Involvement of the smeAB multidrug efflux pump in resistance to plantant microbials and contribution to nodulation competitiveness in *Sinorhizobium meliloti*. *Appl. Envi-ron. Microbiol.* 77, 2855–2862.
37. Emms, D. & Kelly, S. (2019) OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* 20, 238. <https://doi.org/10.1186/s13059-019-1832-y>
38. Emmerich, R., Hennecke, H., & Fischer, H. M. (2000). Evidence for a functional similarity between the two-component regulatory systems RegSR, ActSR, and RegBA (PrrBA) in alpha-Proteobacteria. *Archives of microbiology*, 174(5), 307–313. <https://doi.org/10.1007/s002030000207>
39. Fahraeus, G. (1957). The infection of clover root hairs by nodule bacteria studied by a simple glass slide technique. *Journal of general microbiology*, 16(2), 374–381. <https://doi.org/10.1099/00221287-16-2-374>
40. Fernández-López, M., D'Haese, W., Mergaert, P., Verplancke, C., Promé, J., Van Montagu, M. & Holsters, M. (1996). Role of nodI and nodJ in lipo-chitooligosaccharide secretion in *Azorhizobium caulinodans* and *Escherichia coli*. *Mol Microbiol.* 20(5):993-1000. doi: 10.1111/j.1365-2958.1996.tb02540.x. PMID: 8809752.
41. Fischer H. M. (1994). Genetic regulation of nitrogen fixation in rhizobia. *Microbiological reviews*, 58(3), 352–386. <https://doi.org/10.1128/mr.58.3.352-386.1994>
42. Fonseca, M., Peix, A., de Faria, S., Mateos, P., Rivera, L., Simões-Araujo, J., et al. (2012). Nodulation in *Dimorphandra wilsonii* Rizz. (Caesalpinioideae), a threatened species native to the Brazilian Cerrado. *PloS One*, 7(11), e49520. <https://doi.org/10.1371/journal.pone.0049520>
43. Francia, M., Varsaki, A., Garcillan-Barcia, M., Latorre, A., Drainas, C. & de la Cruz F. (2004) A classification scheme for mobilization regions of bacterial plasmids. *FEMS Microbiol. Rev.* 28:79-100.
44. Ghyselincq, J., Pfeiffer, S., Heylen, K., Sessitsch, A., & De Vos, P. (2013). The effect of primer choice and short read sequences on the outcome of 16S rRNA gene based diversity studies. *PloS one*, 8(8), e71360. <https://doi.org/10.1371/journal.pone.0071360>

45. Gonzalez-Pasayo, R. & Martinez-Romero, E. (2000). Multiresistance genes of *Rhizobium* metli CFN42. *Mol. Plant Microbe Interact.* 13, 572–577
46. Hall, T. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series.* 41, 95-98.
47. Haskett, T., Terpolilli, J., Bekuma, A., O'Hara, G., Sullivan, J., Wang, P., Ronson, C. & Ramsay, J. (2016) Assembly and transfer of tripartite integrative and conjugative genetic elements. *PNAS*, 113, 12268–12273.
48. Hofacker, I., Fekete, M. & Stadler, P. (2002) Secondary structure prediction for aligned RNA sequences. *J Mol Biol.* 2002 vol. 319 (18) pp.3724-32
49. Hossain, M. & Lundquist, P. (2016) Nodule inhabiting non-rhizobial bacteria and their influence in growth of selected plants in Bangladesh. *BioResearch Communications* 2:134–138
50. Jabbouri, S., Relić, B., Hanin, M., Kamalaprija, P., Burger, U., Promé, D., Promé, J. & Broughton, W. (1998) *noI* and *noE* (HsnIII) of *Rhizobium* sp. NGR234 are involved in 3-O-carbamoylation and 2-O-methylation of Nod factors. *J Biol Chem.* 15;273(20):12047-55. doi: 10.1074/jbc.273.20.12047.
51. Jabbouri, S., Fellay, R., Talmont, F., Kamalaprija, P., Burger, U., Relić, B., Promé, J. C., & Broughton, W. J. (1995). Involvement of *nodS* in N-methylation and *nodU* in 6-O-carbamoylation of *Rhizobium* sp. NGR234 nod factors. *The Journal of biological chemistry*, 270(39), 22968–22973. <https://doi.org/10.1074/jbc.270.39.22968>
52. Jakubowski, S., Kerr, J., Garza, I., Krishnamoorthy, V., Bayliss, R., Waksman, G., & Christie, P. (2009). *Agrobacterium* VirB10 domain requirements for type IV secretion and T pilus biogenesis. *Molecular microbiology*, 71(3), 779–794. <https://doi.org/10.1111/j.1365-2958.2008.06565.x>
53. Jimenez-Vicente, E., Yang, Z. Y., Martin Del Campo, J., Cash, V., Seefeldt, L., & Dean, D. R. (2019). The *NifZ* accessory protein has an equivalent function in maturation of both nitrogenase MoFe protein P-clusters. *The Journal of biological chemistry*, 294(16), 6204–6213. <https://doi.org/10.1074/jbc.RA119.007905>
54. Juhas, M., van der Meer, J., Gaillard, M., Harding, R., Hood, D. & Crook, D. (2009). Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS microbiology reviews*, 33(2), 376–393. <https://doi.org/10.1111/j.1574-6976.2008.00136>.
55. Nguyen, L., Schmidt, H., von Haeseler, A., & Minh B. (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol. Biol. Evol.*, 32:268-274
56. Langille, M., Hsiao, W. & Brinkman F. (2010) Detecting genomic islands using bioinformatics approaches. *Nat Rev Microbiol.* 8(5):373-82. doi: 10.1038/nrmicro2350. PMID: 20395967.
57. Lindemann, A., Pessi, G., Schaefer, A. L., Mattmann, M., Christensen, Q., Kessler, A., Hennecke, H., Blackwell, H., Greenberg, E., & Harwood, C. (2011). Isovaleryl-homoserine lactone, an unusual branched-chain quorum-sensing signal from the soybean symbiont *Bradyrhizobium japonicum*. *Proceedings of the National Academy of Sciences of the United States of America*, 108(40), 16765–16770. <https://doi.org/10.1073/pnas.1114125108>
58. Jaiswal, S. & Dakora, F. (2019). Widespread distribution of highly adapted *Bradyrhizobium* species nodulating diverse legumes in Africa, *Front. Microbiol.*, 10, 310
59. Kiss, J., Szabó, M., Hegyi, A., Douard, G., Praud, K., Nagy, I., Olasz, F., Cloeckaert, A., & Doublet, B. (2019). Identification and Characterization of *oriT* and Two Mobilization Genes Required for Conjugative Transfer of *Salmonella* Genomic Island 1. *Frontiers in microbiology*, 10, 457. <https://doi.org/10.3389/fmicb.2019.00457>

