



Diversidad de genes bacterianos relacionados con promoción del crecimiento vegetal

Miguel García Hidalgo

Máster de Bioinformática y Bioestadística
Microbiología, biotecnología y biología molecular

Paloma Pizarro Tobías

Profesora Consultora

David Merino Arranz

Carles Ventura Rojo

Profesores responsables de la asignatura

02 de enero de 2018



Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-

SinObraDerivada [3.0 España de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

FICHA DEL TRABAJO FINAL

Título del trabajo:	<i>Diversidad de genes bacterianos relacionados con promoción del crecimiento vegetal</i>
Nombre del autor:	<i>Miguel García Hidalgo</i>
Nombre del consultor/a:	<i>Paloma Pizarro-Tobías</i>
Nombre del PRA:	<i>David Merino Arranz y Carles Ventura Rojo</i>
Fecha de entrega (mm/aaaa):	<i>01/2018</i>
Titulación:	<i>Máster en Bioestadística y Bioinformática</i>
Área del Trabajo Final:	<i>Microbiología, biotecnología y biología molecular</i>
Idioma del trabajo:	<i>Castellano</i>
Palabras clave	<i>Phylogeny, Plant Growth Promoting Rhizobacteria (PGPR), Rhizosphere.</i>

Resumen del Trabajo (máximo 250 palabras): *Con la finalidad, contexto de aplicación, metodología, resultados i conclusiones del trabajo.*

El estudio de las propiedades de las bacterias promotoras del crecimiento vegetal (Plant Growth Promoting Rhizobacteria, PGPR) ha conseguido tal avance en los últimos años que ya son consideradas como una herramienta verosímil para la mejora de suelos y la biofertilización. No obstante, para una correcta administración de bacterias en un suelo o cultivo, se deben conocer los diferentes efectos beneficiosos de cada cepa y su compatibilidad para que mantengan su equilibrio de forma sostenible. A partir de datos bibliográficos, se consideran los principales géneros bacterianos promotores del crecimiento vegetal según algunas de sus funciones más importantes, la fijación de nitrógeno (Fe nitrogenasa y Mo-Fe nitrogenasa), producción de ácido indolacético (piruvato decarboxilasa e indolepiruvato decarboxilasa) disminución de concentración en etileno (ACC-deaminasa) en la rizosfera, promoción del desarrollo radicular (nitrito reductasa) y la promoción de la resistencia sistémica inducida (Induced Systemic Resistance, ISR) (acetolactato decarboxilasa, acetolactato sintasa y diacetil reductasa). Las proteínas involucradas en estos procesos son comparadas según las especies en las que se encuentran para conocer su diversidad, evolución y coocurrencia. Con el estudio filogenético de las cepas se discute si la diversificación de las distintas funciones ha seguido una evolución asociada a la otorgada por los marcadores *rpoD* y *gyrB*. Por último se estudia la relación entre los distintos grupos bacterianos (alfa, beta y gammaproteobacteria, firmicutes y Gram+ con alto [G+C]) con los principales cultivos de interés humano (Poáceas, Fabáceas, Brasicáceas, Malváceas y Solanáceas).

Abstract (in English, 250 words or less):

In recent times, Plant Growth Promoting Rhizobacteria (PGPR) research has developed to be considered as a plausible key tool to soil improvement and biofertilization. The examination of the capabilities of each PGPR strain and their compatibility is essential to reach an optimum knowledge for the accurate management in a determinate soil or crop at equilibrium. This should ensure the sustainable optimization of bacterial functions. From bibliographical data, the main PGPR genera are considered according to their strengths, nitrogen fixation (Fe nitrogenase and Mo-Fe nitrogenase), indolacetic acid production (pyruvate decarboxylase and indolepyruvate decarboxylase), ethylene reduction (ACC-deaminase), root promotion (nitrite reductase) and Induced Systemic Resistance triggering (acetolactate decarboxylase, acetolactate synthase and diacetyl reductase). Involved proteins in those processes are compared under their native species to know their diversity, evolution and co-occurrence. Besides, strains are phylogenetically distributed in light of *rpoB* and *gyrB* markers to know whether functional diversification is associated. Finally, the relationship between bacterial groups (Alpha, beta and gamma proteobacteria, firmicutes and gram+ high [G+C]) is studied according the main crops for humans (Poaceae, Fabaceae, Brassicaceae, Malvaceae and Solanaceae).

Índice

1. Contexto y justificación del Trabajo.....	1
1.1 Introducción.....	1
1.2 Objetivos del Trabajo.....	3
1.3 Enfoque y método seguido.....	3
1.4 Planificación del Trabajo.....	5
1.5 Breve resumen de productos obtenidos.....	8
1.6 Breve descripción de los otros capítulos de la memoria.....	8
2. Diversidad de genes bacterianos relacionados con promoción del crecimiento vegetal.....	10
2.1. CAPÍTULO 1. LAS BACTERIAS PGPR. PRINCIPALES MECANISMOS DE PROMOCIÓN DE CRECIMIENTO VEGETAL.....	11
2.1.1 Nitrogenasa.....	13
2.1.2 Indolpiruvato Descarboxilasa.....	14
2.1.3 ACC Desaminasa.....	15
2.1.4 Nitrito Reductasa.....	16
2.1.5 Ruta de 2-3-butanodiol.....	17
2.2. CAPÍTULO 2. RELACIÓN FILOGENÉTICA DE PGPRs.....	18
2.2.1 NITROGENASA.....	25
Proteína Fe.....	26
Proteína Mo Fe Alfa.....	27
Proteína Mo Fe Beta.....	28
2.2.2. INDOLPIRUVATO DESCARBOXILASA.....	29
2.2.3. ACC DEAMINASA.....	30
2.2.4. NITRITO REDUCTASA.....	31
2.2.5. RUTA DE 2-3-BUTANODIOL.....	32
Acetolactato Sintasa.....	32
Acetolactato Descarboxilasa.....	32
Diacetil Reductasa.....	33
2.2.6. MARCADORES FILOGENÉTICOS.....	34
gyrB.....	35
rpoB.....	35
2.3. CAPÍTULO 3. ESTUDIO COMBINADO DE LAS RELACIONES ENTRE MECANISMOS PROMOTORES DEL CRECIMIENTO VEGETAL, DIVERSIDAD DE PGPRs Y CULTIVOS. .	37
3. Conclusiones.....	40
4. Glosario.....	41
5. Bibliografía.....	42
6. Anexos.....	47
Anexo I. Tablas de cepas.....	47
Anexo II. Árboles filogenéticos.....	54
Anexo III. Tabla de citas.....	56
Anexo IV. <i>Script</i> de R.....	57

Lista de figuras

Figura 1. Mecanismos Directos e Indirectos de interacción Planta-PGPR	12
Figura 2. Complejo nitrogenasa	13
Figura 3. Estructura química de Nitrogenasa Mo-Fe de <i>Azotobacter vinelandii</i>	13
Figura 4. Síntesis de Ácido Indolacético a partir de Triptófano mediante Piruvato descarboxilasa	14
Figura 5. Indolpiruvato descarboxilasa de <i>Enterobacter cloacae</i>	14
Figura 6. Reacción catalizada por la enzima ACC desaminasa	15
Figura 7. ACC desaminasa con sustrato de <i>Pseudomonas</i> sp	15
Figura 8. Nitrito Reductasa de Cobre de <i>Geobacillus thermodenitrificans</i>	16
Figura 9. Reacción catalizada por la enzima Nitrito Reductasa	16
Figura 10. Acetolactato Sintasa de <i>Klebsiella pneumoniae</i>	17
Figura 11. Acetoin Reductasa de <i>Burkholderia cenocepacia</i>	17
Figura 12. Procedimiento Neighbour Joining	20
Figura 13. Coocurrencia de genes estudiados	25
Figura 14. Distribución de citas según cultivos y grupo de bacterias	38

Tablas

Tabla 1. Lista de Cepas introducidas en el estudio y genes incluidos para cada cepa	21
Tabla 2. Presencia de genes por grupo de estudio	24
Tabla 3. Coocurrencia de genes estudiados	24

Árboles filogenéticos

Árbol Filogenético 1. Fe Protein - <i>nifH</i>	26
Árbol Filogenético 2. Fe-Mo Alpha - <i>nifD</i>	27
Árbol Filogenético 3. Mo - Fe Beta - <i>nifK</i>	28
Árbol Filogenético 4. Piruvato descarboxilasa - <i>ipdC/ppdC</i>	29
Árbol Filogenético 5. ACC deaminasa - <i>acdS</i>	30
Árbol Filogenético 6. Nitrito Reductasa - <i>nirK</i>	31
Árbol Filogenético 7. Acetolactato Sintasa - <i>budB</i>	32
Árbol Filogenético 8. Acetolactato Descarboxilasa - <i>budA</i>	32
Árbol Filogenético 9. Diacetil Reductasa - <i>budC</i>	33
Árbol Filogenético 10. Comparativa de la filogenia según marcadores filogenéticos	36

1. Contexto y justificación del Trabajo

1.1 Introducción

El crecimiento de la población mundial requiere cada vez de mayor cantidad de recursos, siendo uno de los problemas principales la necesidad de alimentos [1]. Se piensa que, en el año 2050, la demanda mundial puede aumentar hasta en un 100%. Para evitar el colapso se requieren estrategias que, siendo respetuosas con el medio ambiente, ayuden a mejorar la eficiencia de los cultivos.

La agricultura tradicional ha contado con la aplicación de abonos inorgánicos como herramienta principal para la promoción del crecimiento vegetal [2]. De este modo, el avance en investigación tanto en las fórmulas de estos abonos como la selección genética de los cultivos, ha logrado alcanzar tasas de producción en los cultivos que hasta hace años eran impensables [3], llegando al punto en el que la única restricción pueda ser la genética de la especie como ocurre en cultivos hidropónicos.

En este sentido, la sobreexplotación del suelo en monocultivos, asociada a la falta de estrategias de mantenimiento como la rotación de cultivos o la utilización de abonos orgánicos, ha provocado un gran empobrecimiento en la cantidad de nutrientes. Ante este agotamiento de los suelos, existe una alta dependencia en los abonos inorgánicos para poder mantener las producciones [4].

Sin embargo, ante las evidencias sobre los efectos contaminantes del uso de fertilizantes nitrogenados y fosfatados, ha aumentado el interés en medios alternativos para una fertilización adecuada que sea respetuosa con el medio ambiente[5]. En los últimos años, se tiende a la sustitución de estos abonos minerales por abonos orgánicos reduciéndose problemas de bloqueo de nutrientes, o la lixiviación de nitratos y fosfatos que conllevan grandes efectos contaminantes[6], [7].

Un paso más allá es compensar la pérdida de seres vivos que habitan en el suelo. Los microorganismos que ocupan este nicho son importantes en el crecimiento vegetal participando en aspectos como la producción de nutrientes accesibles para las plantas o generando competencia con microorganismos patógenos [8]. De estos microorganismos que se encuentran en el suelo, son de gran interés los formadores de la rizosfera [9], [10]. Esta zona cercana a la raíz tiene con un contenido característico de microorganismos gracias a los exudados producidos por las raíces de las plantas que promueven esta interacción [11]. El estudio actual sobre rizobacterias promotoras del crecimiento vegetal (PGPRs) es cada vez más avanzado conociendo aspectos precisos de cuáles son los procesos en los que cada cepa tiene mayor efecto. Actualmente, ya se realizan experiencias de biofertilización [12] para favorecer un contenido bacteriano específico para cada tipo de cultivo, suelo o clima, mejorando así la producción.

Las PGPR se caracterizan por favorecer el crecimiento vegetal tanto a través de métodos directos como de indirectos. Entre los mecanismos directos se encuentran la mejora de contenido en nutrientes del suelo mediante la fijación de nitrógeno o por la solubilización de fosfatos y sulfatos, el favorecimiento en la absorción de hierro mediante la producción

de sideróforos y la liberación de fitohormonas. Entre los indirectos, la secreción de antibióticos, enzimas líticas, la competencia frente al crecimiento de microorganismos patógenos o la eliminación de sustancias que pueden ser tóxicas para la planta [2], [6], [8], [13], [14].

Para la aplicación de PGPR como biofertilizantes es necesaria una precisa comprensión de los mecanismos que promueven el crecimiento vegetal y el papel que desempeña cada cepa. En este sentido es importante el conocimiento de la patogenicidad potencial de cada cepa, no solo para el cultivo en el que se aplica sino para el ulterior consumo de la planta asociada.[15]

A pesar del auge en el estudio de PGPRs, todavía hay mucho camino por recorrer hasta alcanzar un nivel de conocimiento que permita descubrir qué cepas son necesarias aplicar en determinado suelo o según las necesidades del cultivo. Un adecuado biofertilizante será aquel que aglutine las cepas necesarias para completar las necesidades de la planta en su ciclo completo de vida y compensar los desajustes físico-bio-químicos del suelo. Para lograr este objetivo, se deberán conocer tanto las probables interacciones ecológicas entre las cepas con el resto de elementos vivos en ese ambiente edáfico [16], [17], como las características que aportará cada una de las cepas para el crecimiento de la planta[18].

En el presente trabajo se estudian comparativamente las proteínas participantes en la fijación de nitrógeno (Fe nitrogenasa y Mo-Fe nitrogenasa), producción de ácido indolacético (piruvato decarboxilasa e indolepiruvato decarboxilasa) en la disminución de la concentración de etileno (ACC-deaminasa) en la rizosfera, promoción del desarrollo radicular (nitrito reductasa) y la inducción de resistencia sistémica (ISR) (acetolactato decarboxilasa, acetolactato sintasa y diacetil reductasa) de los principales géneros englobados en PGPR[2], [13], [19].

Tras la obtención de resultados por medio de las comparativas dentro de cada proteína, se procederá a relacionar con la filogenia estudiada según los marcadores filogenéticos *rpoB*[20] y *gyrB*[21], [22], con los que se diferencia la filogenia de las cepas.

Por último, se estudia la posible relación existente entre las cepas encontradas en cada uno de los principales cultivos de interés humano.

1.2 Objetivos del Trabajo

1. Identificar las principales bacterias PGPR
 - a. Enumerar las principales PGPRs según su contribución al crecimiento vegetal.
 - b. Citar las funciones beneficiosas de cada especie de microorganismo en la promoción de crecimiento vegetal.
 - c. Identificar los genes involucrados en cada característica.
 - d. Deducir las características que deben cumplir PGPRs para su potencial uso como biofertilizante.
2. Comparar las proteínas características.
 - a. Relacionar las proteínas involucradas en cada una de las funciones de estudio.
 - b. Relacionar nodos comunes según proteínas.
 - c. Deducir las relaciones filogenéticas entre las distintas enzimas estudiadas.
3. Comparar filogenia PGPR según marcadores filogenéticos.
 - a. Obtener filogenia con los distintos marcadores (*rpoB*, *gyrB*).
 - b. Contrastar con filogenia con las relaciones obtenidas en el objetivo 2.
4. Diagnosticar las relaciones entre PGPR y planta.
 - a. Relacionar PGPRs que colaboran con determinado cultivo.

1.3 Enfoque y método seguido

Para llevar a cabo un estudio sobre las posibles relaciones filogenéticas entre microorganismos PGPR y su relación con los cultivos en que se encuentran, la metodología más directa, aunque también más costosa, sería la obtención de muestras de suelo proveniente de distintos cultivos para la secuenciación y estudio de las mismas [23], [24] o la inoculación de distintas PGPR en cultivos hidropónicos para analizar el desarrollo en cada uno de ellos[25].

Sin embargo, debido a la gran cantidad de bibliografía respecto al tema [2], [7], [13], [14], [26], y con las secuencias genéticas y proteicas disponibles a través de bases de datos, es posible realizar un acercamiento a los resultados que se obtendrían con el estudio directo. De este modo, se evita el trabajo de toma de muestras, selección, cultivo y secuenciación, simplificándolo en la obtención de muestras a través de los bancos de datos.

Mediante el buscador de *The National Center for Biotechnology Information* (NCBI) [27] se obtienen los accesos de las secuencias proteicas de las enzimas de estudio guardadas en las bases de datos de *genBank* [27] y *RefSeq* [28], [29]. Para una comparativa proteica se ha mostrado más adecuado el estudio de la secuencia de aminoácidos en contraposición de la nucleotídica para evitar ruido que pueda generar los cambios de base que no tendrían efecto en la proteína final [30]. De este modo, se realiza la comparativa preliminar mediante el programa *Basic Local Alignment Search Tool for Proteins* (BLASTp)[31].

Para conocer la coocurrencia de genes en las distintas cepas, se estudia el número de veces en que cada gen comparte genoma con otro como se describe en *material y métodos* del capítulo 2.2. *Capítulo 2.*

Relación Filogenética de PGPRs.

El trabajo se concentrará en la comparación de estas muestras a través de R y los paquetes *Analysis of Phylogenetics and Evolution* (ape) [32] y *bioconductor* [33] con herramientas basadas en alineamientos múltiples de secuencias, msa [34] siguiendo el método CLUSTAL [34], [35].

Para la generación de los árboles filogenéticos, partiendo del alineamiento anterior, se calculan las respectivas matrices de distancias para, a continuación, generar el árbol filogenético siguiendo el método de los vecinos más cercanos.

En el caso de los marcadores filogenéticos el método seguido varía ligeramente mediante la corrección del alineamiento para la detección de falsas homologías. Es muy importante no perder la idea de que son marcadores filogenéticos, por lo que la semejanza/diferencia detectada en ellos indica la verdadera relación filogenética de las cepas que los contienen.

Tras esta corrección del alineamiento, se calcula la matriz a partir de la distancia de *Hamming*[36] y la generación del árbol mediante la función *nj* de *ape*.

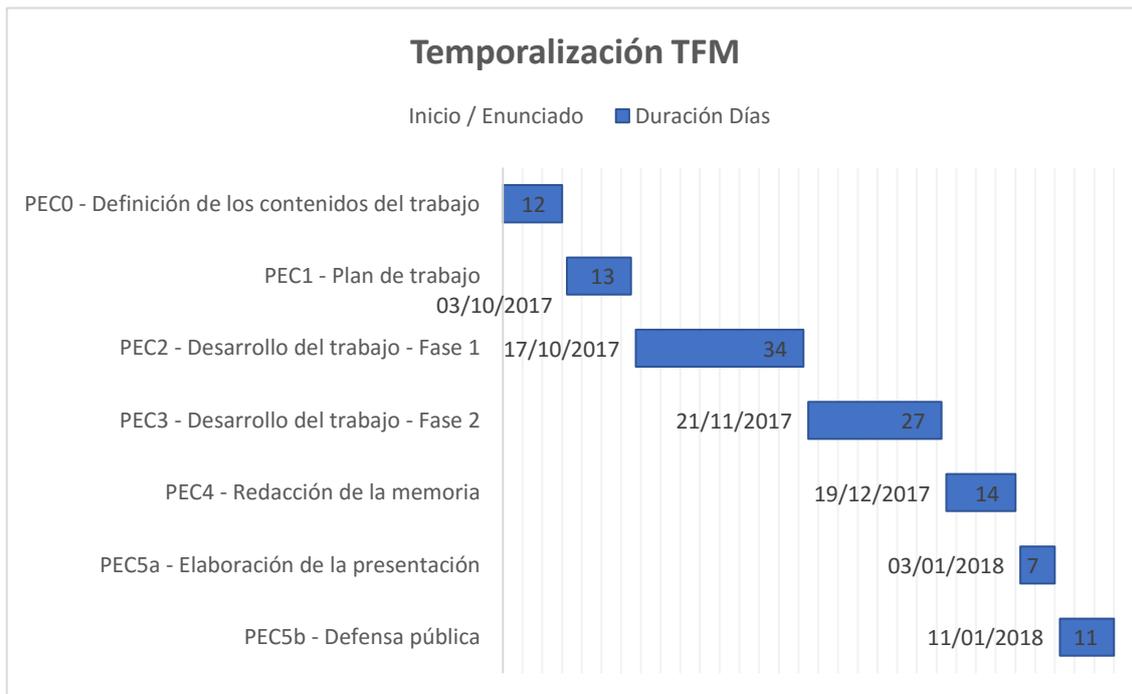
1.4 Planificación del Trabajo

Los recursos utilizados son los siguientes:

1. Plan docente de la asignatura: *Diversidad de genes bacterianos relacionados con promoción del crecimiento vegetal*.
2. Bibliografía:
 - a. Propuesta por la tutora [23], [24]
 - b. Bibliografía a través de NCBI [27], [37]
3. Secuencias genéticas a través de *genbank*[27] y *RefSeq*[29].
4. BLAST[31]
5. Programa R Software [38]
 - a. ape[32]
 - b. Bioconductor[39]
 - c. msa[34]
 - d. seqinr[40]
 - e. rentrez[41]
 - f. decipher[42]

A continuación, se incluye un cuadro de Gantt en el que figuran los períodos para redacción de la memoria y presentación del Trabajo Fin de Máster.

Posteriormente, se presenta un cuadro con los hitos y su temporalización para conseguir los objetivos propuestos.



CUADRO DE PLANIFICACIÓN DEL TRABAJO

OBJETIVO	HITO	RECURSO	FECHA COMPLETO	PEC*	RIESGO
IDENTIFICAR LAS PRINCIPALES BACTERIAS PGPR			20 Octubre	2	
<p>Enumerar las principales PGPRs según su contribución al crecimiento vegetal</p> <p>Citar las funciones beneficiosas de cada especie de microorganismo en la promoción de crecimiento vegetal</p> <p>Identificar los genes involucrados en cada característica</p> <p>Deducir las características que deben cumplir PGPRs para su potencial uso como biofertilizante.</p>	Cuadro resumen con las PGPRs más conocidas en la bibliografía, funciones principales, genes promotores de la función y cultivos asociados a cada género	<p>Bibliografía</p> <p>[23], [24], [27], [43], [44]</p>			<p>Se encuentra en la bibliografía aportada para el TFM y accesible a través de <i>NCBI</i></p> <p>Puede que algunas cepas no hayan sido caracterizadas en alguna propiedad de interés. Se optará por especies y cepas que hayan sido descritas en I.</p>
COMPARAR LAS PROTEÍNAS CARACTERÍSTICAS				2-3	
<p>Relacionar las proteínas involucradas en cada una de las funciones de estudio</p> <p>Relacionar nodos comunes según proteínas</p> <p>Deducir las relaciones filogenéticas entre las distintas enzimas estudiadas</p>	<p>Obtención de secuencias genéticas</p> <p>Comparativa de las secuencias.</p> <p>Generación de árboles</p> <p>Generación de gráfico de coocurrencia</p>	<p>Genbank [28]</p> <p>Refseq [29]</p> <p>BlastP [31]</p> <p>rentrez[41]</p> <p>Ape [32]</p>	<p>27 de Octubre</p> <p>8 de Noviembre</p> <p>16 de Noviembre**</p>	2	<p>En el caso de las proteínas, al ser varias y distintas puede que conlleve demasiado tiempo. Si es así, se elegirán las de mayor interés en el crecimiento vegetal.</p>

CUADRO DE PLANIFICACIÓN DEL TRABAJO (Cont.)

OBJETIVO		HITO	RECURSO	FECHA COMPLETO	PEC*	RIESGO
COMPARAR FILOGENIA PGPR SEGÚN MARCADORES					2-3	
	Obtener filogenia con los distintos marcadores (<i>rpoB</i> , <i>gyrB</i>) Contrastar con filogenia con las relaciones obtenidas en el objetivo 2	Obtención de marcadores Comparativa según marcadores Generación de árboles filogenéticos	Genbank[28] APE[32] MSA[34] Bioconductor[39]	28 de Noviembre 5 de Diciembre 14 de Diciembre		Adecuación al propósito. Si se tratan de demasiadas especies para la comparación se adecuará con las más importantes.
DIAGNOSTICAR LAS RELACIONES ENTRE PGPR Y PLANTA				3		
	Relacionar PGPRs que colaboran con determinado cultivo	Estudio comparativo de los gráficos y árboles anteriores. Relación entre ellos. Relación con los cultivos asociados.	Bibliografía [2], [13], [25], [45]–[49] Bioconductor[39]	20 de Diciembre	3	Puede que la bibliografía y las interacciones descritas sean insuficientes para un resultado óptimo. Se buscarán posibles relaciones con otros organismos participantes en crecimiento vegetal (AM)[50]

*Se presentarán la consecución de HITOS en las PEC correspondientes

**Se deja un espacio de tiempo entre hito y entrega de la PEC para revisiones o, en el mejor de los casos, adelantar trabajo.

1.5 Breve resumen de productos obtenidos

La memoria incluye, en forma de capítulos, descripción de las bacterias PGPR, las relaciones entre distintos microorganismos PGPR según las proteínas que participan en los procesos de promoción del crecimiento vegetal, las relaciones filogenéticas según marcadores específicos y, por último, un resumen con las posibles relaciones filogenéticas y los cultivos a los que estas PGPR se encuentran asociadas.

Asimismo, se realiza una presentación con los resultados gráficos obtenidos y el procedimiento desarrollado hasta su consecución.

Por último, una autoevaluación en la que se considerará el desarrollo del trabajo, la consecución de objetivos y los aspectos más importantes en los que se han encontrado dificultades. Igualmente se comentarán cuáles han sido las capacidades desarrolladas en el transcurso del proyecto.

1.6 Breve descripción de los otros capítulos de la memoria

Capítulo 1. Las Bacterias PGPR. Principales mecanismos de Promoción de crecimiento vegetal.

