
Caracterización genómica y desarrollo de un sistema de edición génica a través de CRISPR/Cas9 de bacterias promotoras del crecimiento de plantas aisladas de *Saccharum officinarum* para una mejora en la fijación de nitrógeno

Andrés De Jesús Say Agosto



UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ciencias y Humanidades



Caracterización genómica y desarrollo de un sistema de edición génica a través de CRISPR/Cas9 de bacterias promotoras del crecimiento de plantas aisladas de *Saccharum officinarum* para una mejora en la fijación de nitrógeno

Trabajo de graduación en modalidad de Tesis presentado por
Andrés De Jesús Say Agosto para optar al grado académico de
Licenciado en Bioquímica y Microbiología

Guatemala,
2023

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ciencias y Humanidades



Caracterización genómica y desarrollo de un sistema de edición génica a través de CRISPR/Cas9 de bacterias promotoras del crecimiento de plantas aisladas de *Saccharum officinarum* para una mejora en la fijación de nitrógeno

Trabajo de graduación en modalidad de Tesis presentado por
Andrés De Jesús Say Agosto para optar al grado académico de
Licenciado en Bioquímica y Microbiología

Guatemala,
2023

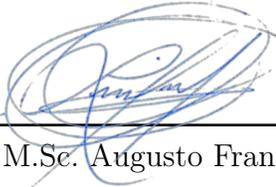
Vo.Bo.:

(f) 
M.Sc. Isabella García Caffaro

Tribunal examinador:

(f) 
M.Sc. Isabella García Caffaro

(f) 
Ph.D. Dalia Lau-Bonilla

(f) 
M.Sc. Augusto Franco

Fecha de aprobación: Guatemala, 12 de diciembre del 2023.

El presente trabajo es parte de un proyecto denominado “Mejoramiento de la salud de suelo de monocultivo a través de la optimización de procedimiento de compostaje y manejo de microbiología de suelos” desarrollado por el Centro de Estudios en Biotecnología y el Ingenio Magdalena. El propósito del estudio surge de la necesidad de entender a mayor profundidad las características promotoras de crecimiento de bacterias promotoras del crecimiento vegetal aisladas de suelo de cultivo de caña de azúcar.

Quisiera agradecer a todas las personas e instituciones que me apoyaron durante el desarrollo del presente trabajo de graduación y durante mi carrera profesional, principalmente:

A mi asesora principal, M.Sc. Isabella García Caffaro, coordinadora del proyecto, por su constante apoyo y dedicación durante el diseño y ejecución de esta investigación. Sus excelentes consejos y su orientación han sido fundamentales para mi mejora como científico.

A mi segunda asesora, la Dra. Dalia Lau-Bonilla, quien además de haberme apoyado durante todo este trabajo, ha sido mi catedrática por varios años. Agradezco sobremanera su mentoría, su pasión por la educación y en la ciencia han sido un pilar fundamental para mi desarrollo académico y como futuro científico. Agradezco los años de enseñanza y su apoyo constante en todos los proyectos que he desarrollado.

Al Ingenio Magdalena y al Centro de Estudios en Biotecnología por darme la oportunidad de participar en este proyecto de investigación y por proporcionar las herramientas para poder llevar a cabo el trabajo.

A M.Sc. Augusto Franco, por ser un excelente catedrático, asimismo por apoyarme en el uso y análisis de las herramientas bioinformáticas empleadas en este trabajo. Adicionalmente, quisiera agradecer a M.Sc. Andrés Grajeda por su asesoría para llevar a cabo la metodología de secuenciación.

A M.Sc. Miguel Morales y Marie Cosenza. Su orientación y apoyo desde los primeros años me han ayudado a mejorar en notablemente personal y académicamente. Les agradezco por creer siempre en mí. Asimismo, quisiera agradecerles por su apoyo, junto al Dr. Diego Archila, por asesorarme en el área de biología molecular.

A mis padres, Emilio Say y Vanesa Agosto de Say, por ser mi principal ejemplo de constancia y esfuerzo. Por apoyarme en todos los proyectos que realizo y por enseñarme a siempre dar lo mejor de mí mismo. Asimismo, quisiera agradecer a mis hermanos, Emilio y Monika, quienes siempre han creído en mí y me han alentado durante toda mi carrera profesional, su desarrollo en sus propias carreras ha sido un gran ejemplo para mí.

Por último, quisiera darle las gracias a mis todos mis compañeros y amigos de carrera. Especialmente a Melissa Caballeros, Ximena Porras, Silvia Vela y Josué Castañón por todo su apoyo desde el comienzo de este trabajo y por todas las conversaciones que nos mantenían entusiasmados durante nuestros grandes periodos en el laboratorio.

Prefacio	IV
Lista de figuras	VII
Lista de Cuadros	VIII
Resumen	IX
Abstact	X
1. Introducción	1
2. Objetivos	3
2.1. Objetivo general	3
2.2. Objetivos específicos	3
3. Justificación	4
4. Marco contextual	6
5. Marco teórico	7
5.1. Caña de azúcar	7
5.2. Fertilizantes	8
5.2.1. Fertilizantes químicos	8
5.2.2. Biofertilizantes	8
5.3. Salud del suelo	11
5.3.1. Composición y microorganismos del suelo	11
5.3.2. Bacterias promotoras del crecimiento vegetal	13
5.4. Genómica	21
5.4.1. Secuenciación de nueva generación (NGS)	21
5.4.2. Caracterización genómica	21
5.5. Algoritmos de análisis bioinformático y ensamblaje de genomas	22
5.5.1. Evaluación de calidad de secuencias	22
5.5.2. Ensamblaje de genoma	23
5.6. Ingeniería genética	27

6. Metodología	30
6.1. Diseño, enfoque y tipo de investigación	30
6.2. Población y muestra	30
6.3. Variables	30
6.4. Hipótesis	31
6.4.1. Hipótesis nula	31
6.4.2. Hipótesis alternativa	31
6.5. Metodología	31
6.5.1. PCR de colonia del gen <i>16S</i>	31
6.5.2. Extracción de ADN de bacterias	32
6.5.3. Amplificación del gen <i>16S</i>	32
6.5.4. Amplificación del gen <i>nifH</i>	33
6.5.5. Secuenciación del gen <i>16S</i>	33
6.5.6. Identificación de la especie bacteriana	33
6.5.7. Selección de bacterias candidatas	33
6.5.8. Secuenciación de genoma completo	34
6.5.9. Análisis bioinformático	34
6.5.10. Desarrollo de mecanismo de mejoramiento por ingeniería genética	39
7. Resultados	41
7.1. Identificación de cepas bacterianas por <i>16S rRNA</i>	41
7.2. Selección de bacterias para la secuenciación de genoma completo (WGS)	42
7.3. Análisis bioinformático de WGS de bacterias	43
7.4. Identificación de genes asociados a la fijación de nitrógeno	49
7.5. Identificación de genes asociados a biodisponibilidad de fósforo	50
7.6. Identificación de genes asociados a la producción de ácido indol-3-acético (IAA)	52
7.7. Identificación de genes asociados a la producción de sideróforos	53
7.8. Identificación de genes reportados para el control de <i>Fusarium</i>	54
7.9. Identificación de factores de virulencia	55
7.10. Sistema de edición génica propuesto para la optimización en la fijación de nitrógeno	57
8. Discusión	62
9. Conclusiones	70
10.Recomendaciones	71
11.Bibliografía	72
12.Anexos	88
Anexo A. Análisis bioinformático	88
Anexo B. Sistema de clonación	93
Anexo C. Metadatos de secuenciación	93
Anexo D. Protocolos	97
Anexo E. Información suplementaria	113

Lista de figuras

Figura 1.	Tipos de formulaciones biológicas	10
Figura 2.	Diversidad de abundancia de bacterias y hongos en el suelo	12
Figura 3.	Diversidad y abundancia de protistas y arqueas en el suelo	13
Figura 4.	Interacciones en el microbioma impacta la salud del suelo	14
Figura 5.	Los procesos biogeoquímicos son afectados por el microbioma	15
Figura 6.	Ciclo del nitrógeno bacteriano	17
Figura 7.	Ciclo del fósforo	18
Figura 8.	Mecanismos de solubilización	19
Figura 9.	Efecto de tratamientos con IAA	20
Figura 10.	Ensamblaje <i>de Novo</i>	24
Figura 11.	Esquemática del sistema <i>Velvet</i>	25
Figura 12.	Implementación del algoritmo de <i>SPAdes</i>	26
Figura 13.	Ruta de interferencia por el sistema CRISPR/Cas9 tipo II	29
Figura 14.	Metodología general de secuenciación y análisis bioinformático	35
Figura 15.	Abundancia de géneros en suelo de <i>Saccharum officinarum</i>	41
Figura 16.	Amplificación de <i>nifH</i> de las bacterias seleccionadas	43
Figura 17.	Determinación de genes para la fijación de nitrógeno. Se muestran los genes identificados en las cepas estudiadas	49
Figura 18.	Determinación de genes para aumentar la biodisponibilidad de fósforo	51
Figura 19.	Genes identificados en la ruta de síntesis de IAA	52
Figura 20.	Estructura y distribución de genoma de <i>B. subtilis</i>	57
Figura 21.	Plásmido inducible por ácido salicílico	58
Figura 22.	Plásmido inducible por OC6	59
Figura 23.	Plásmido inducible por OCH14	60
Figura 24.	Sistema de edición KO-CRISPR para el gen <i>nifL</i>	61
Figura 25.	Alineamiento de <i>B. tropicus</i> 1AB11 con genoma de referencia por FastANI	88
Figura 26.	Alineamiento de <i>P. paraglycinae</i> 2MO6 con genoma de referencia por FastANI	89
Figura 27.	Alineamiento de <i>P. paraglycinae</i> 2MO7 con genoma de referencia por FastANI	90
Figura 28.	Alineamiento de <i>B. subtilis</i> 2MO12 con genoma de referencia por FastANI	90
Figura 29.	Alineamiento de <i>P. paraglycinae</i> 2MO14 con genoma de referencia por FastANI	91
Figura 30.	Alineamiento de <i>A. baumannii</i> 4MO3 con genoma de referencia por FastANI	91
Figura 31.	Alineamiento de <i>P. canadensis</i> 2AP13 con genoma de referencia por FastANI	92
Figura 32.	Alineamiento de <i>P. canadensis</i> 2AP16 con genoma de referencia por FastANI	92
Figura 33.	Propuesta de clonación del gen <i>nifA</i> en los plásmidos inducibles	93
Figura 34.	Calidad de bibliotecas de las muestras 1-10	96
Figura 35.	Calidad de bibliotecas de las muestras 11-20	96

Lista de cuadros

Cuadro 1.	Definición de las variables de investigación	30
Cuadro 2.	Receta para la amplificación del gen <i>16S</i> por PCR de colonia	31
Cuadro 3.	Condiciones de amplificación de <i>16S</i> para PCR de colonia	32
Cuadro 4.	Receta para amplificación de <i>16S</i>	32
Cuadro 5.	Condiciones de amplificación de <i>16S</i>	33
Cuadro 6.	Bacterias candidatas con capacidad de promover crecimiento vegetal	42
Cuadro 7.	Calidad de secuencias crudas caracterizadas por <i>fastQC</i>	44
Cuadro 8.	Calidad de secuencias filtradas por <i>Trimmomatic</i>	45
Cuadro 9.	Calidad de los ensamblajes con <i>ABYSS</i>	46
Cuadro 10.	Calidad de los ensamblajes con <i>Velvet</i>	46
Cuadro 11.	Calidad de los ensamblajes con <i>SPAdes</i>	47
Cuadro 12.	Determinación de genoma de referencia por <i>GTDB-Tk</i>	48
Cuadro 13.	Genes identificados para la producción de sideróforos	53
Cuadro 14.	Genes identificados para el control potencial de <i>Fusarium</i>	54
Cuadro 15.	Genes asociados a factores de virulencia	55
Cuadro 16.	Identificación con <i>GTDB-Tk</i> de la muestra 1AB11	88
Cuadro 17.	Identificación con <i>GTDB-Tk</i> de la muestra 1AP23	89
Cuadro 18.	Identificación con <i>GTDB-Tk</i> de la muestra 2MO6	89
Cuadro 19.	Identificación con <i>GTDB-Tk</i> de la muestra 2MO7	89
Cuadro 20.	Identificación con <i>GTDB-Tk</i> de la muestra 2MO12	90
Cuadro 21.	Identificación con <i>GTDB-Tk</i> de la muestra 2MO14	90
Cuadro 22.	Identificación con <i>GTDB-Tk</i> de la muestra 4MO3	91
Cuadro 23.	Identificación con <i>GTDB-Tk</i> de la muestra 4MO3	91
Cuadro 24.	Identificación con <i>GTDB-Tk</i> de la muestra 2AP16	92
Cuadro 25.	Identificación con <i>GTDB-Tk</i> de la muestra 2AP12	92
Cuadro 26.	Pureza y concentración para las bibliotecas genómicas	94
Cuadro 27.	Preparación de bibliotecas genómicas	95

La diversidad de microorganismos cumple un rol indispensable en la salud del suelo. Las bacterias promotoras del crecimiento vegetal (PGPB) han sido ampliamente estudiadas por su potencial uso para sustituir la aplicación de agroquímicos y desarrollar de esta manera sistemas de cultivo más sostenibles. Para esto, es necesario entender a profundidad las características promotoras del crecimiento vegetal, para poder optimizarlas a través de técnicas de biología molecular. Por esta razón, el propósito del presente estudio fue llevar a cabo una identificación y selección de cepas candidatas para secuenciación de genoma completo de bacterias aisladas de suelo de cultivo de *saccharum officinarum* para llevar a cabo una caracterización de genes implicados en características promotoras del crecimiento vegetal mediante herramientas bioinformáticas. Asimismo, llevar a cabo un diseño de un sistema de edición génica a través de tecnología CRISPR para optimizar la ruta de fijación de nitrógeno. Se determinó varios genes implicados en las rutas de fijación de nitrógeno, biosíntesis de ácido indol-3-acético, sideróforos y de compuestos que presentan potencial para el control de patógenos fúngicos. Las cepas estudiadas no presentaron un potencial de ser patogénicas para humanos. Por último, se propuso una transformación a través de plásmidos inducibles para clonar el gen nativo *nifA* y un *knock-out* del gen *nifL* a través de CRISPR/Cas9 para aumentar la fijación de nitrógeno de *B. subtilis*. Se recomienda llevar a cabo una secuenciación con mayor profundidad y ampliar la identificación de genes implicados en la fijación de nitrógeno mediante PCR.

Palabras clave: *Bacterias promotoras del crecimiento vegetal, caracterización de genes, análisis bioinformático, edición génica, CRISPR*

The diversity of microorganisms plays an indispensable role in soil health. Plant-growth-promoting bacteria (PGPB) have been widely studied for their potential use to substitute the application of agrochemicals and thus develop more sustainable cropping systems. For this, it is necessary to understand in depth the plant growth-promoting characteristics to optimize them through molecular biology techniques. For this reason, the present study aimed to carry out an identification and selection of candidate strains for whole genome sequencing of bacteria isolated from *saccharum officinarum* cultivation soil to carry out a characterization of genes involved in plant growth-promoting characteristics using bioinformatics tools. In addition, carry out a design of a gene editing system through CRISPR technology to optimize the nitrogen fixation pathway. Several genes involved in nitrogen fixation pathways, indole-3-acetic acid biosynthesis, siderophores and compounds with potential for fungal pathogen control were determined. The strains studied did not present a potential to be pathogenic to humans. Finally, a transformation through inducible plasmids was proposed to clone the native *nifA* gene and a knock-out of the *nifL* gene through CRISPR/Cas9 to increase nitrogen fixation of *B. subtilis*. Further sequencing and expanded identification of genes involved in nitrogen fixation by PCR is recommended.

Keywords: *Plant-growth-promoting bacteria, gene characterization, bioinformatics analysis, gene editing, CRISPR*

Los sistemas de monocultivo y la aplicación constante de agentes químicos como fertilizantes y pesticidas impactan la salud del suelo y disminuyen su diversidad microbológica (Tripathi *et al.*, 2020). Los microorganismos del suelo se encargan de llevar a cabo varios procesos metabólicos que aumentan la biodisponibilidad de muchos nutrientes y mejoran la capacidad de asimilación de estos por las plantas (Fall *et al.*, 2013); por lo tanto, al verse afectados por compuestos químicos, el rendimiento de cultivo se ve afectado (Xu *et al.*, 2020). A pesar del efecto negativo de los compuestos químicos a la microbiota, no es posible solamente dejar de utilizarlos, ya que permiten mantener una producción alta y así abastecer la demanda mundial (Gouda *et al.*, 2018). Además, el abuso de pesticidas puede llevar al desarrollo de resistencia, asimismo, el uso descontrolado de fertilizantes puede llegar a tener repercusiones en la salud humana (Bijay-Singh *et al.*, 2021). Por lo anterior, surge una necesidad de desarrollar alternativas que fomenten el crecimiento de las plantas y las protejan de organismos patógenos, al mismo tiempo que no afecten de forma negativa a la salud humana y a la salud del suelo de cultivo a largo plazo (Bijay-Singh *et al.*, 2021).

Las bacterias de suelo cumplen un rol sumamente importante en los ecosistemas, estas actúan directamente en varios procesos biogeoquímicos que modulan la disponibilidad de varios nutrientes necesarios para todos los organismos (Fierer, 2017). Por ejemplo, el nitrógeno, uno de los compuestos más importantes para todos los seres vivos, se encuentra en su mayoría como nitrógeno gaseoso (N_2), por lo tanto, existen bacterias fijadoras de nitrógeno que convierten el nitrógeno gaseoso en compuestos como amoníaco, el cuál puede ser asimilado por las plantas (mantelin , 2004). De igual manera, las bacterias están implicadas en la biodisponibilidad de otros compuestos como el fósforo (P) y potasio (K), y pueden formar relaciones mutualistas con las plantas (Phares *et al.*, 2022). Particularmente existe un grupo de bacterias que se caracteriza por fomentar el crecimiento de plantas conocidas como (PGPB, por sus siglas en inglés), estas han sido estudiadas como biofertilizantes e incluso como biocontroladores (Sokol *et al.*, 2022; Fierer, 2017).

A pesar de que se han realizado estudios para utilizar las PGPB como biofertilizantes o bioformulaciones que fomenten el crecimiento de un cultivo particular (Tahat *et al.*, 2020). Es importante mencionar que también es posible realizar modificaciones genéticas en estas bacterias para potenciar sus capacidades y mejorar el rendimiento de los cultivos. Para esto, primero es necesario entender qué genes se encuentran en las bacterias y como pueden ser modificados, cabe mencionar que se ha demostrado que las PGPB tienen mejores capacidades si son aisladas del suelo donde quieren ser utilizadas (Williams, Liede, Fahy, & al., 2020; Yadav *et al.*, 2020). El presente estudio pretende realizar una identificación de bacterias de suelo, para luego llevar a cabo un proceso

de secuenciación de nueva generación (WGS) con el propósito de estudiar los genes involucrados en rutas de interés como fijación de nitrógeno en las bacterias que presenten mayor potencial, con el objetivo de desarrollar un sistema mutagénesis dirigida a través de CRISPR/Cas9.

2.1 Objetivo general

Realizar una caracterización y análisis genómico de cepas bacterianas para la promoción del crecimiento de plantas y desarrollo de un sistema de edición génica a través de CRISPR/Cas9 para optimizar la ruta de fijación de nitrógeno

2.2 Objetivos específicos

- Seleccionar cepas candidatas para la fijación de nitrógeno para secuenciación de genoma completo
- Seleccionar el algoritmo de ensamblaje con mayor calidad para cada una de las muestras secuenciadas
- Identificar genes de interés para la promoción del crecimiento de la caña de azúcar
- Identificar posibles genes patogénicos
- Identificar genes involucrados en la ruta de fijación de nitrógeno
- Plantear una ruta para la sobreexpresión de genes de fijación de nitrógeno por medio de la herramienta CRISPR

Los sistemas de monocultivo y la aplicación constante de agentes químicos como fertilizantes y pesticidas impactan la salud del suelo y disminuyen su diversidad microbológica (Tripathi *et al.*, 2020). Los microorganismos del suelo se encargan de llevar a cabo varios procesos metabólicos que aumentan la biodisponibilidad de muchos nutrientes y mejoran la capacidad de asimilación de estos por las plantas (Fall *et al.*, 2013); por lo tanto, al verse afectados por compuestos químicos, el rendimiento de cultivo se ve afectado (Xu *et al.*, 2020). A pesar del efecto negativo de los compuestos químicos a la microbiota, no es posible solamente dejar de utilizarlos, ya que permiten mantener una producción alta y así abastecer la demanda mundial (Gouda *et al.*, 2018). Además, el abuso de pesticidas puede llevar al desarrollo de resistencia, asimismo, el uso descontrolado de fertilizantes puede llegar a tener repercusiones en la salud humana (Bijay-Singh *et al.*, 2021). Por lo anterior, surge una necesidad de desarrollar alternativas que fomenten el crecimiento de las plantas y las protejan de organismos patógenos, al mismo tiempo que no afecten de forma negativa a la salud humana y a la salud del suelo de cultivo a largo plazo (Bijay-Singh *et al.*, 2021).

Las bacterias de suelo cumplen un rol sumamente importante en los ecosistemas, estas actúan directamente en varios procesos biogeoquímicos que modulan la disponibilidad de varios nutrientes necesarios para todos los organismos (Fierer, 2017). Por ejemplo, el nitrógeno, uno de los compuestos más importantes para todos los seres vivos, se encuentra en su mayoría como nitrógeno gaseoso (N_2), por lo tanto, existen bacterias fijadoras de nitrógeno que convierten el nitrógeno gaseoso en compuestos como amoníaco, el cuál puede ser asimilado por las plantas (mantelin , 2004). De igual manera, las bacterias están implicadas en la biodisponibilidad de otros compuestos como el fósforo (P) y potasio (K), y pueden formar relaciones mutualistas con las plantas (Phares *et al.*, 2022). Particularmente existe un grupo de bacterias que se caracteriza por fomentar el crecimiento de plantas conocidas como (PGPB, por sus siglas en inglés), estas han sido estudiadas como biofertilizantes e incluso como biocontroladores (Sokol *et al.*, 2022; Fierer, 2017).

A pesar de que se han realizado estudios para utilizar las PGPB como biofertilizantes o bioformulaciones que fomenten el crecimiento de un cultivo particular (Tahat *et al.*, 2020). Es importante mencionar que también es posible realizar modificaciones genéticas en estas bacterias para potenciar sus capacidades y mejorar el rendimiento de los cultivos. Para esto, primero es necesario entender qué genes se encuentran en las bacterias y como pueden ser modificados, cabe mencionar que se ha demostrado que las PGPB tienen mejores capacidades si son aisladas del suelo donde quieren ser utilizadas (Williams, Liede, Fahy, & al., 2020; Yadav *et al.*, 2020). El presente estudio pretende realizar una identificación de bacterias de suelo, para luego llevar a cabo un proceso

de secuenciación de nueva generación (WGS) con el propósito de estudiar los genes involucrados en rutas de interés como fijación de nitrógeno en las bacterias que presenten mayor potencial, con el objetivo de desarrollar un sistema mutagénesis dirigida a través de CRISPR/Cas9.

La necesidad de desarrollar compuestos biológicos para el mantenimiento de los productos agrícolas en Guatemala llevó al desarrollo de un proyecto de investigación conformado por el Centro de Estudios en Biotecnología (CEB) de la Universidad del Valle de Guatemala (UVG) y el Ingenio Magdalena. El propósito de la investigación es realizar productos biológicos para el desarrollo de un sistema de mantenimiento sostenible de los cultivos de caña de azúcar, así como llevar a cabo formulaciones que potencialmente puedan ser utilizados en otros cultivos. Adicionalmente, el Ingenio Magdalena busca ser pionero en la investigación y diseño de productos para la sostenibilidad agrícola en Guatemala, para proveer en el mercado alternativas a los compuestos químicos convencionales.

El Centro de Estudios en Biotecnología tiene experiencia en el desarrollo de investigación para el desarrollo de métodos y productos industriales. Por lo tanto, en el año 2021, se propuso un proyecto de investigación entre el Ingenio Magdalena y el Centro de estudios en biotecnología titulado “Mejoramiento de la salud del suelo de monocultivo a través de la optimización de procedimiento de compostaje y manejo de microbiología de suelos”, para estudiar la dinámica de la microbiota del suelo, aislar y caracterizar bacterias y hongos para evaluar sus capacidades promotoras del crecimiento con el propósito de desarrollar una formulación que actúe como biofertilizante, biopesticida y entre otras aplicaciones potenciales. El presente trabajo es una continuación de la tesis titulada “Análisis del microbioma bacteriano de suelos utilizados para el cultivo de caña de azúcar, en Guatemala mediante metagenómica del gen 16S ARN ribosomal y caracterización de las bacterias promotoras del crecimiento que se encuentran en ellos” llevada a cabo por María Valeria Ochoa Elías (Ochoa, 2022).

5.1 Caña de azúcar

La caña de azúcar, *Saccharum officinarum*, es una planta de cultivo que es utilizada para la producción principal de azúcar a nivel industrial, además, puede utilizarse también para la producción de combustible renovable a base de etanol (James & Tate, 2004). Este cultivo es particularmente importante en Guatemala debido a que representa alrededor de 6% de las exportaciones anuales totales (Villagran, 2020).

La alta demanda de azúcar ha llevado al desarrollo de sistemas de monocultivo que afecta el desarrollo de la caña (Martíni et al., 2020); esto repercute en degradación química, que implica una disminución de disponibilidad de nutrientes como el carbono, nitrógeno, fósforo, potasio y entre otros nutrientes necesarios para el crecimiento de plantas (Tetteh, 2015; Umrit et al., 2014). Asimismo, se fomenta la degradación física, que son alteraciones tales como porosidad y organización estructural de la tierra, lo que limita el acceso y paso de nutrientes y agua (Canisares et al., 2019; Barros et al., 2004). Por último, la sobreexplotación de la tierra también promueve la degradación de calidad biológica; esto sucede porque las bacterias comensales y simbióticas del suelo sufren un cambio en la dinámica poblacional y muchas especies no pueden proliferar (Martíni et al., 2020; Lehman et al., 2015).