60. Krysciak, D., Grote, J., Rodriguez Orbegoso, M., Utpatel, C., Förstner, K. U., Li, L., Schmeisser, C., Krishnan, H. B., & Streit, W. R. (2014). RNA sequencing analysis of the broad-host-range strain *Sinorhizobium fredii* NGR234 identifies a large set of genes linked to quorum sensing-dependent regulation in the background of a *tral* and *ngl* deletion mutant. *Applied and environmental microbiology*, 80(18), 5655–5671. <https://doi.org/10.1128/AEM.01835-14>
61. Larrea, D., de Paz, H., Arechaga, I., de la Cruz, F. & Llosa, M. (2013) Structural independence of conjugative coupling protein TrwB from its Type IV secretion machinery. *Plasmid*,70(1), 146-53. doi: 10.1016/j.plasmid.2013.03.006.
62. Lee C., Babic A., Grossman A. (2010). Autonomous plasmid-like replication of a conjugative transposon. *Mol Microbiol* 75: 268–79.
63. Letunic I and Bork P (2021) *Nucleic Acids Res* doi: 10.1093/nar/gkab301 Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation (full article)
64. Li X., Xie Y., Liu M., Tai M., Sun J., Deng Z. y Ou H. (2018) oriTfinder: a web-based tool for the identification of origin of transfers in DNA sequences of bacterial mobile genetic elements. *Nucleic Acids Research*, 46(W1):W229–W234. DOI: 10.1093/nar/gky352
65. Lindemann, A., Koch, M., Pessi, G., Muller, A., Balsiger, S., Hennecke, H. et al. (2010). Host-specific symbiotic requirement of BdeAB, a RegR controlled RND-type efflux system in *Bradyrhizobium japonicum*. *FEMS Microbiol. Lett.* 312, 184–191.
66. Lindström, K. & Mousavi, S. (2019). Effectiveness of nitrogen fixation in rhizobia. *Microbial Biotechnology*. 13. 10.1111/1751-7915.13517
67. Liu, M., Li, X., Xie, Y., Bi, D., Sun, J., Li, J., Tai, C., Deng Z., y Ou H. (2019) ICEberg 2.0: an updated database of bacterial integrative and conjugative elements. *Nucleic Acids Research*, DOI: 10.1093/nar/gky1123.
68. Llosa M., Gomis-Ruth F., Coll M. et al. (2002) Bacterial conjugation: a two-step mechanism for DNA transport. *Mol Microbiol* 45:1–8.
69. Long, S. R. (1996). *Rhizobium Symbiosis: Nod Factors in Perspective*. *THE PLANT CELL ONLINE*, 8(10), 1885–1898. doi:10.1105/tpc.8.10.1885
70. López-Lara, I. M., Kafetzopoulos, D., Spaink, H. P., & Thomas-Oates, J. E. (2001). Rhizobial NodL O-acetyl transferase and NodS N-methyl transferase functionally interfere in production of modified Nod factors. *Journal of bacteriology*, 183(11), 3408–3416. <https://doi.org/10.1128/JB.183.11.3408-3416.2001>
71. López-López, A., Negrete-Yankelevich, S., Rogel, M., Ormeño-Orrillo, E., Martínez, J. & Martínez-Romero, E. (2013). Native bradyrhizobia from Los Tuxtlas in Mexico are symbionts of *Phaseolus lunatus* (lima bean). *Syst Appl Microbiol* 36:33–38. doi:10.1016/j.syapm.2012.10.006
72. Martínez-Romero, E. (2009). Coevolution in *Rhizobium*-legume symbiosis? *DNA Cell Biol* 28: 361-370. *DNA and cell biology*. 28. 361-70. 10.1089/dna.2009.0863.
73. Martínez-Romero, E., & Rosenblueth, M. (1990). Increased Bean (*Phaseolus vulgaris* L.) Nodulation Competitiveness of Genetically Modified *Rhizobium* Strains. *Applied and environmental microbiology*, 56(8), 2384–2388. <https://doi.org/10.1128/aem.56.8.2384-2388.1990>
74. Martens M, Dawyndt P, Coopman R, Gillis M, De Vos P, Willems A. Advantages of multilocus sequence analysis for taxonomic studies: a case study using 10 housekeeping genes in the genus *Ensifer*. *Int J Syst Evol Microbiol*. 2008;58:200–14
75. McLeod S., Burrus V. y Waldor M. (2006) Requirement for *Vibrio cholerae* integration host factor in conjugative DNA transfer. *J Bacteriol* 188:5704–11.

76. Minh, B. Q., Nguyen, M. A., & von Haeseler, A. (2013). Ultrafast approximation for phylogenetic bootstrap. *Molecular biology and evolution*, 30(5), 1188–1195. <https://doi.org/10.1093/molbev/mst024>
77. Nouwen N, Arrighi J., Cartieaux F., Chaintreuil C., Gully D., Klopp C., et al.(2017) The role of rhizobial (NifV) and plant (FEN1) homocitrate synthases in *Aeschynomene*/photosynthetic *Bradyrhizobium* symbiosis. *Sci Rep.* 7:1–10.
78. Ormeño-Orrillo, E., Martínez-Romero, E., y Zúñiga-Dávila, D. (2020). Identification of the symbiosis island of *Bradyrhizobium paxllaeri* LMTR 21T. *Brazilian journal of microbiology: [publication of the Brazilian Society for Microbiology]*, 51(2), 527–529. <https://doi.org/10.1007/s42770-019-00164-5>
79. Ormeño-Orrillo, E., Zúñiga, Dávila, D., Martínez-Romero, E. (2015). Biodiversity of nitrogen-fixing nodule bacteria associated with lima bean (*Phaseolus lunatus* L.) in its domestication centers, p 91–102. In Araújo ASF, Lopes ÁCA, Gomes RLF (ed), *Phaseolus lunatus: diversity, growth and production* Nova Science Publishers, New York, NY.
80. Osorio, C. R., Marrero, J., Wozniak, R. A., Lemos, M. L., Burrus, V., y Waldor, M. K. (2008). Genomic and functional analysis of ICEPdaSpa1, a fish-pathogen-derived SXT-related integrating conjugative element that can mobilize a virulence plasmid. *Journal of bacteriology*, 190(9), 3353–3361. <https://doi.org/10.1128/JB.00109-08>
81. Overbeek, R., Olson, R., Pusch, G., Olsen G., et al. (2014) The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.*
82. Parker, M. (2015).The Spread of *Bradyrhizobium* Lineages Across Host Legume Clades: from *Abarema* to *Zygia*. *Microb Ecol* 69, 630–640. <https://doi.org/10.1007/s00248-014-0503-5>
83. Parte, A. (2018). LPSN — List of Prokaryotic names with Standing in Nomenclature (bacterio.net), 20 years on. *International Journal of Systematic and Evolutionary Microbiology*, 68, 1825-1829; DOI: 10.1099/ijsem.0.002786
84. Paudel, D., Liu, F., Wang, L., Crook, M., Maya, S., Peng, Z., Kelley, K., Ané, J. M., & Wang, J. (2020). Isolation, Characterization, and Complete Genome Sequence of a *Bradyrhizobium* Strain Lb8 From Nodules of Peanut Utilizing Crack Entry Infection. *Frontiers in microbiology*, 11, 93. <https://doi.org/10.3389/fmicb.2020.00093>
85. Pavlovic, G., Burrus, V., Gintz, B., Decaris, B., & Guédon, G. (2004). Evolution of genomic islands by deletion and tandem accretion by site-specific recombination: ICES_{t1}-related elements from *Streptococcus thermophilus*. *Microbiology (Reading, England)*, 150(Pt 4), 759–774. <https://doi.org/10.1099/mic.0.26883-0>
86. Peix, A., Ramírez-Bahena, M. H., Flores-Félix, J. D., Alonso de la Vega, P., Rivas, R., Mateos, P. F., Igual, J. M., Martínez-Molina, E., Trujillo, M. E., & Velázquez, E. (2015). Revision of the taxonomic status of the species *Rhizobium lupini* and reclassification as *Bradyrhizobium lupini* comb. nov *International journal of systematic and evolutionary microbiology*, 65(Pt 4), 1213–1219. <https://doi.org/10.1099/ijse.0.000082>
87. Perret, X., Staehelin, C., Broughton, W. (2000) Molecular basis of symbiotic promiscuity. *Microbiol Mol Biol Rev*, 64, 180–201.
88. Preisig, O., Zufferey, R., & Hennecke, H. (1996). The *Bradyrhizobium japonicum* fixGHIS genes are required for the formation of the high-affinity cbb3-type cytochrome oxidase. *Archives of microbiology*, 165(5), 297–305. <https://doi.org/10.1007/s002030050330>
89. Pritchard, L., Glover, R., Humphris, S., Elphinstone, J. & Toth, I. (2016) Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens, *Anal. Methods*, 8 , 12 —24