Descripción de las bacterias consideradas PGPR con las principales características que hacen de ellas un recurso adecuado para la biofertilización [13], [24]. Mecanismos promotores del crecimiento vegetal[7] y ePGPR e iPGPR[44].

Se describen también las proteínas involucradas en la fijación de nitrógeno (Fe nitrogenasa y Mo-Fe nitrogenasa), producción de ácido indolacético (piruvato decarboxilasa e indolepiruvato decarboxilasa) disminución de contenido en etileno (ACC-deaminasa), promoción radical (nitrito reductasa) y la promoción de la Resistencia Sistémica Inducida (acetolactato decarboxilasa, acetolactato sintasa y diacetil reductasa)[13], [18], [19], [51], [52].

Capítulo 2. Relación filogenética de PGPRs.

Estructurado en forma de *Introducción, Material y Metodos, Resultados y Discusión*.

Obtención de cepas de estudio, secuencias, generación de alineamientos y cálculo de distancias entre las secuencias obtenidas de cada una de las proteínas de interés. Se generan los gráficos descriptivos de las relaciones filogenéticas de las distintas cepas estudiadas y la coocurrencia de genes codificantes de proteínas promotoras del crecimiento vegetal. Se generan las predicciones filogenéticas en base a cada una de las proteínas de estudio.

Se exponen y comparan los árboles filogenéticos obtenidos en relación al estudio de relojes moleculares *rpoB* y *gyrB* [20], [22] obtenidos a través de *genbank* [28], [53] y *RefSeq* [29].

Capítulo 3. Estudio combinado de las relaciones entre mecanismos promotores del crecimiento vegetal, diversidad de PGPRs y cultivos.

Compendio de los capítulos anteriores en el que se realiza la discusión sobre las relaciones entre distintos grupos de microorganismos y cultivos en los que se encuentran normalmente.

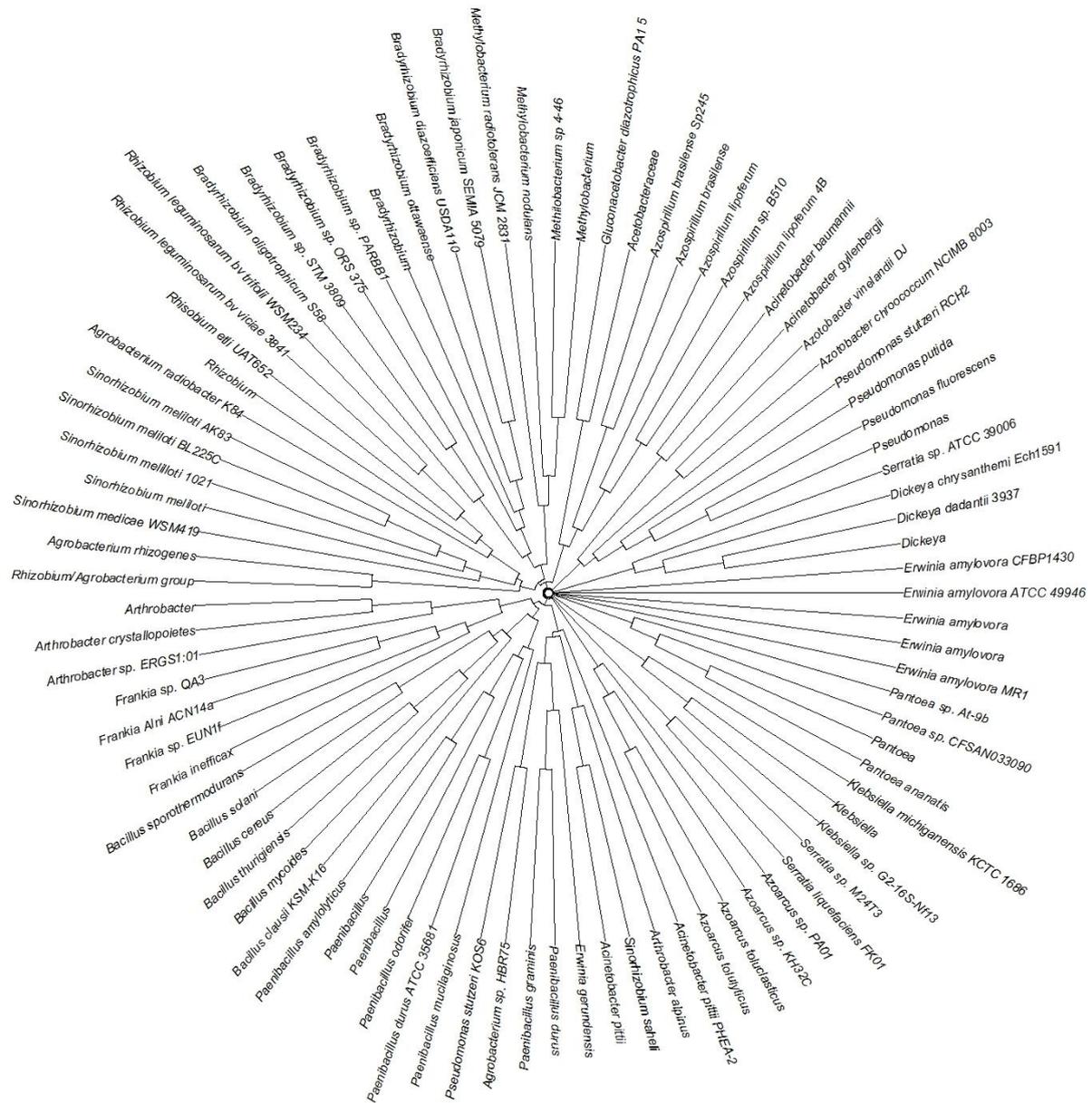
Si existe información bibliográfica suficiente, se podrá realizar un estudio n-KK con el que poder agrupar géneros de PGPRs con determinados cultivos.

Bibliografía.

Anexos.

- Anexo I. Tablas de cepas.
- Anexo II. Árboles filogenéticos.
- Anexo III. Tabla de citas.
- Anexo IV. *Script* de R.

2. Diversidad de genes bacterianos relacionados con promoción del crecimiento vegetal



Filogenia según el marcador *rpoB*.

2.1. CAPÍTULO 1.

LAS BACTERIAS PGPR. PRINCIPALES MECANISMOS DE PROMOCIÓN DE CRECIMIENTO VEGETAL.

La fitotecnia se ha centrado en la mejora de los rendimientos de cultivos mediante distintos métodos. Globalmente, se pueden agrupar en métodos mecánicos, como el laboreo del terreno para evitar la aparición de especies que compitan con la de cultivo, culturales, como la rotación de cultivos y, sobre todo en el estudio de métodos químicos como aplicación de fitosanitarios o fertilizantes[54]. A medida que las investigaciones han ido aportando luz sobre determinados aspectos culturales se ha ido modulando la forma de mejorar el rendimiento de la producción con el mantenimiento de la mayor fuente de riqueza de la agricultura, el medio ambiente. De este modo, el agricultor es cada vez más consciente de las ventajas que posee la utilización de técnicas menos agresivas con el entorno para lograr un adecuado rendimiento a largo plazo[6].

El contenido en los distintos nutrientes y su estructura son aspectos esenciales para el mantenimiento del equilibrio de un suelo. Un cultivo sostenible será aquel que tras un periodo de cultivo vuelva a estar en las mismas, o mejores, condiciones que cuando se inició [55], [56]. Siguiendo este fundamento, la fertilización inorgánica cada vez tiene más en cuenta el estudio previo del suelo o el cultivo para el que está destinado. No obstante, hoy en día no es suficiente con este aporte de nutrientes que vayan a ser absorbidos por la planta. La degradación es tal que se necesita de la recuperación de estos suelos favoreciendo procesos que promuevan su equilibrio y regulación. Con el aporte de fertilizantes orgánicos se intenta no solo el aporte de nutrientes, sino mejorar la textura del suelo y su capacidad de retención de agua. De esta forma se favorece el mantenimiento de las relaciones entre los seres vivos que componen el suelo[50], [57], [58].

Además del contenido inorgánico del suelo, el contenido orgánico es esencial. La existencia de seres vivos y de restos de algunos de ellos es un aspecto importante para la regulación y productividad del suelo [59]. La rizosfera, zona alrededor de la raíz con una composición diferenciada del resto del suelo, centra los estudios recientes en el análisis de la productividad de cultivos[8]. La comunicación entre los seres vivos del suelo es cada vez más estudiada en base a sustancias como aminoácidos, ácidos orgánicos o proteínas que permiten las relaciones entre plantas y microorganismos cercanos[16], [60]. Gracias a los exudados de las raíces, se altera la población de microorganismos que rodean a esa planta promoviendo determinados grupos en detrimento de otros.

Los microorganismos presentes en la rizosfera generan beneficios tanto directamente aprovechables para la planta como para la mejora de la calidad del suelo. Dentro de estos microorganismos presentes en la rizosfera destacan, además de componentes del reino fungi formadores de micorrizas, las bacterias promotoras del crecimiento vegetal (PGPR)[13], [57]. Su estudio detallado ha permitido conocer su implicación en distintos procesos que mejoran el rendimiento de la planta. No obstante, no ha habido un estudio más pormenorizado hasta la última década, cuando el concepto de rizosfera ha tomado mayor importancia[8]. A pesar de que las PGPR componen cerca del 5% de los microorganismos presentes en la rizosfera[61], el estudio de la respuesta de vegetales inoculados por estas bacterias ha despertado su interés como una posible

herramienta en la agricultura. Además, mediante técnicas moleculares se han podido identificar las distintas cepas que favorecen el crecimiento de la planta.

No todas las bacterias PGPR interaccionan del mismo modo con la planta cercana. Así, las ePGPR son de vida libre, colonizando la planta en el espacio extracelular, bien en la rizosfera, en la superficie de la raíz (rizoplano), o en espacios intercelulares. Mientras, aquellas que habitan dentro de las células de la raíz de la propia planta son denominadas iPGPR[19]. Entre los géneros encuadrados en ePGPR se encuentran, *Agrobacterium*, *Arthrobacter*, *Azotobacter*, *Azospirillum*, *Bacillus*, *Burkholderia*, *Erwinia*, *Flavobacterium*, *Pseudomonas* y *Serratia*. *Allorhizobium*, *Azorhizobium*, *Bradyrhizobium*, *Mesorhizobium* o *Rhizobium* crecen en el interior de la raíz de la planta, formando parte del grupo denominado iPGPR[8].

Las PGPR afectan al crecimiento de la planta a distintos niveles[2], [13], [14]. Por un lado, los mecanismos directos mejoran la disponibilidad de nutrientes fijando nitrógeno atmosférico, solubilizando fósforo o potasio, y segregando sustancias como ácido indolacético, ACC deaminasa u otras sustancias volátiles que promueven la Resistencia Sistémica Adquirida[62], [63]. Por otro lado, mediante la competencia con otros microorganismos o la producción de metabolitos secundarios, mejoran el crecimiento de la planta de forma indirecta, [64].

En este trabajo se estudian varios de los métodos directos que emplean las bacterias PGPR para el crecimiento del vegetal, la **fijación de nitrógeno**, producción de **ácido indolacético**, participación en la ruta de **etileno**, producción de **óxido nítrico** y producción de elementos **volátiles** que desatan la **resistencia sistémica adquirida**. A continuación, se aproximan las proteínas de estudio que participan en dichos procesos y las principales cepas en las que se ha registrado esta actividad.

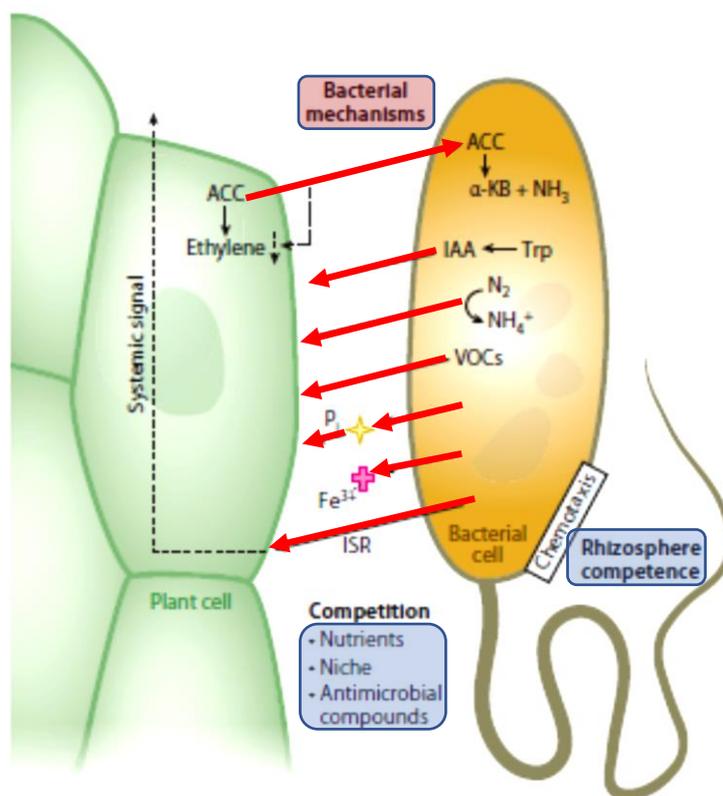


Figura 1. Mecanismos Directos e Indirectos de interacción Planta-PGPR. Esquema representativo de mecanismos directos (destacados en rojo) e indirectos (azul). A partir de Bulgarelli, 2013.

2.1.1 Nitrogenasa

La fijación biológica de nitrógeno atmosférico es un proceso clave en el ciclo de nutrientes de todo ecosistema. El nitrógeno es componente principal de los aminoácidos, por lo que es requerido, además de para otras funciones, para toda la estructura proteica de la planta [65]. La disponibilidad de nitrógeno en el suelo es limitada, por lo que mejorar la fijación biológica de N_2 atmosférico supone un gran avance en la fertilización del suelo. La nitrogenasa es una familia de proteínas encargada de catalizar la reacción N_2 en NH_4 . Existen diversas estructuras y homólogos de nitrogenasas [66].

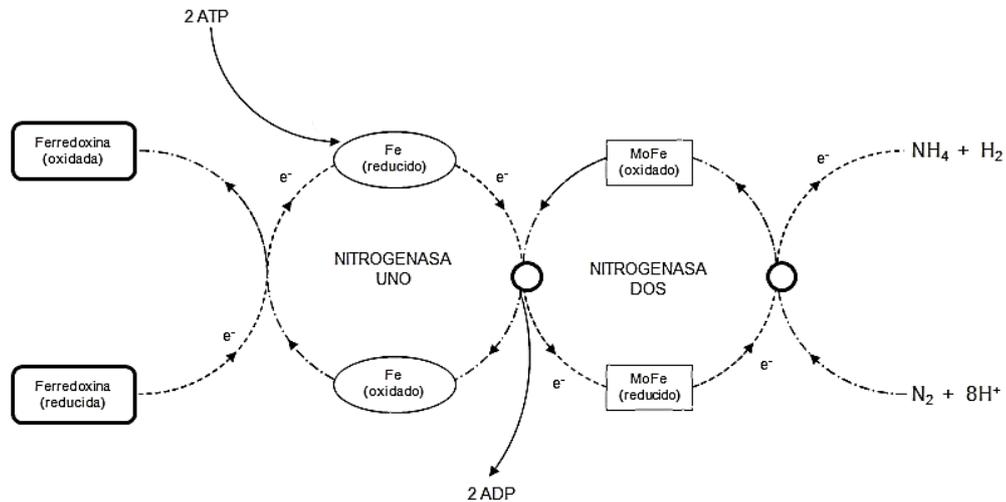


Figura 2. Complejo nitrogenasa, Nitrogenasa uno (Fe) y Nitrogenasa dos (Mo-Fe). Reacción de transformación de N_2 en NH_4 . Nitrogenasa

La Molibdeno (Mo) Nitrogenasa es la enzima mejor caracterizada del grupo, Figura 3 [27]. Está formada por dos subunidades metalosulfuro, la hierro-proteína (Fe prot, codificada por *nifH*) y la Molibdeno-hierro proteína (Mo-Fe prot). La subunidad Fe prot es un homodímero que contiene un lugar de unión para MgATP en cada una de sus subunidades. La parte catalítica de la proteína se concentra en la subunidad Me-Fe prot. Se trata de un tetrámero formado por dos cadenas α , codificadas por el gen *nifD* y dos β , cuya información se encuentra en el gen *nifK*[67].

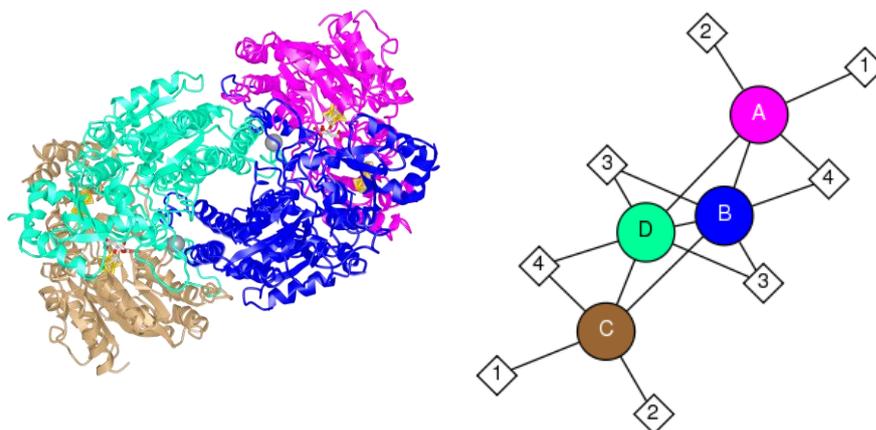


Figura 3. Estructura química de Nitrogenasa Mo-Fe de *Azotobacter vinelandii*, estado oxidado. Compuesta por Fe-prot (A y C) y Fe-Mo prot (B y D). 1: Ácido Homocítrico, 2: Grupo Fe-Mo-S, 3: Ca^{2+} y 4: Grupo Fe-S.

2.1.2 Indolpiruvato Descarboxilasa

Las bacterias promotoras del crecimiento vegetal pueden actuar a múltiples niveles. Uno de ellos es participar en la mediación del crecimiento mediante la producción de fitohormonas. Una de las hormonas producidas por las PGPR de forma generalizada es el Ácido Indolacético (IAA) [2]. Esta fitohormona está involucrada en múltiples funciones de proliferación celular como la germinación, aumento de la proporción de xilema, o el desarrollo radical y procesos vegetativos como mejorar la toma de nutrientes por la planta.

El principal precursor del Ácido Indolacético es el triptófano. Este aminoácido, que ha sido encontrado en exudados de la raíz, a través de la vía de Indol-3-piruvato y de Indol-3-acetaldehído es transformado a IAA por un gran número de bacterias. En este proceso la enzima Indolpiruvato descarboxilasa, *Figura 5* [27], codificada por el gen *ipdC*, transforma 3-indol-piruvato en indol-3-acetaldehído[68], [69].

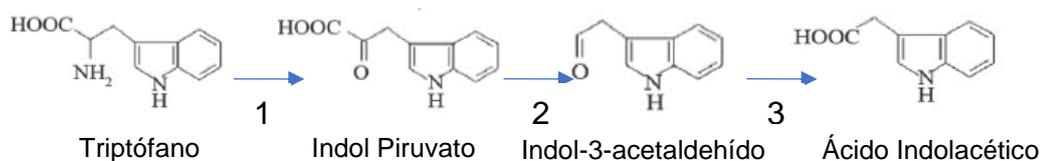


Figura 4. Síntesis de Ácido Indolacético a partir de Triptófano mediante Piruvato descarboxilasa. 1:Aminotransferasa, 2:Indol-3-piruvato descarboxilasa (IPDC), 3:Indol-3-acetaldehído deshidrogenasa.

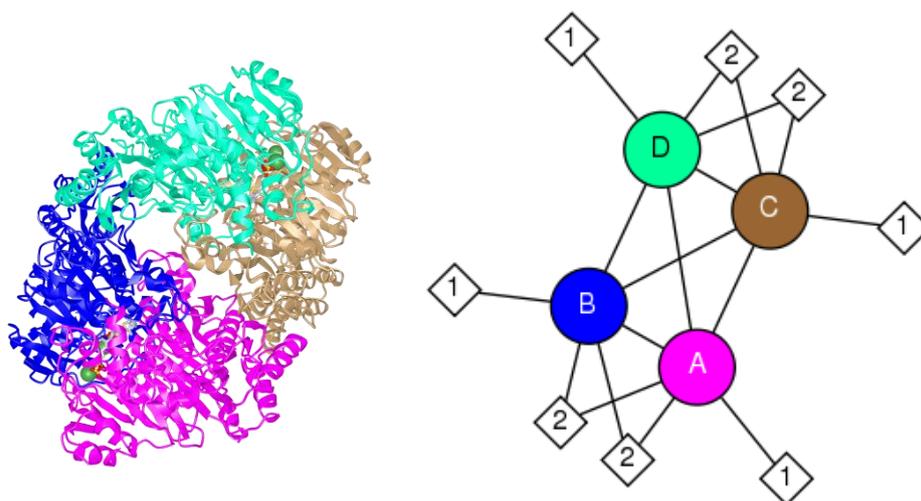


Figura 5. Indolpiruvato descarboxilasa de *Enterobacter cloacae*. Estructura de "dímero de dímeros" con estructura β - α - β . Dímero (A-B) unido con dímero (B-C). 1: Mg^{2+} , 2:Tiamina difosfato.

2.1.3 ACC Desaminasa

La producción de etileno es necesaria a lo largo de la vida de la planta, pues interviene en procesos como la elongación de la raíz o la maduración de fruto. No obstante, su producción excesiva está vinculada a situaciones de estrés en la planta provocando la senescencia de hojas o la abscisión del fruto.

Las PGPRs interfieren en la producción de etileno por la planta mediante la acción de la enzima 1-aminociclopropano-1-carboxílico desaminasa (ACC desaminasa), *Figura 7* [27]. La actividad de esta enzima impide el paso del ácido 1-aminociclopropano-1-carboxílico (ACC) en Etileno ya que desamina el compuesto para producir α -cetobutirato y amonio.

ACC deaminasa, codificada por el gen *acdS*, pertenece a las enzimas Dependientes de Piridoxal-Fosfato (PLP)[52], [70].

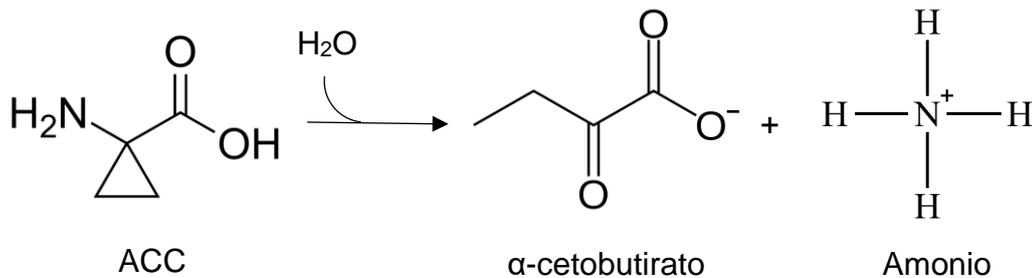


Figura 6. Reacción catalizada por la enzima ACC desaminasa.

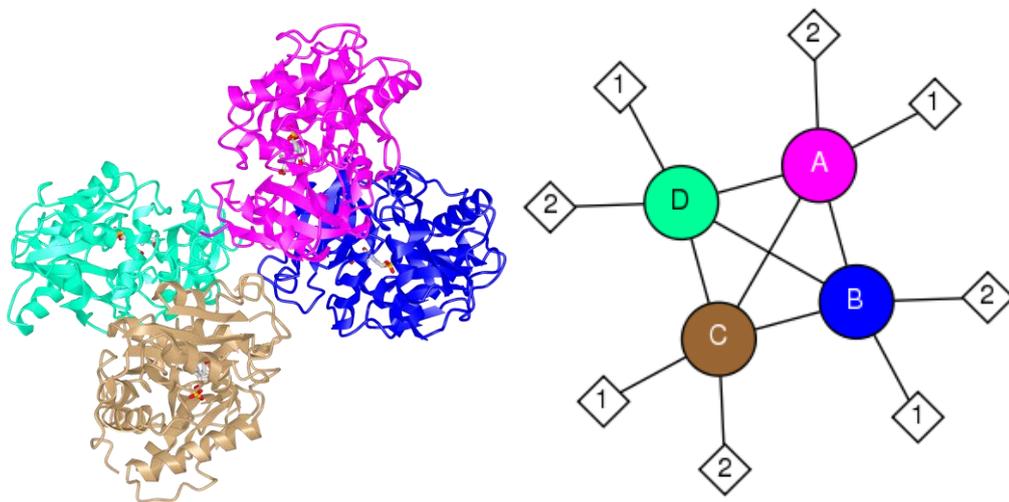


Figura 7. ACC desaminasa con sustrato de Pseudomonas sp. Formada por cuatro subunidades (ABCD). 1: Piridoxal-5-fosfato, 2: Ácido- 1-aminociclopropanecarboxílico.

2.1.4 Nitrito Reductasa

El óxido nítrico (NO) tiene un papel fundamental en el desarrollo radicular de las plantas, participa en la señalización celular a través de su interacción con auxinas en procesos de crecimiento y es, además, clave en la respuesta ante estrés abiótico o biótico[71].