Otros factores que pueden afectar el rendimiento de la caña de azúcar, es la infección por diversos patógenos. Particularmente, se ha reportado varias especies de hongos como *Fusarium fujikuroi* y *Fusarium moniliformae* que han sido reportados como agentes causantes de Pokkah boeng, una enfermedad fúngica característica de caña de azúcar prevalente en china, la cual causa una pérdida de coloración de la planta y necrosis del tejido (Vishwakarma et al., 2013; Yao et al., 2020). Asimismo, se ha reportado que *Fusarium sacchari* también es un patógeno potencial para la caña de azúcar, este se caracteriza por fomentar marchitez en el tallo y induciendo daños que pueden llevar a necrosis en los cultivos (Viswanathan et al., 2017). A pesar de que la mayoría de las enfermedades prevalentes de caña de azúcar son causadas por hongos; esta también puede ser infectada por patógenos virales como el virus del mosaico de la caña de azúcar (SCMV) y el virus de la hoja amarilla de la caña de azúcar (SCYLV); los cuales pueden causar una destrucción devastadora en el cultivo (Strachan et al., 2022; Lu et al., 2021; Holkar et al., 2020).

5.2 Fertilizantes

5.2.1 Fertilizantes químicos

Los fertilizantes son compuestos químicos ampliamente utilizados en los campos para mejorar el rendimiento de los cultivos; estos se caracterizan por proveer los nutrientes necesarios para apoyar el crecimiento de las plantas (Savci, 2012). Los nutrientes principales que presentan los fertilizantes químicos son el nitrógeno (N), el fósforo (P), sulfuro (S) y el potasio (K) (Marshner, 2012). Existen siete fertilizantes secos producidos ampliamente; la urea 46-0-0, nitrato de amonio, sulfato de amonio, fosfato diamónico 1846-0 (DAP), fosfato monoamónico (MAP), superfosfato triple 0-46-0 (GTSP) y fosfato monoamónico sulfurado (MAP-S), estos pueden estar complementados con Zinc y potasio (Fall *et al.*, 2013). Por el otro lado, existe otra categoría de fertilizante que es relevante en la agricultura, estos son los fertilizantes líquidos, que normalmente están compuestos por una mezcla de nitrato de amonio y urea, estos pueden estar complementados con amonio anhídrido y ácido superfosfórico (Fall *et al.*, 2013; IFIA, 2000).

La alta demanda mundial de productos alimenticios ha desarrollado una necesidad latente en el uso de compuestos que permitan mantener el cultivo sano para cumplir con el rendimiento, adicionalmente, cabe mencionar que se ha estimado que la producción agrícola debe de aumentar aproximadamente un 70 % para cumplir con las necesidades globales en las próximas décadas (FAO, 2011). Por lo tanto, para mantener el rendimiento requerido, la agroindustria ha desarrollado sistemas de monocultivo, que utilizan constantemente fertilizantes (Bindraban *et al.*, 2012). Como se mencionó previamente, el monocultivo tiene un impacto negativo a largo plazo en los cultivos, esto combinado con la aplicación persistente de los fertilizantes tiene un efecto devastador en la salud del suelo (Tripathi *et al.*, 2020; Xu *et al.*, 2020). A pesar que los fertilizantes proveen de nutrientes necesarios para el crecimiento de las plantas, estos pueden tener un impacto negativo si no se utilizan con moderación (Xu *et al.*, 2020). Su aplicación constante fomenta un desbalance de nutrientes, lo que puede afectar directamente a la planta, asimismo, debido a que algunos compuestos de los fertilizantes presentan un pH bajo, pueden llevar a la acidificación del suelo, afectando la disponibilidad de los nutrientes (Goulding, 2016). La acidificación también afecta a la biodiversidad del microbioma del suelo; al afectar la dinámica ecológica de bacterias, se disminuye la tasa de nutrientes disponibles, ya que muchas de estas bacterias se encargan de aumentar la biodisponibilidad de nitrógeno, fósforo, potasio, entre otros nutrientes (Tahad *et al.*, 2020; Dubey *et al.*, 2019).

Los fertilizantes también han sido criticados debido a que pueden causar contaminación en los cuerpos de agua. En muchos casos, un mal manejo del sistema de aplicación de fertilizantes ha llevado a la contaminación de ríos, afectando fuertemente los ecosistemas acuáticos (Bijay-Singh *et al.*, 2021). el aumentar los nutrientes como el nitrógeno (N), Fósforo (P) y potasio (K) procedente de los fertilizantes desencadena la eutrofización, lo que no solamente afecta la dinámica ecológica, sino que puede llevar a la disminución de ciertos nutrientes como lo es el oxígeno o incluso la luz solar (Srivastav, 2020). Por último, cabe mencionar que, debido a la regulación de producción y aplicación de estos compuestos químicos, los precios de mercado tanto del producto químico como el producto agrícola han comenzado a aumentar en sobremanera, lo que enfatiza la importancia de desarrollar productos alternativos que no presenten los mismos problemas potenciales y sean más costo-efectivos a largo plazo.

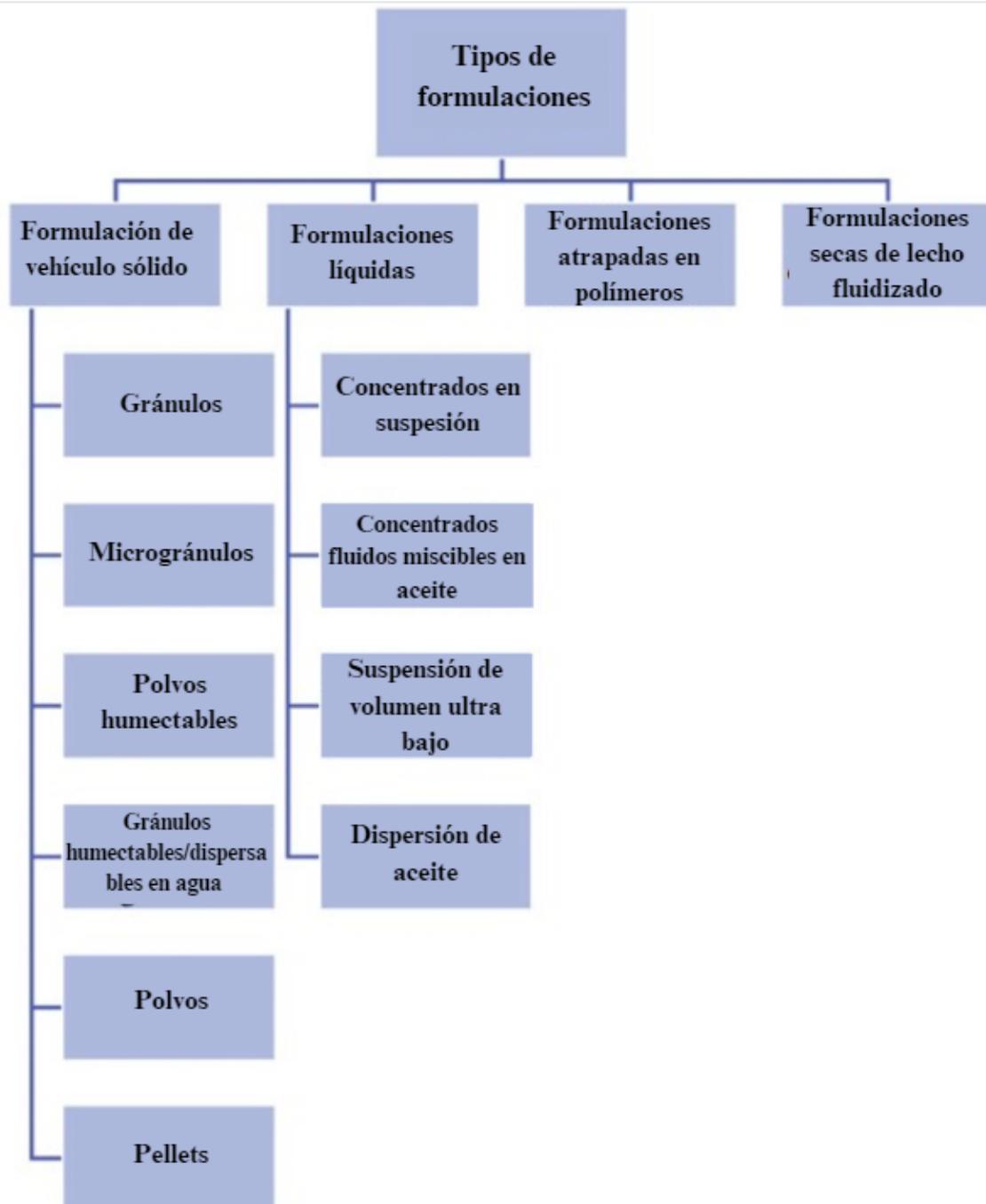
5.2.2 Biofertilizantes

La agricultura tradicional cada vez es menos sostenible debido a que la demanda alimentaria aumenta cada año y los métodos tradicionales que emplean monocultivos, fertilizantes y pesticidas químicos disminuyen el rendimiento de los cultivos a largo plazo (Ehrlich y Harte, 2015). Por lo tanto, el desarrollo de nuevas formulaciones que permitan mantener el rendimiento de cultivo, pero sin afectar la salud del suelo y el medio ambiente es de extrema urgencia (Maçik *et al.*, 2020).

La definición de biofertilizantes ha cambiado rápidamente a lo largo de los años, no obstante, se puede definir a grandes rasgos como un microorganismo o un conjunto de microorganismos que presenta propiedades promotoras de crecimiento vegetal (Schenk *et al.*, 2009). Esto implica que aumenta la biodisponibilidad de los nutrientes necesarios para que una planta pueda crecer y desarrollarse, además, este conjunto de microorganismos debe presentarse en una formulación lista para ser utilizada por los agricultores. Asimismo, los biofertilizantes también pueden tener compuestos que contribuyan a las actividades benéficas de los microorganismos del suelo. Estas formulaciones aprovechan las capacidades biológicas de los microorganismos a favor de los sistemas de cultivo (Schütz *et al.*, 2018; Maçik *et al.*, 2020). A pesar de que los biofertilizantes representan una solución prometedora, existen algunas desventajas que aún deben ser dirigidas como un tiempo limitado de almacenamiento, una sensibilidad mayor a la temperatura y complicaciones asociadas a su transporte; adicionalmente, debido a que está compuesto mayormente por microorganismos, es necesario aplicar una mayor concentración del producto para que las plantas puedan tener los nutrientes necesarios (Sahu *et al.*, 2016).

La efectividad de los biofertilizantes depende de varios factores, principalmente la composición del suelo donde se aplicará, por lo que es necesario desarrollar formulaciones que contengan el ingrediente activo en un transportador apropiado y aditivos que permitan proveer estabilidad y protección a los microorganismos (Maçik *et al.*, 2020; Satish *et al.*, 2022). Estas bioformulaciones deben de asistir en la introducción del microorganismo al sitio diana y ayudar en el mejoramiento de sus capacidades metabólicas para alcanzar una tasa máxima de nutrientes al ser inoculada con las plantas (Mahanty *et al.*, 2016; Maçik *et al.*, 2020). Se han desarrollado formas de transportar la formulación al suelo, en la *figura 1* se muestra un diagrama de los cuatro tipos principales y de que están compuestos.

Figura 1
Tipos de formulaciones biológicas



Nota. Adaptado de Mączek et al., (2020)

5.3 Salud del suelo

5.3.1 Composición y microorganismos del suelo

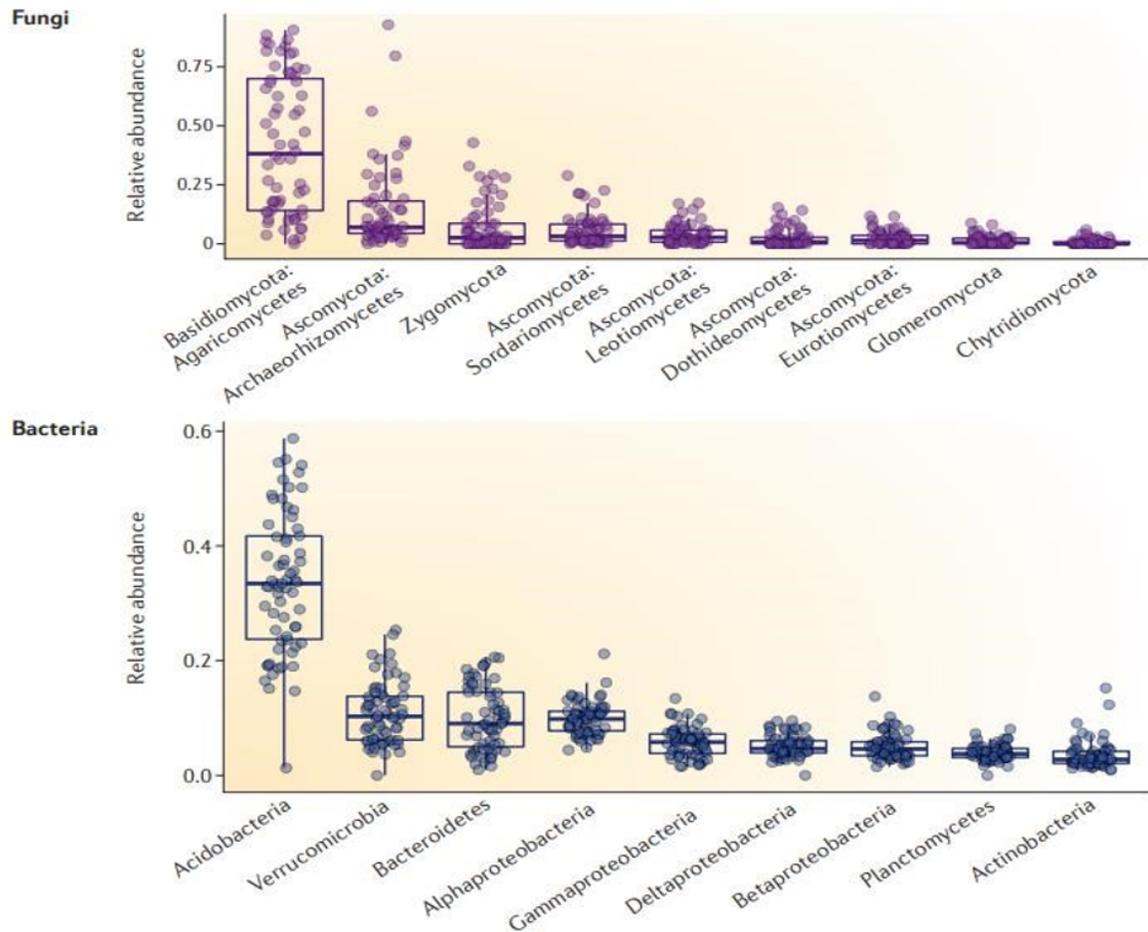
Los suelos están compuestos por una gran variedad de factores, como minerales, materia orgánica y una gran variedad de macroorganismos y microorganismos; esta última característica es de particular interés, ya que representa una gran diversidad genética y metabólica (Leff *et al.*, 2015); de hecho, se ha determinado que un solo gramo de tierra puede contener miles de representantes de taxones microbianos, incluyendo bacterias, hongos y virus (Fierer, 2017). El conjunto de todos los microorganismos que constituyen una comunidad junto a sus características genéticas que en el suelo se les conoce como microbioma de suelo. Los microorganismos del suelo han sido ampliamente estudiados por su capacidad de moldear los ciclos biogeoquímicos de elementos importantes para los ecosistemas (Sokol *et al.*, 2022); sus características metabólicas influyen en sobremedida la interacción de todos los organismos del suelo, como bacterias, hongos, plantas, entre otros (Islam *et al.*, 2020).

Para estudiar con precisión el suelo, es importante entender que este no es solamente un ambiente, más bien está compuesto por varios ambientes, de los cuales todos y cada uno presenta comunidades microbianas distintas que presentan una capacidad de influenciarse entre ellos (Fierer, 2017). El microbioma es altamente complejo, por lo tanto, se ha demostrado que las condiciones que componen un ambiente de suelo son altamente variables, por lo que la composición nativa de un suelo particular de una región no será similar a otra (Delgado-Baquerizo *et al.*, 2019; Sokol *et al.*, 2022). Los factores que se estudian comúnmente en el suelo son la concentración y biodisponibilidad de carbono orgánico, nitrógeno, fósforo, entre otros nutrientes; el pH y la salinidad del suelo; y las características físicas que lo componen como la porosidad (Fierer, 2017). Cada uno de estos factores se ve afectado por muchas variables como lo es el clima, los macroorganismos y microorganismos, factores de estrés abióticos como una escasez de agua y la actividad humana (Maçik *et al.*, 2020).

Existe una gran diversidad de microorganismos en el suelo, estos cumplen un rol dentro de la compleja dinámica ecológica (Compant *et al.*, 2019). Por lo tanto, la caracterización de microorganismos de suelo ha sido un foco de investigación para el desarrollo de nuevos métodos de agricultura sostenible. Como se mencionó previamente, no existe una microbiota típica que pueda ser aplicada a todos los tipos de suelo que se encuentran en regiones similares, de hecho, se ha determinado que la composición de la microbiota puede variar incluso en muestras de suelo que han sido tomadas de regiones geográficamente cercanas, esta variación se atribuye a las características fisicoquímicas de los compuestos presentes en el suelo y a los mismos microorganismos que han podido proliferar (Wei *et al.*, 2019; Van Nuland *et al.*, 2021). Se ha caracterizado una gran variedad de bacterias, hongos, arqueas y protistas en el suelo. Los grupos más frecuentes de hongos son representantes de Basidiomycota, particularmente Agaricomycetes; también se ha encontrado Archeorhizomycetes, zigomycota, Sordariomycetes y Leotiomycetes. En el caso de bacterias, existe una abundancia mayor de Acidobacteria, Bacteroidetes, Alpha-proteobacteria y Gammaproteobacteria (*figura 2*) (Fierer *et al.*, 2017).

Figura 2

Diversidad de abundancia de bacterias y hongos en el suelo

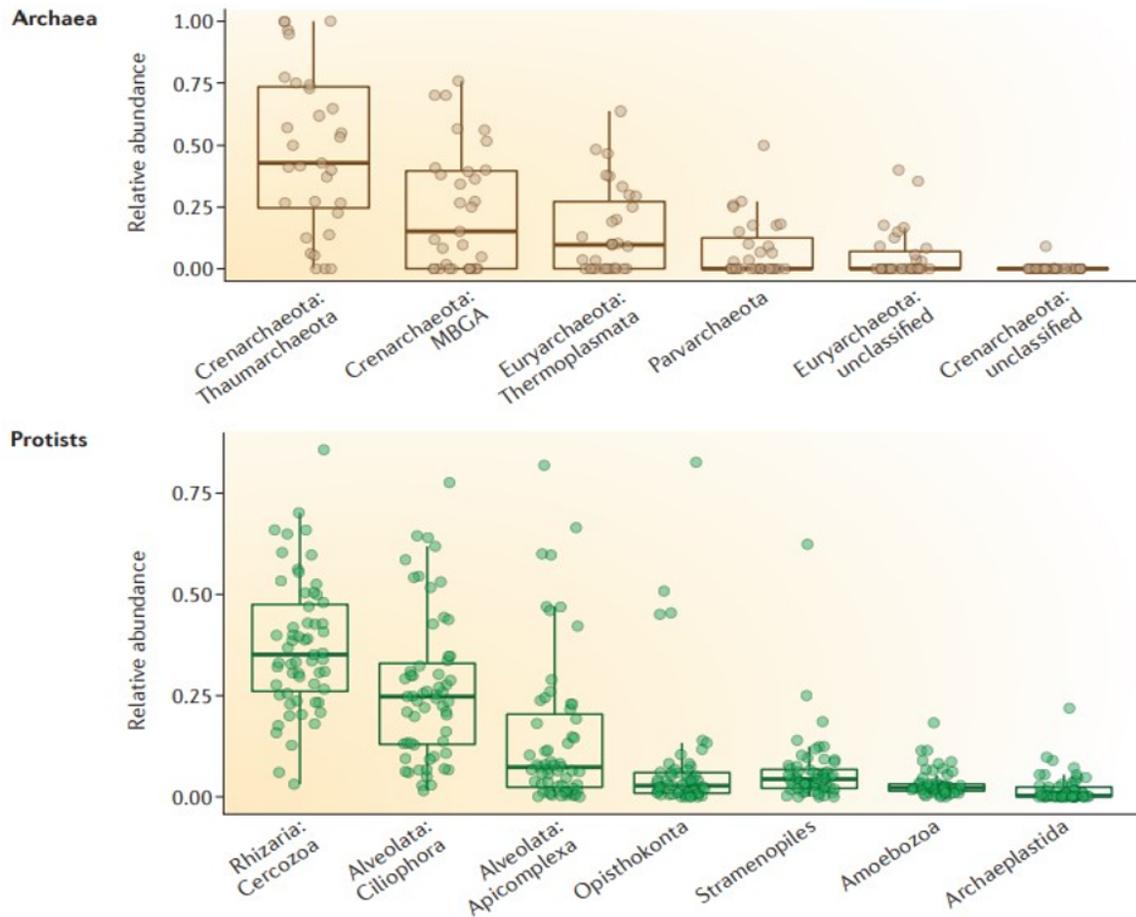


Nota. Adaptado de Fierer et al., (2017)

La presencia y abundancia relativa de arqueas y protistas suele variar mucho más, sin embargo, se ha caracterizado dentro de las arqueas (*figura 3*), la clase *Thaymarchaeota* posee una alta abundancia relativa, así como también *Rhizaria* dentro de los protistas (Fierer *et al.*, 2017). Es importante mencionar que el cambio en la variedad y abundancia de estos microorganismos también se ha asociado a la presencia de diferentes plantas, es decir, existen ciertos taxones que presentan algún tipo de relación simbiótica con una especie en particular, asimismo, se ha determinado también que puede existir codependencia (Dastogeer *et al.*, 2020; Sánchez-Cañizares *et al.*, 2017).

Figura 3

Diversidad y abundancia de protistas y arqueas en el suelo



Nota. Adaptado de Fierer et al., (2017)

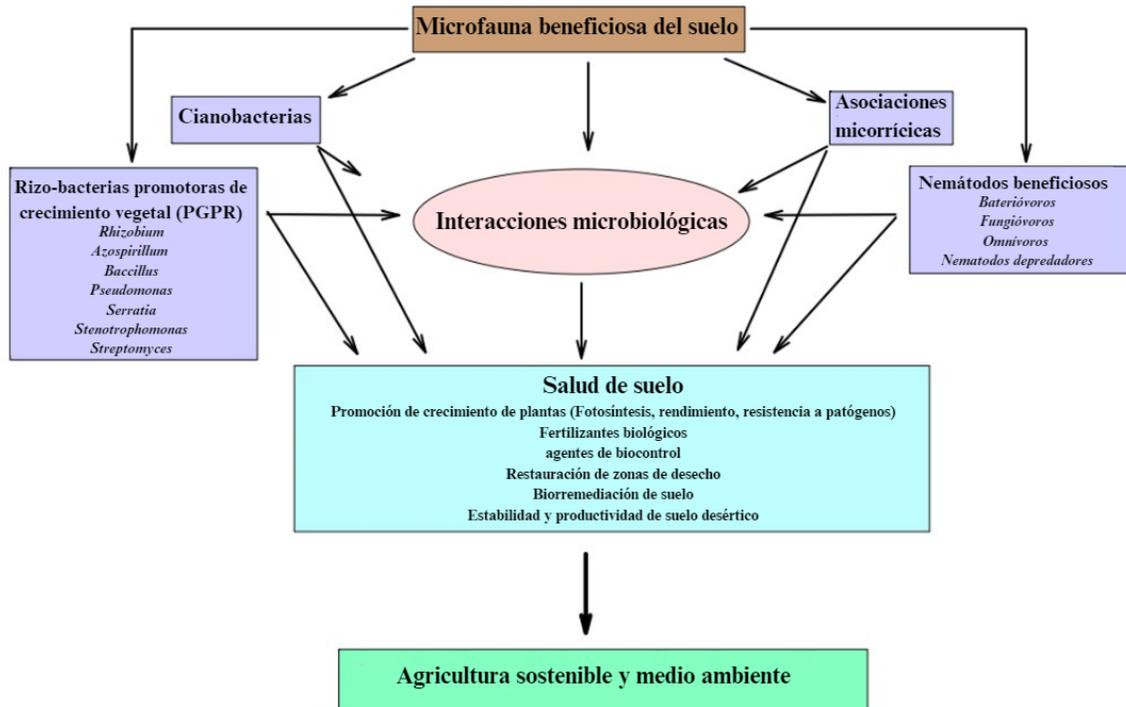
A pesar de esfuerzos extensivos en caracterizar los microorganismos que componen la microbiota del suelo, y las condiciones abióticas que influyen en las proporciones de las poblaciones, aún sigue siendo complicado dilucidar completamente toda la diversidad de bacterias y arqueas en el suelo (Fierer *et al.*, 2017). Se han desarrollado métodos para permitir la realización de una mejor estimación. Entre estos métodos podemos mencionar la detección de *relic DNA*, material genético ancestral de organismos que se ha preservado en el suelo y fósiles (Carini *et al.*, 2016). Asimismo, existen métodos de secuenciación metagenómica *shotgun*, que permite analizar regiones o genes específicos de muestra con varios organismos (Hasan *et al.*, 2014).