90. Quesada-Vincens, D., Fellay, R., Nasim, T., Viprey, V., Burger, U., Prome, J. C., Broughton, W. J., & Jabbouri, S. (1997). Rhizobium sp. strain NGR234 NodZ protein is a fucosyltransferase. *Journal of bacteriology*, 179(16), 5087–5093. <https://doi.org/10.1128/jb.179.16.5087-5093.1997>
91. Ramsay, J., Firth, N. (2017). Diverse mobilization strategies facilitate transfer of non-conjugative mobile genetic elements. *Curr Opin Microbiol*, 38,1-9. doi: 10.1016/j.mib.2017.03.003.
92. Roche, P., Maillat, F., Plazanet, C., Debelle, F., Ferro, M., Truchet, G., Promé, J. C., & Dénarié, J. (1996). The common nodABC genes of Rhizobium meliloti are host-range determinants. *Proceedings of the National Academy of Sciences of the United States of America*, 93(26), 15305–15310. <https://doi.org/10.1073/pnas.93.26.15305>
93. Rogel, M. A., Hernández-Lucas, I., Kuykendall, L. D., Balkwill, D. L., & Martínez-Romero, E. (2001). Nitrogen-fixing nodules with Ensifer adhaerens harboring Rhizobium tropici symbiotic plasmids. *Applied and environmental microbiology*, 67(7), 3264–3268. <https://doi.org/10.1128/AEM.67.7.3264-3268.2001>
94. Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., & Huelsenbeck, J. P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic biology*, 61(3), 539–542. <https://doi.org/10.1093/sysbio/sys029>
95. Richter, M., & Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences of the United States of America*, 106(45), 19126–19131. <https://doi.org/10.1073/pnas.0906412106>
96. Rubio LM, Ludden PW.(2008) Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *Annu Rev Microbiol*. 62:93-111. doi: 10.1146/annurev.micro.62.081307.162737. PMID: 18429691.
97. Shvarev, D.; Nishi, C. N.; Maldener, I. Two DevBCA-like ABC transporters are involved in the multidrug resistance of the cyanobacterium Anabaena sp. PCC 7120. *FEBS Lett*. 2019, 593, 1818– 1826, DOI: 10.1002/1873-3468.13450
98. Sudhir Kumar, Glen Stecher, Michael Li, Christina Knyaz, and Koichiro Tamura (2018) MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* 35:1547-1549
99. Tao, J., Wang, S., Liao, T. et al. (2021). Evolutionary origin and ecological implication of a unique nif island in free-living Bradyrhizobium lineages. *ISME J* <https://doi.org/10.1038/s41396-021-01002-z>
100. Toledo, I., Lloret, L., & Martínez-Romero, E. (2003). Sinorhizobium americanus sp. nov., a new Sinorhizobium species nodulating native Acacia spp. in Mexico. *Systematic and applied microbiology*, 26(1), 54–64. <https://doi.org/10.1078/072320203322337317>
101. Varghese, J., Mukherjee, S., Ivanova, N., Konstantinidis, K., Mavrommatis, K., Kyrpides, N. & Pati, A. (2015) Microbial species delineation using whole genome sequences. *Nucleic Acids Res*. 43(14):6761–71.
102. Versalovic, J., Koeuth, T. & Lupski, R. (1991). Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Research*, 19(24), 6823–6831. doi:10.1093/nar/19.24.6823
103. Vincent, J. (1970). *A Manual for the Practical Study of Root Nodule Bacteria*, Blackwell Scientific Publications, Oxford.
104. Vinuesa, P., Silva, C., Werner, D., & Martínez-Romero, E. (2005). Population genetics and phylogenetic inference in bacterial molecular systematics: the roles of migration and

- recombination in Bradyrhizobium species cohesion and delineation. *Molecular phylogenetics and evolution*, 34(1), 29–54. <https://doi.org/10.1016/j.ympev.2004.08.020>
105. Wang J, Li Q, Xu S, Zhao W, Lei Y, Song C et al (2018) Traits-based integration of multi-species inoculants facilitates shifts of indigenous soil bacterial community. *Front Microbiol* 9:1–13. <https://doi.org/10.3389/fmicb.2018.01692>
106. Wang E., Martínez- Romero J. y Martínez-Romero E. (1999). Genetic diversity of rhizobia from *Leucaena Leucocephala* nodules in Mexican soils. *Molecular Ecology*. 8. 711 - 724. [10.1046/j.1365-294X.1999.00608.x](https://doi.org/10.1046/j.1365-294X.1999.00608.x).
107. Wardell, G., Hynes, M., Young, P., & Harrison, E. (2022). Why are rhizobial symbiosis genes mobile?. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 377(1842), 20200471. <https://doi.org/10.1098/rstb.2020.0471>
108. Wick, R., Judd, L., Gorrie, C. & Holt, K. (2017) Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13(6): e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>
109. Weisburg, W., Barns, S., Pelletier, D. & Lane D. (1991) 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* 173:697–703
110. Wernersson, R. & Pedersen, A. (2003). RevTrans - Constructing alignments of coding DNA from aligned amino acid sequences. *Nucl. Acids Res.*, 2003, 31(13), 3537-3539.
111. Wisniewski, J., Traore, D., Bannam, T., Lyras, D., Whisstock, J., & Rood, J. I. (2016). TcpM: a novel relaxase that mediates transfer of large conjugative plasmids from *Clostridium perfringens*. *Molecular microbiology*, 99(5), 884–896. <https://doi.org/10.1111/mmi.13270>
112. Yao, Y., Sui, X., Zhang, X., Wang, E. & Chen W.(2015). *Bradyrhizobium erythrophlei* sp. nov. and *Bradyrhizobium ferriligni* sp. nov., isolated from effective nodules of *Erythrophleum fordii*. *Int J Syst Evol Microbiol* ;65(Pt 6):1831-1837. doi: 10.1099/ijs.0.000183. PMID: 25754551.
113. Zahran, H. (2017). Plasmids impact on rhizobia-legumes symbiosis in diverse environments. *Symbiosis* 73, 75–91. <https://doi.org/10.1007/s13199-017-0476-5>
114. Zhang, R., Ou, H. Y., Gao, F., & Luo, H. (2014). Identification of Horizontally-transferred Genomic Islands and Genome Segmentation Points by Using the GC Profile Method. *Current genomics*, 15(2), 113–121. <https://doi.org/10.2174/1389202915999140328163125>
115. Zulkower V & Rosser S. (2020) DNA Features Viewer: a sequence annotation formatting and plotting library for Python. *Bioinformatics*. 36:4350–2.