La enzima Nitrito Reductasa de Cobre, *Figura 8* [27], reduce el nitrito a óxido nítrico en el proceso de desnitrificación. Está formada por un homotrímero codificado por el gen *nirK*, en el que cada uno de sus elementos posee centros de cobre implicados en el transporte de electrones y otros asociado al centro activo [72].



Figura 9. Reacción catalizada por la enzima Nitrito Reductasa.

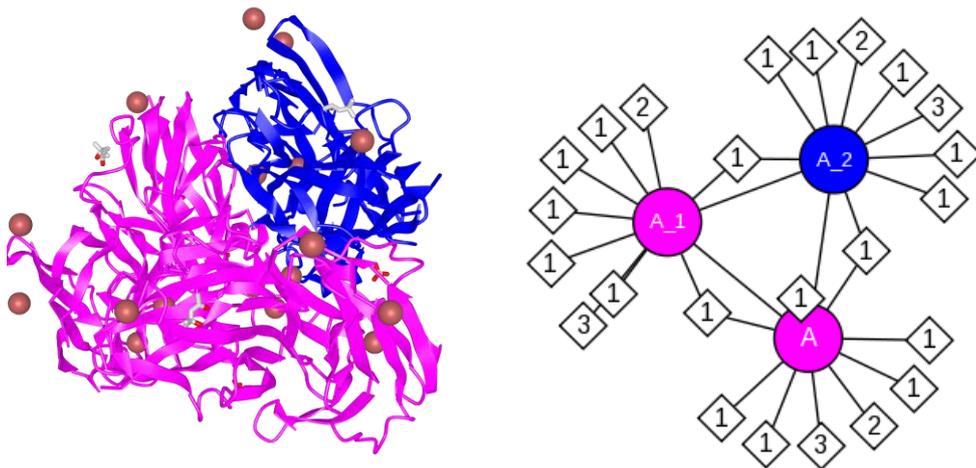


Figura 8. Nitrito Reductasa de Cobre de Geobacillus thermodenitrificans. Homotrímero (A-A_1-A_2). 1: Cu²⁺, 2: (4S)-metil-2,4-pentanediol, 3: SO₄²⁻

2.1.5 Ruta de 2-3-butanodiol

La Respuesta Sistémica Adquirida son distintas vías de transducción de señales que acaban por generar una respuesta en la planta que la protege contra ciertos patógenos.

Los genes *budA*, *budB* y *budC* codifican para el operon responsable de la producción de los compuestos volátiles 3-hidroxi-butanona (acetoina) y 2-3-butanodiol que promueven el crecimiento vegetal[24], [73].

La Acetolactato Sintasa, *Figura 10* [27], *budB*, convierte piruvato en acetolactato que, a través de Acetoin Decarboxilasa, *budA* lo transforma en 3-hidroxi-butanona. Tras este paso, el producto generado puede ser liberado por la PGPR o bien, mediante la enzima Diacetil Reductasa, *Figura 11* [27], *budC*, ser transformada en 2,3-butanodiol.

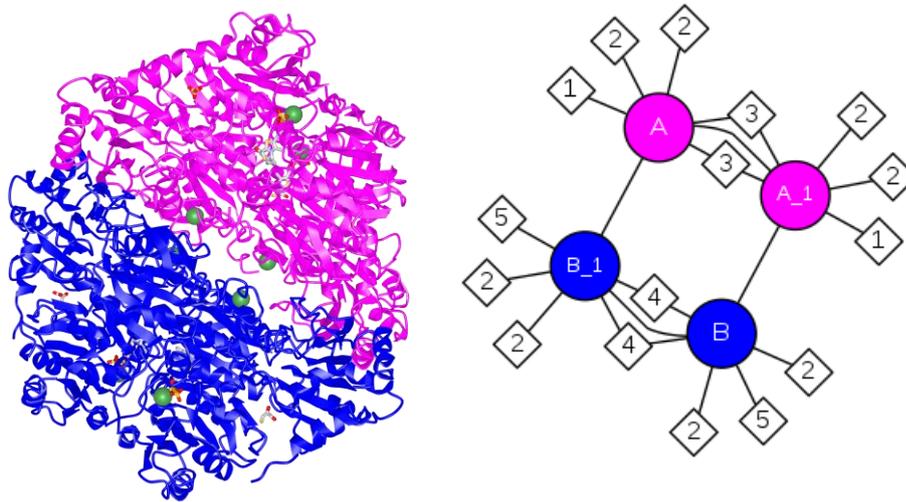


Figura 10. Acetolactato Sintasa de *Klebsiella pneumoniae*. Cuatro proteínas (AA_1BB_1). 1: PO_4^{2-} , 2: Mg^{2+} , 3: Complejo fluorado, 4: Tiamina difosfato, 5: Ácido 3-fluoro-2-oxopropanoico

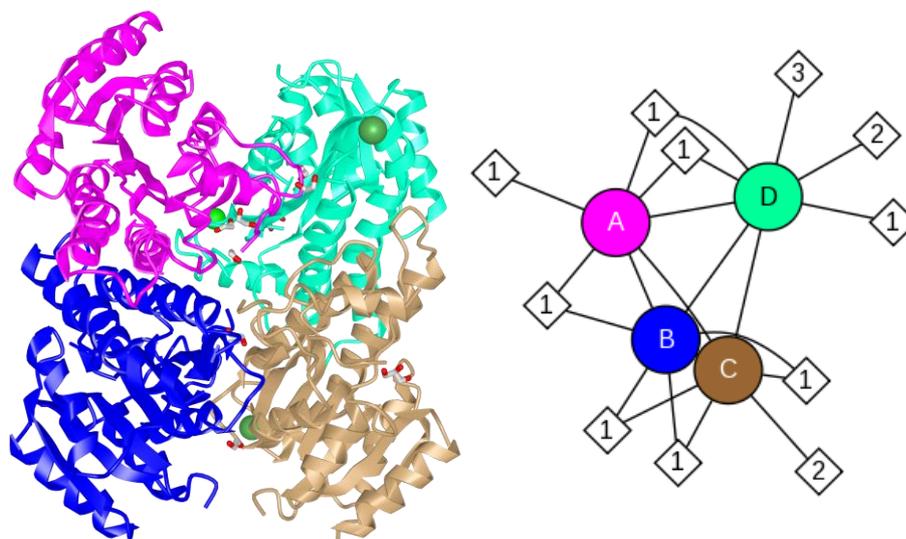


Figura 11. Acetoin Reductasa de *Burkholderia cenocepacia*. Tetrámero (ABCD). 1: 1-2 etanodiol, 2: Glicerol, 3: Mg^{2+} , 4: Cl.

2.2. CAPÍTULO 2. RELACIÓN FILOGENÉTICA DE PGPRs.

El estudio de la diversidad de los seres vivos requiere de su clasificación para que, en base a determinadas características comunes, se generen agrupamientos que faciliten el descubrimiento de relaciones ancestrales entre las distintas especies. Con el estudio filogenético se comparan características compartidas por individuos de distintas especies y se clasifican según el grado de semejanza.

A lo largo de la historia, el análisis filogenético de los seres vivos ha sido elaborado mediante la comparación de caracteres identificables por el ojo humano, bien a simple vista o con la utilización de distintos instrumentos ópticos. Además, para el estudio de los elementos compartidos por seres vivos que se encuentran fuera del alcance de resolución óptica, como el caso de las bacterias, se han utilizado otros métodos como cultivos en distintos medios o en diversas condiciones ambientales.

El avance en técnicas bioquímicas ha permitido alcanzar la capacidad técnica para realizar estudios filogenéticos que sobrepasan esos caracteres fenotípicos como la resistencia a ciertos antibióticos o el crecimiento en determinados rangos de temperatura. Más aún, la secuenciación genética permite el estudio filogenético desde su origen, intercambios o mutaciones en las secuencias de nucleótidos que forman el ADN, la llamada filogenia molecular [32]. La similitud entre secuencias biológicas suele corresponder a similitud en su estructura o función.

Según el dogma central de la biología molecular, la herencia transmitida en forma de ADN codificante es transcrita a ARN y posteriormente traducida a Proteína. A pesar de que la información contenida en el ADN es transcrita a ARN mensajero con la sustitución de Timina por Uracilo, en el siguiente paso existe cierto nivel de variabilidad. La traducción de ARN a proteína tiene cierto nivel de tolerancia por el que un mismo aminoácido puede estar codificado por varios codones distintos, se trata de un sistema “degenerado”. Este sistema de protección permite que, aún pudiendo sufrir algún tipo de mutación en los pasos anteriores, la proteína generada no sufrirá modificación alguna y podrá cumplir su función de forma adecuada. El cambio genético que derive en la alteración del aminoácido codificado, y su permanencia a través de siguientes generaciones será una nueva ramificación evolutiva de la proteína codificada. Es este el motivo por el cual, para estudios filogenéticos de ADN codificante, se recomienda utilizar la semejanza entre proteínas y no entre secuencias de nucleótidos.

Material y métodos

Las secuencias de estudio se han elegido mediante el repaso de la bibliografía en la que se revisan los principales géneros englobados como PGPR [8], [24], [26], [74]. Los géneros más estudiados pertenecen a *Azospirillum*, *Gluconacetobacter*, *Pseudomonas* y *Rhizobium*, aunque algunos gram + están también considerados como *Bacillus* o *Paenibacillus*. No obstante, el análisis a partir de medios bibliográficos, como es el caso, contiene el sesgo de aquellos géneros más estudiados. Así, la mayoría de las cepas consideradas como PGPR serán aquellas que promueven la fijación de nitrógeno, ya que la concentración de nitrógeno absorbible por la planta se ha concebido históricamente como el mayor factor limitante [64].

La Tabla 1 muestra los datos de las proteínas de interés de cada una de las cepas estudiadas. A pesar de estar incluidos en la bibliografía, para el posterior análisis se han seleccionado únicamente cepas que puedan ser utilizadas como biofertilizante, excluyéndose los organismos patógenos.

Con el programa *Protein Blast*, BLASTP [31] se lanzó una búsqueda con los principales géneros y se obtuvo el código de acceso para *National Center for Biotechnology Information* (NCBI) [27] de las proteínas de interés, Anexo 3.1. Posteriormente, mediante la utilización de R Software [38] se realiza el resto de obtención y tratamiento de datos. El *script* utilizado forma parte del Anexo IV. *Script* de R. Las funciones *entrez_search*, *entrez_fetch* del paquete *rentrez* [41] permiten la búsqueda y descarga de las secuencias de interés a partir de la base de datos de NCBI [27]. Los datos fuente son *genbank* y *refseq* [28], [29]. Se generó un bucle que, lanzando la búsqueda mediante el identificador de cada proteína, descargó la secuencia y su referencia de forma automática para guardarla en un archivo *fasta*. En el Anexo I. *Tablas de cepas*, se incluyen las tablas con los datos identificativos y referencias por proteína de interés.

Para conocer la distribución de los genes dentro de un mismo genoma, coocurrencia, se lanzó la función *crossprod* del paquete *base* de R [38]. Tras calcular el número de veces en que coexisten y se generó un gráfico (*Figura 13*) con la ayuda del paquete *igraph*[75].

Con las secuencias descargadas se ejecutó un alineamiento CLUSTALW [35] con la función *msa* [34] para, posteriormente calcular la matriz de distancias con la función *dist.alignment* del paquete *seqinr* [40].

El método CLUSTALW se encuadra en los métodos de alineamiento progresivo. Estos calculan alineamientos de pares entre las secuencias y elige el mejor alineamiento, progresivamente irá añadiendo más secuencias. Con la implementación del algoritmo de Feng y Doolittle [76] se cumplen tres pasos clave, primero, se genera un alineamiento global dos a dos donde las puntuaciones de similitud generan la matriz de distancias. Segundo, se genera un árbol guía a partir de esta matriz y, por último, se crea el alineamiento múltiple haciendo alineamiento de pares según las distancias. El número de secuencias totales sería una Variación sin Repetición de n secuencias tomadas de dos en dos (I).

$$(I) \quad V_{n,2} = \frac{n!}{2*(n-2)!}$$

La conversión de similitud a distancia según Feng y Doolittle utilizada en CLUSTALW utiliza la similitud entre secuencias, $S_{real(ij)}$ (II), la media de similitudes calculadas para las dos secuencias aleatorizadas, $S_{rand(ij)}$, y la media de las similitudes identidad calculada, $S_{iden(ij)}$ (III). Como resultado se obtiene la distancia entre secuencias, $D_{(ij)}$, (IV).

$$(II) \quad S_{iden(ij)} = \frac{S_{real(ii)} + S_{real(jj)}}{2}$$

$$(III) \quad S_{eff(ij)} = \frac{S_{real(ij)} + S_{rand(ij)}}{S_{iden(ij)} + S_{rand(ij)}} * 100$$

$$(IV) \quad D_{ij} = -\ln S_{eff}(ij)$$

Mediante otro tipo de alineamientos como MUSCLE [34], con el mismo paquete, se corroboró la calidad del alineamiento, pues no existían variaciones en los resultados.

Con la matriz de distancias se generaron los distintos árboles que se comentan en los apartados de este capítulo. Para ello, se utilizó la función *nj* que usa el método de vecinos más cercanos del paquete *ape*[32]. Este método parte de un árbol en el que todos los tips, cepas en nuestro caso, están conectados a un solo nodo formando una estrella, todos son vecinos. Posteriormente, se hacen las comparaciones entre los vecinos para conocer cuál es la pareja más cercana que se unirán mediante un nodo interno nuevo. Una vez generado el nodo, se repite el proceso, considerando la distancia del nuevo nodo generado con respecto al resto de tips[77].

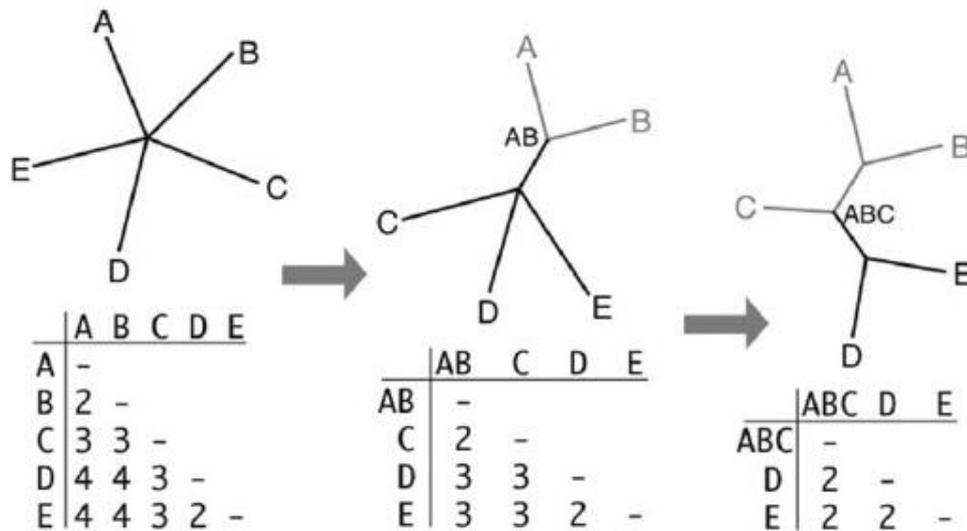


Figura 12. Procedimiento Neighbour Joining. Extraído de Dunn, 2011.

Resultados

En la Tabla 1 se muestran las cepas de interés obtenidas a partir de la bibliografía. Mediante un código de colores se diferencia el grupo al que pertenece cada uno de los géneros según la descripción de cada cepa para facilitar la comprensión de los árboles generados.

Se contabilizan un total de 34 alfaproteobacteria, 4 betaproteobacteria, 36 gammaproteobacteria y, por último, 15 y 11 cepas pertenecientes al grupo de firmicutes y gram positivas con alto contenido en G+C respectivamente.

Tabla 1. Lista de Cepas introducidas en el estudio y genes incluidos para cada cepa. *NP = No publicado.

	Firmicutes		α Proteobacteria		β Proteobacteria		γ Proteobacteria		Gram + Alto [G+C]		Autor
	nifH	nifD	nifK	ipdC/ppdC	acdS	nirK	budA	budB	budC		
<i>Acinetobacter baumannii</i> ABNIH24					+						Snitkin,E.S.
<i>Acinetobacter gyllenbergii</i> MTCC 11365					+						Singh,N.K.
<i>Acinetobacter pittii</i>				+	+						Brasiliense,D.M.
<i>Acinetobacter pittii</i> PHEA-2									+		Zhan,Y.
<i>Acinetobacter</i> sp. P7-29	+										Arshad,M.
<i>Acinetobacter</i> sp. P8-1	+										Arshad,M.
<i>Acinetobacter</i> sp. P8-9	+										Arshad,M.
<i>Agrobacterium rhizogenes</i>					+	+					Xiang,T.
<i>Agrobacterium</i> sp.	+										Jha,B.
<i>Agrobacterium</i> sp. P7-7	+										Arshad,M.
<i>Agrobacterium</i> sp. SUL3				+	+	+					Jones,K.J.
<i>Agrobacterium tumefaciens</i>	+				+	+					Poret-Peterson,A.T.
<i>Arthrobacter</i> sp. PAMC 25486				+							Jung,J.-H.
<i>Arthrobacter alpinus</i>					+						See-Too,W.S.
<i>Arthrobacter crystallopoietes</i> BAB-32					+						Joshi,M.N.
<i>Arthrobacter</i> sp. CCBAU 10948	+										Luo,W.
<i>Arthrobacter</i> sp. ERGS1:01					+						Kumar,R.
<i>Arthrobacter subterraneus</i>						+					Varghese,N.
<i>Azoarcus</i> sp. KH32C	+	+	+	+		+					Nishizawa,T.
<i>Azoarcus</i> sp. PA01]				+							Junghare,M.
<i>Azoarcus toluclasticus</i>	+			+							NP*
<i>Azoarcus tolulyticus</i>	+	+	+	+							Varghese,N.
<i>Azospirillum brasilense</i>					+	+			+		NP*
<i>Azospirillum brasilense</i> Sp245	+	+	+	+		+					Pothier,J.F.
<i>Azospirillum lipoferum</i>				+	+	+					NP*
<i>Azospirillum lipoferum</i> 4B	+	+	+						+		Wisniewski-Dye,F.
<i>Azospirillum</i> sp. B510	+	+	+		+	+					Kaneko,T.
<i>Azotobacter chroococcum</i> NCIMB 8003	+	+	+		+						Robson,R.L.
<i>Azotobacter vinelandii</i> DJ	+	+	+		+						Setubal,J.C.
<i>Bacillus cereus</i>				+	+		+		+		Hall,Neil.
<i>Bacillus mycoides</i>					+		+				Bleich,R.M.
<i>Bacillus solani</i>						+					Liu,B.

	nifH	nifD	nifK	ipdC/ppdC	acdS	nirK	budA	budB	budC	Autor
Bacillus sporothermodurans				+						Krawczyk,A.O.
Bacillus spp	+	+	+							NP*
Bacillus thuringiensis				+	+		+			Hollensteiner,J.
Bradyrhizobium canariense	+	+	+		+					Chekireb,D.
Bradyrhizobium japonicum				+	+	+				NP*
Bradyrhizobium japonicum SEMIA 5079	+	+	+			+				Siqueira,A.F.
Bradyrhizobium oligotrophicum S58				+	+	+				Okubo,T.
Bradyrhizobium ottawaense	+	+	+							Yu,X.
Bradyrhizobium sp. ORS 375				+	+	+				Mornico,D.
Bradyrhizobium sp. PARBB1							+			Kojadinovic,M.
Bradyrhizobium sp. STM 3809				+	+	+				Mornico,D.
Dickeya chrysanthemi	+	+	+	+	+					Chan,K.-G.
Dickeya dadantii 3937							+	+		Glasner,J.D.
Dickeya solani				+	+					Chan,K.-G.
Dickeya zeae				+	+					Lucas,S.
Erwinia amylovora				+				+		Lagonenko,A.L.
Erwinia amylovora ATCC 49946							+			Sebahia,M.
Erwinia amylovora CFBP 1232	+									Mann,R.A.
Erwinia amylovora CFBP1430					+					Smits,T.H.
Erwinia amylovora MR1	+			+						Mann,R.A.
Erwinia gerundensis									+	Blom,Jochen.
Frankia alni					+					Normand,P.
Frankia alni ACN14a	+	+	+							Normand,P.
Frankia inefficax				+	+					Lucas,S.
Frankia sp. EUN1f	+	+	+		+					Lucas,S.
Frankia sp. QA3	+	+	+		+					Lucas,S.
Gluconacetobacter diazotrophicus PA1 5	+	+	+	+						Giongo,A.
Gluconacetobacter sp. SXCC-1				+						Du,X.J.
Klebsiella michiganensis KCTC 1686				+						Shin,S.H.
Klebsiella sp. 1_1_55	+	+	+				+			Ward,D.
Methylobacterium nodulans ORS 2060	+	+	+		+					Lucas,S.
Methylobacterium radiotolerans								+		Adelskov,J.
Methylobacterium sp. 4-46	+	+	+		+					Copeland,A.
Methylobacterium sp. WSM2598	+	+	+		+					NP*
Paenibacillus amylolyticus						+				Akuzawa,S.

	nifH	nifD	nifK	ipdC/ppdC	acdS	nirK	budA	budB	budC	Autor
Paenibacillus ATCC 35681	+	+	+							Halim,M.A.
Paenibacillus barengoltzii G22						+				Earl,A.
Paenibacillus durus	+	+	+							den Bakker,H.C.
Paenibacillus graminis	+	+	+							den Bakker,H.C.
Paenibacillus mucilaginosus KNP414				+						Lu,J.J.
Paenibacillus odorifer	+	+	+							den Bakker,H.C.
Paenibacillus polymyxa				+						Phi,Q.T.
Paenibacillus sp. P1XP2									+	Adelskov,J.
Pantoea ananatis									+	Sheibani-Tezerji,R.
Pantoea ananatis LMG 20103				+			+	+		De Maayer,P.
Pantoea sp. At-9b	+	+	+		+					Lucas,S.
Pantoea sp. CFSAN033090				+						Gonzalez-Escalona,N.
Pectobacterium carotovorum	+	+	+	+			+	+		Panda,P.
Pseudomonas fluorescens AU11706					+	+			+	NP*
Pseudomonas putida				+	+	+		+		Guo,J.
Pseudomonas sp. 34 E 7				+	+					Chandra,Govind.
Pseudomonas sp. 37 R 15				+	+					Chandra,Govind.
Pseudomonas sp. B4(2012c)	+									Agrawal,K.
Pseudomonas stutzeri KOS6	+		+							Grigoryeva,T.V.
Pseudomonas stutzeri RCH2						+				Lucas,S.
Rhizobium etli	+	+	+	+	+	+				Bustos,P.
Rhizobium gallicum				+	+	+				Bustos,P.
Rhizobium leguminosarum	+	+	+	+	+	+				Lucas,S.
Rhizobium leguminosarum bv. viciae 3841								+		Young,J.P.
Serratia liquefaciens FK01				+						Taira,E.
Serratia sp. ATCC 39006	+	+	+		+			+		Hampton,H.G.
Serratia sp. M24T3					+					Proenca,D.N.
Sinorhizobium medicae				+	+	+				Reeve,W.
Sinorhizobium meliloti				+	+	+				Barnett,M.J.
Sinorhizobium meliloti 1021						+				Barnett,M.J.
Sinorhizobium meliloti AK83	+	+	+							Galardini,M.
Sinorhizobium meliloti BL225C	+	+	+		+	+				Galardini,M.
Sinorhizobium saheli	+	+	+		+					Yan,H.

En la Tabla 2 se muestra el porcentaje de contenido por grupo por cada gen. El grupo de alfaproteobacteria lidera presencia para cualquiera de los genes, excepto para el grupo *budABC*. Sin embargo, como se ha mencionado anteriormente, puede deberse al sesgo por el uso de bibliografía que, históricamente, ha estado más detenida en el estudio de Rhizobacterias por su capacidad de fijación de nitrógeno. Es relevante también que la mayor parte de elementos volátiles de la ruta de 2-3 butanodiol se encuentra en el grupo de gammaproteobacterias.

Tabla 2. Presencia de genes por grupo de estudio. Porcentaje de cepas de cada grupo que contienen cada uno de los genes. Graduación de color con mínimos en verde, máximos en rojo.

Grupo	nifH	nifD	nifK	ipdC / ppdC	acdS	nirK	budA	budB	budC
Alfaproteobacteria	41%	47%	45%	36%	45%	70%	11%	25%	25%
Betaproteobacteria	7%	6%	6%	10%	0%	4%	0%	0%	0%
Gammaproteobacteria	32%	22%	24%	36%	34%	11%	56%	75%	50%
Firmicutes	11%	16%	15%	13%	6%	11%	33%	0%	25%
Gram+alto [G+C]	9%	9%	9%	5%	15%	4%	0%	0%	0%

Como se observa en la Tabla 3, el gen con mayor presencia es *acdS*, en 47 cepas distintas, mientras los de menor son *budB* y *budC* que únicamente aparecen en 8 de las cepas de estudio. La mayor cantidad de coocurrencias se da entre los genes formadores de nitrogenasa, algo lógico de pensar al tratarse de una proteína polimérica. Existe también una coocurrencia alta entre los genes *acdS*, *ipdC/ppdC* y *nirK* codificantes de las enzimas ACC Desaminasa, Indolpiruvato Descarboxilasa y Nitrito Reductasa.

Tabla 3. Coocurrencia de genes estudiados.