5.3.2 Bacterias promotoras del crecimiento vegetal

Las bacterias de suelo tienen la capacidad de realizar varios procesos metabólicos que permiten aumentar la biodisponibilidad de nutrientes del suelo, asimismo, pueden fomentar que exista una interacción con otros microorganismos para que las plantas puedan crecer de mejor manera (Gamalero y Glick *et al.*, 2011). Adicionalmente algunas bacterias tienen capacidades potenciales fungicidas y pesticidas, por lo que son de gran interés para la investigación biotecnológica para la aplicación agroindustrial (Gutiérrez-Santa *et al.*, 2020).

Figura 4

Interacciones en el microbioma impacta la salud del suelo

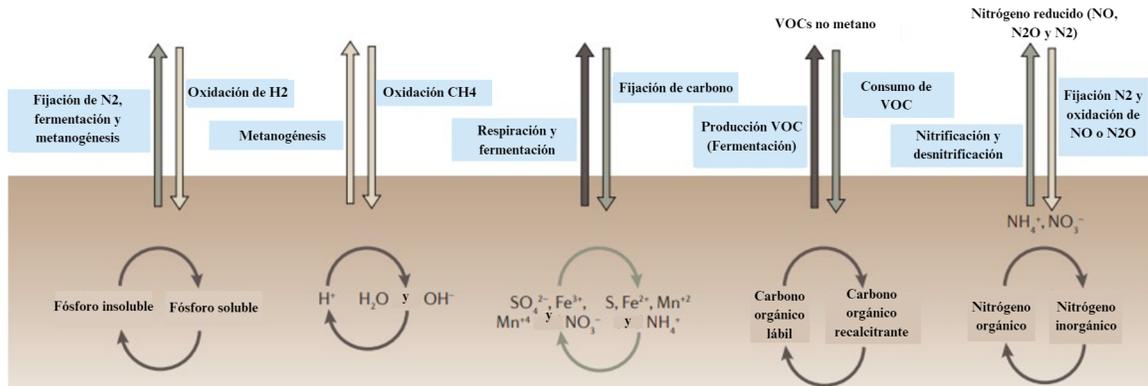


Nota. Adaptado de Tahat et al., (2020)

Las bacterias que poseen la capacidad de fomentar el crecimiento vegetal se denominan bacterias promotoras del crecimiento vegetal (PGPB, por sus siglas en inglés). Estos microorganismos son capaces de formar interacciones con las plantas del suelo y proveer los nutrientes, hormonas y enzimas; algunas de estas bacterias son capaces de colonizar la rizosfera, región del suelo donde se encuentran las raíces de la planta, estas bacterias pueden aumentar la biodisponibilidad de nutrientes y protegerla de patógenos potenciales; a estas bacterias se le atribuye el nombre de rizobacterias promotoras del crecimiento vegetal (PGPR, por sus siglas en inglés) (Ahemad Kibret, 2014). Como se observa en la *figura 4*, la presencia de microorganismos del suelo tiene una relación directa con la salud del suelo (Tahat *et al.*, 2020). La compleja red de interacciones de estos microorganismos es capaz de afectar casi todos los aspectos del suelo, y por ende, repercute en el crecimiento de la planta (Ramakrishna *et al.*, 2019).

Figura 5

Los procesos biogeoquímicos son afectados por el microbioma



Nota. Adaptado de Fierer et al., 2017

Características promotoras

Las PGPB median varios procesos biogeoquímicos, fijan nitrógeno, realiza fermentación y metanogénesis; además, aumentan la solubilidad de ciertos compuestos como el fósforo y el potasio. Además, estas bacterias aumentan la concentración de carbono orgánico a través de fijación, cabe mencionar que, debido a su diverso metabolismo, también puede proveer a la planta de varios compuestos como manganeso, hierro y sulfatos (figura 5) (Fierer et al., 2017; Ramakrishna et al., 2019).

1. Metabolismo de nitrógeno: El nitrógeno es uno de los elementos más abundantes en el planeta, sin embargo, se encuentra mayormente en estado gaseoso (N_2). El ciclo biogeoquímico del nitrógeno es el proceso en donde este elemento pasa por diferentes formas para ser procesado, se divide en dos fases importantes, la fase abiótica y la biótica. El ciclo comienza por la fijación del nitrógeno gaseoso en amoníaco, luego sigue la nitrificación, que sucede al oxidar el amoníaco para la formación de nitratos; estos son asimilados por los organismos y se procesan, por último, sucede la desnitrificación, la cual reduce los nitratos para la producción de nitrógeno gaseoso (Stüeken et al., 2016). Es importante mencionar esto porque la fijación de nitrógeno, nitrificación y desnitrificación son llevadas a cabo por bacterias del suelo (Jetten, 2008). El nitrógeno es importante para el desarrollo de todos los organismos, necesario para la síntesis de proteínas, el metabolismo de los aminoácidos y la formación de ácidos nucleicos como el ATP (Solomon et al., 2015). Es especialmente importante para las plantas debido a que se utiliza para la síntesis de clorofila y moléculas precursoras, fundamental para la fotosíntesis (Luo et al., 2020).

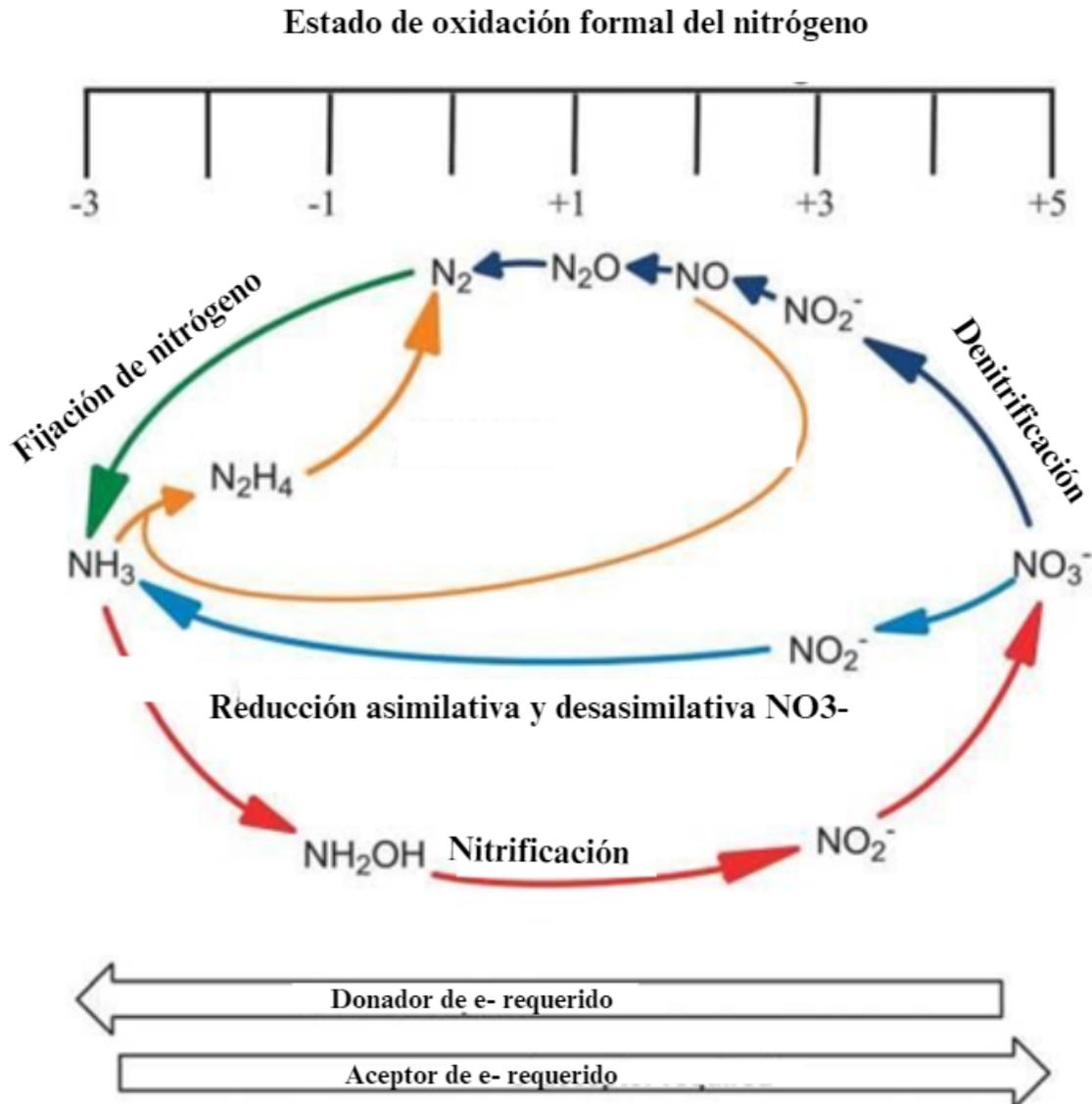
Una de las habilidades de las PGPB es el aumento de la biodisponibilidad del nitrógeno, ya que pueden realizar los procesos metabólicos implicados en el metabolismo del nitrógeno, permitiendo proveer amoníaco y nitratos para que las plantas del suelo puedan asimilarlos y crecer (Souza et al., 2014). Las bacterias que llevan a cabo la fijación de nitrógeno lo realizan a través de un sistema llamado complejo nitrogenasa que está compuesto principalmente por dinitrogenasa y dinitrogenasa reductasa, se ha determinado que los genes necesarios para catalizar estas reacciones se encuentran dentro de un cluster de genes denominados *nifHDK* (Minerdi et al., 2001; Zehr y Turner, 2001). Algunas nitrogenasas alternas se han encontrado como isoformas, particularmente dos, codificadas del gen *nifH*. Una tercera proteína se ha determinado como contraparte de la proteína Mo necesaria para el funcionamiento del complejo nitrogenasa codificada por el gen *nifG* (Zehr y Turner, 2001). Algunas bacterias que han sido identificadas como PGPR con la capacidad de fijar nitrógeno son *Allorhizobium* sp., *Rhizobium* sp., y *Azotobacter* sp. Es importante mencionar que varios estudios

han determinado que genes se encuentran involucrados en cada uno de los procesos. Además, se ha estudiado ampliamente el operón *nif HDK* lo que permite llevar a cabo métodos de edición génica para poder desarrollar cepas que tengan capacidades aumentadas en esta ruta (Sparacino-watkins *et al.*, 2015).

Existen tres complejos nitrogenasas ampliamente descritas las cuales varían dependiendo del conjunto de genes que la conforman (Zhao *et al.*, 2006; Chanderban *et al.*, 2023). El complejo nitrogenasa dependiente de molibdeno está comprendido por los genes *nif*, la nitrogenasa dependiente de vanadio esta codificada por el cluster *vnf* y la nitrogenasa dependiente de hierro por los genes *anf* (Chanderban *et al.*, 2023). Son obstante, el complejo codificado por los genes *nif* es generalmente el más estudiado. La estructura de los genes se divide principalmente por las secuencias codificantes para las subunidades biocatalíticas, los genes *nifH*, *nifD* y *nifK* (Rubio, 2008; Cyratti y Rubio, 2008). Asimismo, también es importante realizar la distinción de los genes implicados en la biosíntesis del cofactor de hierro-molibdeno, mediado por *nifY*, *nife*, *nifN*, *nifX*, *nifU*, *nifS*, *nifV*, *nifB* y *nifQ* (Rubio, 2008). El complejo, además, necesita de un flujo de electrones, el cual está mediado generalmente por los genes *nifJ* y *nifF*. Por último, todo el operón es regulado por dos genes, los cuales se expresan al alza dependiendo de la concentración de oxígeno o amoníaco; el gen *nifA* codifica para una proteína activadora, el cual fomenta la transcripción de los genes *nif* mientras que *nifL* codifica para una proteína represora que inhibe la expresión de *nifA*, disminuyendo así la tasa de fijación de nitrógeno (Zhang *et al.*, 2023; Martinez-Argudo *et al.*, 2004).

A pesar de que la fijación de nitrógeno es sumamente importante para que las plantas asimilen el elemento, también es necesario tomar en cuenta que algunas especies utilizan los intermediarios producidos por la amonificación y nitrificación (Ferreti *et al.*, 2021). En la *figura 6* se muestra los intermediarios del nitrógeno de la desnitrificación y nitrificación, estos también pueden ser utilizados por las plantas.

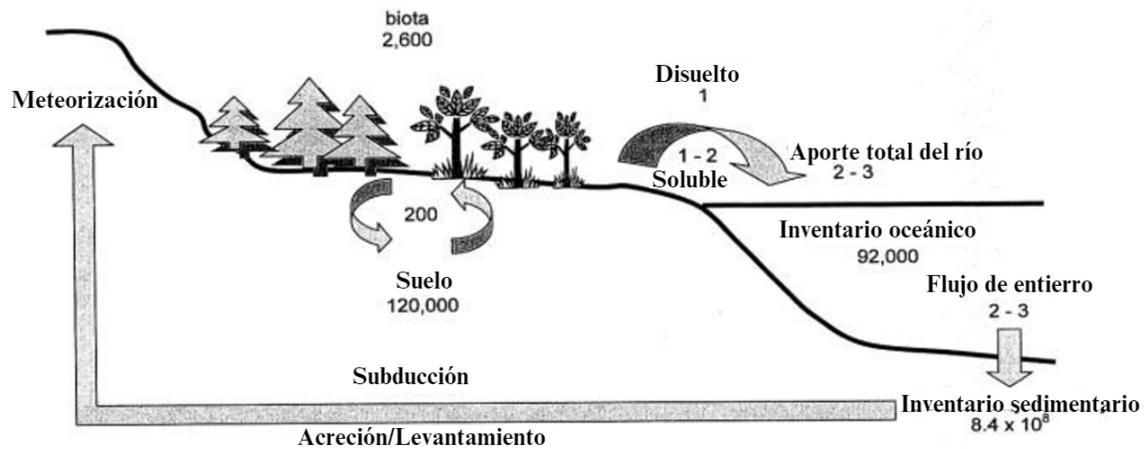
Figura 6
Ciclo del nitrógeno bacteriano



Nota. Adaptado de Sparacino-watkins et al., 2015

2. Biodisponibilidad de fósforo: El fósforo es un elemento importante para el desarrollo de todos los organismos, es un elemento precursor de ácidos nucleicos, algunas enzimas, puede funcionar como cofactor y es esencial para la formación de fosfolípidos. Es un elemento sumamente importante para la formación de ATP, por lo que es necesario para llevar a cabo todos los procesos metabólicos dentro de la célula (Solomon *et al.*, 2015). En plantas, tiene un rol en la fotosíntesis para la formación de nicotinamida adenina dinucleótido fosfato (NADP⁺), una molécula necesaria para la fase lumínica. Al igual que el nitrógeno, este elemento pasa por un ciclo biogeoquímico, sin embargo, este no es tan dinámico. El fósforo se encuentra principalmente en superficies de rocas sedimentadas, este se filtra hacia el suelo, comúnmente como ion fosfato (McClain y Sharkey, 2019). Debido a que la filtración no es un proceso rápido, el fósforo suele ser uno de los nutrientes más escasos en la mayoría de los suelos (*figura 57*), además, puede presentarse en forma de fósforo inorgánico, que no es tan fácilmente asimilado por las plantas (Hinsiger, 2001).

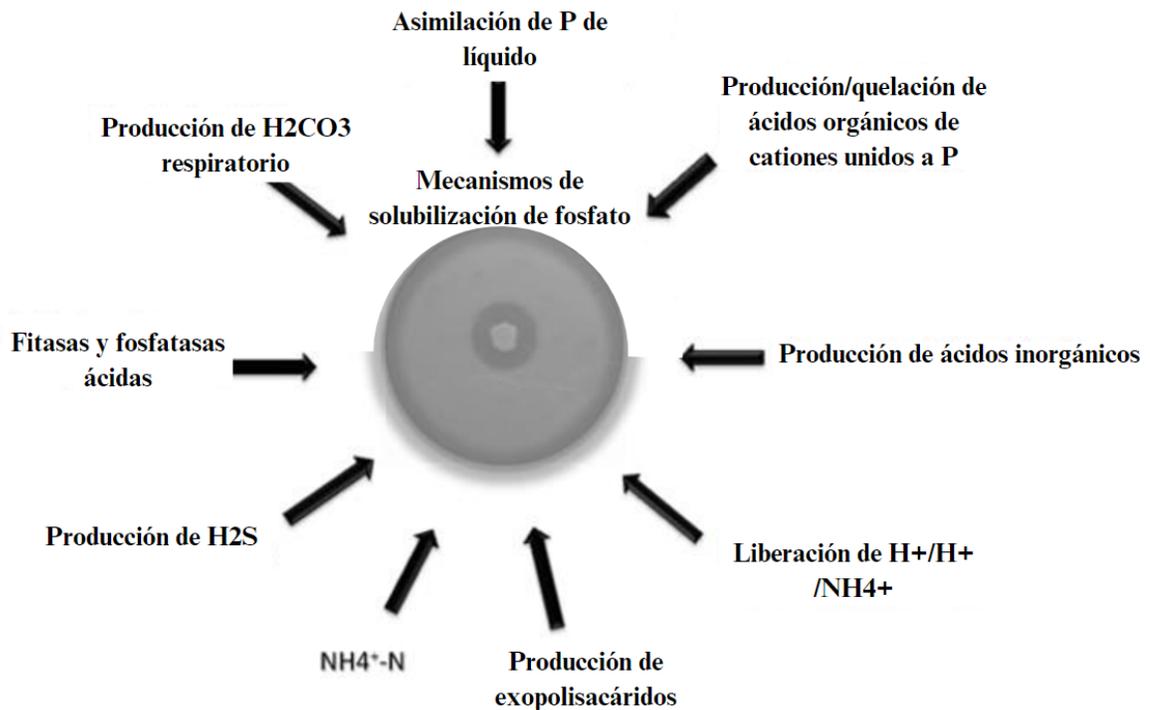
Figura 7
Ciclo del fósforo



Nota. Adaptado de Filippelli, 2002

Las bacterias también son capaces de aumentar la biodisponibilidad del fósforo, se ha determinado que algunas PGPR pueden facilitar la obtención de este nutriente para la planta a través de la solubilización del fosfato (Zaidi *et al.*, 2009). La solubilización de fosfatos se puede llevar a cabo por medio de dos mecanismos principales: por medio de enzimas y por medio de ácidos orgánicos (figura 8). Los microorganismos que solubilizan fosfato a través de la producción de enzimas como las fosfatasas, que permiten degradar compuestos que presentan fósforo en su estructura; y aquellos microorganismos que secretan ácidos orgánicos para disolver el fósforo inorgánico (Billah *et al.*, 2019; Ahemad, 2014). Algunas PGPB que han sido reportadas como solubilizadoras de fosfato son: *Agrobacterium* sp., *Burkholderia* sp., y *Rhizobium* sp.

Figura 8
Mecanismos de solubilización



Nota. Adaptado de Zaidi et al., 2009

3. Producción de fitohormonas: Las fitohormonas son compuestos que son esenciales en el desarrollo y crecimiento vegetal, pueden afectar la respuesta a estímulos de las plantas e influir en sus tropismos, influyen directa o indirectamente en el desarrollo de la raíz y pueden fomentar el desarrollo de las estructuras reproductivas de la planta (Fahad *et al.*, 2014). Algunos estudios sugieren que las fitohormonas también pueden mediar en los mecanismos de defensas de las plantas (Checker *et al.*, 2018). Algunas fitohormonas de interés son: Las citoquininas, el etileno y las auxinas. Una de las auxinas más conocidas es el ácido indol-3-acético (IAA, por sus siglas en inglés); esta fitohormona fomenta la división celular en las raíces y estimula el desarrollo del fruto, asimismo, posee un rol importante en la mediación en los mecanismos de comunicación con la microbiota del suelo (Duca *et al.*, 2020).

Figura 9*Efecto de tratamientos con IAA**Nota. Adaptado de Myo et al., 2019*

La producción de IAA se ha determinado en algunas PGPB, al ser agregadas a las plantas son capaces de aumentar sobremanera el crecimiento de las plantas y mejorar las interacciones entre el cultivo y los microorganismos de la rizosfera (Ferier *et al.*, 2017). Como se puede observar en la *figura 9*, el tratamiento de una planta expuesta a diferentes concentraciones de una formulación de IAA, tiene efectos significativos en el desarrollo de la planta (Myo *et al.*, 2019).

Existen varios genes implicados para la biosíntesis de IAA por bacterias. Primero, es necesario llevar la síntesis del precursor más importante, el cual es el L-triptófano (Naveed *et al.*, 2015). Para esto, muchas rutas parten del crosmato, en donde los genes *trpE/trpA*, *trpD*, *trpE*, *trpC* y *trpB* median la ruta de biosíntesis (Merino *et al.*, 2008; Gollnick *et al.*, 2005). Posteriormente, Los genes *iaaM* e *iaaH* codifican para la triptófano-2-monooxigenasa y la indol-acetamida hidrolasa, las cuales median la conversión de L-triptófano en ácido-indolacético (Duca *et al.*, 2020). algunas bacterias son capaces de sintetizar IAA partiendo de Indol-3piruvato, el cual puede ser producido a partir del L-triptófano a través de la enzima triptófano aminotransferasa (AAT) o bien por la producción de indol a través de la enzima triptofanasa (*tnaA*) para luego sintetizar el compuesto, el cual posteriormente es convertido por la enzima ácido-indol-3-pirúvico descarboxilasa (*ipdC*) (Sergeeva *et al.*, 2007; Jasmin *et al.*, 2014; Duca *et al.*, 2020).

Factores de virulencia

Las bacterias son capaces de producir compuestos o estructuras denominados factores de virulencia los cuales permiten evadir el sistema inmune de los hospederos y causar enfermedades (Leitão, 2020). Las bacterias patogénicas tienen factores de virulencia que se dividen en su capacidad de adherirse, producir exotoxinas, producir exoenzimas, sistemas de entrega, capacidad de modular o evadir el sistema inmune y motilidad (Wu *et al.*, 2008).

Las exotoxinas pueden ser polipéptidos o complejos heteroméricos proteicos que causan daño

o facilitan la interacción de un patógeno con una célula huésped (Sastalla *et al.*, 2016). El tipo de endotoxina depende de cada bacteria, algunos de los genes implicados en la producción por *Pseudomonas*, *Bacillus* y *Acinetobacter* pueden ser *plc*, *tox*, *alo* y *nhe* (Chen *et al.*, 2005). La adherencia es de igual forma un factor de virulencia importante, principalmente porque permite a la bacteria anclarse a un tejido, lo cual focaliza la infección y dificulta en algunos casos su tratamiento (Nowrouzian *et al.*, 2006), estos suelen estar codificados por los genes *pil*, *fim*, *tsa*, *chp*, *fap*, *rop* y *bsl* (Chen *et al.*, 2005). Asimismo, la capacidad de realizar biofilm también puede ser considerado como un factor de virulencia debido a que proporciona una resistencia a la bacteria patogénica (Phillips y Schultz, 2012), los grupos de genes principales en *Pseudomonas*, *Bacillus* y *Acinetobacter* son *ade*, *pga*, *csu*, *alg* y *muc* (Chen *et al.*, 2005).

5.4 Genómica

5.4.1 Secuenciación de nueva generación (NGS)

El primer proceso de secuenciación de ADN se publicó en 1977, llamado Secuenciación de Sanger, donde utiliza el fundamento de secuenciación por síntesis (SBS, por sus siglas en inglés) de la cadena de ADN marcada radiactivamente que es complementaria a la cadena molde utilizando el método de cadena de didesoxinucleótidos, para luego determinar los fragmentos a través de una electroforesis en gel de poliacrilamida (Pervez, Hasnain, Abbas, Moustafa, & Aslam, 2022). Los avances para mejorar esta técnica ayudaron a incorporar nuevos métodos como la reacción en cadena de la polimerasa (PCR), el marcado del ADN con fluorescencia y la disponibilidad de enzimas de buena calidad para modificar el ADN fueron fundamentales para llevar a cabo la secuenciación del primer genoma humano en 2001 (Venter, Adams, Myers, y *et al.*, 2001).

Desde 2005 se han propuesto diferentes métodos de secuenciación de nueva generación (NGS, por sus siglas en inglés) los cuales, proponen novedosos mecanismos para la determinación del orden las bases, la alineación de secuencias y el ensamblaje del genoma (Metzker, 2010). Algunas de las ventajas de las NGS en comparación con las técnicas tradicionales, es la capacidad de secuenciar múltiples genes e identificar millones de variantes o mutaciones simultáneamente. La cantidad de ADN que se necesita para obtener esta información y realizar el análisis es mínima y el tiempo de respuesta es mucho más rápido, además, realizan una mejor comprensión de rutas moleculares potencialmente relacionadas a una enfermedad y posibles opciones de tratamiento, entre otras aplicaciones (Rabbani, Nakaoka, Akhondzadeh, Tekin, y Mahdih, 2016). Estos nuevos procedimientos como NGS utilizan tintes fluorescentes como reemplazo del radiomarcador para el ADN y la electroforesis de matriz capilar en vez de la electroforesis en gel de poliacrilamida basada en gel, lo que ayudó en la automatización de la técnica aumentando la capacidad de secuenciar miles de millones de nucleótidos en una sola ejecución (McCombie, McPherson, y Mardis, 2019).

Actualmente, la NGS sigue optimizando nuevas estrategias para utilizarlas en diferentes áreas como la secuenciación del genoma completo (WGS, por sus siglas en inglés), la secuenciación del exoma completo (WES/ES, por sus siglas en inglés), la llamada de variantes (VC, por sus siglas en inglés), la secuenciación dirigida (TS, por sus siglas en inglés) y la secuenciación del transcriptoma (TCS, por sus siglas en inglés) (Xuan, Yu, Qing, Guo, y Shi, 2013), los cuales pueden ser utilizados para diferentes organismos como bacterias, arqueas, hongos, plantas, animales, etc. también se ha aplicado para en la identificación del material genético de varias enfermedades como el cáncer.