9. MATERIAL SUPLEMENTARIO

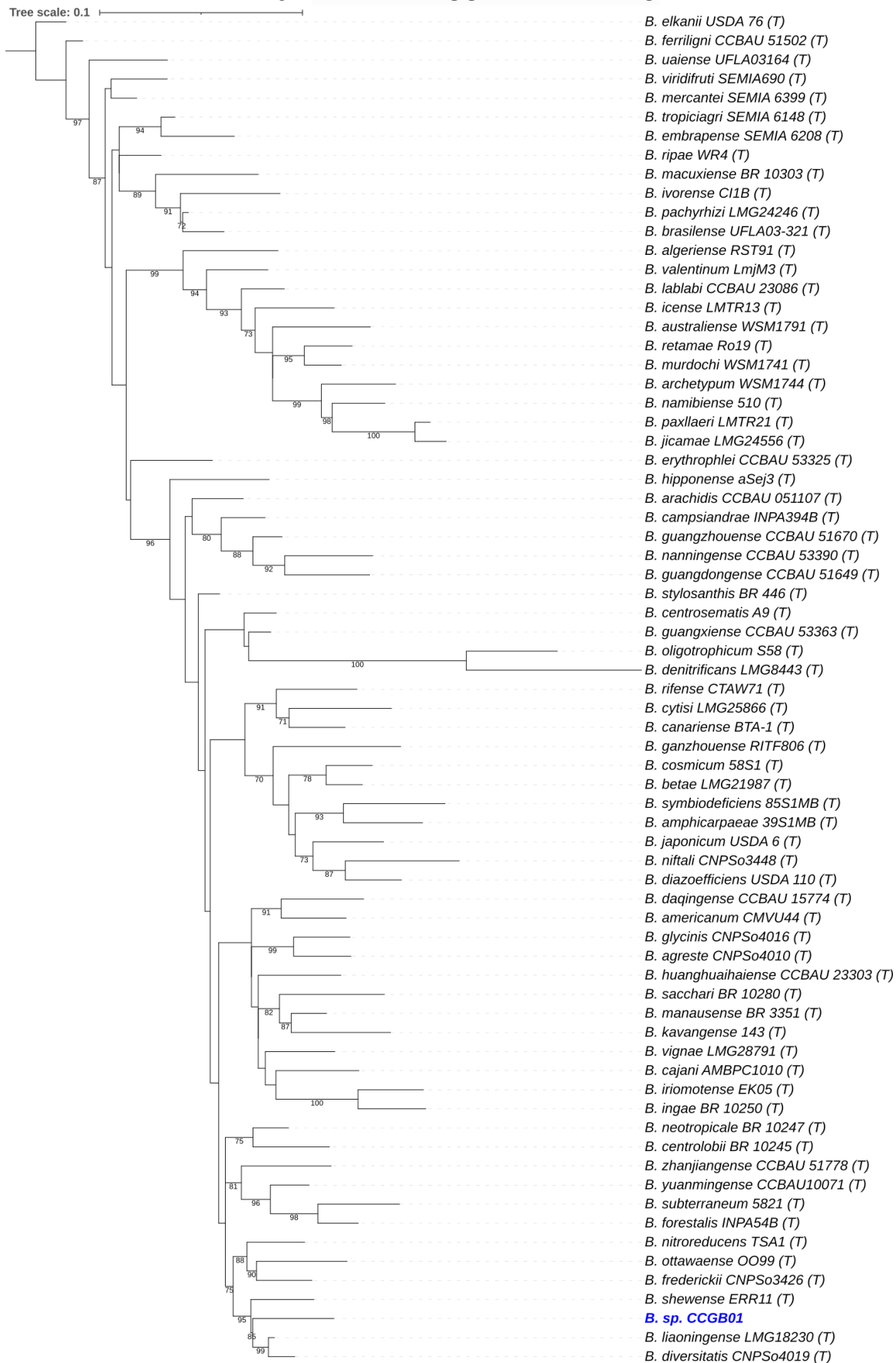


Figura S1. Árbol filogenético de máxima verosimilitud del gen *recA* de *Bradyrhizobium* sp. CCGB01 y cepas tipo del género sin enraizar con un modelo de sustitución GTR+I+G. La longitud del alineamiento fue de 375 pb. Se muestran los valores de UFBoot (≥ 70) de 1000 pseudo réplicas; el soporte representa directamente la probabilidad de que el clado sea confiable y a partir de valores ≥ 95 los clados comienzan a ser confiables (Minh et al., 2013).

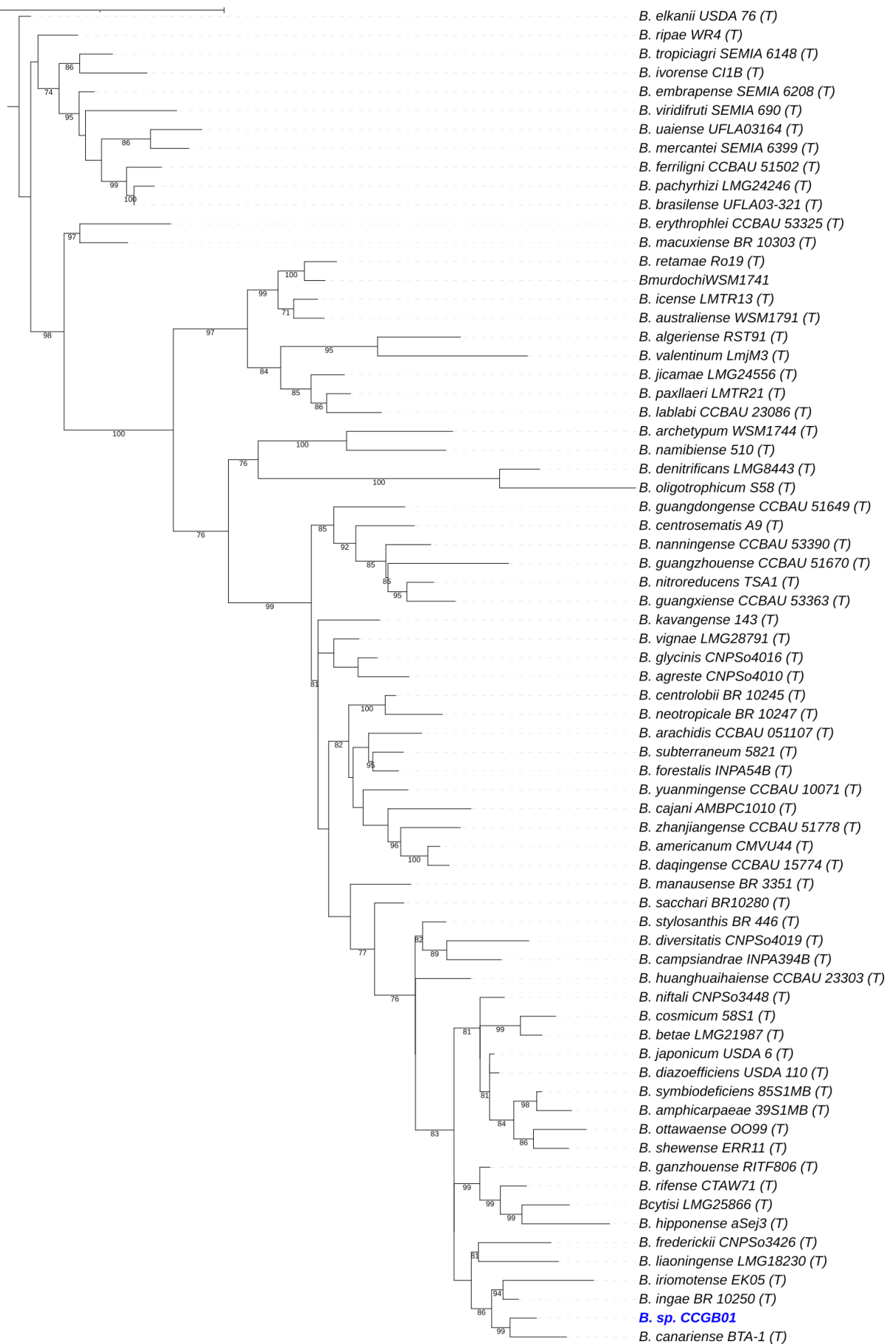


Figura S2. Árbol filogenético de máxima verosimilitud del gen *glnII* de *Bradyrhizobium* sp. CCGB01 y cepas tipo del género sin enraizar con un modelo de sustitución GTR+I+G. La longitud del alineamiento fue de 504 pb. Se muestran los valores de UFBoot (≥ 70) de 1000 pseudo réplicas; el soporte representa directamente la probabilidad de que el clado sea confiable y a partir de valores ≥ 95 los clados comienzan a ser confiables (Minh et al., 2013).