Gen	nifH	nifD	nifK	ipdC / ppdC	acdS	nirK	budA	budB	budC
nifH	44	33	34	12	21	11	2	2	1
nifD	33	32	32	9	18	9	2	2	1
nifK	34	32	33	9	18	9	2	2	1
ipdC/ppdC	12	9	9	39	26	21	4	4	1
acdS	21	18	18	26	47	23	3	2	3
nirK	11	9	9	21	23	27	0	1	2
budA	2	2	2	4	3	0	9	3	1
budB	2	2	2	4	2	1	3	8	0
budC	1	1	1	1	3	2	1	0	8

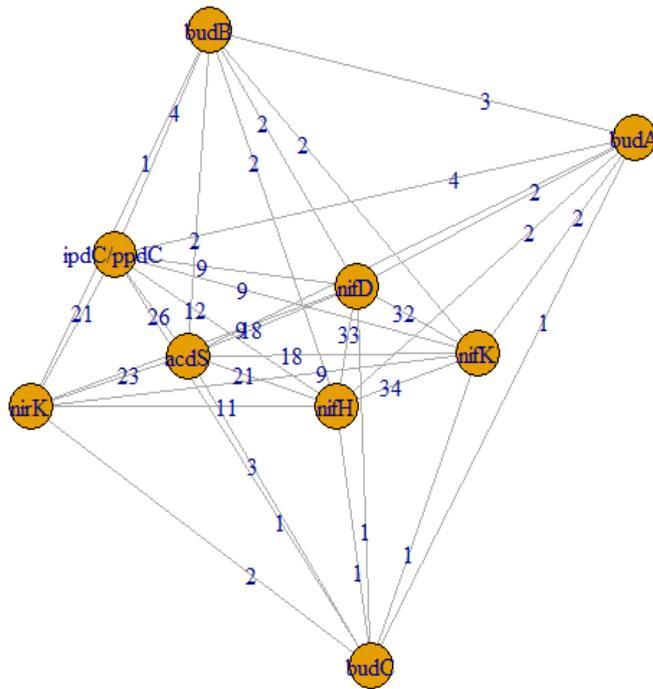
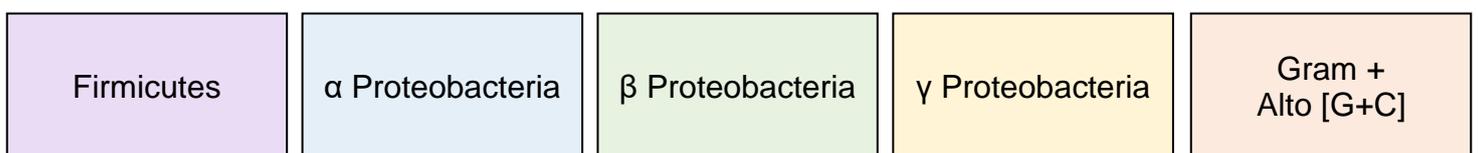


Figura 13. Coocurrencia de genes estudiados.

Para una adecuada evaluación de los árboles filogenéticos generados se necesita la verificación de su significación biológica. A pesar de esto, el método de vecinos más cercanos no suele desviarse de otros modelos más complejos, por lo tanto más lentos, basados en caracteres como máxima parsimonia y máxima similitud [32].

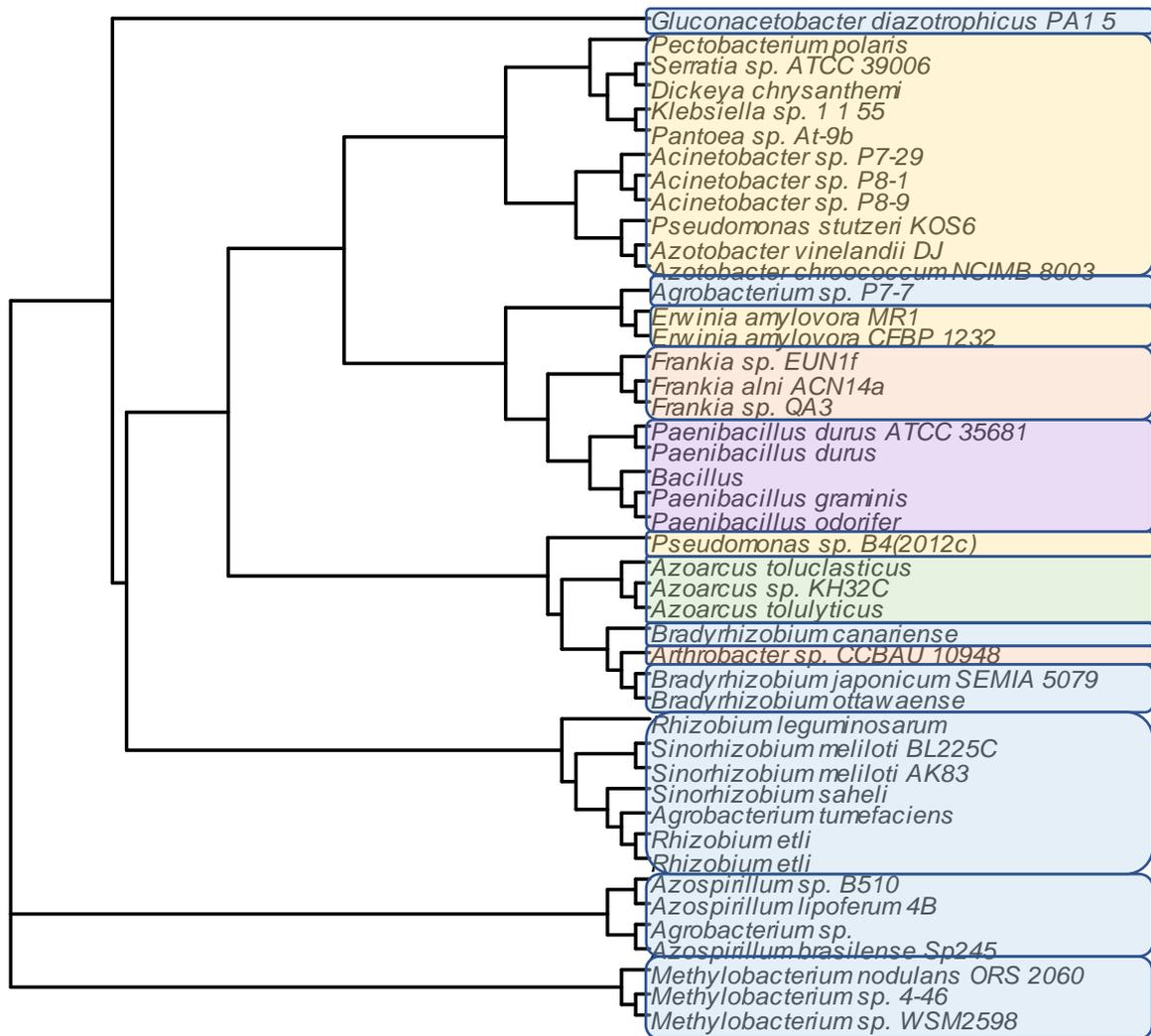
A continuación, se presentan los árboles filogenéticos generados para cada una de las proteínas de estudio. Del mismo modo que en la Tabla 1 se clasifican por colores la pertenencia a los distintos grupos bacterianos, en los árboles se utiliza el mismo código de colores.



2.2.1 NITROGENASA

Se han estudiado de forma separada los genes responsables de la fijación de nitrógeno. Se estudian las proteínas codificadas por los genes *nifD* y *nifK*, para la proteína Molibdeno-Hierro (*Mo-Fe Protein*), mientras que *nifH* codifica para la proteína de Hierro (*Fe Protein*).

A continuación, se comenta cada uno de los árboles según los genes codificantes. Sin embargo, de forma global según un compendio de las proteínas de cada gen, se puede decir que las primeras divisiones en el árbol se corresponden con cepas pertenecientes al grupo de alfa proteobacteria, siendo así evidente la separación entre los géneros simbiotes (Rhizobiales) con respecto aquellos de vida libre.



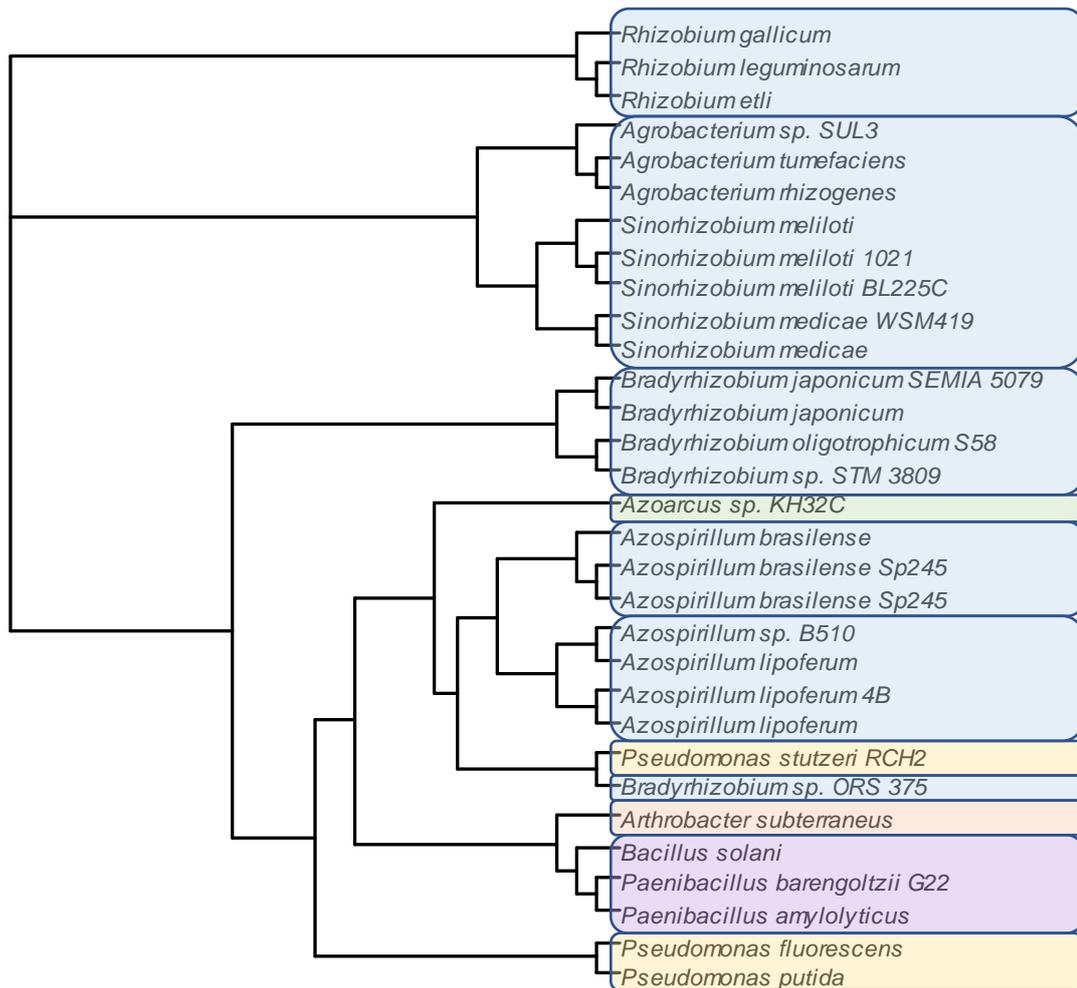
Árbol Filogenético 1. Fe Protein - nifH

En el árbol generado, en el primero de los nodos, se diferencian tres ramas. Se separan del resto de grupos los géneros *Methylobacterium* por un lado, y *Azospirillum* y *Agrobacterium* por otro. A su vez, en la separación interna de *Azospirillum* con *Agrobacterium*, este último género es más semejante a *Azospirillum brasilense* Sp245 que a *Azospirillum lipoferum* 4B y *Azospirillum* sp. B510.

En la rama más poblada, se separa *Gluconacetobacter diazotrophicus* PA15 del resto. En el siguiente nodo, el grupo de *Rizhobium* y *Sinorhizobium* se diferencian. Dentro de este grupo de iGPR se encuentra también *Agrobacterium tumefaciens*, aparentemente más cercana a *Rhizobium* que al grupo descrito de sus congéneres.

En el siguiente nodo, se divide el grupo de *Bradyrhizobium* – *Arthrobacter* - *Azoarcus* - *Pseudomonas* sp B4(2012c) del resto. La siguiente división es más equilibrada en cuanto a número de cepas, se genera un grupo de *Azotobacter* - *Pseudomonas* - *Acinetobacter* con *Pantoea* – *Klebsiella* – *Dyckeya* – *Serratia* – *Pectobacterium* por un lado, y *Agrobacterium* – *Erwinia*, *Frankia* – *Bacillus* - *Paenibacillus* por otro.

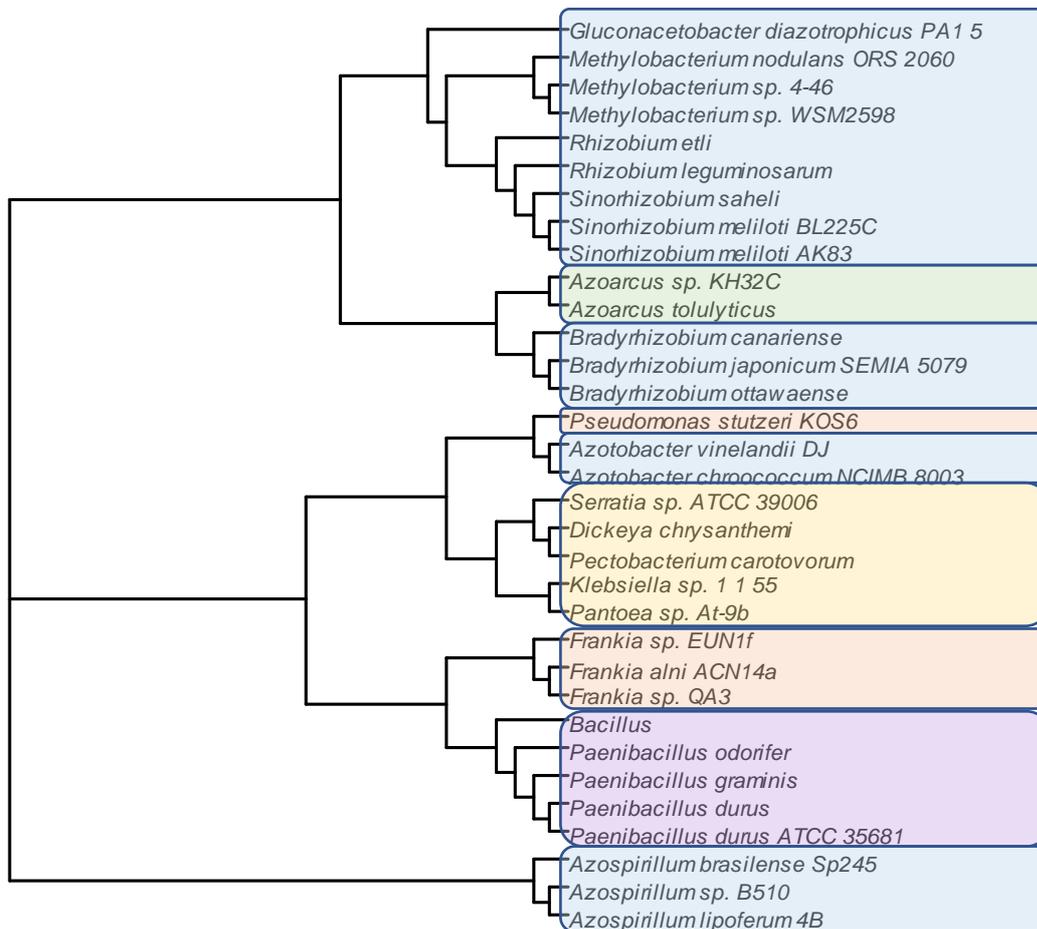
En resumen, destaca la cepa *Agrobacterium tumefaciens* que se encuentra en el grupo de *Rhizobium* mientras su congénere está clasificada junto con *Azospirillum*.



Árbol Filogenético 2.Fe-Mo Alpha - nifD

La primera separación se genera entre el grupo de *Rhizobium* por un lado y *Agrobacterium*-*Sinorhizobium* por otro, del resto de géneros. En el siguiente nodo *Bradyrhizobium* se separa. En la otra rama *Pseudomonas fluorescens* y *P. putida* se alejan del grupo formado por *Azospirillum*-*Azoarcus*-*Bradyrhizobium*-*Pseudomonas stutzeri* RCH2 y *Arthrobacter*-*Bacillus*-*Paenibacillus*.

En este caso de la Subunidad Alfa de la proteína Hierro-Molibdeno, aparece un grupo formado por *Pseudomonas stutzeri* RCH2 y *Bradyrhizobium* sp ORS375 descrito como cepas más cercanas a los géneros *Azospirillum* y a una distancia de 4 y 5 nodos respecto a *Pseudomonas* y *Bradyrhizobium*.



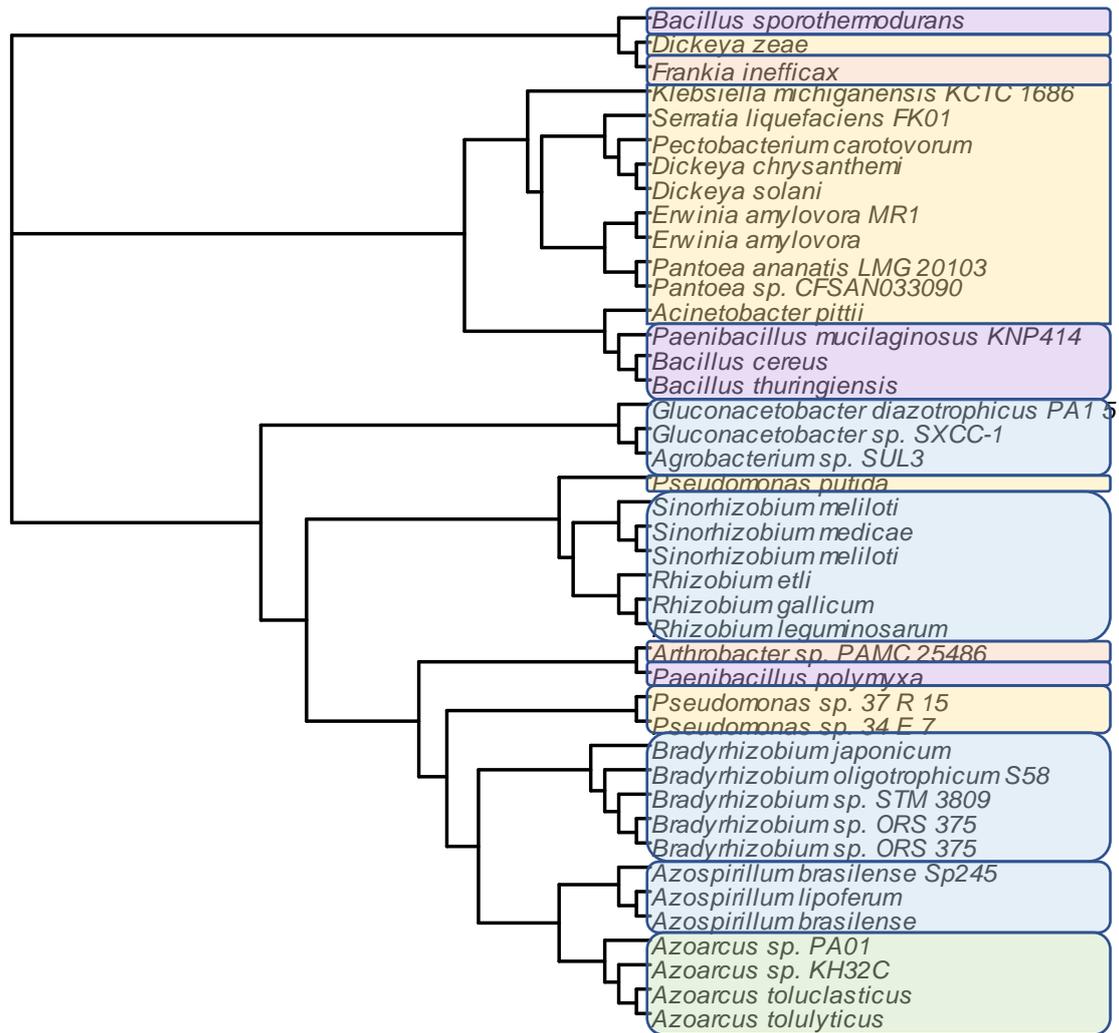
Árbol Filogenético 3.Mo - Fe Beta - nifK

Se detecta una separación inicial en tres grupos. El género *Azospirillum* en una rama única, mientras las otras dos contienen un número de cepas equilibrado. En la rama superior se separan *Azoarcus* – *Bradyrhizobium* por un lado y *Gluconacetobacter* y *Methylobacterium* - *Rhizobium* - *Sinorhizobium* por otro. En la rama central se separa el grupo de *Paenobacillus* - *Bacillus* - *Frankia* de *Pseudomonas*-*Azotobacter*, *Serratia* - *Dyckeya* - *Pectobacterium* y *Klebsiella* - *Pantoea*.

La semejanza en cada una de las proteínas del complejo nitrogenasa tiene ciertas variaciones entre los grupos de bacterias. Así, en la Fe-proteína, el primer distanciamiento es del género *Methylobacterium* y *Azospirillum* con el resto y el género *Bradyrhizobium* se encuentra más cercano a las cepas de gammaproteobacterias y gram+. Esta cercanía de *Bradyrhizobium* también se da en la subunidad alfa de la proteína Fe-Mo. Sin embargo, respecto a la subunidad beta, la cepa de alfa-proteobacterias más cercana a otros grupos bacterianos son del género *Azotobacter*.

En general se describe la singularidad de rizobiales con respecto al resto ya que describen mayor distanciamiento con respecto a beta y gammaproteobacterias y gram+.

2.2.2. INDOLPIRUVATO DESCARBOXILASA



Árbol Filogenético 4. Piruvato descarboxilasa - ipdC/ppdC

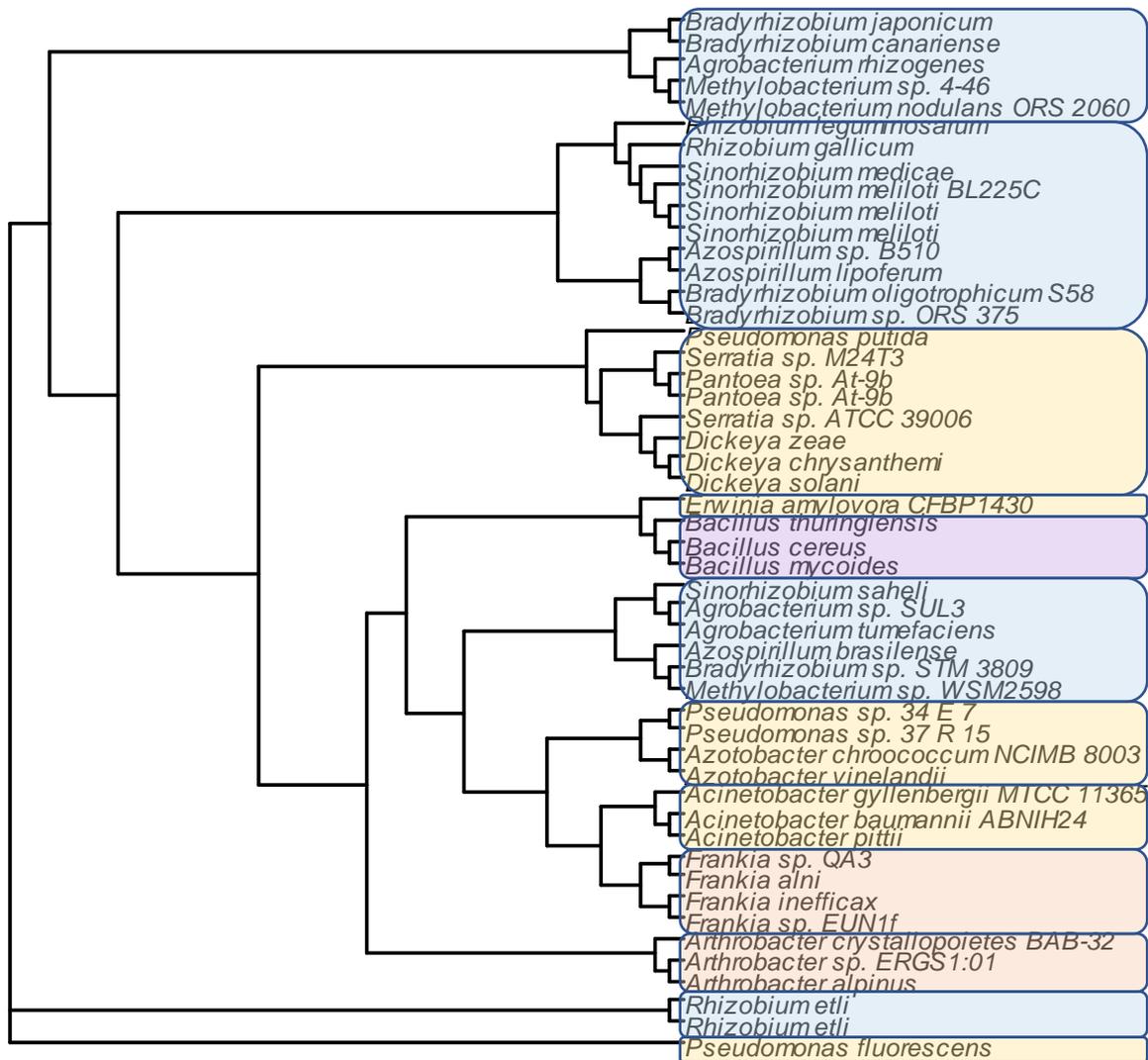
Bacillus sporothermodurans y *Dickeya inefficax*-*Frankia* forman un pequeño grupo distinto al del resto de *Bacillus* que se agrupan con *Paenibacillus* y *Acinetobacter pittii* dentro de la rama de *Klebsiella*, *Pectobacterium*-*Dickeya* y *Erwinia*-*Pantoea*. El grupo más numeroso es el formado por la rama situada en la parte inferior. *Gluconacetobacter*-*Agrobacterium* se separan en el siguiente nodo del resto. Posteriormente se diferencian *Pseudomonas putida* y *Sinorhizobium*-*Rhizobium* del grupo de *Arthrobacter* - *Paenibacillus* y *Pseudomonas* con *Bradyrhizobium*, *Azospirillum*-*Azoarcus*.