5.4.2 Caracterización genómica

Debido a los avances que ha tenido el estudio del genoma completo, se ha optado por estrategias novedosas que permiten un análisis más eficiente, rápido y de alto rendimiento, dentro de los cuales se puede mencionar los métodos de secuenciación del ADN. Estos se basan en determinar el orden de

las cuatro bases de nucleótidos (Adenina, Timina, Citosina y Guanina) que componen un fragmento de ADN, por lo que, permite identificar toda la información genética en un segmento específico (Heather y JM, 2016).

La genómica se describe como el estudio del genoma completo de un organismo y se obtienen grandes cantidades de datos a través de secuenciación de nueva generación (NGS). Estos datos sirven para comprender las implicaciones de la composición genética, para la identificación y caracterización de genes. Lo que permite estudiar rutas moleculares, determinar mutaciones puntuales y evaluar la regulación génica. (Williams, Liede, Fahy, y al., 2020). El análisis de caracterización genómica nos permite delimitar dicha información, con el objetivo de especificar los datos necesarios para nuestro estudio.

En el caso de procariontes, la caracterización genómica se refiere a la colección completa de datos que describan el perfil genético de un cultivo puro (Trüper y Krämer, 1981), el cual, nos permite investigar la diversidad genética, perfiles de virulencia, genes de resistencia antimicrobiana, clúster de genes de adaptación, cambios filogenéticos, variantes génicas, etc. (Lyu, y otros, 2021). Existen varios estudios han utilizado la caracterización genómica para aplicarla en diferentes áreas, como por ejemplo: La identificación de bacterias causantes de enfermedades transmitidas por los alimentos, como *Salmonella enterica* (Lyu, y otros, 2021). Esta técnica también se ha utilizado en los hallazgos de cepas de interés comercial debido a su potencial probiótico como *Lactobacillus fermentum* (Brandt, Nethery, O’Flaherty, & Barrangou, 2020) y en la comparación de bacterias fijadoras de nitrógeno (Medina-Córdoba, Chande, Rishishwar, y *et al.*, 2021).

5.5 Algoritmos de análisis bioinformático y ensamblaje de genomas

5.5.1 Evaluación de calidad de secuencias

Tras un proceso de secuenciación de nueva generación, los equipos retornan archivos en formato FASTQ.GZ en las dos direcciones de la hélice de ADN (Forward y Reverse), siendo este un formato comprimido de FASTQ. Antes de realizar cualquier análisis o de llevar a cabo el ensamblaje del genoma, es necesario realizar una evaluación del grado de calidad de las secuencias (Cock *et al.*, 2009). Los algoritmos que evalúan la calidad de secuencia realizan un conjunto de estadísticas que permiten evaluar si las secuencias son aptas para llevar a cabo análisis avanzados (Brown *et al.*, 2017). Los parámetros utilizados comúnmente son valores como frecuencia relativa de los nucleótidos y posiciones que presenten baja calidad. La herramienta utilizada comúnmente para llevar a cabo estos análisis es FASTQC (Andrews, 2010; <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) la cual evalúa los siguientes parámetros: calidad de secuencia por pase, contenido de N (siendo N’ una representación de que en esa base particular puede encontrarse una A, G, C o T), valores de calidad por secuencia, contenido de GC por secuencia, distribución de longitud de la secuencia, niveles de duplicación, secuencias sobrerrepresentadas, contenido de adaptadores y contenido de Kmero (valor de una secuencia con “k” nucleótidos consecutivos) (De sena Brandine *et al.*, 2021; Brown *et al.*, 2017).

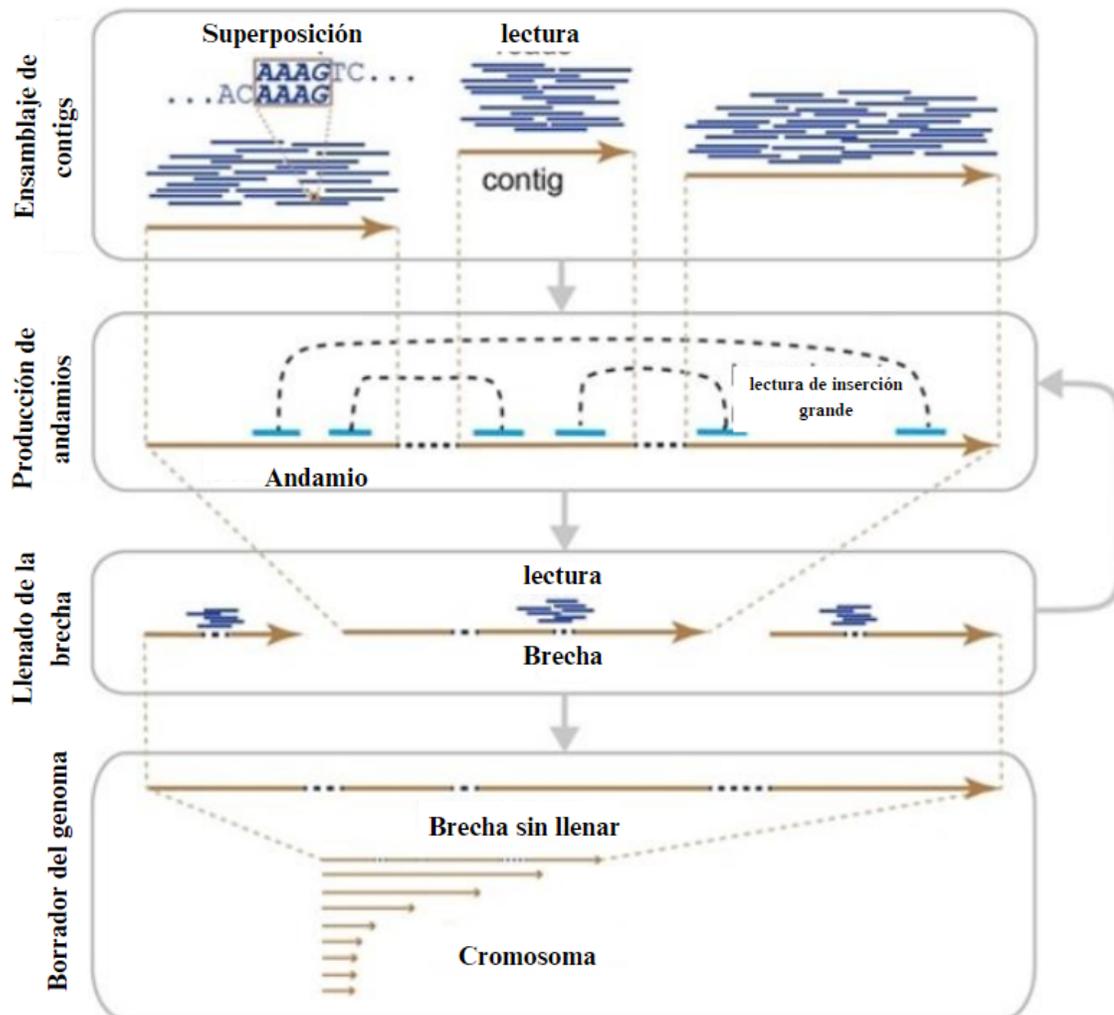
Los análisis de calidad de secuencias crudas normalmente no son capaces de superar todas las pruebas, indicando que deben de ser procesadas para eliminar aquellas secuencias de baja calidad que pueden comprometer el ensamblaje y por ende los análisis, para esto generalmente se lleva a cabo un proceso denominado *trimming* o recorte (Chen *et al.*, 2014). Trimmomatic (Bolger *et al.*, 2014; <http://www.usadellab.org/cms/?page=trimmomatic>) es un programa utilizado para realizar el preprocesamiento de secuencias, este realiza una serie de pasos para cortar y filtrar las secuencias, con el propósito de mejorar su calidad (Bolger *et al.*, 2014; Sewe *et al.*, 2022). Luego de realizar los cortes, es necesario repetir la evaluación de calidad de los resultados del preprocesamiento para determinar si las secuencias cumplen con las estadísticas para llevar a cabo el ensamblaje.

5.5.2 Ensamblaje de genoma

Para estudiar con más profundidad los resultados, es necesario realizar un ensamblaje de un genoma utilizando las lecturas de las secuencias. Un ensamblaje consiste en realizar una reconstrucción de un genoma con base en fragmentos de lecturas de secuenciación que se alinean para encontrar porciones idénticas entre diferentes fragmentos (Chateau y Salson, 2022; Henson *et al.*, 2012). Existen varios que pueden afectar el ensamblaje, como la cobertura, este es el promedio de número de veces que un nucleótido se secuencia en una posición dada (Sims *et al.*, 2014). Durante la secuenciación, se generan secuencias aleatorias, por lo que es posible que existan fragmentos que no puedan secuenciarse. Por lo tanto, se estima que un ensamblaje confiable debe tener una cobertura de 30X-50X para minimizar el riesgo de regiones no secuenciadas (Chateau y Salson, 2022).

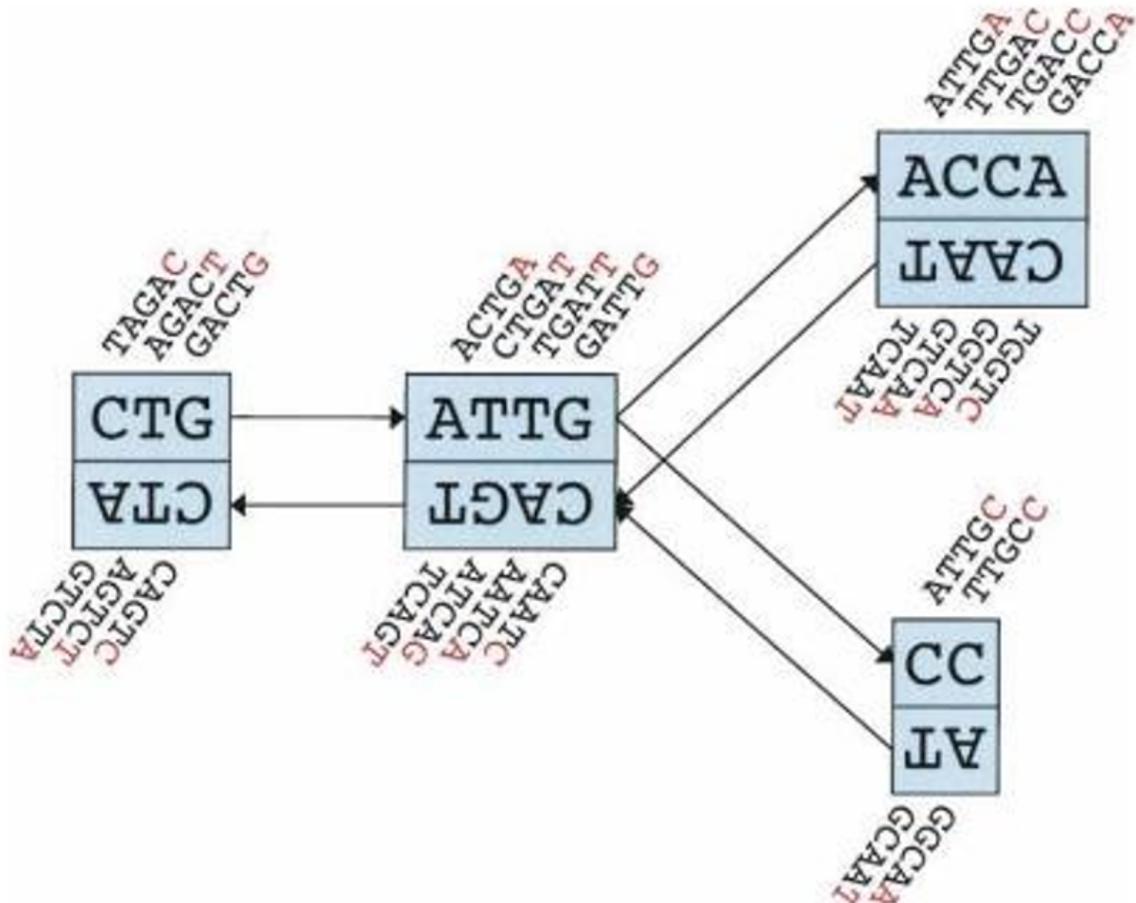
Existen dos tipos principales de ensamblajes; aquellos que utilizan un genoma de referencia para realizar el algoritmo y los ensamblajes de genomas de los cuales no se tiene información suficiente para proveer una referencia, los procesos donde no se requiere un archivo de referencia son llamados ensamblajes *De Novo* (Silva *et al.*, 2013; Lischer & Shimizu, 2017). El ensamblaje *De Novo* se puede comparar con la construcción de un gran rompecabezas con millones de piezas, de las cuales pueden existir muchas secuencias repetidas que pueden causar problemas en el algoritmo, por lo tanto, existen varios algoritmos que se utilizan para tener un mejor ensamblaje (Liao *et al.*, 2019). Estos generalmente toman en cuenta factores como el tamaño del genoma, la presencia de secuencias repetitivas, el porcentaje de N's en las lecturas y entre otros (Sohn y Nam, 2018; Henson *et al.*, 2012). A pesar que cada algoritmo tiene su propia complejidad, los procesos de ensamblaje *De Novo* comprenden tres grandes pasos: la generación de *contigs* (una secuencia de ADN ensamblaje en base de superponer fragmentos de ADN), la construcción de andamios que ampliamente se conocen como *Scaffolds*, y el proceso de rellenar las secciones donde no fue posible identificar con una alta probabilidad la base, mejor conocido como *gap filling* (Sohn y Nam, 2018).

Figura 10
Ensamblaje de Novo



Nota. Adaptado de Sohn y Nam, 2018

Uno de los algoritmos de ensamblaje más conocidos es *Velvet* (Zerbino y Birney, 2008), este se caracteriza por utilizar un principio de grafos De Brujin Euleriano bidireccional (Sohn y Nam, 2018). Este algoritmo realiza el ensamblaje, elimina errores y maneja las secuencias repetidas; primero, el algoritmo de corrección de error une las secuencias complementarias, posteriormente, el algoritmo de las secuencias repetidas evalúa las regiones donde hay superposición (Zerbino y Birney, 2008). En el algoritmo, se analiza que secuencias son iguales en base a un Kmero calculado por las secuencias crudas, cada análisis representa un nodo, el cual es comparado con el nodo reverso complementario como se puede ver en la *figura 11*.

Figura 11*Esquemática del sistema Velvet**Nota. Adaptado de Zerbino & Birney , 2008*

Otro programa que se utiliza para realizar ensamblados de genomas es St. Petersburg genome Assembler *SPAdes* (Bankevich *et al.*, 2012; <https://cab.spbu.ru/software/spades/>). Este es un ensamblador que también se basa en un principio de grados De Bruijn Euleriano de dos capas y se utiliza particularmente para resultados de secuenciación de una sola célula, por lo que refleja resultados satisfactorios en genomas bacterianos (Sohn y Nam, 2018). El algoritmo de ensamblaje tiene cuatro etapas principales: la construcción del grafo de ensamblaje a través de la simplificación De Bruijn, particularmente se utiliza un grafo con tamaños múltiples en el cual se implementa métodos que pueden detectar y remover lecturas quiméricas, y realizar histogramas de distancia; posteriormente se lleva a cabo al ajuste de los bimeros-k, en este paso se estima la distancia entre los Kmeros en el genoma a través de los histogramas de distancia; luego se construye el grafo de ensamblaje pareado basado en el enfoque PDBG; y por último se realiza la construcción de *contigs* (Bankevich *et al.*, 2012; Prjibelski *et al.*, 2020).

Por último, *ABYSS* (*Assembly by short sequences*; <https://github.com/bcgsc/abyss>) (Simpson *et al.*, 2009) es un algoritmo de ensamblaje basado en la distribución de grafos De Bruijn implementado a través de C++ y utiliza el protocolo de comunicación entre nodos MPI (Simpson *et al.*, 2009). Primero, se realiza la construcción del sistema de grafos de Bruijn en donde las secuencias con N 's se descartan; posteriormente, las secuencias l -meros se superponen utilizando un marco de lectura con una longitud k ; posteriormente se realiza la lectura de errores antes de realizar la construcción de los *contigs*, para esto el algoritmo elimina los vértices creados por errores de secuenciación, por último, se realiza la unión de los *contigs* a través de la información pareada (Simpson *et al.*, 2009; Jackman *et al.*, 2017).

Calidad de ensamblaje con *QUAST*

Luego de realizar un ensamblaje, es necesario evaluar la calidad de los resultados de cada algoritmo utilizado. *QUAST* es una herramienta para evaluar métricas de calidad de ensamblaje tales como: tamaño de los *contigs*, lo que comprende el tamaño de estos, la longitud del *contig* más largo, el total de bases en el ensamblaje y los valores Nx , que se refiere a la longitud del *contig* más largo de tal manera que es posible ensamblar un $X\%$ del genoma (Gurevich *et al.*, 2013). Asimismo, también es posible evaluar variaciones estructurales y errores de ensamblaje o *misassemblies*, en donde se muestra el número de estos errores con base en la definición de plataforma, la cual define estos puntos como posiciones en los *contigs* en donde una región se alinea sobre 1kb de la región próxima, o en su defecto, se solapan por una longitud mayor a 1kb (Gurevich *et al.*, 2013; Mikheenko *et al.*, 2018). Por último, también muestra métricas de representación del genoma si se confiere un archivo de referencia; en este se puede observar la fracción del genoma, el rango de duplicación, el porcentaje GC y el número de genes, tanto completos como secuencias parciales (Gurevich *et al.*, 2013).

Anotación del genoma

Para poder realizar un estudio funcional del genoma se utilizan algoritmos para llevar a cabo una anotación computacional, esto se refiere al proceso de identificar características de interés en un archivo de ensamblaje para determinar regiones de secuencias codificantes, secuencias de RNAs y secuencias de péptidos señal (Richardson y Watson, 2012; Seeman, 2014). Uno de los programas más usados ampliamente para anotar genomas de organismos procariontes es *prokka*, el cual es una herramienta que permite llevar a cabo anotaciones automatizadas utilizando como único parámetro necesario el archivo del ensamblaje; cabe mencionar que este programa está especializado en realizar anotaciones de genomas ensamblados por algoritmos *De novo* (Seeman, 2014). Para realizar la identificación de estas características, *prokka* utiliza diferentes herramientas para realizar el escaneo del genoma (Seeman, 2014); para determinar secuencias codificantes utiliza el programa desarrollado por Hyatt. (2010) denominado *Prodigal*; para determinar secuencias de ARN ribosomal utiliza *RNAmmer* (Lagesen *et al.*, 2007); en el caso de secuencias de ARN de transferencia utiliza el programa de Laslett & Canback (2004), *Aragorn*. Asimismo, también es capaz de identificar secuencias de péptido-señal por *signalP* (Petersen *et al.*, 2011) y secuencias no codificantes de ARN por el programa *Infernal* (Kolbe y Eddy, 2011).

5.6 Ingeniería genética

La ingeniería genética comprende las tecnologías que llevan a cabo una manipulación de genes, clonación y producción de sistemas recombinantes (Nicholl, 2002). Los avances en los últimos años han permitido desarrollar tecnologías a base de biología molecular para poder diseñar y ejecutar modificaciones genéticas con el propósito de estudiar mecanismos moleculares, aumentar

las capacidades de un organismo o tratar alguna enfermedad (Thomason, 2014; Christou, 2013; Kershaw *et al.*, 2013). Las técnicas de ingeniería genética han tomado un auge particular en la biotecnología debido a que ofrecen un gran potencial de desarrollar sistemas biológicos que puedan producir o degradar compuestos de interés, por lo tanto, se estudian las rutas moleculares asociadas a la producción de la proteína o al proceso metabólico puntual con el objetivo de aumentar las capacidades (Duan *et al.*, 2015).

Existen varias formas de realizar modificaciones genéticas las cuales aprovechan los principios de procesos biológicos para producir el fenotipo de interés (Nicholl, 2002). Por ejemplo, para el estudio de secuencias específicas en células madre embrionarias existen estudios que evalúan el efecto de la presencia o ausencia de un gen particular a través de la recombinación homóloga (Lanigan *et al.*, 2020); la cual es un proceso celular en el cual se repara los daños de doble hebra a través de un molde para mantener la integridad genómica (San filipino *et al.*, 2018). Dependiendo del gen que se desea estudiar o de la modificación que se desee realizar, es necesario evaluar el mejor método para llevar a cabo el proceso. Es importante mencionar que existen varios tipos de modificaciones genéticas, generalmente se resaltan tres: la inserción de una secuencia de ADN (*Knock-in*), la delección de una secuencia de ADN (*Knock-out*) y el reemplazo de una secuencia con un gen exógeno (Lanigan *et al.*, 2020; Rezazade y Dehghani, 2022; Eisener-Dorman *et al.*, 2009).

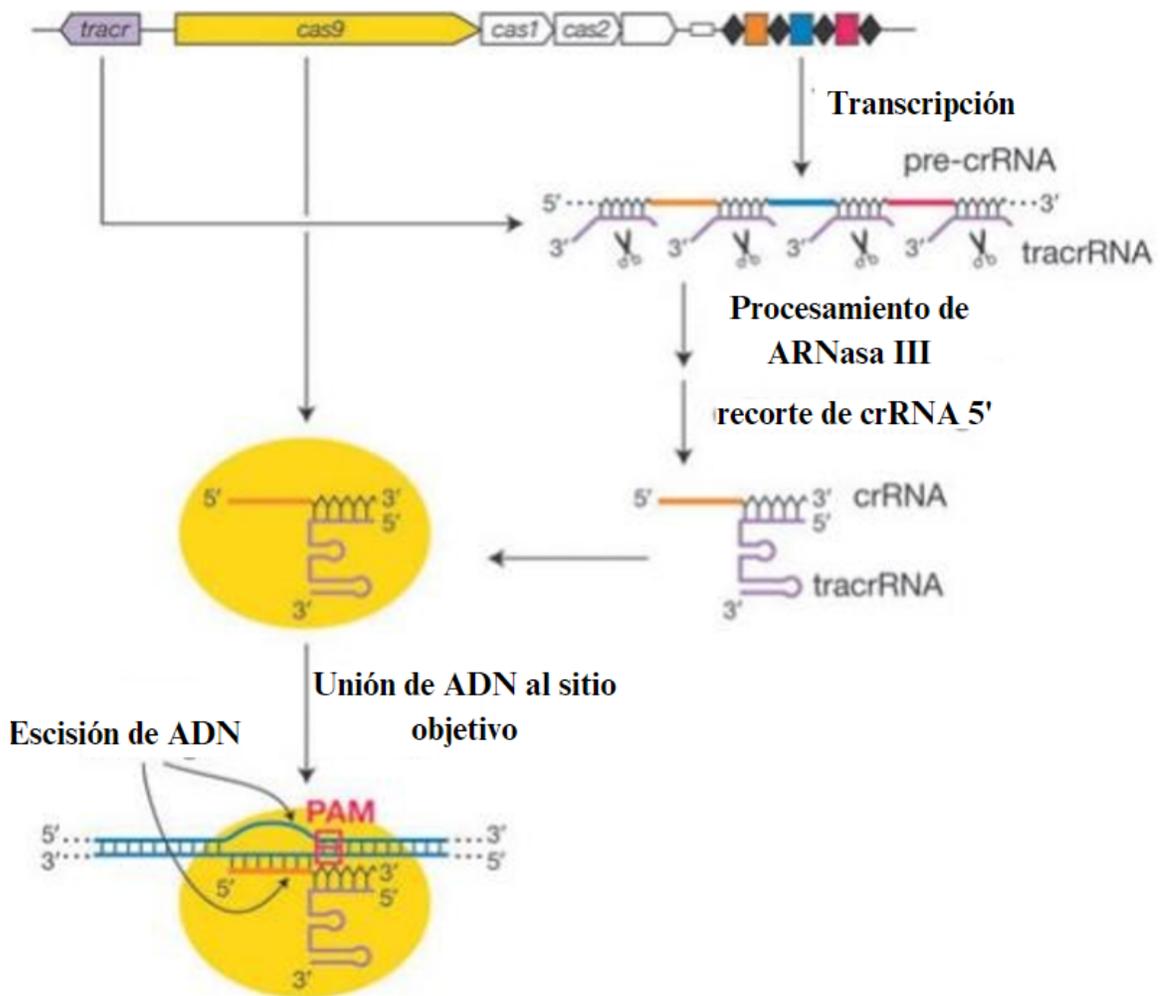
Una de las técnicas más utilizadas para edición génica es a través del uso del sistema CRISPR/Cas9 (*Clustered regularly interspaced palindromic repeats*), identificado por primera vez en la cepa K12 de *E. coli* en 1987, no obstante, fue hasta en el año 2005 que se evidenció realmente cual podría ser su función (Koonin y Makarova, 2019). Las secuencias *spacer* dentro del loci de CRISPR-*cas* de bacterias y arqueas surgieron debido a la infección de virus, las cuales le permiten a estos microorganismos tener un tipo de memoria celular de infecciones previas. En este caso, se reconoce las secuencias de ADN viral y las proteínas Cas inician la destrucción del material debido a que son endonucleasas (Koonin y Makarova, 2019; Sontheimer y Barrangou, 2015).

Los mecanismos de CRISPR/Cas9 son altamente diversos funcional y estructuralmente, estos pueden ser categorizados principalmente en tres tipos (Hryhorowicz *et al.*, 2017; Koonin *et al.*, 2017): los sistemas tipo I, sistemas tipo II y sistema tipo III. Los sistemas tipo I presentan una función de endonucleasa y helicasa las cuales trabajan sincrónicamente para identificar y romper el ADN (Sontheimer y Barrangou, 2015). Los sistemas tipo II son los más utilizados en edición génica debido a que solo utilizan una endonucleasa conocida como proteína Cas9 la cual lleva a cabo la identificación y rotura del ADN (Mir *et al.*, 2017). El sistema tipo III es capaz de identificar y dirigirse a ADN y ARN, sin embargo, su mecanismo aún debe ser estudiado con más profundidad (Sontheimer y Barrangou, 2015).