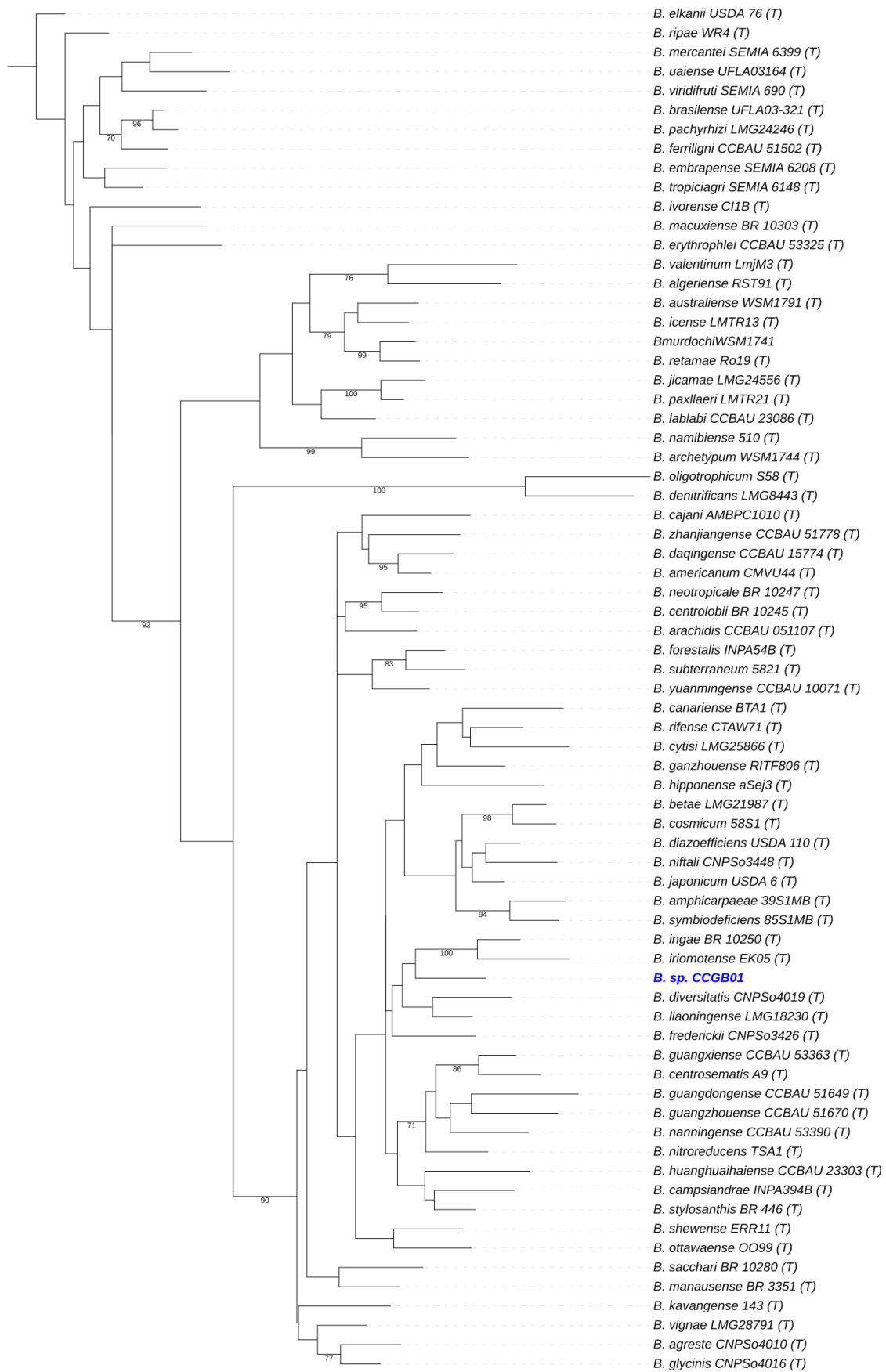


Figura S3. Árbol filogenético de máxima verosimilitud de los genes concatenados *glnII* y *recA* de *Bradyrhizobium* sp. CCGB01 y cepas tipo del género sin enraizar con un modelo de sustitución GTR+I+G. La longitud de los alineamientos fue de *glnII*: 504 pb y *recA*: 377 pb. La barra muestra las sustituciones esperadas por sitio. Se muestran los valores de UFBoot (≥ 70) de 1000 pseudo réplicas; el soporte representa directamente la probabilidad de que el clado sea confiable y a partir de valores ≥ 95 los clados comienzan a ser confiables (Minh et al., 2013).