Destaca la primera separación de *Dickeya zeae* con respecto a *D. chrysanthemi* y *D. solani*, más cercanas a *Erwinia* y *Pantoea*. La cepa *Paenibacillus mucilaginosus* KNP414 forma parte de la rama central afín al género *Bacillus* mientras *Paenibacillus polymyxa* se agrupa dentro de la rama cercana a *Arthrobacter* - *Rhizobium* - *Bradyrhizobium* - *Azoarcus*.

Tras el tercer nodo, *Pseudomonas putida* aparece cercana a los géneros *Rhizobium*-*Sinorhizobium* y separada de las cepas *Pseudomonas* sp.34E7 y *P. putida* 37R15 más cercanas a *Bradyrhizobium* - *Azoarcus*.

En general, se aprecia la diferencia entre gammaproteobacterias, más cercanas a firmicutes, y alfa proteobacterias, más similares entre ellas.

2.2.3. ACC DEAMINASA



Árbol Filogenético 5. ACC deaminasa - acdS

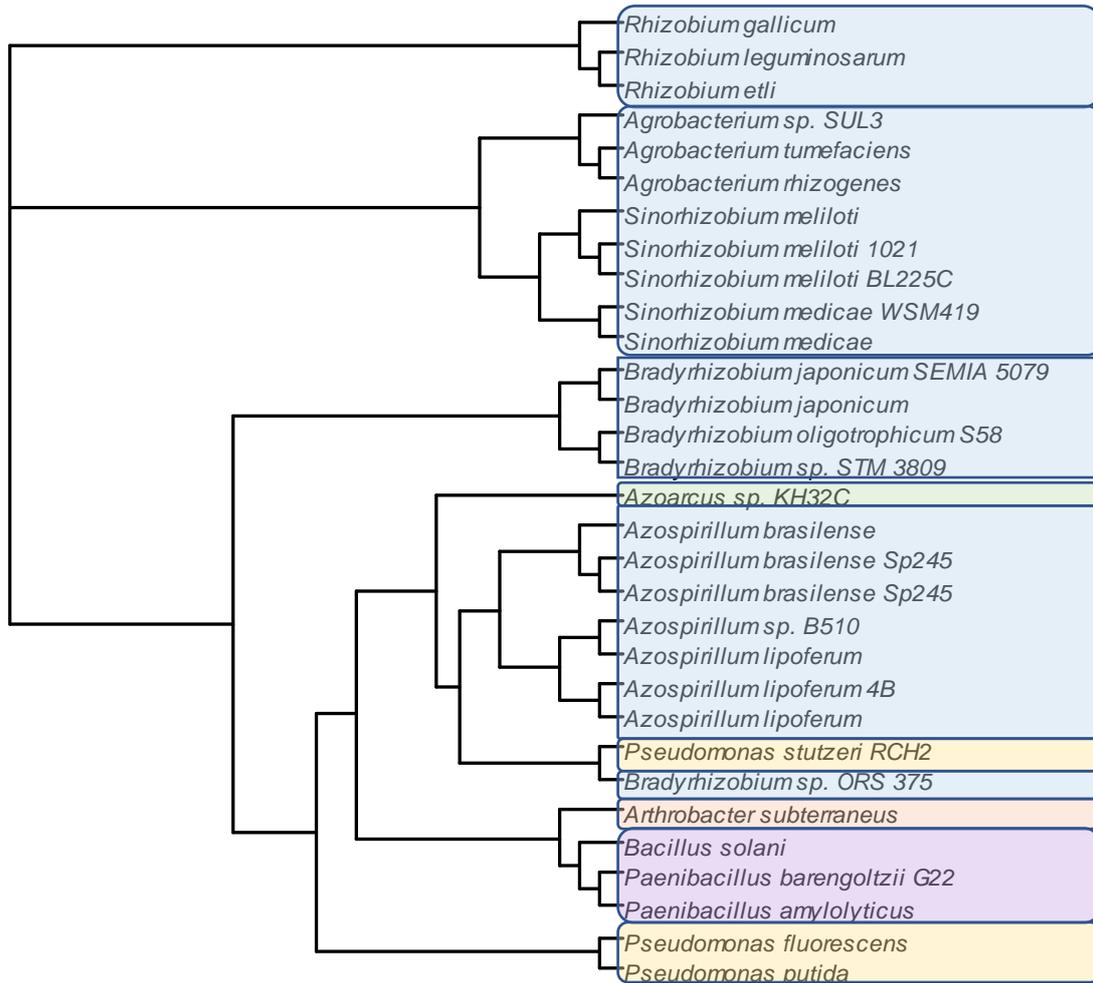
En base al estudio de la secuencia de aminoácidos de la proteína ACC deaminasa, encargada de mediar en la disminución de síntesis de etileno, *Rhizobium etli* y *Pseudomonas fluorescens* se encuentran separadas del resto.

Tras el primer nodo, *Bradyrhizobium japonicum* y *Bradyrhizobium canariense* forman un grupo conjuntamente con *Agrobacterium rhizogenes* y *Methylobacterium nodulans* ORS 2060 y *M. sp* 4-46 separado del resto.

En el segundo nodo se separan *Rhizobium* - *Sinorhizobium*, *Bradyrhizobium*-*Azospirillum* del resto. Tras esta bifurcación el grupo formado por *Dickeya*-*Serratia*-*Pantoea* en conjunto con *Pseudomonas putida* se separa de los grupos que en el siguiente nodo se separan en *Arthrobacter* y *Erwinia* - *Bacillus*, *Sinorhizobium*-*Agrobacterium*, más lejanos a *Azotobacter*-*Pseudomonas* y *Acinetobacter* - *Frankia*.

Existe discrepancia en el género *Pseudomonas*, mientras *Pseudomonas fluorescens* se encuentra separada del resto, *Pseudomonas putida* aparece cercana a los géneros *Serratia* y *Pantoea*.

2.2.4. NITRITO REDUCTASA



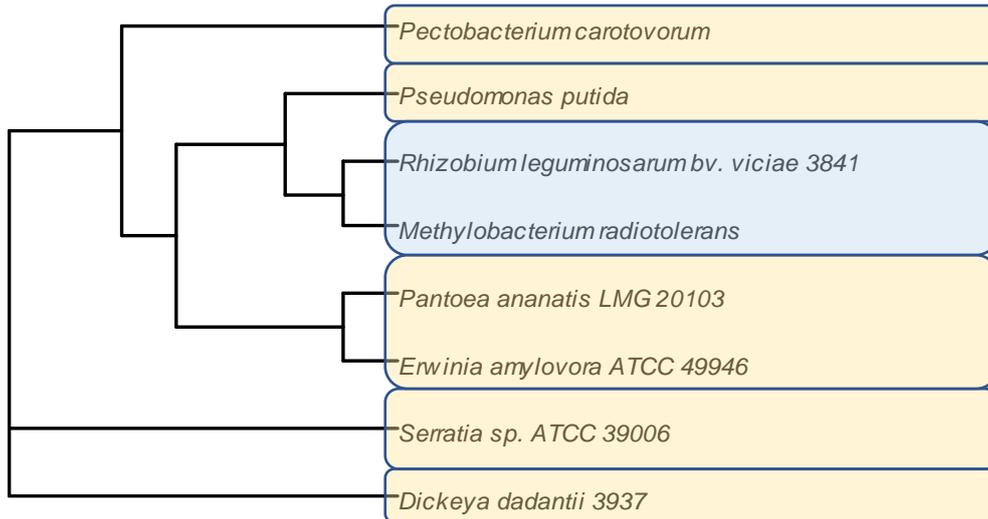
Árbol Filogenético 6. Nitrito Reductasa - nirK

En el árbol generado a partir de la secuencia de aminoácidos de Nitrito Reductasa, los géneros *Rhizobium* y *Sinorhizobium* - *Agrobacterium* aparecen separados del resto de cepas de estudio. En la rama con mayor número de cepas, se separa el género *Bradyrhizobium* del resto, en el siguiente nodo el género *Pseudomonas* se distancia. En el grupo restante *Arthrobacter* - *Bacillus*-*Paenibacillus* aparecen alejados de la rama formada por *Azoarcus* y *Azospirillum*-*Pseudomonas*-*Bradyrhizobium*.

Pseudomonas stutzeri RCH2 aparece más cercana al grupo de *Azospirillum* que al resto de sus congéneres.

2.2.5. RUTA DE 2-3-BUTANODIOL

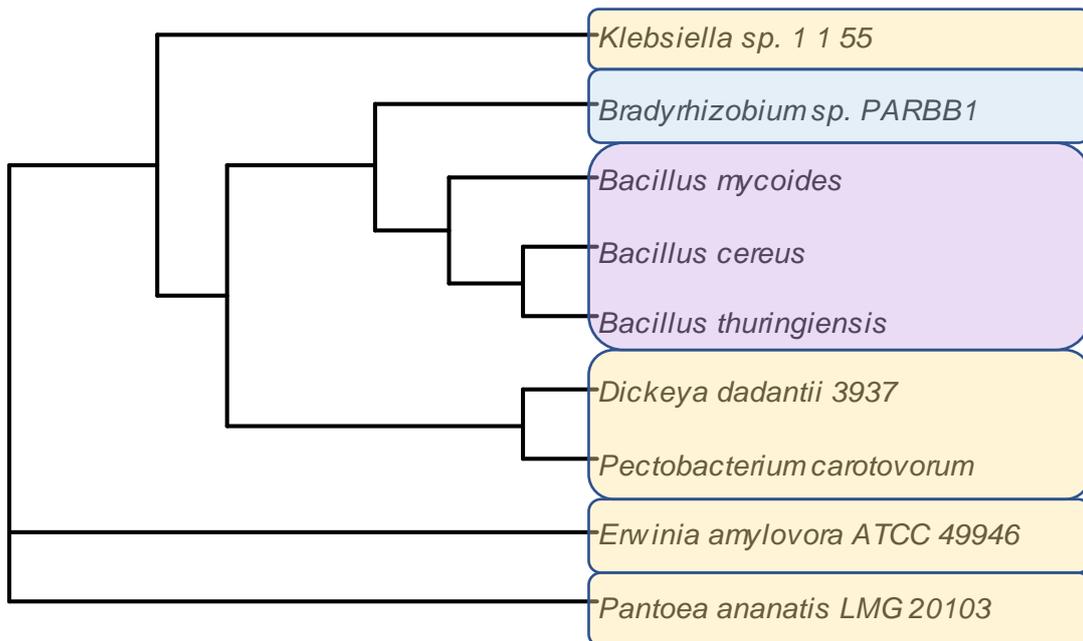
Acetolactato Sintasa



Árbol Filogenético 7. Acetolactato Sintasa - budB

Serratia y *Dickeya* aparecen distanciadas del resto de cepas en el árbol. El grupo principal se separa *Pectobacterium* respecto al resto. En el siguiente nodo, *Pseudomonas putida* forma un grupo junto *Rhizobium-Methylobacterium* separado de *Pantoea - Erwinia*.

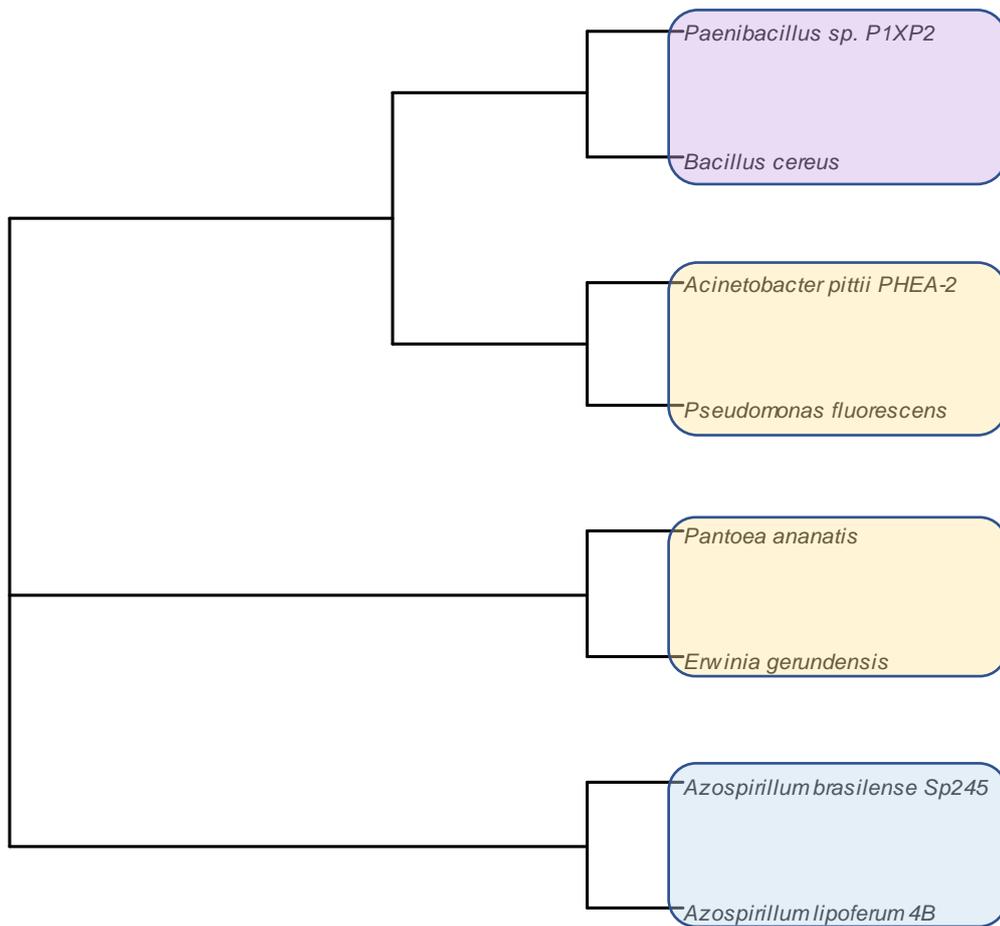
Acetolactato Descarboxilasa



Árbol Filogenético 8. Acetolactato Descarboxilasa - budA

Con respecto a la proteína Acetolactato Descarboxilasa codificada por el gen *budA* se *Erwinia* y *Pantoea* se encuentran alejadas del resto. En el grupo principal se separa *Klebsiella* para en un siguiente nodo formarse dos ramas separadas compuestas por un grupo *Dickeya - Pectobacterium* y *Bradyrhizobium-Bacillus*.

Diacetil Reductasa



Árbol Filogenético 9. Diacetil Reductasa - budC

En la comparativa de la proteína codificada por *budC*, Diacetil Reductasa, existe separación del género *Azospirillum*, por un lado, y *Pantoea ananatis* - *Erwinia gerundensis*, por otro. La rama que contiene mayor número de géneros cercanos entre sí está formada por *Bacillus*-*Paenibacillus* y *Acinetobacter*-*Pseudomonas*.

Con el estudio filogenético se pretenden comparar las distribuciones obtenidas a partir de las proteínas responsables de la promoción vegetal con la filogenia según sus marcadores filogenéticos.

Los genes estudiados para el análisis filogenético son dos. Por un lado el marcador *gyrB*, que codifica para la subunidad beta de la proteína DNA girasa [22], responsable de solventar el problema de tensión molecular generado durante la replicación de ADN. Por otro lado *rpoB*, que porta la información para la subunidad beta de la ARN polimerasa bacteriana [78], encargada de la transcripción.

Material y métodos

A partir de los nombres de las cepas de la Tabla 1 se lanzó una búsqueda a través de la función *entrez_search* y *entrez_link* del paquete *rentrez* para la descarga de los IDs correspondientes de las proteínas codificadas con *gyrB* y *rpoB*.

Se descargaron las secuencias mediante la función *entrez_fetch* desde la base de datos “protein” de NCBI [27], [41]. Una vez descargadas, se realiza la lectura de los archivos *fasta* generados y la reedición de los nombres asignados con las funciones *sub* y *gsub* para que el identificador sea el nombre de la cepa. En algunas secuencias, al no seguir el mismo esquema que la mayoría, el nombre se asigna directamente con *R*. Ver Anexo IV. *Script* de R.

A continuación, se alinearon las secuencias mediante la función *AlignSeqs* del paquete *decipher* y se efectuó una mejora del resultado mediante la función *StaggerAlignment*. Tras esta corrección, se calculó la matriz de distancias mediante la función *stringDist* utilizando como método la *distancia de Hamming*. Este método se basa en el estudio de secuencias de la misma longitud contabilizando el número de posiciones no coincidentes. Las distancias calculadas, p , son las disimilaridades encontradas entre ambas secuencias de longitud N . Las distancias brutas son normalizadas (V) con la división entre la longitud de la secuencia obteniendo el valor p' .

Posteriormente, se aplica una corrección para aminoácidos con el método de de Jukes-Cantor (VI), donde s es el número de residuos distintos posibles, para aminoácidos, 20:

$$(V) \quad p' = \frac{p}{N}$$

$$(VI) \quad d = -\frac{s-1}{s} \ln\left(1 - \frac{s}{s-1} p'\right)$$

Tras la corrección, se genera el árbol definitivo a partir de la matriz de distancias corregida con la función *nj* del paquete *ape*. Por último, para una mejor comprensión de los resultados se resaltan los distintos grupos de bacterias con el programa *Microsoft Word 2016*.

Resultados

gyrB

En los seis primeros nodos del Árbol Filogenético 10. Comparativa de la filogenia según marcadores *gyrB* (azul) - *rpoB* (rojo), en *gyrB* (azul) hay separaciones sucesivas de los grupos *Pantoea*, *Erwinia*, *Klebsiella-Enterobacter*, *Pseudomonas-Azotobacter* y *Acinetobacter* del resto de cepas. Esto indica la separación entre γ -proteobacterias del resto y una mayor cercanía entre las demás cepas de estudio.

A continuación, el grupo de α -proteobacterias conjuntamente con β -proteobacteria se separa del grupo formado por firmicutes y gram +. A pesar de esta separación aparentemente clara, aparecen cepas de los géneros *Methylobacterium*, *Sinorhizobium*, *Agrobacterium*, *Erwinia*, *Pantoea* o *Acinetobacter* que no obstante, se separan en el siguiente nodo del grupo de firmicutes y gram +.

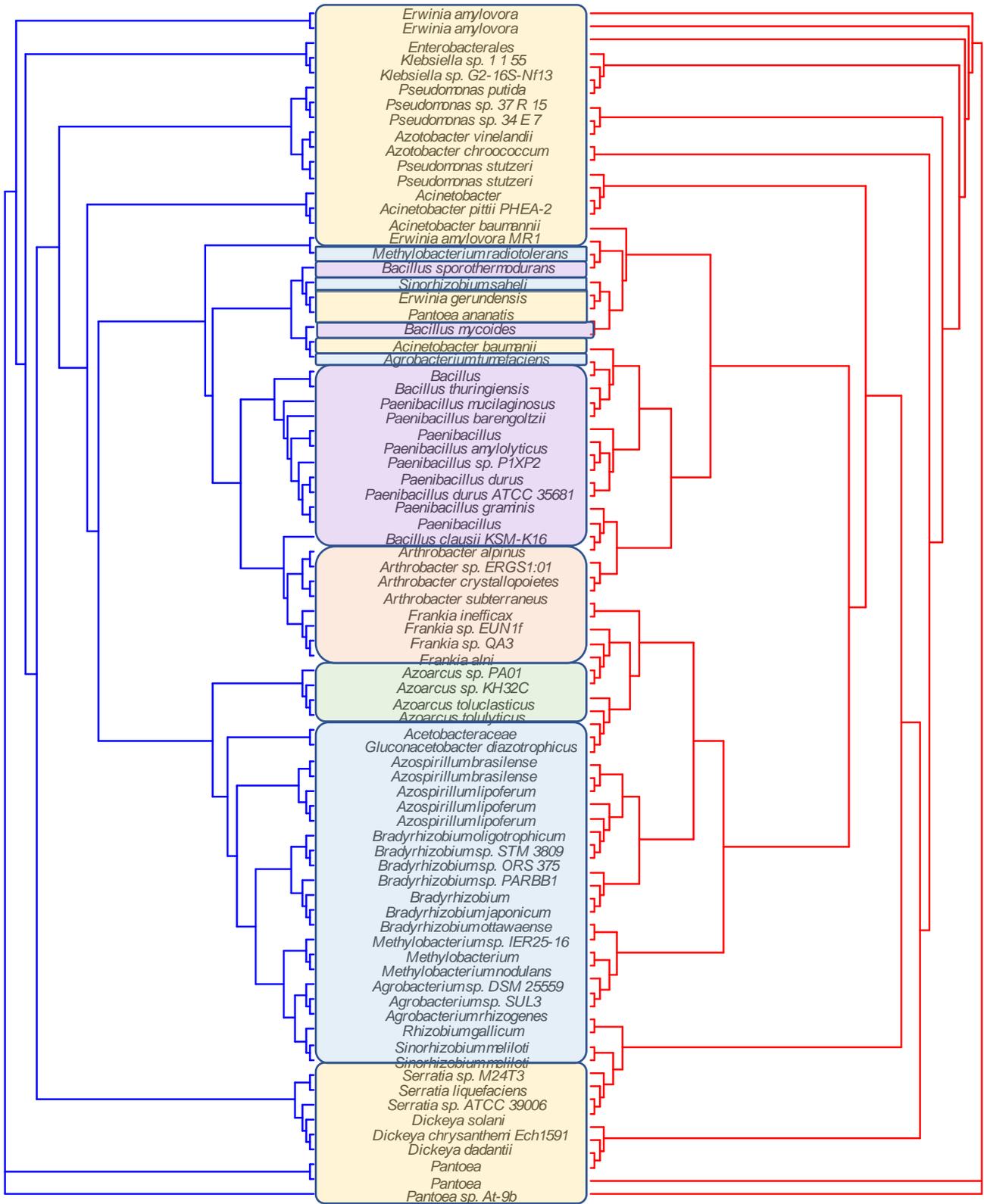
rpoB

Como en el caso de *gyrB*, en los primeros nodos de Árbol Filogenético 10 (en rojo), se da la separación de los géneros pertenecientes al grupo de γ -proteobacteria, *Erwinia*, *Klebsiella*, *Pantoea*, *Serratia*, *Dyckeya*, *Acinetobacter* y *Pseudomonas*. Tras esta separación, el grupo de *Azoarcus*, β -proteobacteria, se distancia del resto. Es en el siguiente nodo cuando se da una separación casi total en dos ramas, solamente *Sinorhizobium saheli* y *Agrobacterium sp HBR75* rompen esta división entre α -proteobacteria del grupo de Rhizobiaceae y la sección dominada por Gram +, Firmicutes y Actinobacterias. A esta última rama de Gram + aparecen cercanas las siguientes cepas que no corresponderían con este grupo, γ -proteobacterias como *Acinetobacter pittii*, *Erwinia* o *Pseudomonas stutzeri KOS6* o α -proteobacterias como *Sinorhizobium saheli* o *Agrobacterium sp HBR75*.

Discusión

Las cepas del grupo de gammaproteobacteria muestran mayor distancia con respecto a las cepas del resto de grupos. Tras ese primer distanciamiento, se diferencian las gram+ del grupo de alfa y betaproteobacterias. El estudio con ambos genes presenta un grupo situado entre la separación de gram+ del resto. Según los marcadores, este grupo intermedio estaría compuesto por *Erwinia amylovora MR1*-*Methylobacterium radiotolerans*, *Sinorhizobium saheli* - *Erwinia gerundensis* - *Pantoea ananitis*, *Acinetobacter baumannii* - *Agrobacterium tumefaciens* y, conjuntamente, *Bacillus sporothermodurans* y *Bacillus mycoides*.

Entre las pequeñas diferencias de filogenia con el estudio de los marcadores, según *rpoB*, aparece mayor cercanía entre *Rhizobium gallicum*, *Sinorhizobium meliloti*, alfaproteobacterias, con *Serratia sp.*, gammaproteobacteria, que la que se muestra con el marcador *gyrB*. Del mismo modo, con *rpoB* se revela mayor cercanía entre el genero *Frankia* con el grupo de betaproteobacterias de *Azoarcus* y alfaproteobacterias rizobiales, con quien comparte la fijación de nitrógeno en simbiosis.