Los sistemas de CRISPR/Cas9 tipo II presentan tres diferentes subtipos; II-A, II-B, y II-C basado en la ocurrencia y las secuencias de los genes del loci de la proteína Cas (Mir *et al.*, 2017; Sontheimer y Barrangou, 2015). El primer subtipo, A, presenta el gen *csn2* que codifica una proteína involucrada en la adquisición de nuevos espaciadores. El subtipo B posee el gen *cas4* en lugar del *csn2*, de los cuales ninguno está codificado en los sistemas II-C. La mayoría de los sistemas tipo II han sido ampliamente estudiados en *S. thermophilus*, *S. pyogenes* y *N. meningitidis*, en donde fue posible caracterizan los motivos adyacentes protoespaciadores (PAM), que son pequeñas secuencias de ADN que son reconocidas por la *Cas9* y se utiliza para indicar en donde debe de unirse y realizar el corte o modificación en la secuencia de ADN (*figura 13*) (Sontheimer & Barrangou, 2015; Mir *et al.*, 2017; Koonin *et al.*, 2017).

Figura 13

Ruta de interferencia por el sistema CRISPR/Cas9 tipo II



Nota. Adaptado de Sontheimer y Barrangou, 2015

6.1 Diseño, enfoque y tipo de investigación

El diseño experimental de la presente investigación es de carácter cuasi experimental aleatorizado ya que se está evaluando la salud del suelo y la presencia de genes de rutas promotoras de crecimiento. La variable independiente es la cepa de bacteria analizada.

6.2 Población y muestra

Las muestras bacterianas del suelo de cultivo de caña de azúcar fueron recolectadas por la tesista Valeria Ochoa. El procedimiento de recolecta y aislamiento de bacterias se presenta en el trabajo “Análisis del microbioma bacteriano de suelos utilizados para el cultivo de caña de azúcar, *Saccharum officinarum*, en Guatemala mediante metagenómica del gen 16S ARN ribosomal y caracterización de las bacterias promotores del crecimiento que se encuentran en ellos”. En la presente investigación los sujetos de estudio son las bacterias aisladas y seleccionadas.

6.3 Variables

Cuadro 1

Definición de las variables de investigación

Variable	Definición	Tipo	Unidades de medición
Presencia o ausencia del gen	Presencia del gen dentro del genoma bacteriano	Cualitativa	Si/No
Nombre del gen	Gen de interés para promoción del crecimiento	Cualitativa	-

Nota. Elaboración propia

6.4 Hipótesis

6.4.1 Hipótesis nula

La caracterización genómica de bacterias de suelo de cultivo de caña de azúcar permitirá estudiar rutas promotoras del crecimiento vegetal para el desarrollo de un sistema de edición génica a través de CRISPR/Cas9.

6.4.2 Hipótesis alternativa

La caracterización genómica de bacterias de suelo de cultivo de caña de azúcar no presentará información relevante para el desarrollo de un sistema de edición génica a través de CRISPR/Cas9 para aumentar sus capacidades promotoras de crecimiento vegetal.

6.5 Metodología

6.5.1 PCR de colonia del gen *16S*

Se inoculará una colonia de cada bacteria en un microtubo de 0.2 mL utilizando un asa estéril. Posteriormente, debe de incubarse a 90°C por 10 minutos para irrumpir con la integridad de la pared y membrana. Posteriormente se realizará el *Master mix* utilizando el buffer 2x *GoTaq® Hot Start colorless Master mix* de PROMEGA. Se realizará una receta con un volumen final de 15 uL por reacción utilizando el Buffer mencionado anteriormente, agua de grado molecular y los primers de amplificación estándar del gen *16s*, **27f** y **1492r**; los componentes de reacción se muestran a continuación:

Cuadro 2

Receta para la amplificación del gen 16S por PCR de colonia

Componentes	Concentración inicial	Concentración final	Volumen agregado (μ L)
Buffer 2x GoTaq®	2x	1x	7.5
Primer 27f	10 μ M	1.5 μ M	1
Primer 1492r	10 μ M	1.5 μ M	1
H2O	-	-	5.5

Nota. Elaboración propia

La amplificación del gen se realizará utilizando el equipo PTC-100TM *Programmable thermal Controller* MJ Research Inc. Utilizando la siguiente configuración:

Cuadro 3*Condiciones de amplificación de 16S para PCR de colonia*

Etapa	Temperatura (°C)	Tiempo (min)	Ciclos
Desnaturalización inicial	95	5:00	1
Desnaturalización	95	0:4	
Alineamiento	53	0:4	30
Extensión	72	1:5	
Extensión final	72	15:00	1
Almacenamiento	4	indefinido	1

Nota. Elaboración propia

Los productos de PCR serán visualizados a través de electroforesis en gel a 1%, utilizando Buffer TBE 1x, el marcador molecular *TrackIt™ 100 bp DNA Ladder* de ThermoFischer y *GelRed Nucleic Acid Stain 1000x DMSO* de SigmaAldrich a un voltaje de 100v.

6.5.2 Extracción de ADN de bacterias

primero se inoculará una colonia de bacteria en un microtubo de 1.5 mL con 1000 µL de caldo LB y se dejará un ON a 37°C con una agitación a 150 rpm. Para la extracción de ADN de bacterias se utilizará el protocolo de extracción de ADN de cultivos bacterianos en suspensión del kit *QIAmp® DNA Mini and Blood Mini Handbook* de QIAGEN utilizando los pasos para obtener ADN genómico no libre de ARN.

6.5.3 Amplificación del gen 16S

Para la amplificación del gen, primero se normalizará la concentración del ADN extraído de todas las muestras a 50 ng/µL. Posteriormente se utilizará el buffer 2x *GoTaq® Hot Start colorless Master mix* de PROMEGA, agua de grado molecular y los primers **27f** y **1492r**. los componentes de reacción se muestran a continuación:

Cuadro 4*Receta para amplificación de 16S*

Componentes	Concentración inicial	Concentración final	Volumen agregado (µL)
Buffer 2x GoTaq®	2x	1x	7.5
Primer 27f	10 µM	1.5 µM	1
Primer 1492r	10 µM	1.5 µM	1
ADN	50 ng/mL	100 ng	2
H2O	-	-	3.5

Nota. Elaboración propia

La amplificación del gel se realizará utilizando el equipo PTC-100TM *Programmable thermal Controller* MJ Research Inc. Utilizando la siguiente configuración:

Cuadro 5*Condiciones de amplificación de 16S*

Etapa	Temperatura (°C)	Tiempo (min)	Ciclos
Desnaturalización inicial	95	5:00	1
Desnaturalización	95	0:4	
Alineamiento	53	0:4	30
Extensión	72	1:5	
Extensión final	72	15:00	1
Almacenamiento	4	indefinido	1

Nota. Elaboración propia

Los productos de PCR serán visualizados a través de electroforesis en gel a 1%, utilizando Buffer TBE 1x, el marcador molecular *TrackIt™ 100 bp DNA Ladder* de ThermoFischer y *GelRed Nucleic Acid Stain 1000x DMSO* de SigmaAldrich a un voltaje de 100v.

6.5.4 Amplificación del gen *nifH*

Para realizar la amplificación del gen *nifH*, de igual manera se realizó una normalización de la concentración de ADN extraído a 50 ng/uL. Se realizó una reacción de 15 uL utilizando el buffer 2x *GoTaq® Hot Start colorless Master mix* de PROMEGA, agua de grado molecular y los primers descritos en el SOP de amplificación de *nifH* realizado por Ochoa, (2022) (ver anexos).

6.5.5 Secuenciación del gen *16S*

Previo a la secuenciación del gen se determinará la concentración del producto de PCR a través de espectrofotometría UV/Visible utilizando el equipo *NanoDrop One/One Microvolume UV/Vis* de ThermoFischer Scientific. La secuenciación se realizará a través de Macrogen Inc. *Online Sequencing order System* para el amplicón de *16s* en una sola dirección. Los resultados retornados son un archivo FASTA con el resultado de la secuenciación y su respectivo cromatograma.

6.5.6 Identificación de la especie bacteriana

Para la identificación de la especie bacteriana se utilizarán dos métodos para realizar una referencia cruzada. Se utilizará la herramienta *Nucleotide BLAST* para hacer un alineamiento de los FASTA de cada bacteria; asimismo, se realizará una búsqueda utilizando la herramienta *SILVAngs* para análisis de genes de ARN ribosomal de la base de datos de SILVA. Se anotó las especies más probables a base del Valor E, el porcentaje de identidad, la cobertura y los puntajes totales.

6.5.7 Selección de bacterias candidatas

Se evaluará el potencial de las bacterias con base en sus capacidades bioquímicas, dichas capacidades fueron determinadas en un estudio previo según (Ochoa, 2022). Las características son las siguientes: Fijación de nitrógeno, solubilización de fosfatos, solubilización de potasio, producción de ACC deaminasa, producción de proteasas, solubilización de silicatos, productores de sideróforos, producción de lipasas, producción de ácido indol acético, producción de quinasas y producción gluconasas. Se seleccionará las bacterias que fueron positivas para la mayoría de las pruebas.

6.5.8 Secuenciación de genoma completo

El procedimiento de preparación para realizar la secuenciación de genoma completo (WGS) se llevará a cabo utilizando el kit *Illumina DNA Prep Reference Guide*. Previo a la secuenciación, se realizará una extracción de ADN como fue descrito previamente.

Preparación de bibliotecas genómicas

Para la preparación de bibliotecas pes necesario primero realizar la tagmentación del ADN genómico, luego es necesario llevar a cabo una limpieza post-tagmentación para posteriormente amplificar el ADN. Por último, se realizará un gel de electroforesis de agarosa 1x para visualizar que la preparación se ha llevado a cabo correctamente y poder calcular la molaridad. El procedimiento detallado se encuentra en la sección de anexos.

Purificación y secuenciación de bibliotecas

Luego de realizar la preparación de bibliotecas es necesario llevar a cabo una limpieza con el propósito de purificar las bibliotecas y que las lecturas sean apropiadas. Por último, debe de prepararse el *pool* de las bibliotecas; para esto, se cuantificará la concentración de ADN y realizar los cálculos para tener una concentración de 2 M por cada una de las muestras. La secuenciación se llevó a cabo utilizando el equipo *iSeq 100 System* de Illumina.

6.5.9 Análisis bioinformático

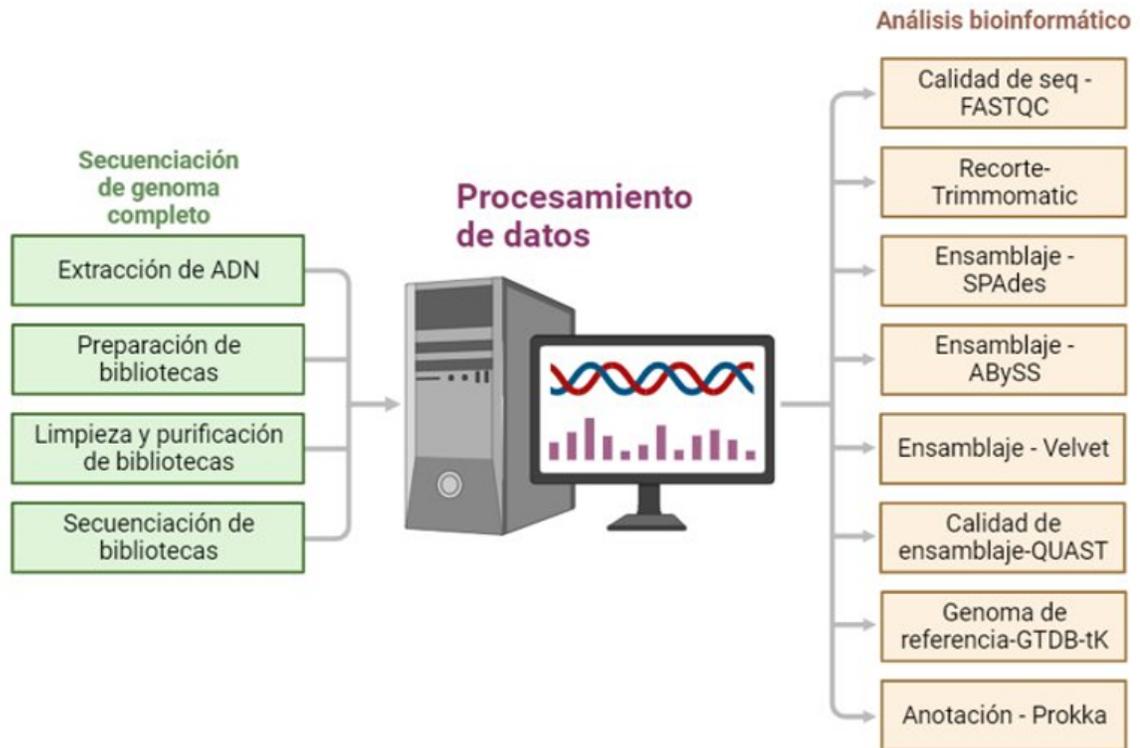
Para realizar el análisis de las secuencias, el ensamblaje del genoma y los análisis computacionales se utilizará una laptop Dell con un sistema operativo Windows 10 Pro 64-bit, un procesador Intel® Core™ i5-1135G7 de 11va generación con 8 nucleos (8 CPU's) y una memoria RAM de 8192 MB. Se utilizará el subsistema de Linux para windows (WSL) a través del sistema operativo de Linux Ubuntu 20.4 LTS. Adicionalmente, para el manejo de ciertos paquetes, se instalará en Ubuntu 20.04, Anaconda versión 1.9.0. Siguiendo el siguiente *pipeline*:

```
- root@name sudo apt-get update
- root@name wget https://repo.anaconda.com/archive/Anaconda3-2021.11-Linux-x86_64.sh
- root@name sha256sum Anaconda3-2021.11-Linux-x86_64.sh
- root@name bash Anaconda3-2021.11-Linux-x86_64.sh
- root@name source ~/.bashrc
```

Para los paquetes que se usarán a través de Anaconda se recomienda sobremanera crear un ambiente individual para cada uno, con el propósito que no exista conflicto por las versiones de las dependencias.

Figura 14

Metodología general de secuenciación y análisis bioinformático



Nota. Elaboración propia

Evaluación de calidad de secuencia

Los archivos retornados por el secuenciador iSeq 100 System son dos archivos por muestra, debido a que son lecturas en dos direcciones, en un formato FASTQ. Para poder evaluar la calidad de las secuencias crudas de la secuenciación se utilizará el paquete FASTQC para obtener un reporte con estadísticas que permite evaluar su calidad para el ensamblaje del genoma. FASTQC se instalará en Ubuntu 20.4 LTS utilizando el siguiente comando:

- **root@name** sudo apt-get update

- **root@name** sudo apt -y install fastqc Para ejecutar el programa se utilizará el siguiente comando:

- **root@name** fastqc *NombreArchivoForward.fastq.gz* *NombreArchivoReverse.fastq.gz*

Este comando retorna un reporte con varias estadísticas para realizar el análisis.

Identificación con *DragenMetagenomics*

Para comenzar a determinar con mayor confianza la taxonomía de las bacterias que se estaban estudiando, con el propósito de confirmar el resultado de la secuenciación del gen *16s*. Se realizará una clasificación utilizando la aplicación *DragenMetagenomics* en la página para análisis *BaseSpace* de Illumina. Los archivos de entrada por cada bacteria son ambos archivos FASTQ.gz

bajo los parametros siguientes:

- Base de datos de referencia: *Extended kraken2*
- Referencia: No aplica
- *Dehosting*: Si aplica
- Identificador del organismo: Bacteria

El programa retorna un arreglo de varios archivos para la determinación del origen de la bacteria en cuestión.

Corte de secuencias no aptas para el ensamblaje

Para las secuencias cuya calidad no sea óptima, se realizará un recorte de las secuencias no aptas utilizando el paquete Trimmomatic para anaconda. La instalación se realizará de la siguiente manera:

- **root@name** conda create —name TrimmomaticEnv
- **root@name** conda activate TrimmomaticEnv
- **root@name** conda install -c bioconda trimmomatic

Para la ejecución del programa se utilizará el siguiente comando:

```
- root@name trimmomatic PE -phred33 NombreArchivoFowardInput.fastq.gz
NombreArchivoReverseInput.fastq.gz NombreArchivoForwardOutputCortado.fastq.gz
NombreArchivoForwardOutputU.fastq.gz NombreArchivoForwardOutputCortado.fastq.gz
NombreArchivoForwardOutputU.fastq.gz SLIDINGWINDOW:4:28 MINLEN:36
```

El programa retorna dos archivos de salida por cada uno de los archivos FASTQ.gz. Se utiliza el archivo recortado para evaluar su calidad utilizando el paquete FASTQC.

Ensamblaje de genoma

Con el propósito de obtener el mejor ensamblaje posible, se utilizarán tres algoritmos distintos para el ensamblaje *De Novo*. Se utilizará velvet, SPAdes y AbySS.

Primer ensamblaje: Para el primer ensamblaje, se utilizará la aplicación *Velvet de novo Assembly* de la página *BaseSpace* de Illumina. Para este programa, los archivos de entrada son ambos archivos FASTQ.gz de cada muestra. El programa retorna un arreglo de archivos reporteros del proceso de ensamblaje, los archivos FASTA resultados del ensamblaje y un análisis de calidad.

Segundo ensamblaje: Para el segundo ensamblaje, se utilizó el algoritmo SPAdes versión 3.13.1 utilizando como archivo de entrada las secuencias FASTQ.gz cortadas por Trimmomatic. Para la instalación del programa se ejecutarán los siguientes comandos:

- **root@name** sudo apt-get update
- **root@name** sudo apt-get -y install spades

Para ejecutar el programa y realizar el ensamblaje se utilizará el siguiente comando:

```
- root@name spades.py -1 NombreArchivoForwardCortado.fastq.gz -2
NombreArchivoReverseCortado.fastq.gz -o DirecciónDeSalida
```

El programa retorna los archivos generados para realizar el ensamblaje, para evaluar la calidad y los archivos FASTA de los *contigs* y los *scaffolds*.

Tercer ensamblaje: Para el tercer algoritmo que se utilizará, es necesario proveer el kmero óptimo para llevar a cabo del ensamblaje. Para esto, se utilizará el paquete Kmergenie, el cual se instalará con el siguiente comando:

```
- root@name conda create --name KmergenieEnv
- root@name conda activate KmergenieEnv
- root@name conda install -c bioconda kmergenie
```

Para la ejecución del programa se utilizará el siguiente comando:

```
- root@name kmergenie NombreArchivotCortado.fastq.gz
```

El programa kmergenie retorna una serie de archivos en donde se evalúa la longitud óptima de cada kmero, mostrando los archivos de normalidad y un reporte final en donde se indica el valor óptimo recomendado.

Para la ejecución del ensamblaje se utilizó el programa AbySS versión 4.2.1, este fue instalado de la siguiente manera:

```
- root@name sudo apt-get update - root@name sudo apt-get -y install abyss
```

El comando para ejecturar el ensamblaje de AbySS debe ser de la siguiente manera:

```
- root@name abyss-pe k=KmerOptimo name=Nombre
in='NombreArchivoForwardCortado.fastq.gz NombreArchivoReverseCortado.fastq.gz'
```

El comando genera una lista de archivos de salida para evaluar el proceso de ensamblaje, asimismo, se retornan los FASTA para los archivos *contigs* y *scaffolds*.

Evaluación de la calidad de ensamblaje

Para evaluar la calidad de ensamblaje se utilizará el paquete QUILT versión 5.2.0, el cual retorna una serie de archivos, entre ellos un reporte donde se muestran varias estadísticas relevantes, tales como el número de *contigs*, *scaffolds*, N50, tamaño del ensamblaje y entre otras. Para instalar el paquete se ejecutará el siguiente comando:

```
- root@name sudo apt-get update
- root@name pip3 install quast
```

Para la ejecución del programa se seguirá el siguiente comando:

```
- root@name quast.py contigs1.fasta scaffolds1.fasta (...) contigsN.fasta scaffoldsN.fasta
-o DirecciónSalida
```

En el programa QUILT, se puede cargar más de un archivo de *contigs* y *scaffolds* por ensamblaje para mostrar las estadísticas y realizar una comparación de todos los archivos. Se retornan varios elementos, entre ellos un HTML en donde se muestran tablas y gráficos sobre las estadísticas del ensamblaje.

Determinación del genoma de referencia *GTDB-Tk*

Para la confirmación del genoma de referencia, se utilizará el paquete GTDB-tk versión, el cual usa el ensamblaje. Para emplear este paquete, se utilizará la laptop debido a que es necesario instalar las bases de datos para realizar los alineamientos. Para la instalación se ejecutará el siguiente comando:

- **root@name** conda create -name gtdbtkEnv
- **root@name** conda activate gtdbtkEnv
- **root@name** conda install -c bioconda gtdbtk
- **root@name** download-db.sh

Posteriormente, se realizará la ejecución del programa para la clasificación de la siguiente manera:

- **root@name** gtdbtk classify -wf -genome_dir *direcciónGenoma* -out_dir *direcciónSalida* -cpus 12 -mash_db *direcciónMashDB* -prefix *nombre* -extension fasta

Este comando retorna una serie de archivos que muestran el porcentaje de similitud con el mejor genoma de referencia. Asimismo, se realizará una indentificación de marcadores genéticos de la bacteria con el siguiente comando:

- **root@name** gtdbtk identify --genome_dir *direcciónGenoma* -out_dir *direcciónSalida* -cpus 12 -mash_db *direcciónMashDB* -prefix *nombre* -extension fasta

Se retornan varios archivos de genes marcadores identificados con la base de datos.

Evaluación de calidad con genoma de referencia *fastANI*

Para evaluar la similitud del genoma de referencia con el ensamblaje, se utilizará el paquete FastANI para obtener un grado de similitud y una comparación visual de regiones conservadas evolutivamente. Se instalará de la siguiente manera:

- **root@name** conda create -name fastANIenv
- **root@name** conda activate fastANIenv
- **root@name** conda install -c bioconda fastani

Para el alineamiento y la obtención del archivo visual se ejecutará el siguiente comando:

- **root@name** fastANI -q *ensamblaje.fasta* -r *genomaReferencia.fasta* -visualize -o *NombreSalida*

Este comando retorna dos archivos; uno de texto con el porcentaje de similitud y un archivo .VISUAL para realizar el diagrama usando python o R. Se utilizará python usando el código proveído en anexos. Se utilizará el siguiente comando:

- **root@name** python visualize.py *ensamblaje.fasta* *genomaReferencia.fasta* *archivo.visual* *nombreSalida.png*

Este comando retorna un archivo .png donde se puede observar la comparación de ambos archivos y sus regiones conservadas.

Corrección de ensamblaje *ragtag*

Luego de obtener el genoma de referencia, se utilizará Ragtag, un paquete que permite realizar ordenamientos del ensamblaje en base a un genoma de referencia para mejorar la calidad. Se instalará de la siguiente manera:

- **root@name** conda create -name ragtagEnv
- **root@name** conda activate ragtag Env
- **root@name** conda install -c bioconda ragtag

Para ejecutar el paquete se utilizará el siguiente comando, donde es necesario el método de reordenamiento, el genoma de referencia y el genoma ensamblado:

- **root@name** ragtag.py correct *GenomaReferencia.fasta ensamblaje.fasta*

Se obtienen varios archivos de los ensamblajes ordenados según el genoma de referencia.

Anotación de genoma *prokka*

Para la anotación del genoma se utilizará la aplicación *Prokka genome Annotation* de la página para análisis *BaseSpace* de Illumina. El archivo de entrada es el ensamblaje final y se utilizará el ID “*BS*”. El programa retorna varios formatos con información relevante para análisis tales como información de los genes y proteínas dentro del genoma.

Determinación de genes asociados a las rutas de interés en cada bacteria

Primero se generará un listado de los genes involucrados en las rutas de interés en cada uno de los géneros de las especies bacterianas analizadas. Para esto, primero se realizará una búsqueda en la literatura donde se reporte la presencia de genes en KEGG y en la literatura. Se generará un listado de los genes reportados con sus respectivos códigos de acceso de “GenBank”.

Posteriormente, se realizará la identificación de los genes reportados a través del archivo *GFF* generado por el programa *prokka* de forma manual. Se buscará el gen por su nombre y código para construir un cuadro de presencia o ausencia de genes.

6.5.10 Desarrollo de mecanismo de mejoramiento por ingeniería genética

Diseño de plásmidos para la transformación bacteriana

Se realizará una búsqueda de plásmidos inducibles por compuestos que no afecten la salud del suelo para poder proponer una transformación bacteriana para mejorar la fijación de nitrógeno. Se buscará un plásmido con un gen reportero y resistencia para poder realizar la selección de las bacterias transformantes. Asimismo, se agregará la secuencia reguladora y la secuencia promotora para poder mediar la inducción del gen que se desea clonar. Se deberá considerar secuencias consenso para el origen de replicación, así como secuencias asociadas a una mejor permanencia del plásmido en la bacteria.