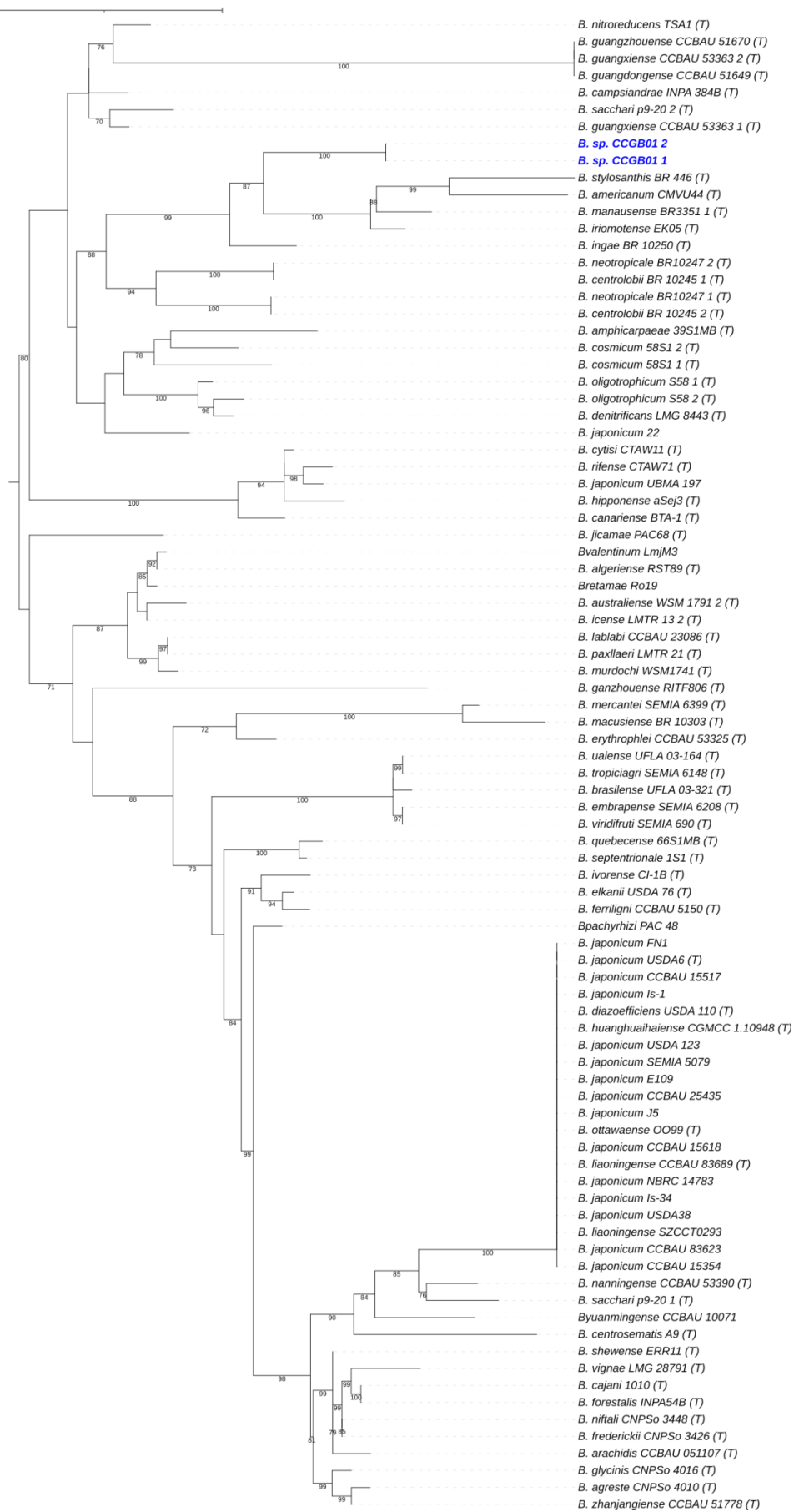


Figura S4. Árbol filogenético de máxima verosimilitud de los genes *nifH* para *Bradyrhizobium* sp. CCGB01 y cepas tipo del género sin enraizar con un modelo de sustitución TrN+I+G4 La longitud del alineamiento fue de 255 pb. Se muestran los valores de UFBboot (≥ 70) de 1000 pseudo réplicas; el soporte representa directamente la probabilidad de que el clado sea confiable y a partir de valores ≥ 95 los clados comienzan a ser confiables (Minh et al., 2013).

Tabla S1. Números de acceso de las secuencias parciales de genes y genomas utilizados en este trabajo. NA=No aplica; ^T= cepa tipo

<i>Strain</i>	<i>16S rRNA</i>	<i>glnII</i>	<i>recA</i>	<i>nodC</i>	<i>nifH</i>	whole genome
<i>B. agreste</i> CNPSo 4010 ^T	MK676048.1	JACCHP010000002.1	JACCHP010000025.1	MK893236.1	JACCHP010000001	GCF_016031625.1
<i>B. algeriense</i> RST89 ^T	FJ546419	NA	NA	FJ348666.1	PYCM01000032	GCF_003062285.1
<i>B. algeriense</i> RST91 ^T	NA	PYCN01000014.1	PYCN01000017.1	NA	NA	NA
<i>B. americanum</i> CMVU44 ^T	KU991833.1	KX012942.1	KC247141.1	KC247130.1	KC247135.1	NA
<i>B. amphicarpaceae</i> 39S1MB ^T	KP768779.1	KP768605.1	KP768721.1	NA	KF615664.1	GCF_002266435.2
<i>B. arachidis</i> CCBAU 051107 ^T	HM107167.1	KF962689.1	KF962707.1	HM107267.1	CP030050	GCF_015291705.1
<i>B. archetypum</i> CNPSo 4013 ^T	MK676065.1	NA	NA	MK893251.1	NA	NA
<i>B. archetypum</i> WSM1744 ^T	NA	JAAVLW010000005.1	JAAVLW010000002.1	NA	JAAVLW010000001	GCF_013114835.1
<i>B. australiense</i> CNPSo 4014 ^T	MK676067.1	NA	NA	MK893253.1	NA	NA
<i>B. australiense</i> WSM1791 ^T	NA	JAAVLX010000004.1	JAAVLX010000002.1	NA	JAAVLX010000005	GCF_013114825.1
<i>B. betae</i> PL7HG1 ^T	AY372184.1	NA	NA	NA	NA	GCF_008932115.1
<i>B. betae</i> LMG21987 ^T	NA	AB353733.1	AB353734.1	NA	NA	NA
<i>B. brasiliense</i> UFLA 03-321 ^T	MPVQ01000039.1	MPVQ00000000	KT793142.1	KT793173.1	MPVQ01000047	GCF_002776695.1
<i>B. cajani</i> AMBPC 1010 ^T	KY349447.1	KY349442.1	KY349440.1	KY349444.1	WQNE01000007	GCF_009759665.1
<i>B. campsiandrae</i> INPA 01394B ^T	KU230298.1	JAANIHO10000002.1	JAANIHO10000001.1	NA	JAATTO010000053	GCF_014529705.1
<i>B. canariense</i> BTA-1 ^T	AJ558025.1	AY386765.1	AY591553.1	AJ560653.1	VSST01000027	NA
<i>B. centrolobii</i> BR10245 ^T	KF927049.1	LUUB01000131	LUUB01000078.1	LUUB01000057.1	LUUB01000073	GCF_001641635.1
<i>B. centrosematis</i> A9 ^T	KC247115.1	KX012940.1	KC247145.1	KC247134.1	KC247139.1	NA
<i>B. cosmicum</i> 58S1 ^T	CP041656.1	CP041656.1	CP041656.1	NA	NA	GCF_007290395.1
<i>B. cytisi</i> CTAW11 ^T	EU561065.2	JN186291.1	JN186293.1	EU597844.2	VSSR01000014/CP041656	GCF_008123515.1
<i>B. daqingense</i> CGMCC 1 10947 ^T	NA	NA	NA	NA	NA	GCF_007830205.1
<i>B. daqingense</i> CCBAU 15774 ^T	KJ184551.1	KF962690.1	KF962708.1	HQ231326.1	NA	NA
<i>B. denitrificans</i> LMG8443 ^T	X66025.1	HM047121.1	EU665419.1	NA	HM047125.1	NA
<i>B. diazoefficiens</i> USDA 110 ^T	BA000040.2	BA000040	BA000040.2	BA000040.2	CP011360	GCF_000011365.1