Árbol Filogenético 10. Comparativa de la filogenia según marcadores *gyrB* (azul) - *rpoB* (rojo)

2.3. CAPÍTULO 3.

ESTUDIO COMBINADO DE LAS RELACIONES ENTRE MECANISMOS PROMOTORES DEL CRECIMIENTO VEGETAL, DIVERSIDAD DE PGPRs Y CULTIVOS.

Hasta ahora se han descrito algunas de las principales características que definen a las bacterias de promoción vegetal y se ha buscado la relación entre las proteínas que participan en ellas. Pero no se debe olvidar que la importancia de la presencia de PGPRs en los cultivos es su labor en la mejora del desarrollo vegetal. Durante el desarrollo de la planta, esta cambia la composición de microorganismos a su alrededor [11], [79], por lo que puede indicar que esta variación se debe a una posible relación entre planta-bacteria que está determinada por las exigencias de la planta [54]. Pero este contenido en PGPRs, no solamente depende de la exigencia de la planta, sino que, en determinados suelos, existen ciertas bacterias predominantes que ejercerán competencia con respecto a las cepas de interés [80]. Ante estas evidencias, la posible utilización de PGPRs como biofertilizantes debe tener en cuenta los requerimientos de la planta en cada momento y la composición del sustrato.

El contenido de rizobacterias en torno a una planta está determinado por los exudados generados por la raíz. Se ha descubierto, incluso, que el mayor contenido de microorganismos se encuentra en la zona más activa del crecimiento de la raíz [79], [81]. En los restos de esta zona de crecimiento se han hallado células con gran cantidad de vesículas que liberarían esos exudados con alto contenido en aminoácidos, azúcares o vitaminas [44], [64] para favorecer la llegada de ciertas cepas.

A pesar de que los mecanismos de interacción entre rizobacterias y plantas no están totalmente definidos, actualmente ya se están utilizando inóculos para la promoción del crecimiento vegetal [44]. Por este motivo, recientemente, se ha aprobado un epígrafe específico sobre “Productos especiales basados en microorganismos” en *Real Decreto 999/2017, de 24 de noviembre, por el que se modifica el Real Decreto 506/2013, de 28 de junio, sobre productos fertilizantes*.

En el presente capítulo, se examina una muestra de las distintas investigaciones realizadas en el estudio de microorganismos que promueven el crecimiento en cultivos específicos.

Material y Métodos

A partir de 192 citas tomadas de De Souza, Vessey 2003; Ahemad, 2014; Pérez-Montaña, 2014 y Nadeem, 2014; [26], [44], [46], [82], [83], se genera una tabla agrupando los registros según familias de cultivos.

En el grupo Poaceae se engloban las citas de bacterias con presencia en arroz (*Oryza sativa*), maíz (*Zea mays*), trigo (*Triticum sp*) y cebada (*Hordeum vulgare*); en Fabaceae se agrupan distintas leguminosas como garbanzo (*Cicer arietinum*), judía (*Phaseolus vulgaris*), cacahuete (*Arachis hypogaea*) y soja (*Glycine max*); en Brassicaceae con la colza (*Brassica napus*) y mostaza india (*Brassica juncea*); por último, en Solanaceae se agrupan las citas de tomate (*Solanum lycopersicum*) y pimiento (*Capsicum sp*).

Con estos datos generados, que se resumen en la Figura 14, y los resultados del capítulo anterior se hace una aproximación a los vínculos entre las distintas cepas con los cultivos.

Resultados

En la Figura 14. Distribución de citas según cultivos y grupo de bacterias. se aprecia que las citas de betaproteobacteria se han encontrado en su totalidad en el grupo de poáceas, mientras alfaproteobacterias tiene su máximo en el grupo de Fabaceae, donde se encuentran todas las leguminosas que hacen simbiosis con Rhizobiales. Los resultados en Brassicaceae son dominados por gammaproteobacterias y firmicutes siendo, además, el único grupo en el que no aparecen citas para alfaproteobacterias. Por otro lado, las citas tomadas en Malvaceae y Solanaceae son residuales y son compartidas por gammaproteobacteria y firmicutes como pareja de alfaproteobacterias.

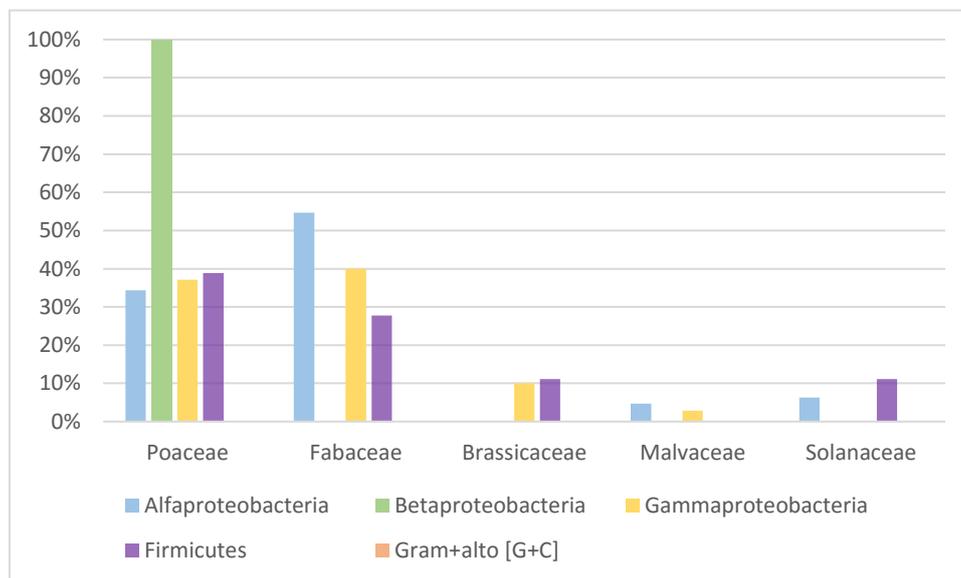


Figura 14. Distribución de citas según cultivos y grupo de bacterias.

Discusión

Con los datos obtenidos del estudio de las proteínas en el Capítulo 2.2, se describe la diversidad existente en las funciones desempeñadas por cada grupo bacteriano. En la mayoría de los casos los genes de interés son compartidos por los distintos grupos. Únicamente, los genes *budABC* no se han descrito en las betaproteobacterias y gram+ de alto contenido [G+C] seleccionada. Por el contrario, sí se muestra el potencial de fijación del nitrógeno por el grupo de alfaproteobacteria y su capacidad de generación de óxido nítrico, el 70% de las bacterias con el gen *nirK* incluidas son de este grupo. En contraposición, las betaproteobacterias y las bacterias gram+ de alto contenido [G+C] son los grupos con menor representación en las características de promoción vegetal estudiadas.

Sin embargo, la importancia para un posible uso de determinada cepa como biofertilizante reside también en la co-expresión de proteínas dentro de un mismo genoma. Con el estudio de genes se aprecia que, a parte de los que forman parte de una sola proteína como la nitrogenasa, los genes codificantes para fitohormonas (*ipdC/ppdc*, *acdS* y *nirK*) son los que presentan mayor coocurrencia.

Por otro lado, a partir del estudio de semejanza entre proteínas, se describe la idea del distanciamiento existente entre el grupo de alfaproteobacterias con el resto de grupos. En la mayoría de las proteínas

estudiadas, la mayor diferencia es la descrita por el grupo de rizobiales con respecto al resto. Sin embargo, con el estudio de marcadores genéticos el mayor distanciamiento es el de gammaproteobacterias con respecto a los demás grupos. Esto indicaría una especialización en ciertas funciones de promoción del crecimiento vegetal, algo razonable ya que no se debe olvidar que las bacterias rizobiales tienen un nicho ecológico más estrecho que el resto de grupos que aparecen en el estudio.

Tras el análisis de la relación entre cultivos y PGPR asociadas, es necesario destacar que los cultivos con más ensayos son aquellos de mayor importancia para el consumo humano. El grupo más destacado sería el formado por poáceas donde se incluyen arroz, maíz, trigo y cebada, seguido de leguminosas (Fabaceae) por lo comentado de la importancia de la fijación de nitrógeno.

A pesar del sesgo existente, se aprecia que la distribución de los distintos grupos bacterianos es más o menos homogénea en porcentaje de citas salvo para el grupo de betaproteobacterias, donde las cepas pertenecientes a *Azoarcus sp* solamente son encontradas en el grupo de las poáceas.

En resumen, es evidente la complejidad de las relaciones entre cultivos y cepas bacterianas, a la vez que también lo es entre las propias cepas. No obstante, contando con el sesgo mencionado, existe una especialización por parte del grupo de alfarproteobacteria con respecto al resto en la promoción del crecimiento vegetal.

3. Conclusiones

Con este trabajo se han estudiado 9 genes de promoción del crecimiento vegetal (*nifHDK*, *ipdC/ppdC*, *acdS*, *nirK* y *budABC*) atendiendo a las proteínas para las que codifican (Nitrogenasa, Indolpiruvato descarboxilasa, ACC Desaminasa, Nitrito Reductasa, Acetolactato Sintasa, Acetoín decarboxilasa y Diacetil Reductasa). Con la comparación entre cepas según cada una de estas proteínas y con el apoyo del estudio mediante marcadores filogenéticos (*gyrB* y *rpoB*), se ha mostrado cierta especialización del grupo de alfaproteobacteria en este nicho ecológico. No obstante, al tratarse de un estudio a partir de citas bibliográficas existe cierto sesgo por la importancia histórica que se le ha dado a rizobiales como bacterias fijadoras de nitrógeno.

De este modo, los objetivos marcados en 1.4 Planificación del Trabajo se han cumplido prácticamente en su totalidad. El aspecto que no se ha podido cumplir de forma adecuada ha sido la relación de PGPRs debido a que no existe, en bibliografía, una tabla específica con el detalle de las condiciones en las que se encuentran las cepas encontradas por cada cultivo con las que aplicar técnica de *machine learning*. No obstante, se ha podido hacer la aproximación recogiendo datos y generar una tabla sencilla que incluye citas de cada cultivo, Anexo III. Tabla de citas. Esta tarea de recogida de datos ha sido, quizás, la más laboriosa. Así, teniendo en cuenta que se trata de un campo algo lejano de mis conocimientos, casi la mitad del tiempo del trabajo ha sido de lectura de bibliografía desde aspectos básicos hasta otros más avanzados. De este modo, la metodología seguida ha sido acorde a la planificación, aunque con cierto retraso hasta la mitad del semestre debido a lo costoso de introducirse en un nuevo tema de estudio. No obstante, con la ayuda de la tutora se ha llegado a un cumplimiento de los objetivos marcados en el inicio de la elaboración del proyecto.

Como se ha mencionado a lo largo del trabajo, en un estudio con fuentes únicamente bibliográficas existe sesgo, más aún en un trabajo de este tipo de estudio de relaciones. Hasta hace pocos años, los estudios se centraban en la descripción de la proteína o de la cepa de interés a partir de cultivos *in vitro* controlados. Sin embargo, gracias al avance en técnicas de ultrasecuenciación y *machine learning* cada vez es mayor el estudio del sistema en su entorno natural, no de sus componentes por separado. A pesar de aumentar la complejidad para la explicación de relaciones, da una idea más aproximada a lo que sucede en la realidad. El siguiente paso en este tipo de estudio sería corroborar estas ideas generadas mediante el estudio de muestras tomadas en campo y con estudio de las variables como fase del ciclo de la planta, tipo de suelo, temperatura, humedad, tipo de cultivo y composición de la rizosfera. De este modo se logrará la tecnificación adecuada para la utilización de fórmulas específicas de PGPRs que ayuden, en conjunción con otros métodos, a la mejora de los rendimientos de los cultivos y la recuperación de suelos sobreexplotados.

4. Glosario

A

Abono

~ inorgánico

Fertilizante basado en composiciones químicas inorgánicas con fórmulas determinadas. Generalmente ricas en nitrógeno, fósforo y potasio, entre otros elementos., 1

~ mineral. Véase Abono inorgánico

~ orgánico

Fertilizante que cuenta con materia orgánica, generalmente residuos, en su composición., 1

B

Biofertilizante

Fertilizante cuya formulación esta basada en el aporte de microorganismos., 2

C

Catalítico

Proceso ~

Actividad por la que se aumenta la velocidad de una reacción química., 15

Codón

Triplete que, en un ARN mensajero, codifica la incorporación de aminoácidos específicos en la biosíntesis de proteínas, 20

Cultivos hidropónicos

Método de cultivo que prescinde de suelo. En su lugar utiliza un sustrato físico y la adición de distintos compuestos según las necesidades de la planta., 1

E

Edáfico

Relativo al suelo., 2

Enzima

Proteína especializada en la catalización de reacciones químicas., 2

~lítica

Enzima especializada en la rotura de enlaces químicos., 2

F

Fenotípico

Carácter ~

Realización visible de la información genética en un determinado ambiente., 20

Fertilizante

Sustancia con alto contenido en nutrientes que tiene como función el mantenimiento o incremento de su contenido en el suelo para el aprovechamiento de las plantas., 12

Fitosanitario

Sustancia destinada a la prevención y/o mitigación de enfermedades en cultivos., 12

Fitotecnia

Materia que investiga los fundamentos biológicos, edáficos, climáticos, sanitarios y técnicos para optimizar la producción de cultivos., 12

L

Lixiviación

Proceso de arrastre por el agua de nutrientes hacia capas profundas de suelo, clave en la contaminación de acuíferos., 1

M

Marcador filogenético

Molécula capaz de describir el tiempo evolutivo entre dos especies., 2

N

Nicho ecológico

Posición físico-funcional de una especie o población en el ecosistema., 1

P

Patogenicidad

Capacidad de provocar enfermedad en un individuo hospedador., 2

R

Rizosfera

Lugar del suelo cercano a la raíz del vegetal con una composición de seres vivos característica., 1

5. Bibliografía

- [1] Food and Agriculture Organization, “How to Feed the World in 2050,” Rome, 2009.
- [2] G. G. Shailendra Singh, “Plant Growth Promoting Rhizobacteria (PGPR): Current and Future Prospects for Development of Sustainable Agriculture,” *J. Microb. Biochem. Technol.*, vol. 7, no. 2, pp. 96–102, 2015.
- [3] Food and Agriculture Organization, “Increasing Crop Production Sustainably,” 2009.
- [4] D. Tilman, C. Balzer, J. Hill, and B. L. Befort, “Global food demand and the sustainable intensification of agriculture.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 108, no. 50, pp. 20260–4, 2011.
- [5] Unión Europea, “Una Agricultura Sostenible para el Futuro que Queremos,” 2012.
- [6] M. A. Altieri, “Ecological Impacts of Industrial Agriculture and the Possibilities for Truly Sustainable Farming.”
- [7] P. Vejan, R. Abdullah, T. Khadiran, S. Ismail, and A. Nasrulhaq Boyce, “Role of plant growth promoting rhizobacteria in agricultural sustainability- A review,” *Molecules*, 2016.
- [8] P. N. Bhattacharyya and D. K. Jha, “Plant growth-promoting rhizobacteria (PGPR): Emergence in agriculture,” *World Journal of Microbiology and Biotechnology*, vol. 28, no. 4, pp. 1327–1350, 2012.
- [9] L. Hiltner, “Über neuere Erfahrungen und Probleme auf dem Gebiete der Bodenbakteriologie unter besonderden berucksichtigung und Brache,” *Arb. Dtsch. Landwirtsch Gesellschaft*, vol. 98, 1904.
- [10] A. Hartmann, M. Schmid, D. van Tuinen, and G. Berg, “Plant-driven selection of microbes,” *Plant Soil*, vol. 321, no. 1–2, pp. 235–257, Aug. 2009.
- [11] L. Mommer, P. Hinsinger, C. Prigent-Combaret, and E. J. W. Visser, “Advances in the rhizosphere: stretching the interface of life,” *Plant Soil*, vol. 407, no. 1–2, pp. 1–8, 2016.
- [12] J. K. Vessey, “Plant growth promoting rhizobacteria as biofertilizers,” *Plant Soil*, vol. 255, no. 2, pp. 571–586, 2003.
- [13] D. Goswami, J. N. Thakker, and P. C. Dhandhukia, “Portraying mechanics of plant growth promoting rhizobacteria (PGPR): A review,” *Cogent Food Agric.*, vol. 2, no. 1, pp. 1–19, 2016.
- [14] D. C. K. Jha and D. M. Saraf, “Plant growth promoting Rhizobacteria (PGPR): a review,” Jan. 2015.
- [15] G. Berg, L. Eberl, and A. Hartmann, “The rhizosphere as a reservoir for opportunistic human pathogenic bacteria,” *Environ. Microbiol.*, vol. 7, no. 11, pp. 1673–1685, 2005.
- [16] C. De-la-Peña and V. M. Loyola-Vargas, “Biotic interactions in the rhizosphere: a diverse cooperative enterprise for plant productivity.,” *Plant Physiol.*, vol. 166, no. 2, pp. 701–19, Oct. 2014.
- [17] V. Artursson, R. D. Finlay, and J. K. Jansson, “Interactions between arbuscular mycorrhizal fungi and bacteria and their potential for stimulating plant growth,” *Environ. Microbiol.*, vol. 8, no. 1, pp. 1–10, Jan. 2006.
- [18] S. Nakkeeran, W. G. D. Fernando, and Z. A. Siddiqui, “Plant Growth Promoting Rhizobacteria Formulations and its Scope in Commercialization for the Management of Pests and Diseases,” in *PGPR: Biocontrol and Biofertilization*, Berlin/Heidelberg: Springer-Verlag, pp. 257–296.

- [19] O. Martínez-Viveros, M. . A. Jorquera, D. . E. Crowley, G. Gajardo, and M. . L. Mora, "MECHANISMS AND PRACTICAL CONSIDERATIONS INVOLVED IN PLANT GROWTH PROMOTION BY RHIZOBACTERIA," *J. soil Sci. plant Nutr.*, vol. 10, no. 3, pp. 293–319, Jul. 2010.
- [20] T. Kumatani, Y. Yoshimi, A. Hitoshi, N. Ae, and M. Aino, "Phylogenetic analyses of plant-growth-promoting rhizobacteria isolated from tomato, lettuce, and Japanese pepper plants in Hyogo Prefecture, Japan," 2009.
- [21] S. G. Bavykin *et al.*, "Use of 16S rRNA, 23S rRNA, and gyrB Gene Sequence Analysis To Determine Phylogenetic Relationships of Bacillus cereus Group Microorganisms," *J. Clin. Microbiol.*, vol. 42, no. 8, pp. 3711–3730, 2004.
- [22] L.-T. Wang, F.-L. Lee, C.-J. Tai, H. Kasai, and C. Fwu-Ling Lee fill, "Comparison of gyrB gene sequences, 16S rRNA gene sequences and DNA–DNA hybridization in the Bacillus subtilis group."
- [23] W.-I. Kim, "Genetic Diversity of Cultivable Plant Growth-Promoting Rhizobacteria in Korea," *J. Microbiol. Biotechnol.*, vol. 21, no. 8, pp. 777–790, Aug. 2011.
- [24] M. Bruto, C. Prigent-Combaret, D. Muller, and Y. Moëgne-Loccoz, "Analysis of genes contributing to plant-beneficial functions in plant growth-promoting rhizobacteria and related Proteobacteria," *Sci. Rep.*, vol. 4, no. 1, p. 6261, 2015.
- [25] M. A. Baset Mia, Z. H. Shamsuddin, Z. Wahab, and M. Marziah, "Effect of plant growth promoting rhizobacterial (PGPR) inoculation on growth and nitrogen incorporation of tissue-cultured Musa plantlets under nitrogen-free hydroponics condition," *AJCS*, vol. 4, no. 2, pp. 85–90, 2010.
- [26] R. De Souza, A. Ambrosini, and L. M. P. Passaglia, "Plant growth-promoting bacteria as inoculants in agricultural soils."
- [27] NCBI Resource Coordinators, "Database Resources of the National Center for Biotechnology Information," *Nucleic Acids Res.*, vol. 45, no. D1, pp. D12–D17, Jan. 2017.
- [28] D. A. Benson, I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and D. L. Wheeler, "GenBank.," *Nucleic Acids Res.*, vol. 33, no. Database issue, pp. D34-8, Jan. 2005.
- [29] N. A. O'Leary *et al.*, "Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation.," *Nucleic Acids Res.*, vol. 44, no. D1, pp. D733-45, Jan. 2016.
- [30] A. P. Arkin and D. C. Youvan, "An algorithm for protein engineering: simulations of recursive ensemble mutagenesis.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 89, no. 16, pp. 7811–5, Aug. 1992.
- [31] S. Altschul *et al.*, "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs," *Nucleic Acids Res.*, vol. 25, no. 17, pp. 3389–3402, Sep. 1997.
- [32] E. Paradis, J. Claude, and K. Strimmer, "APE: Analyses of Phylogenetics and Evolution in R language," *Bioinformatics*, vol. 20, no. 2, pp. 289–290, Jan. 2004.
- [33] B. J. Callahan, K. Sankaran, J. A. Fukuyama, P. J. McMurdie, and S. P. Holmes, "Bioconductor workflow for microbiome data analysis: from raw reads to community analyses," *F1000Research*, vol. 5, p. 1492, Jun. 2016.
- [34] U. Bodenhofer, E. Bonatesta, C. Horejš-Kainrath, and S. Hochreiter, "msa: an R package for multiple sequence alignment," *Bioinformatics*, vol. 31, no. 24, p. btv494, Aug. 2015.
- [35] M. A. Larkin *et al.*, "Clustal W and Clustal X version 2.0," *Bioinformatics*, vol. 23, no. 21, pp. 2947–2948, Nov. 2007.
- [36] R. W. Hamming, "Error Detecting and Error Correcting Codes," *Bell Syst.*

- Tech. J.*, vol. 29, no. 2, pp. 147–160, Apr. 1950.
- [37] “plant growth bacteria - Genome - NCBI.” [Online]. Available: <https://www.ncbi.nlm.nih.gov/genome>. [Accessed: 12-Oct-2017].
- [38] R Core Team, “R: A Language and Environment for Statistical Computing.” R Foundation for Statistical Computing, Vienna, Austria, 2017.
- [39] W. Huber *et al.*, “Orchestrating high-throughput genomic analysis with Bioconductor,” *Nat. Methods*, vol. 12, no. 2, pp. 115–121, Jan. 2015.
- [40] D. Charif and J. R. Lobry, “SeqinR 1.0-2: A Contributed Package to the R Project for Statistical Computing Devoted to Biological Sequences Retrieval and Analysis,” Springer, Berlin, Heidelberg, 2007, pp. 207–232.
- [41] D. Winter, “rentrez: Entrez in R. R package version 1.1.0.” 2017.
- [42] E. Wright, “Decipher.” 2017.
- [43] A. Gupta *et al.*, “Whole genome sequencing and analysis of plant growth promoting bacteria isolated from the rhizosphere of plantation crops coconut, cocoa and arecanut,” *PLoS One*, vol. 9, no. 8, p. e104259, Aug. 2014.
- [44] M. Ahemad and M. Kibret, “Mechanisms and applications of plant growth promoting rhizobacteria: Current perspective,” *J. King Saud Univ. - Sci.*, vol. 26, no. 1, pp. 1–20, 2014.
- [45] J. Vacheron *et al.*, “Plant growth-promoting rhizobacteria and root system functioning,” *Front. Plant Sci.*, vol. 4, 2013.
- [46] F. Pérez-Montaño *et al.*, “Plant growth promotion in cereal and leguminous agricultural important plants: From microorganism capacities to crop production,” *Microbiol. Res.*, vol. 169, no. 5–6, pp. 325–336, 2014.
- [47] R. Gupta, V. S. Bisaria, and S. Sharma, “Effect of Agricultural Amendments on *Cajanus cajan* (Pigeon Pea) and Its Rhizospheric Microbial Communities--A Comparison between Chemical Fertilizers and Bioinoculants.,” *PLoS One*, vol. 10, no. 7, p. e0132770, 2015.
- [48] J. Vacheron, Y. Moënne-Loccoz, A. Dubost, M. Gonçalves-Martins, D. Muller, and C. Prigent-Combaret, “Fluorescent *Pseudomonas* Strains with only Few Plant-Beneficial Properties Are Favored in the Maize Rhizosphere,” *Front. Plant Sci.*, vol. 7, no. August, pp. 1–13, 2016.
- [49] M. A. Matilla *et al.*, “Complete genome of the plant growth-promoting rhizobacterium *Pseudomonas putida* BIRD-1,” *J. Bacteriol.*, vol. 193, no. 5, p. 1290, 2011.
- [50] M. Venkateshwaran, J. D. Volkening, M. R. Sussman, and J. M. Ané, “Symbiosis and the social network of higher plants,” *Current Opinion in Plant Biology*, vol. 16, no. 1. pp. 118–127, Feb-2013.
- [51] R. P. Singh, G. M. Shelke, A. Kumar, and P. N. Jha, “Biochemistry and genetics of ACC deaminase: A weapon to ‘stress ethylene’ produced in plants,” *Front. Microbiol.*, vol. 6, no. SEP, pp. 1–14, 2015.
- [52] D. Blaha, C. Prigent-Combaret, M. S. Mirza, and Y. Moënne-Loccoz, “Phylogeny of the 1-aminocyclopropane-1-carboxylic acid deaminase-encoding gene *acdS* in phytobeneficial and pathogenic Proteobacteria and relation with strain biogeography,” *FEMS Microbiol. Ecol.*, vol. 56, no. 3, pp. 455–470, Jun. 2006.
- [53] M. Land *et al.*, “Insights from 20 years of bacterial genome sequencing.,” *Funct. Integr. Genomics*, vol. 15, no. 2, pp. 141–61, Mar. 2015.
- [54] D. Katiyar, A. Hemantaranjan, and B. Singh, “Advances in Plants & Agriculture Research Plant Growth Promoting Rhizobacteria-an Efficient Tool for Agriculture Promotion,” 1945.
- [55] G. Bacci *et al.*, “Exploring the dynamics of bacterial community composition in soil: the pan-bacteriome approach,” *Antonie Van Leeuwenhoek*, vol. 107, no. 3, pp. 785–797, Mar. 2015.