Diseño para KO del gen *nifL* a través de CRISPR/Cas9

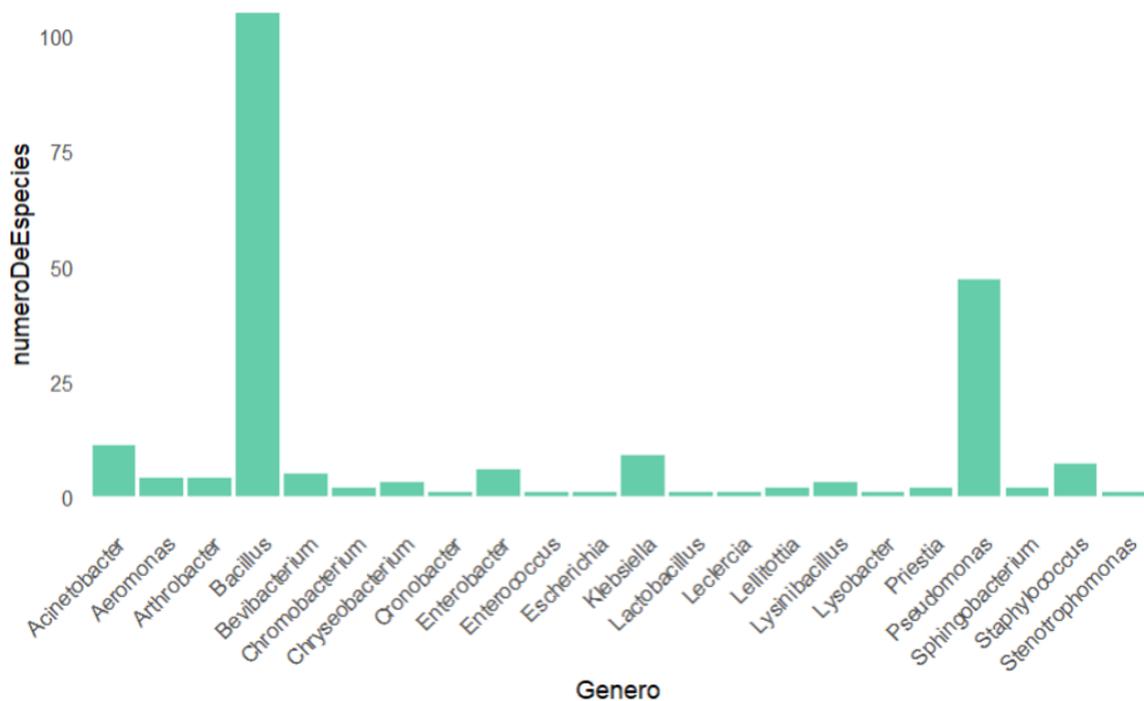
Se realizará un alineamiento multiseuencia de los genes *nifL* anotados en la base de datos de NCBI, para identificar regiones conservadas en el gen entre distintas especies. Se realizará un diseño de un ARN guía para llevar a cabo el KO de la bacteria. Para esto, se utilizará la herramienta de diseño de gRNA en *Benchling*. Se elegirá el genoma de referencia *Bacillus subtilis* con el código de acceso GCA_000009045.1 y se escogerá la secuencia PAM “**NGG**” para la endonucleasa SpCas9. Debido a lo que se desea realizar es un KO, se elegirá un intervalo entre 100-250 nucleótidos de la secuencia para el diseño del gRNA. Por último, se evaluará el resultado en base del sitio de corte, el sentido del gRNA, y los puntajes *on/off-target*, los cuales permiten analizar la eficacia del KO y la probabilidad de cortes inespecíficos respectivamente.

7.1 Identificación de cepas bacterianas por *16S rRNA*

Los resultados de secuenciación del gen *16S* permitieron identificar las especies aisladas de bacterias de suelo de cultivo de caña de azúcar.

Figura 15

Abundancia de géneros en suelo de Saccharum officinarum



Nota. Elaboración propia

Como puede verse en la *figura 15*, existe una gran diversidad caracterizada en el suelo. Se identificaron 22 géneros distintos de los cuales *Bacillus* y *Pseudomonas* son los que tienen mayor número de especies representantes.

7.2 Selección de bacterias para la secuenciación de genoma completo (WGS)

Posterior a la identificación a través de secuenciación del gen de la subunidad del ARN ribosomal *16S*, se utilizó la información descrita por Ochoa, V. (2022) para escoger cuales bacterias se utilizarían para la secuenciación. Se tomó en cuenta las siguientes características: 1. AD: Producción de ACC deaminasa; 2. FN: Capacidad de fijar nitrógeno; 3. GL: Producción de gluconasas; 4. IAA: Producción de ácido-indol-3-acético; 5. LP: Producción de lipasas; 6. PR: Producción de pirrolnitrina; 7. Q: Producción de quitinasas; 8. SD: Producción de sideróforos; 9. Capacidad de solubilización de fosfatos; 10. SI: Producción de silicatos y 10. SP: Solubilización de potasio. En base a estas características, se escogió dentro de 222 bacterias aquellas que poseían la mayor cantidad de estas.

Cuadro 6

Bacterias candidatas con capacidad de promover crecimiento vegetal

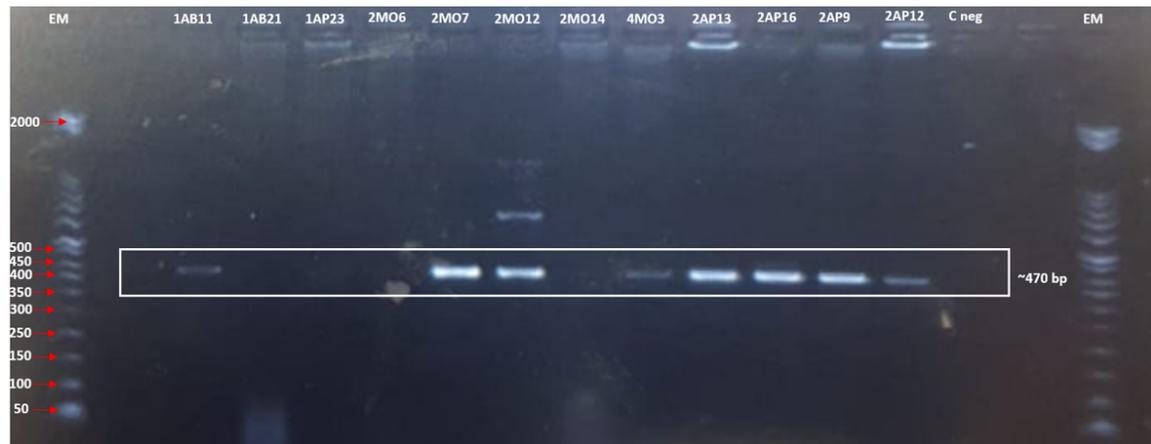
Especie	Características											
	AD	FN	GL	IAA	LP	PN	PR	Q	SD	SF	SI	SP
1AB11	■	■	■	■	■	■	■	■	■	■	■	■
1AP23	■	■	■	■	■	■	■	■	■	■	■	■
2MO6	■	■	■	■	■	■	■	■	■	■	■	■
2MO7	■	■	■	■	■	■	■	■	■	■	■	■
2MO12	■	■	■	■	■	■	■	■	■	■	■	■
2MO14	■	■	■	■	■	■	■	■	■	■	■	■
4MO3	■	■	■	■	■	■	■	■	■	■	■	■
2AP13	■	■	■	■	■	■	■	■	■	■	■	■
2AP16	■	■	■	■	■	■	■	■	■	■	■	■
2AP9	■	■	■	■	■	■	■	■	■	■	■	■

Nota. Elaboración propia

En el *cuadro 6*, se puede evidenciar que todas las bacterias seleccionadas tenían la capacidad de fijar nitrógeno, solubilizar potasio y fosfatos, producir sideróforos para quelar metales, sintetizar pirrolnitrina para el control potencial de patógenos fúngicos, y producir fitohormonas como la ACC desaminasa y ácido-indol-3-acético. Solamente la bacteria 1AP23 se caracterizó por producir quitinasas. Por el otro lado, cinco bacterias presentaron la capacidad de producir gluconasas. Cabe mencionar que, debido al enfoque del estudio, se priorizó la capacidad de fijar nitrógeno sobre las demás características promotoras del crecimiento de plantas.

Figura 16

Amplificación de nifH de las bacterias seleccionadas



Nota. Elaboración propia

Para evaluar a mayor profundidad la capacidad fijadora de nitrógeno de las bacterias seleccionadas, se prosiguió a realizar una PCR del gen *nifH*, el cual codifica para la nitrogenasa del complejo de fijación. Como se menciona en la sección de metodología, se utilizó cebadores degenerados debido a que se ha caracterizado que dentro de este gen existe una alta variabilidad (Coelho *et al.*, 2009). La prueba evidenció que, de las 12 bacterias seleccionadas, solamente cuatro de ellas no fueron positivas para la identificación del gen (ver *figura 16*). Es importante mencionar que no fueron descartadas debido a que diversos estudios han determinado que existen otros genes caracterizados para la fijación de nitrógeno, particularmente los genes *vnf* y *anf*, además las pruebas bioquímicas realizadas por Ochoa, (2022) evidenciaron que sí eran capaces de fijar nitrógeno.

7.3 Análisis bioinformático de WGS de bacterias

Las doce bacterias seleccionadas fueron secuenciadas a través de secuenciación de nueva generación (NGS) a través de *Illumina*. Se obtuvo para cada una de las muestras dos archivos *fastq.gz* de las lecturas en ambas direcciones. Los análisis de calidad determinados por el programa *fastQC* reflejaron que los resultados de las secuencias fueron satisfactorios. No obstante, previo a proseguir con los análisis superiores, se descartó dos muestras, las cepas 1AB21 y 2AP12, debido a que presentaron contaminación por ADN de otra muestra. Los cuadros de resultados del análisis de calidad generados por *FastQC* presentan tres colores; verde si el parámetro evaluado fue exitoso, naranja si hay alguna alerta y rojo, en el caso de no haber cumplido con los requerimientos del parámetro.

Cuadro 7

Calidad de secuencias crudas caracterizadas por *fastQC*

Parámetros de calidad	1AB11		1AP23		2MO6		2MO7		2MO12	
	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2
Estadísticas básicas										
Calidad de secuencia base										
Calidad de secuencia <i>per tile</i>										
Puntuación de calidad por secuencia										
contenido de base por secuencia										
Contenido de GC por secuencia										
Contenido de bases N										
Distribución de longitud de secuencia										
Niveles de duplicación de secuencia										
Secuencias sobrerrepresentadas										
Contenido de adaptadores										

Parámetros de calidad	2MO14		4MO3		2AP13		2AP16		2AP12	
	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2
Estadísticas básicas										
Calidad de secuencia base										
Calidad de secuencia <i>per tile</i>										
Puntuación de calidad por secuencia										
contenido de base por secuencia										
Contenido de GC por secuencia										
Contenido de bases N										
Distribución de longitud de secuencia										
Niveles de duplicación de secuencia										
Secuencias sobrerrepresentadas										
Contenido de adaptadores										

Nota. Elaboración propia

Los análisis de calidad de las secuencias crudas muestran que todas, a excepción de R2 de la muestra 1AB11, presentaron una buena calidad bajo los parámetros de las estadísticas básicas, la calidad de secuencia base y la calidad de secuencia *per-tile*. Por el otro lado, todas las secuencias no cumplieron con los requerimientos para el parámetro de contenido de base por secuencia, esto se debe a que generalmente existe una inclinación dependiendo de la preparación de las bibliotecas genómicas (sato *et al.*, 2019). Asimismo, todas las muestras reflejaron tener alerta o fallo en el parámetro GC por secuencia, esto está asociado a las secuencias sobrerrepresentadas (Schwartz *et al.*, 2010), por lo que este parámetro también mostró una alerta.

Las alertas y fallos en los parámetros de calidad indican que la realización de un ensamblaje puede llevar a errores en la metodología (Chen *et al.*, 2017), asimismo, se podría obtener un resultado erróneo, por lo tanto, las secuencias fueron procesadas con el paquete *trimmomatic* para cortar regiones de las secuencias que general errores en el análisis.

Cuadro 8
Calidad de secuencias filtradas por Trimmomatic

Parámetros de calidad	1AB11		1AP23		2MO6		2MO7		2MO12	
	R1	R2								
Estadísticas básicas										
Calidad de secuencia base										
Calidad de secuencia <i>per tile</i>										
Puntuación de calidad por secuencia										
contenido de base por secuencia										
Contenido de GC por secuencia										
Contenido de bases N										
Distribución de longitud de secuencia										
Niveles de duplicación de secuencia										
Secuencias sobrerrepresentadas										
Contenido de adaptadores										
Parámetros de calidad	2MO14		4MO3		2AP13		2AP16		2AP12	
	R1	R2								
Estadísticas básicas										
Calidad de secuencia base										
Calidad de secuencia <i>per tile</i>										
Puntuación de calidad por secuencia										
contenido de base por secuencia										
Contenido de GC por secuencia										
Contenido de bases N										
Distribución de longitud de secuencia										
Niveles de duplicación de secuencia										
Secuencias sobrerrepresentadas										
Contenido de adaptadores										

Nota. Elaboración propia

La calidad analizada posterior al realizar el procesamiento mostró que fue posible mejorar los parámetros de secuencias sobrerrepresentadas (a excepción de R2 de la muestra 2AP12) y en algunos casos el contenido de GC por secuencia. No obstante, no fue posible mejorar la calidad en base al contenido de base por secuencia, ya que como se mencionó anteriormente, existe una tendencia a generar este error por preparación de las bibliotecas, particularmente en el proceso de tagmentación.

Las secuencias procesadas fueron utilizadas posteriormente para realizar el ensamblaje del genoma. Se utilizó tres algoritmos *De novo*: *Abyss*, *SPAdes* y *velvet*, para evaluar y seleccionar el mejor ensamblaje para el análisis del genoma. Para evaluar la calidad de los ensamblajes se utiliza el programa Quast que calcula los siguientes parámetros de calidad: 1. El número de *scaffolds*, 2. La longitud total del genoma ensamblado, 3. La longitud del *scaffold* requerida de tal manera que el uso de *scaffolds* del mismo tamaño o mayor puede producir el 50 % del ensamblaje (N50), 4. N90, 5. El número mínimo de *scaffolds* requeridos para producir al menos el 50 % del ensamblaje (L50), 6. L90 y porcentaje de GC.

Cuadro 9*Calidad de los ensamblajes con ABySS*

Muestras	Parámetros de calidad de ensamblaje						
	# de Scaffolds	Longitud total	N50	N90	L50	L90	% GC
1AB11	36	5,094,644	320,277	101,338	6	18	35.26
1AP23	157	4,782,408	75,179	19,016	22	66	62.72
2MO6	706	6,438,959	18,644	4,574	100	354	60.44
2MO7	191	6,634,750	102,918	23,058	22	80	60.44
2MO12	1,406	3,844,299	3,914	1,251	299	965	43.60
2MO14	171	6,590,341	107,597	27,744	19	64	60.48
4MO3	52	3,738,673	114,672	38,529	9	29	39.02
2AP13	1,719	6,097,550	5,108	1,707	358	1,142	60.34
2AP16	1,188	6,221,030	8,408	2,633	216	742	60.35
2AP12	35	3,969,172	296,750	74,341	5	17	38.62

Nota. Elaboración propia

Los ensamblajes generados por *AbySS* presentan generalmente un número de *scaffolds* más elevado que *SPAdes* y *velvet*. Asimismo, se puede observar en los cuadros, particularmente en base al parámetro L50, que, en algunas muestras, es necesario utilizar 100 o más *scaffolds* para poder completar el 50% del genoma. Estos parámetros, junto a N50, reflejan que existe una gran fragmentación en los ensamblajes. Sin embargo, particularmente las muestras 1AB11 y 1 AP23 usando este algoritmo presentaron un menor número de *scaffolds* que los ensamblajes con el algoritmo *SPAdes*. Lo que podría indicar que el genoma se encuentra más íntegro.

Cuadro 10*Calidad de los ensamblajes con Velvet*

Muestras	Parámetros de calidad de ensamblaje						
	# de Scaffolds	Longitud total	N50	N90	L50	L90	% GC
1AB11	36	5,094,644	320,277	101,338	6	18	35.26
1AP23	157	4,782,408	75,179	19,016	22	66	62.72
2MO6	706	6,438,959	18,644	4,574	100	354	60.44
2MO7	191	6,634,750	102,918	23,058	22	80	60.44
2MO12	1,406	3,844,299	3,914	1,251	299	965	43.60
2MO14	171	6,590,341	107,597	27,744	19	64	60.48
4MO3	52	3,738,673	114,672	38,529	9	29	39.02
2AP13	1,719	6,097,550	5,108	1,707	358	1,142	60.34
2AP16	1,188	6,221,030	8,408	2,633	216	742	60.35
2AP12	35	3,969,172	296,750	74,341	5	17	38.62

Nota. Elaboración propia

En el caso de la calidad de los ensamblajes con el algoritmo *velvet*, se puede evidenciar que, a grandes rasgos, presenta mejores resultados que *ABySS* para todas las muestras. Presenta un menor número de *scaffolds*, una menor cantidad de ellos para completar 50% del ensamblaje e incluso una mayor longitud, lo que sugiere que presenta un ensamblaje más con menos fragmentación.

Cuadro 11*Calidad de los ensamblajes con SPAdes*

Muestras	Parámetros de calidad de ensamblaje						
	# de Scaffolds	Longitud total	N50	N90	L50	L90	% GC
1AB11	148	5,233,367	958,922	203,949	3	7	35.61
1AP23	2,394	6,757,699	233,377	702	9	1,250	60.91
2MO6	73	6,650,351	191,689	64,337	12	32	60.44
2MO7	72	6,649,642	225,699	69,783	10	28	60.44
2MO12	789	4,063,197	9,045	2,323	136	482	43.54
2MO14	86	6,643,631	231,957	69,783	10	27	60.45
4MO3	27	3,697,400	358,776	89,998	5	12	38.92
2AP13	344	6,336,386	30,392	9,269	59	201	60.34
2AP16	150	6,347,709	74,900	20,579	29	84	60.34
2AP12	19	3,965,519	405,435	191,160	2	8	38.61

Nota. Elaboración propia

Por último, *SPAdes* fue el algoritmo que generó mejores resultados. La mayoría de los ensamblajes tienen el menor número de *Scaffolds* utilizando este algoritmo en comparación con los otros dos. Además, todas las muestras (menos la 1AP23) necesitaron un menor número de *scaffolds* para ensamblar el 50% (L50) y la longitud de estos era mayor (N50). Asimismo, es importante resaltar que todos los ensamblajes con *SPAdes* resultaron en una longitud de ensamblaje mayor.

Posterior a realizar la comparación de los tres ensamblajes, se seleccionó el mejor ensamblaje para cada muestra basado principalmente en los resultados con menos fragmentación, para poder realizar de mejor manera los alineamientos, y por ende, obtener mejores resultados en la anotación del genoma a través de *prokka*. Sin embargo, previo a esto, se determinó el genoma de referencia más apropiado utilizando el paquete *GTDB-Tk* (descrito en la sección de metodología) a través de alineamientos utilizando el método de clasificación *FastANI*. Con este resultado, se puede obtener con mayor certeza la especie bacteriana de la muestra.

Los resultados de *GTDB-Tk* (ver *cuadro 12*) muestran el índice de similitud calculado con *FastANI*, la clasificación realizada por este, en el cual se muestra el dominio, el filo, la clase, orden, familia, género y la especie determinada (de ser el caso). Asimismo, se muestra el radio de referencia en el cual fue determinada la mejor clasificación.

Los resultados de *GTDB-Tk* permitieron identificar de mejor manera el género y especie de las cepas seleccionadas: *Bacillus tropicus* (1AB11), *Pseudomonas sp* (1AP23), *Pseudomonas paraglyciniae* (2MO6, 2MO7 y 2MO14), *Bacillus subtilis* (2MO12), *Acinetobacter baumannii* (4MO2), *Pseudomonas canadensis* (2AP13 y 2AP16) y *Acinetobacter sp* (2AP12). Por lo tanto, las muestras 2MO6, 2MO7 y 2MO14 comparten un genoma de referencia, *P. paraglyciniae*. Asimismo, las muestras 2AP13 y 2AP16 comparten *P. canadensis* como genoma de referencia. Posteriormente, se realizó la anotación del ensamblaje del genoma utilizando el algoritmo *Prokka*, en el cual se obtuvieron varios archivos con todos los genes codificantes caracterizados. Se utilizó el archivo con extensión GFF para determinar los genes de las características de interés.

Cuadro 12

Determinación de genoma de referencia por GTDB-Tk

Muestra	Clasificación	Referencia de fastANI	Radio de referencia de fastANI	Taxonomía de fastANI	fastANI_ani	fastANI_af	método de clasificación	Cobertura (%)
1AB11	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus_A;s__Bacillus_A_tropicus	GCF_00188403 5.1	95.3658	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus_A;s__Bacillus_A_tropicus	96.65	0.914	ani_screen	86.44
1AP23	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas_E;s__Pseudomonas_E_sp003445295	GCF_00767772 5.1	95	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas_E;s__Pseudomonas_E_sp003445295	98.7	0.94	ani_screen	89.025
2M06	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas_E;s__Pseudomonas_E_paraglyciniae	GCF_02160162 5.1	95	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas_E;s__Pseudomonas_E_paraglyciniae	95.77	0.917	ani_screen	79.9
2M07	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas_E;s__Pseudomonas_E_paraglyciniae	GCF_02160162 5.1	95	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas_E;s__Pseudomonas_E_paraglyciniae	95.79	0.918	ani_screen	78.98
2M012	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus_subtilis	GCF_00000904 5.1	95	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus_subtilis	98.55	0.959	ani_screen	89.636
2M014	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas_E;s__Pseudomonas_E_paraglyciniae	GCF_02160162 5.1	95	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas_E;s__Pseudomonas_E_paraglyciniae	95.83	0.911	ani_screen	79.092
4M03	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Acinetobacter;s__Acinetobacter_baumanni	GCF_00975968 5.1	95	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Acinetobacter;s__Acinetobacter_baumanni	97.96	0.934	ani_screen	86.084
2AP13	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas_E;s__Pseudomonas_E_canadensis	GCF_00050321 5.1	95	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas_E;s__Pseudomonas_E_canadensis	99.94	0.963	ani_screen	94.98
2AP16	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas_E;s__Pseudomonas_E_canadensis	GCF_00050321 5.1	95	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas_E;s__Pseudomonas_E_canadensis	99.96	0.965	ani_screen	95.181
2AP13	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Acinetobacter	N/A	N/A	N/A	N/A	N/A	clasificación taxonómica definida por topología y ANI	N/A

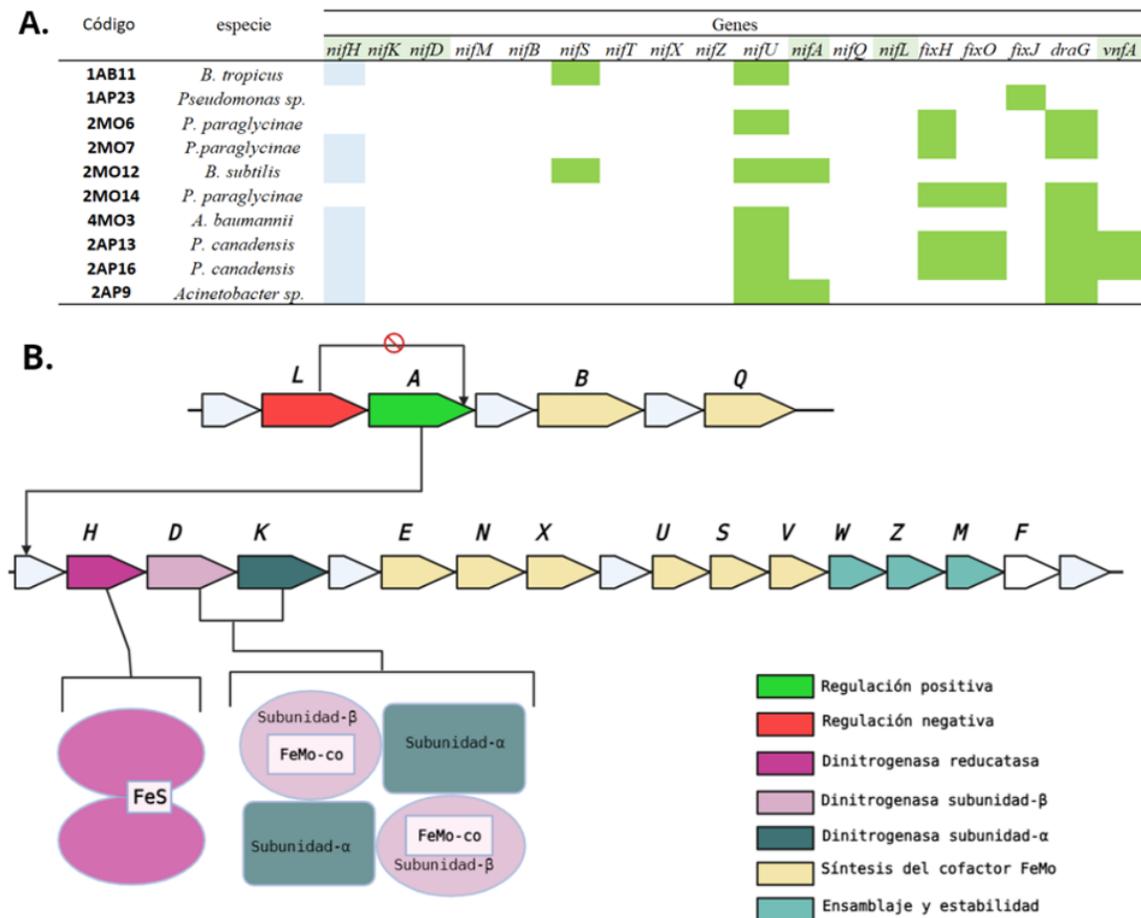
Nota. Elaboración propia

7.4 Identificación de genes asociados a la fijación de nitrógeno

Luego de realizar una búsqueda de los genes implicados en la ruta de fijación de nitrógeno bacteriana en la *Kyoto Encyclopedia of Genes and Genomes* (KEGG) del grupo de genes *nif*, *vnf* y *ndf* se realizó una lista de aquellos codificantes para la formación y estabilidad del complejo nitrogenasa. Posteriormente se buscó manualmente en el archivo GFF cada uno de los genes para la identificación y construcción del operón *nif*.

Figura 17

Determinación de genes para la fijación de nitrógeno. Se muestran los genes identificados en las cepas estudiadas



Nota. (A). El cuadro evidencia en que cepas fue posible identificar los genes en la anotación; los cuadros en verde indican presencia del gen, los genes resaltados (color verde) son aquellos que funcionan para el complejo catalítico y para la regulación del operón, asimismo, los cuadros azules representan las muestras que fueron positivas en el PCR. (B). Esquematización de los genes necesarios para el funcionamiento del complejo nitrogenasa.