<i>B. diversitatis</i> CNPSo 4019 ^T	MK676046.1	JACEGD010000024.1	JACEGD010000018.1	MK893234.1	JACEGD010000022	GCF_016031635.1
<i>B. elkanii</i> USDA 76 ^T	U35000.3	AY599117.1	AY591568.1	KB900701.1	ARAG01000005	GCF_000379145.1
<i>B. embrapense</i> SEMIA 6208 ^T	AY904773.1	LFIP02000001.1	LFIP02000009.1	KP234521.1	LFIP01000065	GCF_001189235.2
<i>B. erythrophlei</i> CCBAU 53325 ^T	KF114645.1	KF114693.1	KF114669.1	KF114576.1	KF114598.1	NA
<i>B. ferriligni</i> CCBAU 51502 ^T	KX683400.1	KJ818099.1	KJ818112.1	KJ818109.1	KJ818108.1	NA
<i>B. forestalis</i> INPA 54B ^T	PGVG01000009.1	PGVG01000024.1	PGVG01000004.1	PGVG01000037.1	PGVG01000016	GCF_002795245.1
<i>B. frederickii</i> CNPSo 3426 ^T	MK672937.2	SPQS01000012.1	SPQS01000004.1	NA	SPQS01000036	GCF_004570865.1
<i>B. ganzhouense</i> RITF 806 ^T	JQ796661.2	JX277110.1	JX277144.1	JX292035.1	JX292065.1	NA
<i>B. glycinis</i> CNPSo 4016 ^T	MK676047.1	JACCHQ010000001.1	JACCHQ010000015.1	MK893235.1	JACCHQ01000018	GCF_016031655.1
<i>B. guangdongense</i> CCBAU 51649 ^T	KC508867.1	KC509023.1	KC509269.1	KX683398.1	CP030052	GCF_004114975.1
<i>B. guangxiense</i> CCBAU 53363 ^T	KC508877.1	KC509033.1	KC509279.1	KX683397.1	CP022220	GCF_004114915.1
<i>B. guangzhouense</i> CCBAU 51670 ^T	CP030053.1	CP030053.1	CP030053.1	NA	CP030054	GCF_004114955.1
<i>B. hipponense</i> aSej3 ^T	MK458611.1	VSTH01000057.1	VSTH01000176.1	VSTH01000123.1	VSTH01000149	GCF_008123965.1
<i>B. huanghuaihaiense</i> CCBAU 23303 ^T	HQ231463.1	KF962691.1	KF962709.1	HQ231507.1	NA	NA
<i>B. huanghuaihaiense</i> CGMCC 1.10948 ^T	NA	NA	NA	NA	VLLA01000047	GCF_007830635.1
<i>B. icense</i> LMTR 13 ^T	KF896156.1	KF896175.1	JX943615.1	KF896159.1	CP016428	GCF_001693385.1
<i>B. ingae</i> BR 10250 ^T	KF927043.1	KF927067.1	KF927061.1	KF927054.1	KF927085.1	NA
<i>B. iriomotense</i> EK05 ^T	AB300992.1	AB300995.1	AB300996.1	AB301000.1	AB300998.1	NA
<i>B. ivorensis</i> CI1B ^T	KX396570.1	CAADFC020000018.1	CAADFC020000009.1	CAADFC020000011.1	CAADFC020000008	GCF_900696085.1
<i>B. japonicum</i> USDA 6 ^T	AB510002	HQ587875.1	AM168341.1	AP012206.1	AP012206.1	GCF_000284375.1
<i>B. jicamae</i> PAC68 ^T	AY624134.1	NA	NA	AB573869.1	LLXZ01000033	GCF_001440395.1
<i>B. jicamae</i> LMG 24556 ^T	NA	FJ428204.1	HQ587415.1	NA	NA	NA
<i>B. kavangense</i> 143 ^T	KP899562.1	KM378446.1	KM378399.1	KT033402.1	NA	NA
<i>Blablabi</i> CCBAU 23086 ^T	LLYB01000065.1	GU433498.1	KF962710.1	LLYB01000064.1	LLYB01000119	GCF_001440475.1
<i>B. liaoningense</i> 2281 ^T	AF208513.1	NA	NA	NA	NA	NA
<i>B. liaoningense</i> CCBAU 05525	NA	NA	NA	AJQC01000616.1	NA	NA

<i>B. liaoningense</i> LMG18230 ^T	NA	AY386775.1	AY591564.1	NA	NA	NA
<i>Blupini</i> USDA3051 ^T	KM114861.1	NA	KM114866.1	KM114864.1	NA	NA
<i>B. macuxiense</i> BR10303 ^T	KX527919.1	LNCU01000062	LNCU01000014.1	LNCU01000070.1	NA	GCF_001542415.1
<i>B. manausense</i> BR3351 ^T	HQ641226.2	LJYG01000112.1	KF785992.1	KF786002.1	LJYG01000013	GCF_001440035.1
<i>B. mercantei</i> SEMIA 6399 ^T	FJ025102.1	KX690621.1	KX690615.1	NA	MKFI01000003	GCF_001982635.1
<i>B. murdochi</i> CNPSo 4020 ^T	MK676062	NA	NA	MK893248.1	NA	NA
<i>B. murdochi</i> WSM1741 ^T	NA	AXAU01000001.1	AXAU01000016.1	NA	AXAU01000021.1	GCF_000472965.1
<i>B. namibiense</i> 510 ^T	KX661401.2	KM378440.1	KM378377.1	KX661399.1	NA	NA
<i>B. nanningense</i> CCBAU 53390 ^T	KC508872.1	LBJC01000005.1	LBJC01000019.1	LBJC01000020.1	LBJC01000002	GCF_004114535.1
<i>B. neotropiale</i> BR10247 ^T	KF927051.1	LSEF01000034.1	LSEF01000061.1	KJ661727.1	LSEF01000026	GCF_001641695.1
<i>B. niftali</i> CNPSo 3448 ^T	MK673807.2	SPQT01000002.1	SPQT01000004.1	MK675806.1	SPQT01000005	GCF_004571025.1
<i>B. nitroreducens</i> TSA1 ^T	AB542368.1	LFJC00000000	LFJC01000003.1	NA	LFJC01000003	GCF_002776695.1
<i>B. oligotrophicum</i> S58 ^T	JQ619230.1	JQ619233.1	JQ619231.1	NA	AP012603	GCF_000344805.1
<i>B. ottawaense</i> OO99 ^T	JN186270.1	HQ587750.1	HQ587287.1	HQ587980.1	NPNY01000047	GCF_002278135.2
<i>B. pachyrhizi</i> PAC48 ^T	AY624135.1	NA	NA	LFIQ01000053.1	LFIQ01000027	GCF_001189245.1
<i>B. pachyrhizi</i> LMG24246 ^T	NA	FJ428201.1	HQ587416.1	NA	NA	NA
<i>B. paxllaeri</i> LMTR21 ^T	AY923031.1	CP042968.1	CP042968.1	KF896160.1	MAXB01000076	GCF_001693515.2
<i>B. retamae</i> Ro19 ^T	KC247085.1	KC247108.1	KF962711.1	KC247112.1	LLYA01000189	GCF_001440415.1
<i>B. rifense</i> CTAW71 ^T	EU561074.2	KF962692.1	KF962712.1	EU597853.2	VSSS01000037	GCF_008123425.1
<i>B. ripae</i> WR4 ^T	MF593081.1	MF593086.1	MF593090.1	MF593106.1	NA	NA
<i>B. sacchari</i> P920 ^T	KF113091.3	NA	NA	KF196792.2	LWIG01000010(56)	GCF_002068095.1
<i>B. sacchari</i> BR10280 ^T	NA	KX065099.1	KX065095.1	NA	NA	NA
<i>B. shewense</i> ERR11 ^T	FMAI00000000	FMAI01000006.1	FMAI01000022.1	FMAI01000039.1	FMAI01000005	GCF_900094605.1
<i>B. stylosanthis</i> BR446 ^T	KU724142.1	LVEM01000001.1	LVEM01000001.1	KU724160.1	LVEM01000003	GCF_001641335.1
<i>B. subterraneum</i> 5821 ^T	KP308152.1	KM378484.1	KM378397.1	NA	NA	NA
<i>B. symbiodeficiens</i> 85S1MB ^T	KP768783.1	CP029427.1	KF615036.1	NA	NA	GCF_002266465.2