- [56] A. L. Fernandez, C. C. Sheaffer, D. L. Wyse, C. Staley, T. J. Gould, and M. J. Sadowsky, "Structure of bacterial communities in soil following cover crop and organic fertilizer incorporation," *Appl. Microbiol. Biotechnol.*, vol. 100, no. 21, pp. 9331–9341, Nov. 2016.
- [57] G. E. D Oldroyd, "Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants," *Nat. Publ. Gr.*, vol. 11, 2013.
- [58] K. J. Field, S. Pressel, J. G. Duckett, W. R. Rimington, and M. I. Bidartondo, "Symbiotic options for the conquest of land," *Trends in Ecology and Evolution*, vol. 30, no. 8, pp. 477–486, Aug-2015.
- [59] L. G. Lioshina, "Cellular and molecular-genetic mechanisms of symbiosis and associative interaction of microorganisms with plants in rhizosphere," *Biopolym. Cell*, vol. 25, no. 1, pp. 27–38, 2009.
- [60] A. Lareen, F. Burton, and P. Schäfer, "Plant root-microbe communication in shaping root microbiomes," *Plant Molecular Biology*, vol. 90, no. 6, 2016.
- [61] M. Cobo-Simón and J. Tamames, "Relating genomic characteristics to environmental preferences and ubiquity in different microbial taxa."
- [62] L. C. Van Loon and P. A. H. M. Bakker, "Induced Systemic Resistance as a Mechanism of Disease Suppression by Rhizobacteria," in *PGPR: Biocontrol and Biofertilization*, Dordrecht: Springer Netherlands, 2005, pp. 39–66.
- [63] P. Vandenkoornhuyse, A. Quaiser, M. Duhamel, A. Le Van, and A. Dufresne, "The importance of the microbiome of the plant holobiont," *New Phytologist*, vol. 206, no. 4, 2015.
- [64] D. Bulgarelli, K. Schlaeppli, S. Spaepen, E. V. L. van Themaat, and P. Schulze-Lefert, "Structure and Functions of the Bacterial Microbiota of Plants," *Annu. Rev. Plant Biol.*, 2013.
- [65] S. K. Upadhyay, A. Devendra, P. Singh, R. Saikia, D. P. Singh, and R. Saikia, "Genetic Diversity of Plant Growth Promoting Rhizobacteria Isolated from Rhizospheric Soil of Wheat under Saline Condition," *Curr. Microbiol.*, vol. 59, no. 5, pp. 489–496, 2009.
- [66] P. C. Dos Santos, Z. Fang, S. W. Mason, J. C. Setubal, and R. Dixon, "Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes," *BMC Genomics*, vol. 13, no. 1, p. 162, May 2012.
- [67] Y. Hu and M. W. Ribbe, "Nitrogenase and homologs.," *J. Biol. Inorg. Chem.*, vol. 20, no. 2, pp. 435–45, Mar. 2015.
- [68] A. Schütz, T. Sandalova, S. Ricagno, G. Hübner, S. König, and G. Schneider, "Crystal structure of thiamindiphosphate-dependent indolepyruvate decarboxylase from *Enterobacter cloacae*, an enzyme involved in the biosynthesis of the plant hormone indole-3-acetic acid," *Eur. J. Biochem.*, vol. 270, no. 10, pp. 2312–2321, May 2003.
- [69] J. Koga, "Structure and function of indolepyruvate decarboxylase, a key enzyme in indole-3-acetic acid biosynthesis," *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.*, vol. 1249, no. 1, pp. 1–13, May 1995.
- [70] R. P. Singh, G. M. Shelke, A. Kumar, and P. N. Jha, "Biochemistry and genetics of ACC deaminase: a weapon to "stress ethylene" produced in plants.," *Front. Microbiol.*, vol. 6, p. 937, 2015.
- [71] L. Sanz *et al.*, "Nitric Oxide Plays a Role in Stem Cell Niche Homeostasis through Its Interaction with Auxin," *PLANT Physiol.*, vol. 166, no. 4, pp. 1972–1984, Dec. 2014.
- [72] X. Qin, L. Deng, C. Hu, L. Li, and X. Chen, "Copper-Containing Nitrite Reductase Employing Proton-Coupled Spin-Exchanged Electron-Transfer and Multiproton Synchronized Transfer to Reduce Nitrite," *Chem. - A Eur. J.*, vol. 23, no. 59, pp. 14900–14910, Oct. 2017.
- [73] C.-M. Ryu *et al.*, "Bacterial volatiles promote growth in Arabidopsis.," *Proc.*

- Natl. Acad. Sci. U. S. A.*, vol. 100, no. 8, pp. 4927–32, Apr. 2003.
- [74] Z. A. Siddiqui, “PGPR: Prospective Biocontrol Agents of Plant Pathogens,” in *PGPR: Biocontrol and Biofertilization*, Berlin/Heidelberg: Springer-Verlag, pp. 111–142.
- [75] M. G. Csardi, “Title Network Analysis and Visualization,” 2015.
- [76] R. F. Doolittle, D. F. Feng, S. Tsang, G. Cho, and E. Little, “Determining divergence times of the major kingdoms of living organisms with a protein clock.,” *Science*, vol. 271, no. 5248, pp. 470–7, Jan. 1996.
- [77] M. Dunn, N. Burenhult, N. Kruspe, S. Tufvesson, and N. Becker, “Aslian linguistic prehistory: A case study in computational phylogenetics,” *Diachronica*, vol. 28, no. 3, pp. 291–323, 2011.
- [78] L. Chistoserdova *et al.*, “Distribution of 2,4-Diacetylphloroglucinol Biosynthetic Genes among the *Pseudomonas* spp. Reveals Unexpected Polyphyletism,” *Reveal. Unexpected Polyphyletism. Front. Microbiol*, vol. 8, no. 8, 2017.
- [79] A. García-Salamanca *et al.*, “Bacterial diversity in the rhizosphere of maize and the surrounding carbonate-rich bulk soil,” *Microb. Biotechnol.*, vol. 6, no. 1, pp. 36–44, 2013.
- [80] P. Prashar, N. Kapoor, and S. Sachdeva, “Rhizosphere: Its structure, bacterial diversity and significance,” *Reviews in Environmental Science and Biotechnology*. 2014.
- [81] S. Dutta and A. R. Podile, “Plant Growth Promoting Rhizobacteria (PGPR): the bugs to debug the root zone,” *Crit. Rev. Microbiol.*, vol. 36, no. 3, pp. 232–244, Aug. 2010.
- [82] J. K. Vessey, “Plant growth promoting rhizobacteria as biofertilizers,” *Plant and Soil*, vol. 255, no. 2. pp. 571–586, 2003.
- [83] S. M. Nadeem, M. Ahmad, Z. A. Zahir, A. Javaid, and M. Ashraf, “The role of mycorrhizae and plant growth promoting rhizobacteria (PGPR) in improving crop productivity under stressful environments,” *Biotechnol. Adv.*, vol. 32, no. 2, pp. 429–448, Mar. 2014.

6. Anexos

Anexo I. Tablas de cepas

nifH - Fe Prot			
Cepa	Acceso	Revista	Año
A.sp. P7-29	SCQ83764.1	Unpublished	Unpublished
A.sp. P8-1	SCQ83766.1	Unpublished	Unpublished
A.sp. P8-9	SCQ83769.1	Unpublished	Unpublished
A.sp.	AOQ25860.1	Unpublished	Unpublished
A.sp. P7-7	SCQ83756.1	Unpublished	Unpublished
A.tumefaciens	AEZ05086.1	Syst. Appl. Microbiol.	2012
A.sp. CCBAU 10948	AET12084.1	Unpublished	Unpublished
A.sp. KH32C	BAL26358.1	Unpublished	Unpublished
A.tolulasticus	WP_018989049.1	Unpublished	Unpublished
A.tolulyticus	WP_076603833.1	Unpublished	Unpublished
A.brasilense Sp245	CCC97492.1	PLoS Genet.	2011
A.lipoferum 4B	CBS86044.1	PLoS Genet.	2011
A.sp. B510	BAI71409.1	DNA Res.	2010
A.chroococcum NCIMB 8003	AJE19640.1	PLoS ONE	2015
A.vinelandii DJ	ACO76403.1	J. Bacteriol.	2009
B.spp	WP_031538364.1	Unpublished	Unpublished
B.canariense	OSJ34399.1	Genome Announc	2017
B.japonicum SEMIA 5079	AHY57040.1	BMC Genomics	2014
B.ottawaense	PAN64902.1	Int. J. Syst. Evol. Microbiol.	2014
D.chrysanthemii	WP_039999404.1	Unpublished	Unpublished
E.amylovora CFBP 1232	CCO95775.1	Unpublished	Unpublished
E.amylovora MR1	CCP05065.1	Unpublished	Unpublished
F.alni ACN14a	CAJ65436.1	Genome Res.	2007
F.sp. EUN1f	EFC83885.1	Unpublished	Unpublished
F.sp. QA3	EIV95132.1	Unpublished	Unpublished
G.diazotrophicus PA1 5	ACI51349.1	Stand Genomic Sci	2010
K.sp. 1_1_55	EFD86074.1	Unpublished	Unpublished
M.nodulans ORS 2060	ACL58883.1	Unpublished	Unpublished
M.sp. 4-46	ACA17917.1	Unpublished	Unpublished
M.sp. WSM2598	WP_018260199.1	Unpublished	Unpublished
P.ATCC 35681	AKG33551.1	Genome Announc	2016
P.durus	WP_046723620.1	Unpublished	Unpublished
P.graminis	AIQ70764.1	Unpublished	Unpublished
P.odorifer	OME57726.1	Unpublished	Unpublished
P.sp. At-9b	ADU72689.1	Unpublished	Unpublished
P.carotovorum	ASY80060.1	Int. J. Syst. Evol. Microbiol.	2017
P.sp. B4(2012c)	AEO14644.2	Unpublished	Unpublished
P.stutzeri KOS6	EWC39832.1	Genome Announc	2013
R.etli	WP_020923547.1	Unpublished	Unpublished
R.etli	ARQ13424.1	Unpublished	Unpublished
R.leguminosarum	WP_062940084.1	Unpublished	Unpublished
S.sp. ATCC 39006	WP_021017294.1	Unpublished	Unpublished

nifH - Fe Prot			
Cepa	Acceso	Revista	Año
S.meliloti AK83	AEG57822.1	Genome Biol Evol	2013
S.meliloti BL225C	AEG07185.1	Genome Biol Evol	2013
S.saheli	OAP39637.1	Unpublished	Unpublished

nifD – Mo-Fe Alpha Subunit			
Cepa	Acceso	Revista	Año
A.sp. KH32C	BAL26359.1	Unpublished	Unpublished
A.tolulyticus	WP_076603832.1	Unpublished	Unpublished
A.brasilense Sp245	CCC97491.1	PLoS Genet.	2011
A.lipoferum 4B	CBS86043.1	PLoS Genet.	2011
A.sp. B510	BAI71408.1	DNA Res.	2010
A.chroococcum NCIMB 8003	AJE19641.1	PLoS ONE	2015
A.vinelandii DJ	ACO76404.1	J. Bacteriol.	2009
B.spp	WP_031538363.1	Unpublished	Unpublished
B.canariense	OSJ33027.1	Genome Announc	2017
B.japonicum SEMIA 5079	AHY57020.1	BMC Genomics	2014
B.ottawaense	PAN63062.1	Int. J. Syst. Evol. Microbiol.	2014
D.chrysanthemi	KGT99095.1	Unpublished	Unpublished
F.alni ACN14a	CAJ65435.1	Genome Res.	2007
F.sp. EUN1f	EFC83884.1	Unpublished	Unpublished
F.sp. QA3	EIV95133.1	Unpublished	Unpublished
G.diazotrophicus PA1 6	ACI51348.1	Stand Genomic Sci	2010
K.sp. 1_1_56	EFD86073.1	Unpublished	Unpublished
M.nodulans ORS 2060	ACL58882.1	Unpublished	Unpublished
M.sp. 4-47	ACA17918.1	Unpublished	Unpublished
M.sp. WSM2599	WP_018260198.1	Unpublished	Unpublished
P.ATCC 35682	AKG35934.1	Genome Announc	2016
P.durus	WP_042206813.1	Unpublished	Unpublished
P.graminis	WP_025703732.1	Unpublished	Unpublished
P.odorifer	AIQ76251.1	Unpublished	Unpublished
P.sp. At-9b	ADU72688.1	Unpublished	Unpublished
P.carotovorum	WP_048259545.1	Unpublished	Unpublished
R.etli	AAB07742.1	J. Bacteriol.	1996
R.leguminosarum	WP_062940085.1	Unpublished	Unpublished
S.sp. ATCC 39006	WP_021017293.1	Unpublished	Unpublished
S.meliloti AK83	AEG57821.1	Genome Biol Evol	2013
S.meliloti BL225C	AEG07186.1	Genome Biol Evol	2013
S.saheli	WP_066877754.1	Unpublished	Unpublished

nifK - Mo-Fe beta			
Cepa	Acceso	Revista	Año
A.sp. KH32C	BAL26360.1	Unpublished	Unpublished
A.tolulyticus	WP_076603831.1	Unpublished	Unpublished
A.brasilense Sp245	CCC97490.1	PLoS Genet.	2011
A.lipoferum 4B	CBS86042.1	PLoS Genet.	2011
A.sp. B510	BAI71407.1	DNA Res.	2010
A.chroococcum NCIMB 8003	AJE19642.1	PLoS ONE	2015
A.vinelandii DJ	ACO76405.1	J. Bacteriol.	2009
B.spp	WP_031538362.1	Unpublished	Unpublished
B.canariense	OSJ33052.1	Genome Announc	2017
B.japonicum SEMIA 5079	AHY57021.1	BMC Genomics	2014
B.ottawaense	PAN63063.1	Int. J. Syst. Evol. Microbiol.	2014
D.chrysanthemi	KGT99096.1	Unpublished	Unpublished
F.alni ACN14a	CAJ65434.1	Genome Res.	2007
F.sp. EUN1f	EFC83883.1	Unpublished	Unpublished
F.sp. QA3	EIV95134.1	Unpublished	Unpublished
G.diazotrophicus PA1 7	ACI51347.1	Stand Genomic Sci	2010
K.sp. 1_1_57	EFD86072.1	Unpublished	Unpublished
M.nodulans ORS 2060	ACL58881.1	Unpublished	Unpublished
M.sp. 4-48	ACA17919.1	Unpublished	Unpublished
M.sp. WSM2600	WP_018260197.1	Unpublished	Unpublished
P.ATCC 35683	AKG35933.1	Genome Announc	2016
P.durus	WP_042207942.1	Unpublished	Unpublished
P.graminis	AIQ70762.1	Unpublished	Unpublished
P.odorifer	AIQ76250.1	Unpublished	Unpublished
P.sp. At-9b	ADU72687.1	Unpublished	Unpublished
P.carotovorum	WP_048259546.1	Unpublished	Unpublished
P.stutzeri KOS6	EWC39830.1	Genome Announc	2013
R.etli	WP_020923545.1	Unpublished	Unpublished
R.leguminosarum	WP_062940086.1	Unpublished	Unpublished
S.sp. ATCC 39006	WP_021017292.1	Unpublished	Unpublished
S.meliloti AK83	AEG57820.1	Genome Biol Evol	2013
S.meliloti BL225C	AEG07187.1	Genome Biol Evol	2013
S.saheli	WP_066877755.1	Unpublished	Unpublished

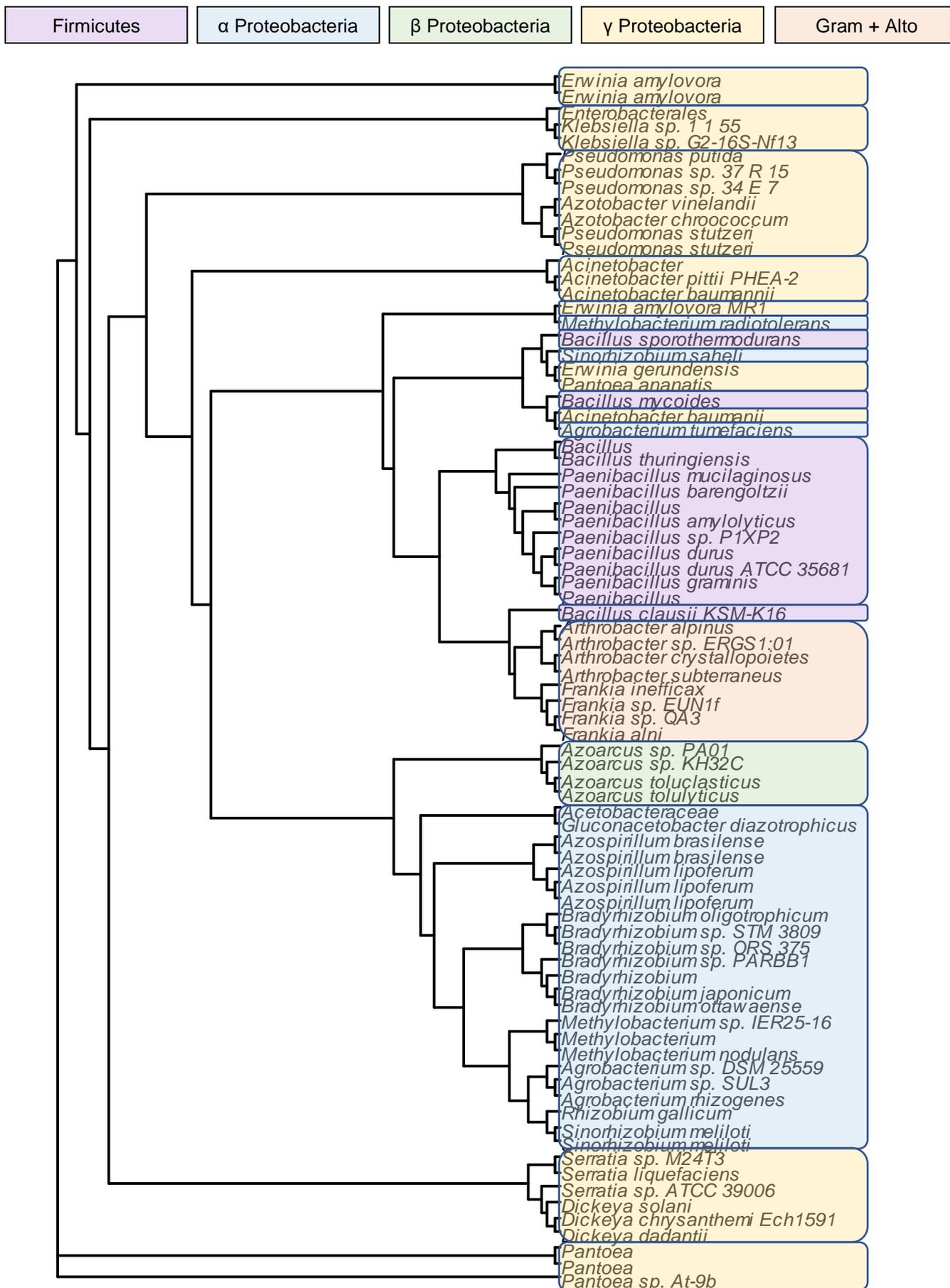
ipdC – Piruvato descarboxilasa			
Cepa	Acceso	Revista	Año
A.pittii	AMX20289.1	Unpublished	Unpublished
A.sp. SUL3	KNY30811.1	Genome Announc	2016
A.sp. PAMC 25486	AIY00516.1	Unpublished	Unpublished
A.sp. KH32C	BAL25295.1	Unpublished	Unpublished
A.sp. PA01]	KON80362.1	Stand Genomic Sci	2015
A.toluclasticus	WP_026295948.1	Unpublished	Unpublished
A.tolulyticus	SIQ60827.1	Unpublished	Unpublished
A.brasilense Sp245	AAT46130.1	Environ. Microbiol.	2005
A.brasilense Sp245 (Plasm)	WP_014239487.1	Unpublished	Unpublished
A.lipoferum	BAB61023.1	Biosci. Biotechnol. Biochem.	2001
B.cereus	CUB46335.1	Unpublished	Unpublished
B.sporothermodurans	KYD05902.1	Genome Announc	2016
B.thuringiensis	OFD10616.1	Unpublished	Unpublished
B.japonicum	WP_038932892.1	Unpublished	Unpublished
B.oligotrophicum S58	BAM86494.1	Appl. Environ. Microbiol.	2013
B.sp. ORS 375	WP_009026768.1	Unpublished	Unpublished
B.sp. ORS 375	CCD91758.1	Unpublished	Unpublished
B.sp. STM 3809	CCE00441.1	Unpublished	Unpublished
D.chrysanthemi	KGT98716.1	Unpublished	Unpublished
D.solani	KHN62512.1	Unpublished	Unpublished
D.zeae	WP_012885658.1	Unpublished	Unpublished
E.amylovora	WP_004158798.1	Unpublished	Unpublished
E.amylovora MR1	CCP07957.1	Unpublished	Unpublished
F.inefficax	WP_013424638.1	Unpublished	Unpublished
G.diazotrophicus PA1 5	CAP54115.1	BMC Genomics	2009
G.sp. SXCC-1	EGG77360.1	J. Bacteriol.	2011
K.michiganensis KCTC 1686	AEX07058.1	J. Bacteriol.	2012
P.mucilaginosus KNP414	AEI43087.1	Genome Announc.	2013
P.polymyxa	ABV24338.1	J. Microbiol. Biotechnol.	2008
P.ananatis LMG 20103	ADD77912.1	J. Bacteriol.	2010
P.sp. CFSAN033090	KOA68224.1	Unpublished	Unpublished
P.carotovorum	ASN84285.1	Unpublished	Unpublished
P.putida	OLS60705.1	Unpublished	Unpublished
P.sp. 34 E 7	CRN02518.1	Molecular Microbiology	2015
P.sp. 37 R 15	CRM14786.1	Molecular Microbiology	2015
R.etli	WP_086084338.1	Unpublished	Unpublished
R.gallicum	WP_018444563.1	Unpublished	Unpublished
R.leguminosarum	WP_047620131.1	Unpublished	Unpublished
S.liquefaciens FK01	GAK27798.1	Unpublished	Unpublished
S.medicae	WP_011970775.1	Unpublished	Unpublished
S.meliloti	ASQ12624.1	Unpublished	Unpublished
S.meliloti	WP_088201732.1	Unpublished	Unpublished

acdS - ACC deaminasa			
Cepa	Acceso	Revista	Año
A.baumannii ABNIH24	EMU51055.1	Genome Res	2013
A.gyllenbergii MTCC 11365	EPH33337.1	Unpublished	Unpublished
A.pittii	ODM01520.1	Unpublished	Unpublished
A.rhizogenes	KEA09258.1	Genome Announc	2014
A.sp. SUL3	WP_082178766.1	Unpublished	Unpublished
A.tumefaciens	PHI00583.1	Unpublished	Unpublished
A.alpinus	ALE93647.1	Unpublished	Unpublished
A.crystallopoietes BAB-32	EMY34629.1	Genome Announc	2013
A.sp. ERGS1:01	ALE04430.1	Unpublished	Unpublished
A.brasilense	OYD80644.1	Unpublished	Unpublished
A.lipoferum	WP_012976927.1	Unpublished	Unpublished
A.sp. B510	BAI75080.1	DNA Res.	2010
A.chroococcum NCIMB 8003	AJE20898.1	PLoS ONE	2015
A.vinelandii DJ	WP_061289107.1	Unpublished	Unpublished
B.cereus	KZD77336.1	Genome Announc	2016
B.mycooides	KZE06807.1	Unpublished	Unpublished
B.thuringiensis	AGQ42751.1	Unpublished	Unpublished
B.canariense	OSJ34166.1	Genome Announc	2017
B.japonicum	KOY09665.1	Genome Announc	2015
B.oligotrophicum S58	BAM89839.1	Appl. Environ. Microbiol.	2013
B.sp. ORS 375	WP_009028524.1	Unpublished	Unpublished
B.sp. STM 3809	WP_008963624.1	Unpublished	Unpublished
D.chrysanthemi	KGT99382.1	Unpublished	Unpublished
D.solani	KHN58195.1	Unpublished	Unpublished
D.zaeae	WP_038906941.1	Unpublished	Unpublished
E.amylovora CFBP1430	CBA21121.1	Mol. Plant Microbe Interact.	2010
F.alni	WP_011605534.1	Unpublished	Unpublished
F.inefficax	WP_013426914.1	Unpublished	Unpublished
F.sp. EUN1f	WP_035952052.1	Unpublished	Unpublished
F.sp. QA3	EIV91108.1	Unpublished	Unpublished
M.nodulans ORS 2060	ACL60323.1	Unpublished	Unpublished
M.sp. 4-46	ACA14842.1	Unpublished	Unpublished
M.sp. WSM2598	WP_018261269.1	Unpublished	Unpublished
P.sp. At-9b	ADU72453.1	Unpublished	Unpublished
P.sp. At-9b	WP_013512282.1	Unpublished	Unpublished
P.fluorescens	KPU59265.1	PLoS ONE 10	2015
P.putida	ADH59751.1	Unpublished	Unpublished
P.sp. 34 E 7	WP_065936098.1	Molecular Microbiology	2015
P.sp. 37 R 15	WP_065953059.1	Molecular Microbiology	2015
R.etli	WP_020923619.1	Unpublished	Unpublished
R.etli	ARQ13419.1	Unpublished	Unpublished
R.gallicum	ABP88062.1	Microb. Ecol.	2009
R.leguminosarum	ABP88088.1	Microb. Ecol.	2009
S.sp. ATCC 39006	ESN63038.1	Genome Announc	2013
S.sp. M24T3	EIC82443.1	J. Bacteriol.	2012
S.medicae	WP_018009224.1	Unpublished	Unpublished
S.meliloti	KKA10522.1	Unpublished	Unpublished
S.meliloti	AJT61603.1	Proc. Natl. Acad. Sci. U.S.A.	2015
S.meliloti BL225C	AEG06520.1	Genome Biol Evol	2013

nirK – Cobre Nitrito Reductasa			
Cepa	Acceso	Revista	Año
A.rhizogenes	AQS63855.1	Genome Announc	2014
A.sp. SUL3	WP_052819230.1	Unpublished	Unpublished
A.tumefaciens	WP_003504637.1	Unpublished	Unpublished
A.subterraneus	SDI14422.1	Unpublished	Unpublished
A.sp. KH32C	WP_015452284.1	Unpublished	Unpublished
A.brasilense	WP_035680909.1	Unpublished	Unpublished
A.brasilense Sp245 (Plasm)	ABY68277.1	Unpublished	Unpublished
A.brasilense Sp245	ABY87187.1	Mol. Plant Microbe Interact.	2008
A.lipoferum	CBS89931.1	PLoS Genet.	2011
A.lipoferum (Plasm)	WP_014249378.1	Unpublished	Unpublished
A.sp. B510	BAI75496.1	DNA Res.	2010
A.sp. B510 (Plasm)	WP_012977327.1	Unpublished	Unpublished
B.solani	KQL19793.1	Unpublished	Unpublished
B.japonicum	WP_039157896.1	Unpublished	Unpublished
B.japonicum SEMIA 5079	AHY54213.1	BMC Genomics	2014
B.oligotrophicum S58	BAM87349.1	Appl. Environ. Microbiol.	2013
B.sp. ORS 375	WP_009028608.1	Unpublished	Unpublished
B.sp. STM 3809	WP_008962965.1	Unpublished	Unpublished
P.amylolyticus	GAS82558.1	Genome Announc	2016
P.barengoltzii G22	EOS56875.1	Unpublished	Unpublished
P.fluorescens	WP_047306321.1	PLoS ONE 10	2015
P.putida	AMK28939.1	Unpublished	Unpublished
P.stutzeri RCH2	AGA86823.1	Unpublished	Unpublished
R.etli	WP_041679288.1	Unpublished	Unpublished
R.gallicum	WP_040116702.1	Unpublished	Unpublished
R.leguminosarum	WP_003590833.1	Unpublished	Unpublished
S.medicae	WP_018210201.1	Unpublished	Unpublished
S.medicae	ABR64876.1	Stand Genomic Sci	2010
S.meliloti	WP_027988776.1	Unpublished	Unpublished
S.meliloti 1021	AAK65339.1	Proc. Natl. Acad. Sci. U.S.A.	2001
S.meliloti BL225C	AEG07313.1	Genome Biol Evol	2013