Elaboración propia

Fue posible localizar los genes *nifS*, *nifU*, *nifA*, *fixH*, *fixO*, *fixJ*, *DraG* en algunas de las bacterias promotoras seleccionadas, particularmente se encontró una mayor incidencia de *nifU* el cual es importante para el ensamblaje del complejo nitrogenasa (Rubio, 2008). Se caracterizó genes para la síntesis de los núcleos de sulfuro férrico (FeS) del complejo, así como proteínas modulares como es el caso de *nifU*. Asimismo, en algunas de las especies se logró identificar genes reguladores,

particularmente *nifA* para la regulación de la nitrogenasa dependiente de molibdeno y también *vnfA* para la regulación del complejo dependiente de vanadio. No se identificó los genes *nifH*, *nifD* y *nifK* en los genomas anotados, sin embargo, en la *figura 17A* se evidencia las bacterias que fueron positivas para el gen *nifH* a través de la PCR realizada. Asimismo, la *figura 17B* muestra un diagrama realizado de los genes necesarios para el funcionamiento correcto de la nitrogenasa dependiendo de molibdeno a través del operón *nif* (Franco-Sierra *et al.*, 2020).

7.5 Identificación de genes asociados a biodisponibilidad de fósforo

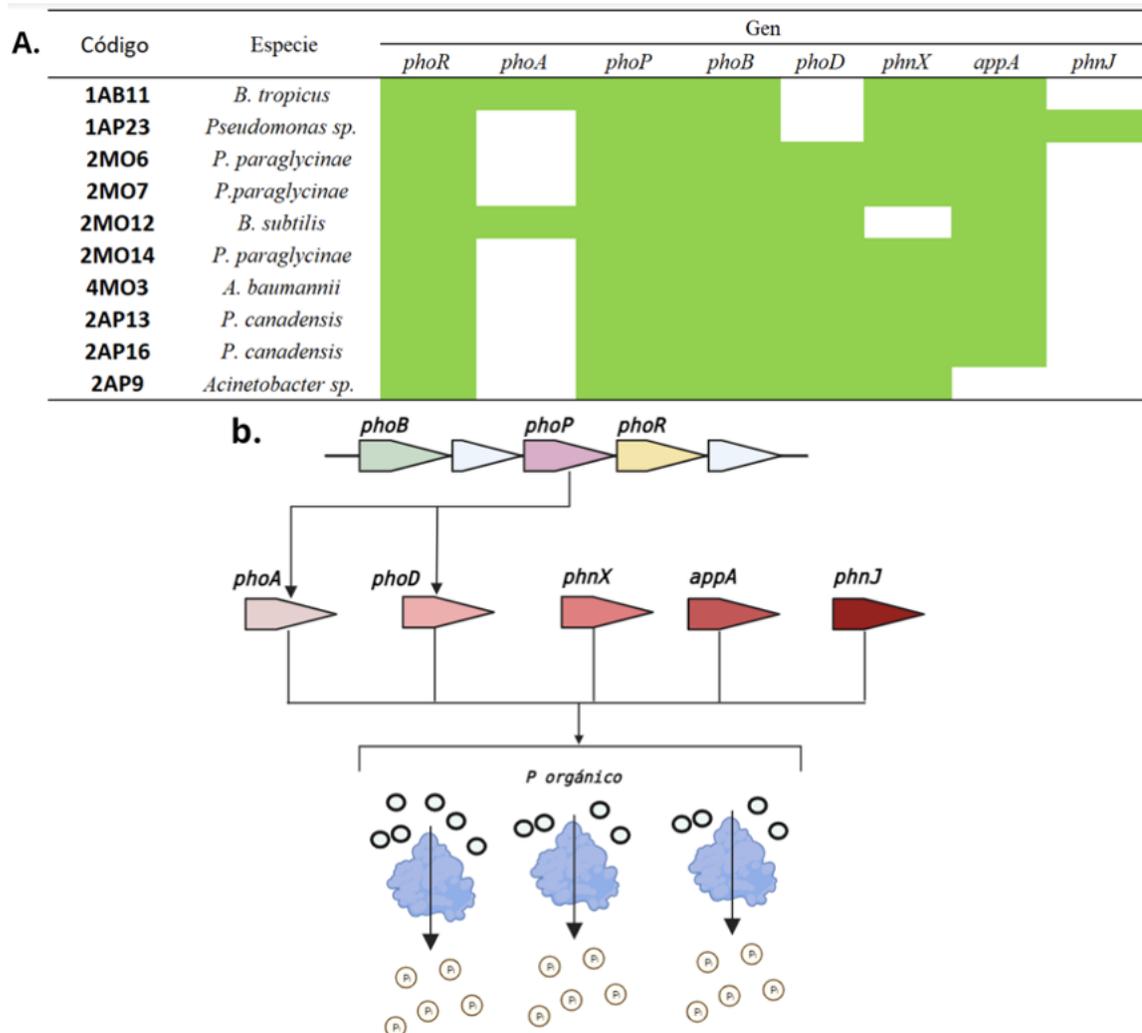
Se realizó de igual manera una anotación manual de los genes encontrados para mejorar la biodisponibilidad de fósforo reportados en la literatura de los cuales se escogieron 8 genes debido a que se ha reportado que codifican a proteínas reguladoras de genes asociados en metabolismo de fósforo y también para la producción de fosfatasa, las cuales son importantes en la solubilización de fosfatos.

En la *figura 18A* se evidencia que todas las especies estudiadas tienen la presencia de los genes *phoR*, *phoP* y *phoB*. Asimismo, se determinó los genes *phnX* y *appA* están presentes en casi todas a excepción de *B. subtilis* y *Acinetobacter sp.* respectivamente. Se logró determinar en menor medida los genes *phoA* y *phnJ* en las especies *B. tropicus*, *B. subtilis* y *Pseudomonas sp.* Así como *phoD* en casi todas las especies a excepción de *B. tropicus*.

La *figura 18B* muestra un diagrama de la ruta para solubilizar fosfatos mediada por los genes determinados en este estudio. Se identificó los genes *PhoR*, *phoP* y *phoB* en todas las cepas estudiadas, los cuales se han caracterizado por regular positivamente operones para solubilización de fosfatos (Liu *et al.*, 2023), particularmente de los genes *phoA* y *phoD* que se han estudiado por codificar fosfatasas (Liang *et al.*, 2020). El gen *phoD* se logró determinar en todas las cepas a excepción de *B. tropicus* 1AB11 y *Pseudomonas sp.* 1AP23, sin embargo, en la cepa 1AB11 sí se identificó el gen *phoA*. De igual manera, se encontró la presencia de otros genes implicados en la solubilización de fósforo, tales como el gen *appA*, en cual solo estuvo ausente en la cepa *Acinetobacter sp.* 2AP9. Asimismo, solamente la cepa *B. subtilis* 2MO12 no presentó el gen *phnX*, mientras que el gen *phnJ* solo se determinó en *Pseudomonas sp.* 1AP23. A pesar de que algunas muestras no tenían la presencia de algunos de estos genes, todas ellas codificaban para al menos un gen regulador y para una fosfatasa.

Figura 18

Determinación de genes para aumentar la biodisponibilidad de fósforo

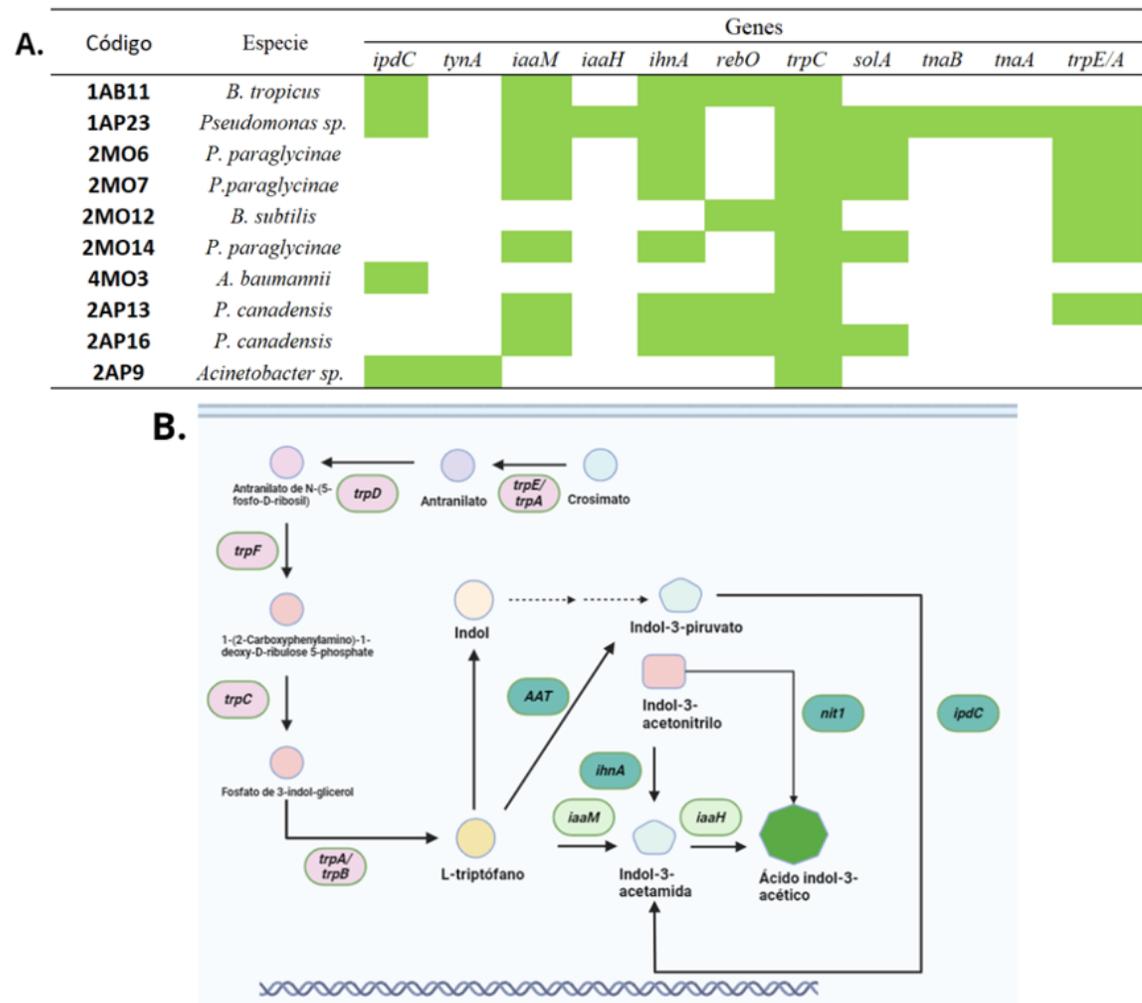


Nota. (A) los cuadros resaltados en verde indican la presencia de los genes implicados en la capacidad de la cepa de fomentar la biodisponibilidad de fósforo. (B) Se muestra un diagrama de la función general de los genes reguladores y los genes que directamente se relacionan a la mineralización.
Elaboración propia

7.6 Identificación de genes asociados a la producción de ácido indol-3-acético (IAA)

Se determinó de igual manera algunos genes implicados en biosíntesis de IAA, fitohormona perteneciente a las auxinas estudiada sobremanera para el crecimiento de plantas.

Figura 19
Genes identificados en la ruta de síntesis de IAA



Nota. (A) Se muestra un cuadro en donde las regiones resaltadas para cada cepa indican la presencia de los genes seleccionados. (B). Muestra un diagrama general de la ruta para sintetizar IAA a través de los genes identificados en las cepas bacterianas.
Elaboración propia

En la figura 19A se muestran los resultados de los genes determinados manualmente de la anotación en base a *prokka*. El ácido indol acético presenta varias rutas de síntesis, de las cuales la gran mayoría parte del L-triptófano. El gen más frecuente determinado fue *trpC*, el cual se encuentra en la ruta de biosíntesis del L-triptófano anotada en *Kegg*. Asimismo, se determinó que varias especies bacterianas tenían la presencia del gen *iaaM*. La especie estudiada que presenta el mayor número de genes determinados en *Pseudomonas sp.*, en la cual solo no fue posible determinar la presencia de los genes *tynA* y *rebO*. Particularmente, se observó que todas las especies de *Pseudomonas* tenían la

presencia del gen *iaaM*, mientras que *Acinetobacter baumannii*, *Acinetobacter sp* y *Bacillus subtilis* poseen distintos genes. Por el otro lado, la figura 19B se muestra una esquematización de la ruta de biosíntesis en base a los genes que fueron determinados en la anotación de *prokka*. Se evidencia que la ruta más directa para la síntesis de IAA a partir del precursor L-triptófano es mediada por *iaaM* e *iaaH*.

7.7 Identificación de genes asociados a la producción de sideróforos

De igual manera, se determinó la producción de algunos sideróforos reportados en la literatura, los cuales son compuestos quelantes que aumentan las capacidades de adquisición de hierro (Ahmed y Holmström, 2014). Se buscó los genes implicados en la producción de los sideróforos enterobactina, bacilibactina y pyochelin (Valdebenito *et al.*, 2006; Miethke *et al.*, 2006; Brandel *et al.*, 2012), así como algunos genes implicados en el transporte y funciones asociadas para el funcionamiento correcto de los mismos.

Cuadro 13

Genes identificados para la producción de sideróforos

especie	Enterobactina						Pyochelin					
	<i>entC</i>	<i>entB</i>	<i>entA</i>	<i>entE</i>	<i>entS</i>	<i>fepA</i>	<i>pfeA</i>	<i>pfeS</i>	<i>fepA</i>	<i>pchA</i>	<i>pchR</i>	<i>pchB</i>
<i>B. tropicus</i>		■			■					■		
<i>Pseudomonas sp.</i>					■	■	■	■	■		■	■
<i>P. paraglycinae</i>					■		■	■			■	
<i>P. paraglycinae</i>					■		■	■			■	
<i>B. subtilis</i>					■				■			
<i>P. paraglycinae</i>					■		■	■			■	
<i>A. baumannii</i>	■	■	■	■	■		■	■			■	
<i>P. canadensis</i>					■							
<i>P. canadensis</i>					■		■	■	■	■	■	■
<i>Acinetobacter sp.</i>					■					■	■	■

especie	Bacilibactina					Transporte y funciones accesorias						
	<i>dhbA</i>	<i>dhbB</i>	<i>dhbC</i>	<i>dhbE</i>	<i>dhbF</i>	<i>yusV</i>	<i>yfhA</i>	<i>yfiZ</i>	<i>sbnA</i>	<i>CirA</i>	<i>yfiY</i>	<i>fiu</i>
<i>B. tropicus</i>	■	■	■	■	■	■	■	■				■
<i>Pseudomonas sp.</i>						■	■					■
<i>P. paraglycinae</i>						■	■		■			■
<i>P. paraglycinae</i>						■	■		■			■
<i>B. subtilis</i>	■	■	■	■	■	■	■	■			■	■
<i>P. paraglycinae</i>						■	■		■			■
<i>A. baumannii</i>	■	■			■	■	■					
<i>P. canadensis</i>				■	■	■	■			■		■
<i>P. canadensis</i>				■	■	■	■					
<i>Acinetobacter sp.</i>									■	■	■	

Nota. Elaboración propia

Los genes asociados para la biosíntesis de enterobactina se encontraron presentes

mayoritariamente en *A. baumannii*, cabe mencionar que la mayoría de las bacterias tienen presente el gen *entS*, y en algunos casos también los genes *entB* y *entS*, sin embargo, no es suficiente para poder realizar correctamente la biosíntesis del compuesto. Por el otro lado, se encontró la *pyochelin* es un sideróforo principalmente caracterizado en especies de *Pseudomonas* (Brandel *et al.*, 2012), como puede observarse en el cuadro, las especies que tienen una mayor cantidad de los genes implicados en su síntesis son la *P. canadensis*, *P. paraglycinae* y *Pseudomonas sp.* Por último, tanto *B. tropicus* y *B. subtilis* tienen presencia de genes implicados en la síntesis de bacilibactina, la cual se caracterizó en el género *Bacillus*.

7.8 Identificación de genes reportados para el control de *Fusarium*

Se realizó una búsqueda corta de algunos genes que han sido estudiados para el control biológicos de *fusarium*, el cual es un hongo fitopatógeno que afecta sobremanera los cultivos de caña de azúcar (Viswakarma *et al.*, 2013). Los genes *prnA*, *prnB*, *prnC* y *prnD* se encuentran implicados en la biosíntesis de la pirrolnitrina (Garbeva *et al.*, 2004). Asimismo, los genes *katA*, *katB* y *katE*, para la producción de la catalasa inducible (Romero-Gutiérrez *et al.*, 2020).

Cuadro 14

Genes identificados para el control potencial de Fusarium

Especie	Genes						
	<i>prnA</i>	<i>prnB</i>	<i>prnD</i>	<i>prnC</i>	<i>katA</i>	<i>katB</i>	<i>katE</i>
<i>B. tropicus</i>					■		■
<i>Pseudomonas sp.</i>						■	■
<i>P. paraglycinae</i>					■	■	■
<i>P.paraglycinae</i>					■	■	■
<i>B. subtilis</i>							■
<i>P. paraglycinae</i>					■	■	■
<i>A. baumannii</i>							■
<i>P. canadensis</i>						■	■
<i>P. canadensis</i>						■	■
<i>Acinetobacter sp.</i>							■

Nota. Elaboración propia

Se puede observar en el cuadro que en ninguna de las especies estudiadas fue posible determinar los genes codificantes para pirrolnitrina, a pesar de que en la caracterización Ochoa, (2022), fue posible identificarlos. Por el otro lado, el gen *katE* se encontró en todas las especies estudiadas. Asimismo, se determinó que todas muestras del género *Pseudomonas* codifican para *katB*. De igual manera, *katA* no fue determinado para todas las especies, solamente para algunos representantes de *Pseudomonas* y las dos bacterias del género *Bacillus*.

7.9 Identificación de factores de virulencia

Se realizó una búsqueda de factores de virulencia para cada género en la *virulence factor Database* VFDB para identificar el potencial de patogenicidad de las bacterias estudiadas.

Cuadro 15
Genes asociados a factores de virulencia

Acinetobacter												
gen	Exotoxinas			Adherencia			Biofilm			modulación inmune		
	plcD	plc1	plc2	pilGMER	fimV	tsaP	adeHGF	pgaABCD	csuAB	lpxABCDM	lpsB	ompA
<i>B. tropicus</i>												
<i>Pseudomonas sp.</i>								C		A/D		
<i>P. paraglycinae</i>										C/D		
<i>P. paraglycinae</i>												
<i>B. subtilis</i>												
<i>P. paraglycinae</i>										C/D		
<i>A. baumannii</i>				E				A/B/C/D		A/D		
<i>P. canadensis</i>										B/C		
<i>P. canadensis</i>										B/C		
<i>Acinetobacter sp.</i>				E				A/B/C/D		A/B		ompA

Pseudomonas										
gen	Exotoxina		Adherencia			Biofilm		modulación inmune		
	plcH	toxA	pilGMER	chpDE	fapAFE	ropS	algABFJWU	mucABCDE	waaACGPF	wbpABM
<i>B. tropicus</i>										
<i>Pseudomonas sp.</i>							B/C A/B/D			
<i>P. paraglycinae</i>		E					A B		A	
<i>P. paraglycinae</i>		E					A B		A	
<i>B. subtilis</i>										
<i>P. paraglycinae</i>							A B		A	
<i>A. baumannii</i>							D		A	
<i>P. canadensis</i>							A/B A/B/D			
<i>P. canadensis</i>							A/B A/B/D			
<i>Acinetobacter sp.</i>							D		A	

Bacillus				
gen	Exotoxina		Adherencia	modulación inmune
	alo	nheABC	bsIA	capABCD
<i>B. tropicus</i>				A/B
<i>Pseudomonas sp.</i>				
<i>P. paraglycinae</i>				
<i>P. paraglycinae</i>				
<i>B. subtilis</i>				A/C
<i>P. paraglycinae</i>				
<i>A. baumannii</i>				A
<i>P. canadensis</i>				D
<i>P. canadensis</i>				D
<i>Acinetobacter sp.</i>				A

Nota. Elaboración propia

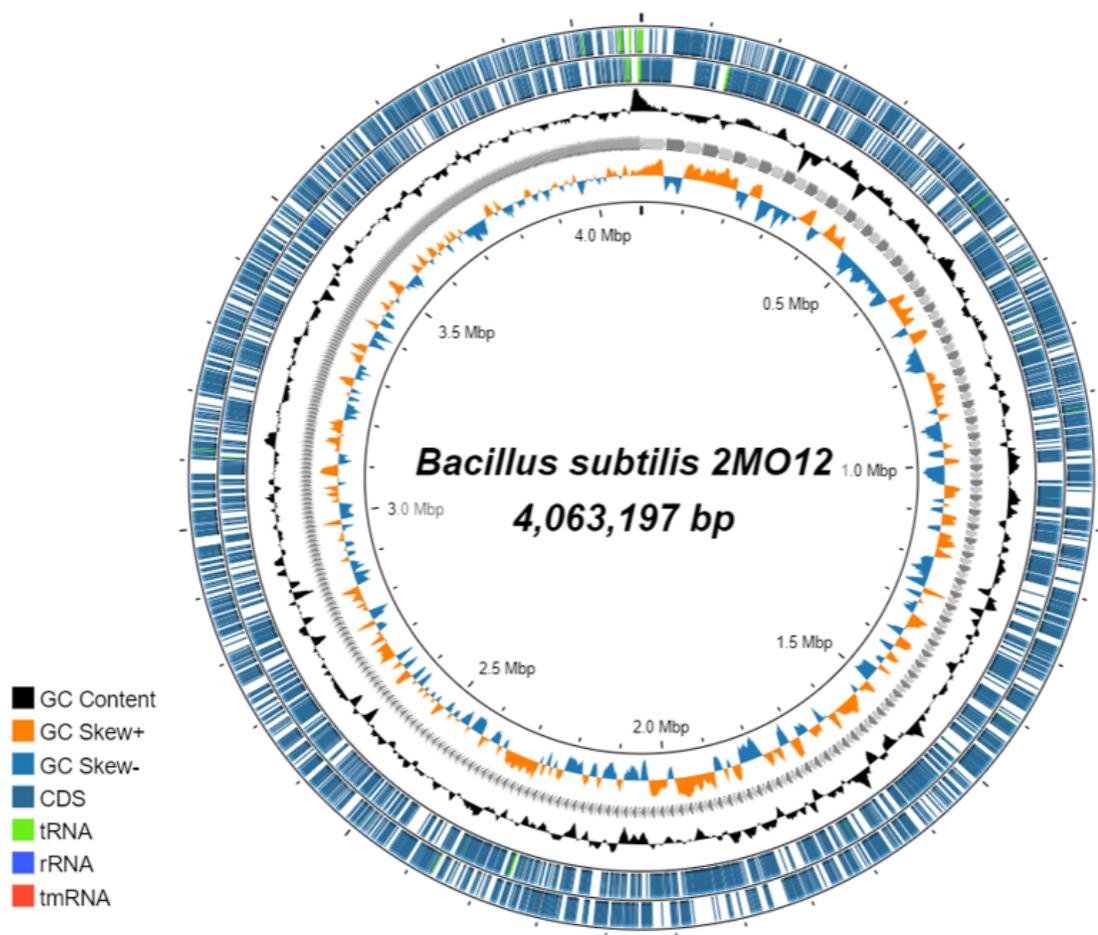
Se realizó una división dependiendo de la función del factor de virulencia, es decir, si este estaba implicado en la producción de alguna endotoxina, si era capaz de producir biofilm para aumentar su supervivencia a distintos factores de estrés, si puede adherirse o su capacidad de resistir o modular la respuesta inmune. No se determinó en ninguna especie genes codificantes de endotoxinas; sin embargo, la mayoría de las especies sí presentaron genes *cap* y *lpx* para llevar a cabo la modulación del sistema inmunológico. De igual manera, todas las bacterias, exceptuando *B. tropicus* y *B. subtilis*, fueron positivas para genes implicados en la formación de biofilm caracterizados para los géneros *Acinetobacter* y *Pseudomonas*. Por último, solamente las dos especies del género *Acinetobacter*, y especies de *P. paraglyciniae* poseían el factor de adherencia *pilE*.

7.10 Sistema de edición génica propuesto para la optimización en la fijación de nitrógeno

El análisis bioinformático evidenció una gran variedad de características asociadas para cada bacteria; es decir, algunas en algunas de ellas se determinó una mayor cantidad de genes asociados a distintas capacidades promotoras, tales como: síntesis de ácido-indol-3-acético, síntesis de sideróforos, síntesis de proteínas para aumentar la biodisponibilidad de fósforo y síntesis de compuestos que pueden utilizarse potencialmente como control biológico. Sin embargo, la especie en la cual se reportó una mayor cantidad de genes asociados a fijación de nitrógeno, y que además ha sido caracterizada como PGPB resultó ser *B. subtilis*, por lo tanto, fue escogida para el desarrollo del sistema de mejoramiento genético para una fijación de nitrógeno aumentada.