<i>B. tropiciagri SEMIA 6148^T</i>	AY904753.1	LFLZ01000067.1	LFLZ01000039.1	KP234520.1	LFLZ01000010	GCF_001189845.1
<i>B. uaiense UFLA 03164^T</i>	VKHP01000010.1	VKHP01000010.1	VKHP01000115.1	VKHP01000001.1	VKHP01000057	GCF_010811875.1
<i>B. valentinum LmjM3^T</i>	JX514883.2	LLXX01000118	LLXX01000017.1	LLXX01000153.1	LLXX01000167	GCF_001440405.1
<i>B. vignae 72^T</i>	KP899563.1	RDQF01000005.1	RDQF01000045.1	KT362339.1	NA	NA
<i>B. vignae LMG 28791^T</i>	NA	NA	NA	NA	RDQF01000014	GCF_900324035.1
<i>B. viridifuturi SEMIA 690^T</i>	FJ025107.1	LGTB01000012	KR149140.1	LGTB01000005.1	LGTB01000010	GCF_001238275.1
<i>B. yuanmigense B071^T</i>	AF193818.1	AY386780.1	NA	NA	NA	NA
<i>B. yuanmingense NBRC 100594^T</i>	NA	NA	NA	AB354633.1	NA	NA
<i>B. yuanmingense CCBAU 10071^T</i>	NA	NA	AM168343.1	NA	FMAE01000036	GCF_900094575.1
<i>B. zhanjiangense CCBAU 51778^T</i>	KC508861.1	CP022221.1	CP022221.1	KC509225.1	CP022221	GCF_004114935.1
<i>B. japonicum E109</i>	NA	NA	NA	NA	CP010313	GCF_000807315.1
<i>B. japonicum Is-34</i>	NA	NA	NA	NA	JRPN01000070	GCF_000773865.1
<i>B. japonicum Is-1</i>	NA	NA	NA	NA	LGUJ01000050	NA
<i>B. japonicum J5</i>	NA	NA	NA	NA	CP017637	GCF_001887695.1
<i>B. japonicum CCBAU 15354</i>	NA	NA	NA	NA	AJPX01000730	GCF_000261545.1
<i>B. japonicum 22</i>	NA	NA	NA	NA	AXVG01000014	NA
<i>B. japonicum CCBAU 15517</i>	NA	NA	NA	NA	AJPY01000630	GCF_000261565.1
<i>B. japonicum CCBAU 83623</i>	NA	NA	NA	NA	AJQB01000611	GCF_000261625.1
<i>B. japonicum CCBAU 15618</i>	NA	NA	NA	NA	AJPZ01000493	GCF_000261585.1
<i>B. japonicum CCBAU 25435</i>	NA	NA	NA	NA	AJQA01000397	GCF_000261605.1
<i>B. japonicum FNI</i>	NA	NA	NA	NA	JGCL01000038	GCF_001038185.1
<i>B. japonicum NBRC 14783</i>	NA	NA	NA	NA	BJNK01000065	GCF_006539645.1
<i>B. japonicum SEMIA 5079</i>	NA	NA	NA	NA	CP007569	GCF_000661935.1
<i>B. japonicum SZCCT0280</i>	NA	NA	NA	NA	NA	GCF_018130365.1
<i>B. japonicum SZCCT0395</i>	NA	NA	NA	NA	NA	GCF_018130735.1
<i>B. japonicum SZCCT0402</i>	NA	NA	NA	NA	NA	GCF_018130795.1

<i>B. japonicum</i> UBMA197	NA	NA	NA	NA	NAFL01000191	GCF_002108935.1
<i>B. japonicum</i> USDA 123	NA	NA	NA	NA	AXVP01000150	GCF_000482525.1
<i>B. japonicum</i> USDA 38	NA	NA	NA	NA	NA	GCF_000472745.1
<i>B. liaoningense</i> SZCCT0293	NA	NA	NA	NA	JAFCKI010000135	GCF_018130525.1
<i>B. liaoningense</i> SZCCT0399	NA	NA	NA	NA	NA	GCF_018130785.1
<i>B. liaoningense</i> CCBAU 83689	NA	NA	NA	NA	AJQD01000643	NA
<i>Bsp</i> MOS001	NA	NA	NA	NA	NA	GCF_004634465.1
<i>B. quebecense</i> 66S1MB ^T	NA	NA	NA	JABWSX010000001.1	KF615665.1	NA
<i>B. septentrionale</i> 1S1 ^T	NA	NA	NA	JAAOLE020000001.1	KF615667.1	NA
<i>B. valentinum</i> bv. <i>lupini</i> LmiM2	NA	NA	NA	KX272811.1	NA	NA
<i>B. retamae</i> bv. <i>lupini</i> LmiH4	NA	NA	NA	KX272805.1	NA	NA
<i>B. sp.</i> GV137	NA	NA	NA	KF483556.1	NA	NA
<i>B. sp.</i> VUPME10	NA	NA	NA	HG940520.1	NA	NA
<i>Nitrobacter</i> sp. 62-13	NA	NA	NA	NA	NA	GCA_001896955.1
<i>N. sp.</i> 62-23	NA	NA	NA	NA	NA	GCA_001897285.1
<i>N. sp.</i> Nb-311A	NA	NA	NA	NA	NA	GCF_000152905.1
<i>N.hamburgensis</i> X14 ^T	NA	NA	NA	NA	NA	GCF_000013885.1
<i>N.vulgaris</i> Ab1	NA	NA	NA	NA	NA	GCF_002028545.1
<i>N.winogradskyi</i> NBRC14297	NA	NA	NA	NA	NA	GCF_006539545.1
<i>N.winogradskyi</i> Nb-255 ^T	NA	NA	NA	NA	NA	GCF_000012725.1
<i>Bosea thiooxidans</i> DSM9653 ^T	AJ250796.1	NA	NA	NA	NA	NA
<i>B. sp.</i> CCGB01*	GCF_024199795.1 / OP185277	GCF_024199795.1	GCF_024199795.1	GCF_024199795.1	GCF_024199795.1	GCF_024199795.1
CCGB12*	OP185278	NA	NA	NA	NA	NA
CCGB20*	OP185279	NA	NA	NA	NA	NA
CCGBR10*	OP185280	NA	NA	NA	NA	NA
CCGBR13*	OP185281	NA	NA	NA	NA	NA
CCGBR17*	OP185282	NA	NA	NA	NA	NA
CCGBR2*	OP185283	NA	NA	NA	NA	NA

CCGBR22*	OP185284	NA	NA	NA	NA	NA
CCGBR26*	OP185285	NA	NA	NA	NA	NA
CCGBR7*	OP185286	NA	NA	NA	NA	NA

*Genoma o gen secuenciados en este trabajo.

Tabla S2. Resumen del análisis de calidad del genoma de *B. sp. CCGB01* anotado con RASTtk v. 1.3.0. El análisis se obtuvo con las herramientas EvalG con una re-implementación del algoritmo checkM (Parks et al., 2014) y el programa EvalCon a través de la plataforma en línea PATRIC v. 3.6.9 (Davis et al., 2020).

Medidas de calidad genómica	Explicación	Valores reportados para el genoma de <i>B. sp. CCGB01</i>
Consistencia gruesa (%)	Representa el porcentaje de roles cuya presencia o ausencia fue correctamente predicha	99
Consistencia fina (%)	Abarca el porcentaje de roles cuyos números de ocurrencia exactos fueron predichos correctamente	95.3
Integridad (%)	Representa el porcentaje de roles universales que aparecen en el genoma	100
Contaminación (%)	Se computariza identificando si los roles universales ocurren más de una vez.	3.8
Genoma de referencia	Se selecciona de manera automática el genoma de mayor calidad del género o especie al que pertenece el genoma que se desea estudiar	<i>Bradyrhizobium sp. BTAi1.</i> Reporte de calidad: https://www.patricbrc.org/view/Genome/288000.8