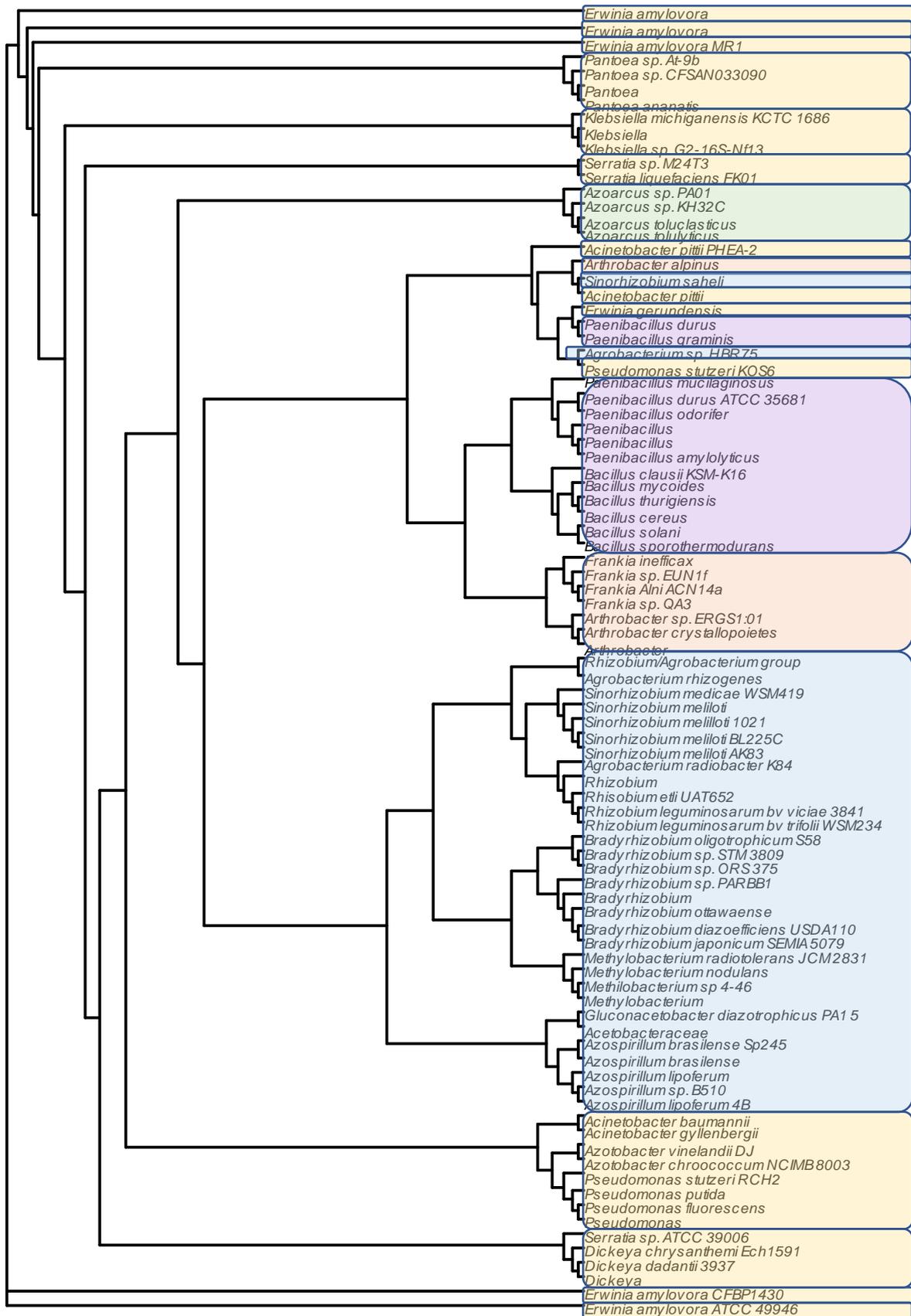
budABC - Ruta de 2-3-butanodiol			
Cepa	Acceso	Revista	Año
A.pittii PHEA-2	ADY81651.1	J. Bacteriol.	2011
A.brasilense	CCC97313.1	PLoS Genet.	2011
A.lipoferum 4B	CBS86092.1	PLoS Genet.	2011
B.cereus	ASZ64678.1	Unpublished	Unpublished
B.cereus	CUB28834.1	Unpublished	Unpublished
B.mycooides	PGR99968.1	Unpublished	Unpublished
B.thuringiensis	PHQ22940.1	Unpublished	Unpublished
B.sp. PARBB1	OYU61772.1	Unpublished	Unpublished
D.dadantii 3937	ADM96873.1	J. Bacteriol.	2011
D.dadantii 3937	ADM96874.1	J. Bacteriol.	2011
E.amylovora	CBJ44688.1	J. Bacteriol.	2010
E.amylovora ATCC 49946	CBJ44687.1	J. Bacteriol.	2010
E.gerundensis	CUU25782.1	Bioinformatics and Systems Biology	2015
K.sp. 1_1_55	EFD86546.1	Unpublished	Unpublished
M.radiotolerans	KZB98327.1	Unpublished	Unpublished
P.sp. P1XP2	KHF27034.1	Unpublished	Unpublished
P.ananatis	CRH37095.1	Austrian Institute of Technology	2015
P.ananatis LMG 20103	ADD79135.1	J. Bacteriol.	2010
P.ananatis LMG 20103	ADD79136.1	J. Bacteriol.	2010
P.carotovorum	ASN87109.1	Unpublished	Unpublished
P.carotovorum	ASN87108.1	Unpublished	Unpublished
P.fluorescens	AJP51914.1	PLoS ONE 10	2015
P.putida	ANY90616.1	Unpublished	Unpublished
R.leguminosarum bv. viciae 3841	CAK11983.1	Genome Biol	2006
S.sp. ATCC 39006	ESN63743.1	Genome Announc	2013

Anexo II. Árboles filogenéticos



Marcadores 1.Árbol filogenético gyrB

Firmicutes	α Proteobacteria	β Proteobacteria	γ Proteobacteria	Gram + Alto
------------	-------------------------	------------------------	-------------------------	-------------



Marcadores 2. Árbol Filogenético con *rpoB*.

Anexo III. Tabla de citas

A partir de de De Souza; Vessey 2003; Ahemad 2014; Pérez-Montaña 2014 y Nadeem 2014; [26], [44], [46], [82], [83],

	Arroz	Maíz	Judía Garbanzo	Alcachofa	Brassica	Algodón	Trigo	Cebada	Tomate	Cacahuete	Soja
Alfaproteobacteria	5	13	27	0	0	3	0	4	4	0	8
Betaproteobacteria	4	0	0	0	0	0	0	0	0	0	0
Gammaproteobacteria	7	10	9	0	7	2	9	0	0	7	12
Firmicutes	15	0	0	0	6	0	0	6	6	0	15
Gram+alto [G+C]	0	0	0	0	0	0	0	0	0	0	0

Anexo IV. Script de R

```
rm(list = ls())
require(ape)
library(xlsx)
library(readxl)
library(reutils)
library(biofiles)
library(rentrez)
library(XML)
library(dplyr)
library(msa)
library(seqinr)
library(tools)

nombres<- read_xlsx("prot_selec.xlsx", sheet = 1)
head(nombres)
nombres<-nombres[which(nombres$Patógena!="Ser humano" | is.na(nombres$
Patógena)), -6] #Selección de no patógenas y exclusión de columna de pa
togenicidad

genes<-vector()
n=1
for(i in unique(nombres$Gen)){
  genes[n]<-i
  n=n+1
}
genes

#identificadores
ids<-vector()
ids<-nombres$Acceso

dir.create("./Txt", showWarnings = FALSE) #Generación de directorios
dir.create("./Fasta", showWarnings = FALSE)
dir.create("./Tablas", showWarnings = FALSE)

spcs<-unique(paste(substr(nombres$Genero,1,1),nombres$Especie, sep="."
))
nombres$contr<-paste(substr(nombres$Genero,1,1),nombres$Especie, sep="
.")
accesos<-matrix(ncol = length(genes), nrow = length(spcs))
accesos<-data.frame(accesos, row.names = spcs)
colnames(accesos)<-genes

for(i in 1:nrow(nombres)){
  n<-nombres$contr[i]
  g<-nombres$Gen[i]
  accesos[n,g]<-nombres[i, "Acceso"]
}
write.xlsx(accesos, "./Tablas/accesos.xlsx")
write.table(accesos, "./Tablas/accesos.txt")
```

```

#PREPARACIÓN PARA ALINEAMIENTO

#Generar FASTA
##Descarga de información

#Tabla con referencias

for(i in 1:nrow(nombres)){
  tryCatch({

    #Tomar datos en xml
    linked_seq2 <- entrez_fetch(ids[i], db="protein", rettype = "gbcd",
retmode = "xml", parsed = TRUE)
    xml_list2 <- xmlToList(linked_seq2)

    #Seleccionar autor
    autor<-xml_list2$INSDSeq$INSDSeq_references$INSReference$INSReference_authors$INSDAuthor
    if(is.null(autor)==FALSE){nombres[which(nombres$Acceso==ids[i]),
"Autor"]<-autor}

    #Seleccionar título
    titulo<-xml_list2$INSDSeq$INSDSeq_references$INSReference$INSReference_title
    if(is.null(titulo)==FALSE){nombres[which(nombres$Acceso==ids[i]),
"Titulo"]<-titulo}

    #Seleccionar revista
    journal<-xml_list2$INSDSeq$INSDSeq_references$INSReference$INSReference_journal
    if(is.null(journal)==FALSE) nombres[which(nombres$Acceso==ids[i]),
"Revista"]<-journal

    #Seleccionar año
    year <-gsub(".*\\((.*)\\).*", "\\1", journal)
    if(is.null(year)==FALSE)nombres[which(nombres$Acceso==ids[i]),
"Año"]<-year

    #Seleccionar referencia pubmed
    pm_id <- xml_list2$INSDSeq$INSDSeq_references$INSReference$INSReference_pubmed
    if(is.null(pm_id)==FALSE) nombres[which(nombres$Acceso==ids[i]),
"Pubmed"]<-pm_id
  }, error=function(e){})
}

View(nombres)

```

```

#Creación de nueva tabla a exportar
tabla_nombres<-nombres[,c(6:3,9:11)]
colnames(tabla_nombres)[1]<-"Cepa"
write.xlsx(as.data.frame(tabla_nombres), "./Tablas/tabla_nombres.xlsx"
, row.names = FALSE)

##Descarga de Secuencias
for(i in 1:length(genes)){

  gen_name<-genes[i] #Asignación del nombre del gen
  prot_name<-nombres[nombres$Gen==gen_name, "Prot"][[1]][1]
#Asignación nombre de la proteína
  prot_name_c<-substr(prot_name, 1,10) #Recortar el nombre completo de
  la proteína

  id<-as.vector(nombres[nombres$Gen==gen_name, "Acceso"])
#Identificador de acceso
  enlace <- entrez_link(db="protein", dbfrom="protein", id=id) #Link de
  acceso
  str(enlace)

  pub_id <- enlace$links$protein_protein_cdart #Número de acceso en
  database "protein"
  str(pub_id)

  linked_seq <- entrez_fetch(id=pub_id, db="protein", rettype =
  "fasta", retmode = "txt") #Descarga de FASTA

  gen_df<-tabla_nombres[tabla_nombres$Gen==gen_name,1:4] #Generación
  de tabla por cada gen estudiado
  gen_df$link<-enlace$links$protein_protein_cdart

  write.table(gen_df, paste("./Txt/",prot_name_c, ".txt", sep = "" ))
#Exportar tabla

  gen_fasta<-paste(substr(gen_name,1,4), ".fasta", sep="") #cambiado
  para ipdc/"
  write(linked_seq, paste("./Fasta/", gen_fasta, sep = "")) #Exportar
  secuencias en FASTA

##Corrección de nombres

  gen_seq<-readAAStringSet(paste("./Fasta/", gen_fasta, sep = ""))
  names(gen_seq)<- sub("^[^[]*", "", names(gen_seq))
  names(gen_seq)<- gsub(".* \\[", "", names(gen_seq))
  names(gen_seq)<- gsub("\\[|\\]", "", names(gen_seq))

##Generar tabla con secuencia

  tabla<-as.data.frame(matrix(nrow = length(names(gen_seq)), row.names
  = names(gen_seq))
  tabla[, "Gen"]<-gen_name
  tabla[, "Proteína"]<-prot_name
  tabla[, "Prot. Num."]<-pub_id
  tabla[, "Long"]<-paste(gen_seq@ranges@width, " aa", sep="")
  tabla<-tabla[order(rownames(tabla)),]

```

```

tabla[, "V1"]<-nombres[nombres$Gen==gen_name, "Acceso"]
colnames(tabla)[1]<-"NCBI id"

write.table(tabla, file = paste("./Tablas/", "Tabla_",
substr(gen_name,1,4), ".txt", sep = "")) #Exportar tabla por grupo
write.csv(tabla, paste("./Tablas/", "Tabla_", substr(gen_name,1,4),
".csv", sep = "")) #Exportar en csv
}

#GENERACIÓN DEL ALINEAMIENTO
nombres<- read_xlsx("prot_selec.xlsx", sheet = 1)
head(nombres)
nombres<-nombres[which(nombres$Patógena!="Ser humano" | is.na(nombres$
Patógena)), -6] #Selección de no patógenas y exclusión de columna de pa
togenicidad

genes<-vector()
n=1
for(i in unique(nombres$Gen)){
  genes[n]<-i
  n=n+1
}
genes

#identificadores
ids<-vector()
ids<-nombres$Acceso

for(i in 1:length(genes)){

  tryCatch({

    gen_name<-genes[i] #Asignación del nombre del gen
    prot_name<-nombres[nombres$Gen==gen_name, "Prot"][[1]][1] #Asignació
n nombre de la proteína
    prot_name_c<-substr(prot_name, 1,10) #Recortar el nombre completo de
la proteína

    gen_fasta<-paste(substr(gen_name,1,4), ".fasta", sep="") #Toma de da
tos de FASTA
    gen_seq<-readAAStringSet(paste("./Fasta/", gen_fasta, sep = "")) #Le
ctura de FASTA

    ##Corrección de nombres

    gen_seq<-readAAStringSet(paste("./Fasta/", gen_fasta, sep = ""))
    names(gen_seq)<- sub("^[^[]*", "", names(gen_seq))
    names(gen_seq)<- gsub(".* \\[", "", names(gen_seq))
    names(gen_seq)<- gsub("\\[|\\]", "", names(gen_seq))

    ##Alineamiento
    gen_alin<-msa(gen_seq)

```

```

#Gráfico de alineamiento

dir.create("./Alin", showWarnings = FALSE)
gen_alin_file<-paste("./Alin/Alin_", substr(prot_name,1,2),"_", substr(gen_name,1,4), ".tex", sep = "" )
gen_alin_fasta<- paste("./Alin/Alin_", substr(prot_name,1,2),"_", substr(gen_name,1,4), ".fasta", sep = "")

msaPrettyPrint(gen_alin, output="tex", askForOverwrite=FALSE, verbose = FALSE, file = gen_alin_file, alFile = gen_alin_fasta)
texi2pdf(gen_alin_file, clean=TRUE)

#Generación de matriz de distancias entre secuencias
gen_alin2 <- msaConvert(gen_alin, type="seqinr::alignment")
gen_dist <- dist.alignment(gen_alin2, "identity")

#Árbol filogenético
dir.create("./Arbol", showWarnings = FALSE)
gen_tree <- nj(gen_dist)
pdf(file = paste("./Arbol/filo_", prot_name,"_", gen_name, ".pdf", sep=""))
plot(gen_tree, main= paste(prot_name, gen_name, sep="-"),use.edge.length=FALSE, cex=0.4)
dev.off()

}, error=function(e){})
}

```

#GENERACIÓN DE FILOGENIA CON MARCADORES

```

gyrB

nombres<-read.xlsx("Tablas/nombres_completo.xlsx", sheetIndex = 1, stringsAsFactors=FALSE)
head(nombres)

nombres[, "Gen.Es"]<-paste(nombres$Genero, nombres$Especie, sep = " ")
marcadores<-data.frame(nombres[, "Gen.Es"])
marc<-data.frame(unique(marcadores$nombres....Gen.Es..))
names(marc)<- "Cepa"
marc$Cepa<-as.character(marc$Cepa)

#Búsqueda id gyrB

for(i in 1:nrow(marc)){
  tryCatch({
    busqueda<-entrez_search("protein", paste("gyrB",marc$Cepa[i], sep = " "))
    gyrB<-busqueda$ids[1]
    if(is.null(gyrB)==FALSE) marc[i,"gyrB"]<-gyrB
    linked_seq <- entrez_fetch(gyrB, db="protein", rettype = "gbc", retmode = "xml", parsed = TRUE)
    xml_list2 <- xmlToList(linked_seq)
    gyrB_def<-xml_list2$INSDSeq$INSDSeq_definition
    if(is.null(gyrB_def)==FALSE) marc[i,"gyrB.title"]<-gyrB_def

    busqueda<-entrez_search("protein", paste("rpoB", marc$Cepa[i], sep = " "))
    rpoB<-busqueda$ids[1]

```

```

    if(is.null(rpoB)==FALSE) marc[i,"rpoB"]<-rpoB
    linked_seq <- entrez_fetch(rpoB, db="protein", rettype = "gbc",
retmode = "xml", parsed = TRUE)
    xml_list2 <- xmlToList(linked_seq)
    rpoB_def<-xml_list2$INSDSeq$INSDSeq_definition
    if(is.null(rpoB_def)==FALSE) marc[i,"rpoB.title"]<-rpoB_def
  }, error=function(e){})
}

write.xlsx(as.data.frame(marc), "../Tablas/tabla_marcadores.xlsx", row.
names = FALSE)

marc<-read.xlsx("../Tablas/tabla_marcadores.xlsx", sheetIndex = 1, stri
ngsAsFactors=FALSE)
for(i in c(2,4)){
  gen_name<-colnames(marc)[i]
  enlace <- entrez_link(db="protein", dbfrom="protein", id=marc[,i]) #
Link de acceso

  pub_id <- enlace$links$protein_protein_cdart #Número de acceso en da
tabase "protein"

  linked_seq <- entrez_fetch(id=pub_id, db="protein", rettype = "fasta
", retmode = "txt")

  #Descarga de FASTA

  gen_fasta<-paste(gen_name, ".fasta", sep="") #cambiado para ipdc/"
write(linked_seq, paste("../Fasta/", gen_fasta, sep = "")) #Exportar
secuencias en FASTA
}

# Elaboración de alineamiento

gen_name<-names(marc)[2]
gen_fasta<-paste(gen_name, ".fasta", sep="")
gen_seq<-readAAStringSet(paste("../Fasta/", gen_fasta, sep = ""))
names<-names(gen_seq)

names(gen_seq)<-names
names(gen_seq2)
names(gen_seq)<-sub("^[^[]*", "", names(gen_seq))
names(gen_seq)<-gsub(".* \\[", "", names(gen_seq))
names(gen_seq)<-gsub("\\[|\\]", "", names(gen_seq))
names(gen_seq)[c(86,87)]<-c("Bacillus mycoides", "Acinetobacter baumani
i")
gen_seq2<-gen_seq[-c(76:78,80,83:85)]

gen_seq<-gen_seq2
alin0<-AlignSeqs(gen_seq, iterations = 0, refinements = 0)
stag<-StaggerAlignment(alin0)
dist<-stringDist(stag, method = "hamming")

Elaboración de árbol gyrB
tree<-nj(dist)
tree$tip.label

```

```

plot(tree, cex=1, type="f", use.edge.length=FALSE, no.margin=TRUE)
nodelabels()

rpoB

gen_name<-names(marc)[4]
gen_fasta<-paste(gen_name, ".fasta", sep="")
gen_seq<-readAAStringSet(paste("./Fasta/", gen_fasta, sep = ""))
# phyDat(gen_fasta, type = "AA")

names(gen_seq)
names(gen_seq)<-sub("^[^[]*", "", names(gen_seq))
names(gen_seq)<-gsub(".* \\[", "", names(gen_seq))
names(gen_seq)<-gsub("\\[|\\]", "", names(gen_seq))
names(gen_seq)[c(68:72,82,83,88,89)]<-c("Pseudomonas fluorescens", "Az
otobacter vinelandii DJ", "Methylobacterium nodulans", "Bacillus cereus
", "Agrobacterium radiobacter K84", "Bacillus thurigiensis", "Frankia A
lni ACN14a", "Sinorhizobium meliloti 1021", "Bradyrhizobium diazoeffi
ciens USDA110")
names(gen_seq)[c(74:79,81)]<-c("Rhizobium leguminosarum bv trifolii WS
M234", "Rhisobium etli UAT652", "Methilobacterium sp 4-46", "Methyloba
cterium radiotolerans JCM 2831", "Bacillus mycoides", "Sinorhizobium m
edicae WSM419", "Rhizobium leguminosarum bv viciae 3841")
alin0<-AlignSeqs(gen_seq, iterations = 0, refinements = 0)
stag<-StaggerAlignment(alin0)

dist<-stringDist(stag, method = "hamming")

Árbol filogenético
tree<-nj(dist)

# tree$tip.label
plot(tree, cex=0.48, type="f", use.edge.length=FALSE, no.margin=TRUE)
plot(tree, cex=1, type="f", use.edge.length=FALSE, no.margin=TRUE)

trk<-tree
trc<-tree2

layout(matrix(1:2, 1, 2), width = c(1.4, 1))

par(mar = c(0, 0, 0, 0))

plot(trk, adj=0.01, cex=0.48, x.lim=110, use.edge.length=FALSE, edge.c
olor= "blue")

axisPhylo()

plot(trc, adj=0.01, show.tip.label= FALSE, use.edge.length=FALSE, edge
.color= "red")

axisPhylo()

```