Figura 20

Estructura y distribución de genoma de B. subtilis



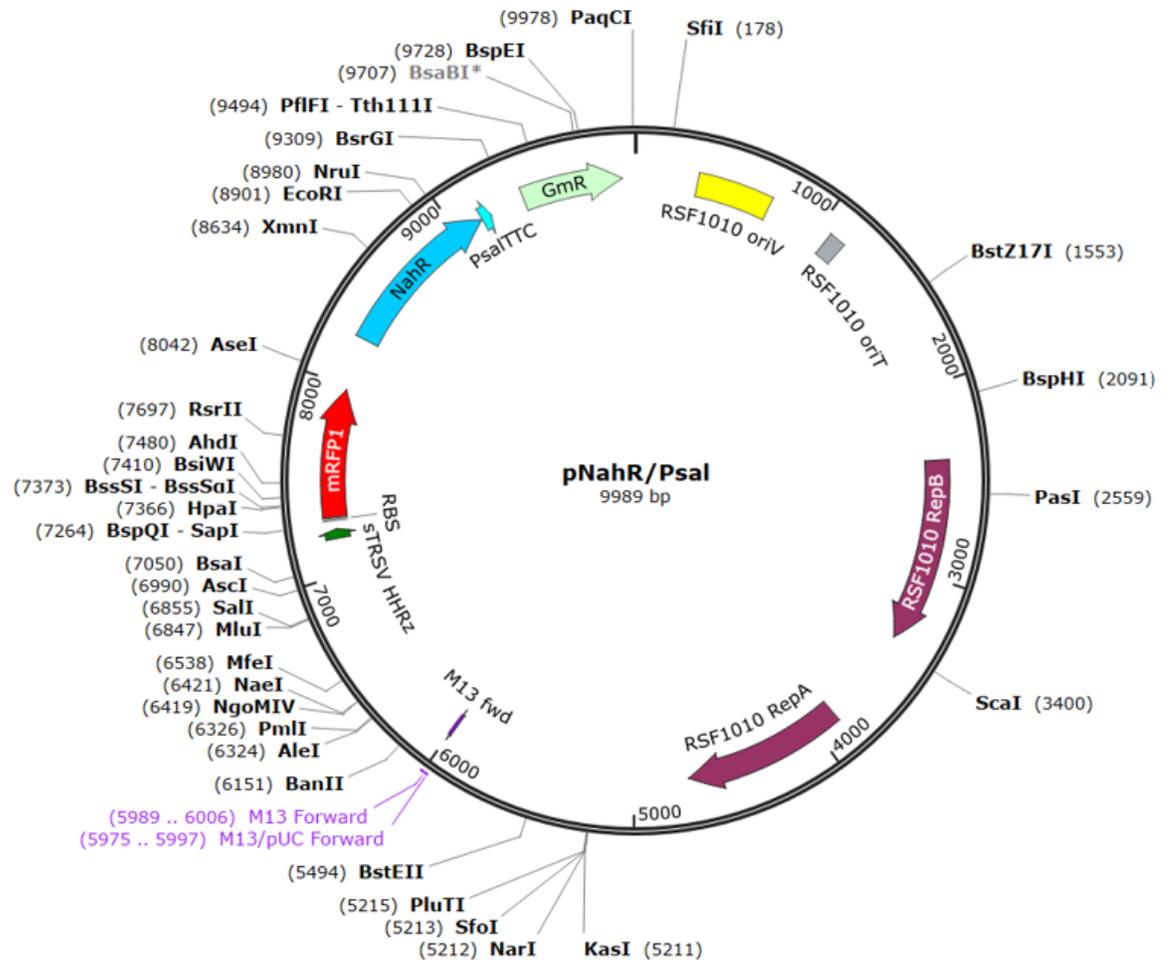
Nota. Elaboración propia

En la figura 20 se presenta la esquematización de la estructura del genoma ensamblado de *B. subtilis*, este está comprendido por un contenido de GC de 43.54% y se determinó una cobertura de casi 90% (ver cuadro 12) en base al genoma de referencia con un tamaño de ensamblaje de 4.06 MB. Se evidencia el contenido de GC en regiones enriquecidas. Asimismo, se muestra la curtosis positiva y negativa GC, la cual indica riqueza de G sobre C, y riqueza de C sobre G respectivamente. En el diagrama de muestra de igual manera las regiones codificantes (CDS), las regiones de ARN de

transferencia (tRNA) y las regiones de ARN ribosomal (rRNA).

Figura 21

Plásmido inducible por ácido salicílico



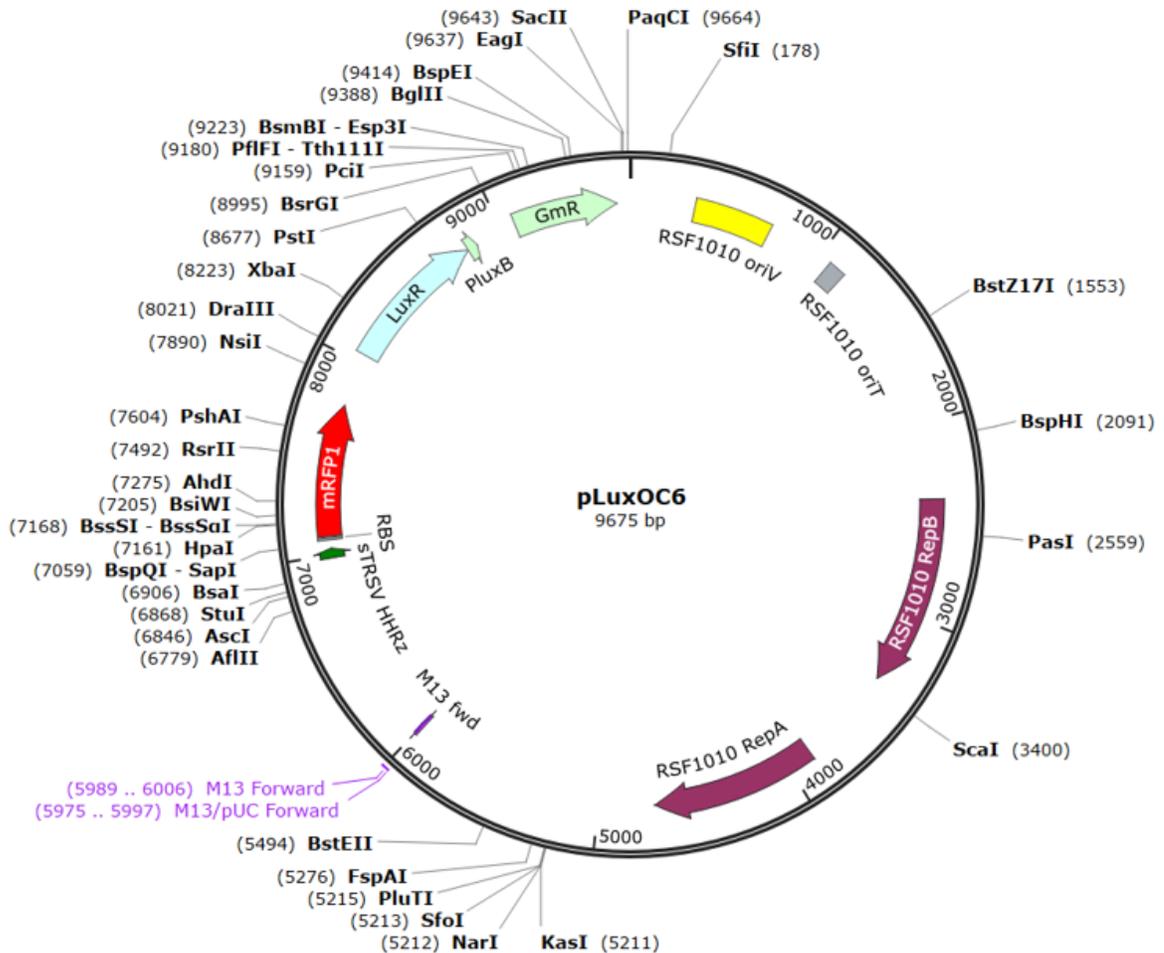
Nota. Elaboración propia

Para mejorar la fijación de nitrógeno utilizando el sistema CRISPR, se propone realizarla una edición génica en dos pasos (ver *figura 31*). Primero, se realizará una clonación del gen nativo *nifA* para la regulación positiva del operón *nif* en tres plásmidos inducibles. Asimismo, se realizará una eliminación del gen de regulación negativa *nifL* a través de un KO por el sistema CRISPR/Cas9. Se llevará a cabo una combinación de técnicas para evaluar cual de estas propuestas presenta un mejor resultado. En la *figura 21* se muestra el constructo del plásmido propuesto pNahR/Psal, en esta se cuenta con una secuencia reguladora NahR y un promotor PsalTCC, el cual es inducible por ácido salicílico. Asimismo, la estructura del plásmido tiene un gen de resistencia a gentamicina, un gen reportero fluorescente *monomeric red fluorescent protein 1* (*mRFP1*), un origen de replicación RSF oriV para la iniciación de la replicación y una región oriT para la transferencia. El plásmido también cuenta con la región codificante de *repA* y *repB*, las cuales se unen al sitio de replicación oriV y regulan el número de copias para una mejor permanencia en las bacterias respectivamente. Por último, el plásmido presenta los sitios para los cebadores *M13* y *M13/pUC* para la clonación de genes. El constructo utilizado fue desarrollado por *Schuster y Reish. (2021)*. La secuencia del regulador NahR fue obtenida de NCBI con el código de acceso NC_016603.1, mientras que la

secuencia del promotor fue obtenida de *Meyer et al.* (2019).

Figura 22

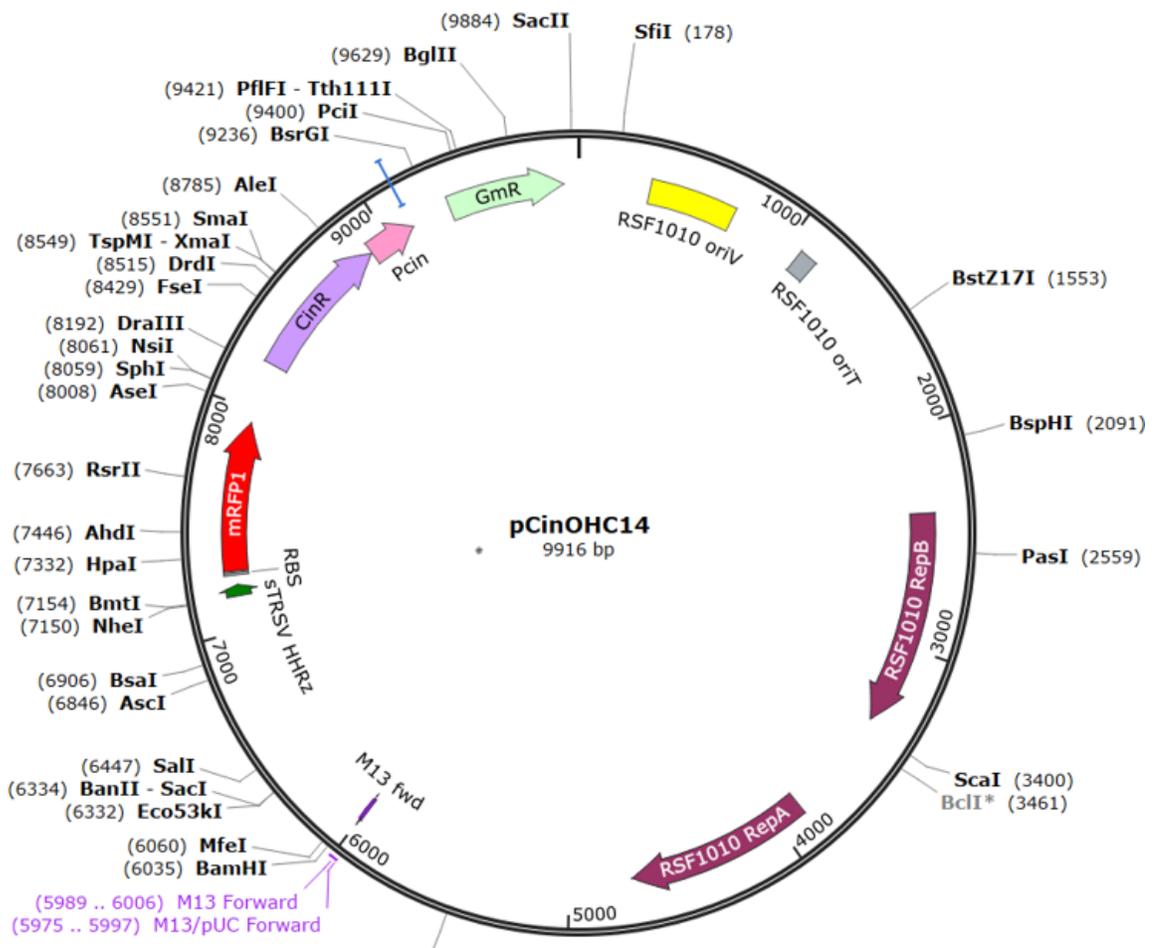
Plásmido inducible por OC6



Nota. Elaboración propia

De igual manera que el plásmido inducible previamente descrito, la estructura general de *PluxOC6* varía solamente en la región con la secuencia reguladora y con el promotor seleccionado. En este caso, se encuentra el regulador *LuxR* y el promotor *PluxB*, los cuales le confieren al constructo la capacidad de inducirse a través de la N-(3-oxohexanoil) homoserina lactona. La secuencia reguladora fue obtenida de NCBI con el código de acceso NC_011186.1, asimismo, la secuencia del promotor *PluxB* fue obtenida de *Meyer et al.* (2019).

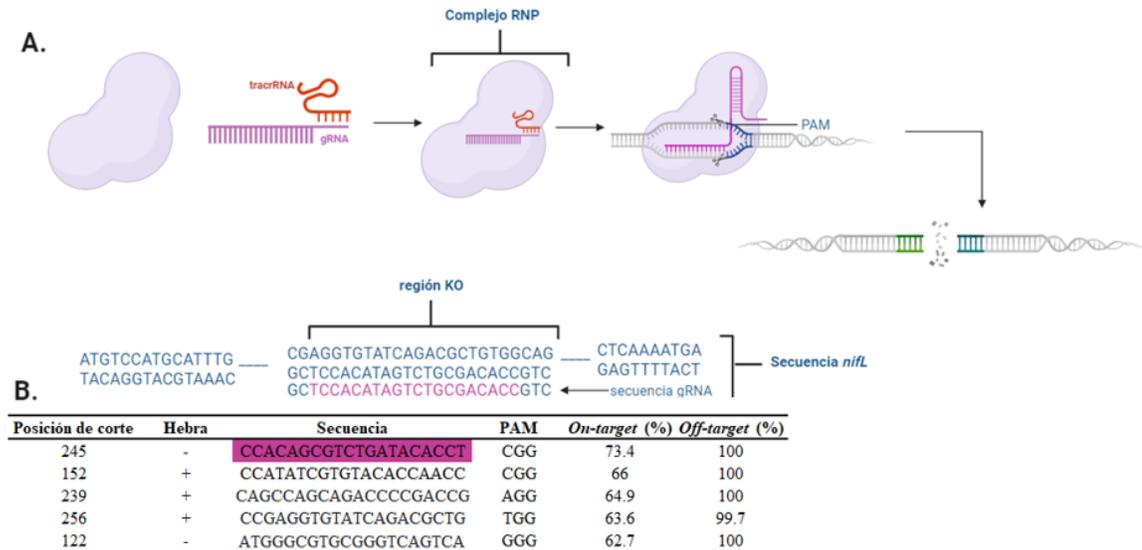
Figura 23

Plásmido inducible por *OCH14*

Nota. Elaboración propia

El presente constructo se caracteriza por tener el regulador *cinR* y el promotor *Pcin*. Al igual que los plásmidos mencionados con anterioridad, el constructo general es el mismo, debido a que de esta manera el protocolo de clonación y transformación sería similar. El plásmido propuesto *PcinOHC14* es inducible por el compuesto hidroxitetradecanoil-homoserina lactona. La secuencia del regulador fue obtenida de NCBI con el código de acceso NC_021505.1 y la secuencia del promotor fue obtenida de Meyer *et al.* (2019).

Figura 24
Sistema de edición KO-CRISPR para el gen *nifL*



Nota. Elaboración propia

Por último, se presenta un esquema del sistema propuesto para realizar una eliminación (KO) del gen regulador *nifL* para la fijación de nitrógeno. Para esto, se diseñó un ARN guía (gRNA) a través de la herramienta *benchling CRISPR design*. En la *figura 24A* se muestra el diagrama en donde se propone realizar un complejo ribonucleoproteína (RNP) conjugando la endonucleasa Cas9, el gRNA y el tracrRNA asociado a la cas9. Al formar el complejo RNP, es posible electroporar la bacteria para realizar el KO. Asimismo, en la *figura 24B* se muestra un cuadro de los mejores cinco gRNA diseñados, en donde se muestra la posición de corte (donde idealmente es entre 100-250 nucleótidos de la secuencia), el sentido del gRNA, la secuencia PAM necesaria; el puntaje *on-target*, que representa la eficiencia de corte de la Cas9; por último, el puntaje *off-target* representa la probabilidad inversa de ediciones inespecíficas por la endonucleasa Cas9, el ARN guía con un mejor puntaje se muestra resaltado en morado. Por lo tanto, en ambos puntajes un valor mayor representa una mejor calidad de diseño del gRNA.

Discusión

Las prácticas de agricultura convencional impactan fuertemente la salud del suelo, principalmente por el uso de agroquímicos para proveer al cultivo de nutrientes y protegerlo de agentes patógenos afectando consecuentemente la diversidad de microorganismos del suelo (Meena *et al.*, 2020). Estos microorganismos son beneficiosos para las plantas debido a que llevan a cabo procesos metabólicos que modulan la biodisponibilidad de nutrientes, por lo tanto, una disminución en su diversidad influye en el crecimiento saludable de las plantas (Ahemad y Kibret, 2014). Lo anterior, ha llevado al desarrollo de inóculos a base de microorganismos para mitigar los daños y reconstituir la salud del suelo, con el objetivo de desarrollar sistemas más sostenibles para mantener los cultivos agrícolas. Para esto, es importante entender la dinámica del suelo del cultivo particular, debido a que un microbioma saludable para una especie de planta no necesariamente es igual al de otra completamente diferente (Banerjee y Van der Heijden, 2023). Asimismo, es necesario estudiar las capacidades promotoras de crecimiento tales como; la fijación de nitrógeno, producción de ácido-indol-3-acético, sideróforos, y el aumento de biodisponibilidad de fósforo llevado a cabo por los microorganismos para utilizar técnicas moleculares para aumentar estos procesos y que sean más efectivos. El propósito de este estudio es llevar a cabo una caracterización de las capacidades promotoras de crecimiento vegetal de bacterias aisladas de suelo de *Saccharum officinarum* a través de un análisis de su genoma completo, asimismo, desarrollar una propuesta de edición génica para aumentar el proceso de fijación de nitrógeno. Este estudio apoyara a que en un futuro se utilicen inóculos a base de microorganismos en el cultivo de la caña para mejorar la calidad del suelo y disminuir la cantidad de fertilizante aplicado en los suelos.

Este estudio partió del trabajo de Ochoa, (2022) en el que se aislaron y evaluaron características promotoras del crecimiento de 222 bacterias. El primer paso de este estudio fue la identificación de las 222 bacterias a través de una secuenciación del gen de ARN ribosomal *16S*, en la *figura 15* se muestra que fue posible identificar hasta 22 géneros, dentro de los cuales los más frecuentes fueron *Bacillus*, *Pseudomonas* y en menor medida *Acinetobacter*. Diversos estudios han obtenido resultados similares en distintas regiones de suelo en diferentes cultivos. El estudio de Beneduzi *et al.*, (2013) determinó en suelo de cultivo de caña de azúcar en Brasil una diversidad similar, donde se identificó varias especies de *Acinetobacter*, *Bacillus*, *Pseudomonas*, *Azorhizobium* y entre otros géneros. De igual manera, Bonato *et al.*, (2021) evidenció que en suelos de cultivo de caña en el trópico había una gran incidencia de los mismos géneros más presentes determinados en

la identificación por 16S.

Las capacidades promotoras de crecimiento de *Bacillus*, *Pseudomonas*, *Klebsiella*, *Acinetobacter* y entre otros géneros identificados han sido estudiados extensivamente. Estos han presentado la capacidad de fijar nitrógeno, la capacidad de sintetizar distintas fitohormonas, producir enzimas para la degradación o conversión de compuestos para aumentar su biodisponibilidad e incluso potencialmente controlar patógenos (Sansinenea, 2019; Santoyo *et al.*, 2012; Gulati *et al.*, 2009). Esto se ve evidenciado incluso en el *cuadro 6*, en este se muestran 12 bacterias escogidas basadas en sus capacidades promotoras de crecimiento, todas ellas son capaces de fijar nitrógeno, producir ACC deaminasa, ácido indol-3-acético, pirrolnitrina, proteasas, sideróforos, solubilizar fosfatos y solubilizar potasio. Por lo tanto, presentan un gran potencial para utilizarse en la formulación de bioinóculos. Para la selección de las bacterias candidatas para llevar a cabo WGS, también se realizó una PCR del gen *nifH* (ver *figura 16*) que codifica a una proteína principal del complejo nitrogenasa convencional dependiente de molibdeno. Como puede observarse, de las 12 bacterias, solamente 4 de ellas no fueron positivas para el gen, y a pesar de esto las pruebas de fijación de nitrógeno fueron positivas. Esto sugiere que podrían codificar a genes de la nitrogenasa dependiente de vanadio regulada por el operón *vnf*, o la nitrogenasa dependiente de hierro regulada por el operón *anf* (Mus *et al.*, 2018), asimismo, es posible que por la variabilidad de estos genes tampoco haya podido determinarse porque falta de conservación (Coelho *et al.*, 2009).

Se llevó a cabo una secuenciación de genoma completo de las bacterias candidatas seleccionadas (ver *cuadro 6*) a través de tecnología de secuenciación de nueva generación (NGS). Al analizar los datos de secuenciación se observó que dos de las doce muestras secuenciadas presentaron contaminación, las bacterias 1AB21 y 2AP12, lo que pudo haberse debido a la presencia de otra cepa que morfológicamente es muy parecida, para evitar esto es posible realizar un estriado de la colonia tomada para la extracción de ADN para poder ratificar que en efecto es una cepa aislada. Por lo tanto, fueron descartadas y se prosiguió con el análisis bioinformático de las lecturas obtenidas. Primero se analizó las secuencias crudas a través del programa *FastQC*, el cual permite ver estadísticas base para determinar si los resultados obtenidos son apropiados para un análisis superior, como puede ser el ensamblaje y anotación de genomas (Andrews, 2010). Se determinó que todas las muestras tenían resultados aceptables, como se muestra en el *cuadro 7*, los parámetros principales como la calidad de secuencia base, la calidad de secuencia *per-tile* y la puntuación de calidad por secuencia fue aceptable en todas las secuencias a excepción de la lectura R2 de la primera muestra. Esto indica que las lecturas obtenidas se encuentran en o por encima de una profundidad de 30X, lo que sugiere que los resultados son fiables (Rieber *et al.*, 2013). Sin embargo, algunos parámetros mostraron una alerta, el resultado de análisis del contenido de GC y el contenido de base por secuencia se pudieron deber a que se ha reportado que existe una tendencia de error en secuenciación en genomas que tienen regiones con una baja proporción de GCs, lo cual también puede estar asociado a la preparación de bibliotecas (Van Dijk *et al.*, 2014). El estudio de Sato *et al.*, (2019) sugiere que es importante tomar en cuenta el porcentaje de GCs para la selección de la tecnología de secuenciación y también para la selección del kit de preparación de bibliotecas, particularmente determinaron que el kit *DNA Flex Library Prep Kit* de *Nextera* no evidenció una introducción de este error por el porcentaje de GCs.

Es importante mencionar que la alerta en el parámetro de secuencias sobrerrepresentadas puede haber impactado en la alerta del parámetro de contenido GC por secuencia, asimismo, la sobrerrepresentación puede llevar a problemas en el ensamblaje, particularmente a errores de ensamblaje (“Misassemblies”), una reducción en la exactitud y también un aumento en la demanda computacional para procesarlo (Schwartz *et al.*, 2010). Por lo tanto, se realizó un procesamiento o “corte” de las secuencias a través de *Trimmomatic*. En el *cuadro 8* se presentan nuevamente los parámetros de calidad de las secuencias, en este caso, posterior al procesamiento. La mayoría de los parámetros mejoraron, en este caso, solo la lectura R2 de la muestra 2AP9 permaneció en alerta en el parámetro de secuencias sobrerrepresentadas. Cabe mencionar que se ha determinado que generalmente las lecturas R2 suelen tener menor calidad debido a errores en la amplificación y por la desaminación de las bases (Schirmer *et al.*, 2015), lo que explicaría por qué la mayoría de las secuencias donde hay alertas, o no cumplieron con los requisitos del parámetro son R2. Por último,

cabe mencionar que, a pesar del procesamiento, el parámetro de contenido de base por secuencia no fue cumplido, la documentación del programa reporta que errores asociados a este parámetro se deben comúnmente a la preparación de bibliotecas, algunos kits de tagmentación están propensos a generar este error debido a que existe una tendencia en los primeros 12 bp de la corrida (Brahamam Bioinformatics-FastQC Project; s.f.). Asimismo, las secuencias sobrerrepresentadas también pueden influir en que este parámetro no pueda ser cumplido.

Debido a que existen diversos algoritmos para generar el ensamblaje y no existe un consenso del mejor algoritmo para el ensamblaje de genomas bacterianos, se llevó a cabo tres distintos ensamblajes con algoritmos de ensamblaje basados en el principio de grafos De Bruijn: ABySS, *velvet* y *SPAdes*. Los parámetros de calidad de los ensamblajes se analizaron y se eligió el que generara mejores resultados para la identificación del mejor genoma de referencia y consecuentemente la anotación del genoma. En los cuadros 9, 10 y 11 se muestran los parámetros de calidad obtenidos para los ensamblajes realizados con ABySS, *SPAdes* y *velvet*; La mayoría de los ensamblajes con este algoritmo se encontraban por encima de los 100 *scaffolds* y por lo general, el número de estos fue más aceptable evidentemente a través de los algoritmos *Velvet* y *SPAdes*, los cuales generaron ensamblajes con un menor número de *Scaffolds*. Esto se puede deber a que estos dos algoritmos realizan el cálculo del *K*mero óptimo para su ensamblaje; *SPAdes* lleva a cabo una realización de grafos incorporando las lecturas pareadas a través de la estructuración de *K*-bimeros (Bankevich *et al.*, 2012). Por otro lado, para llevar a cabo el ensamblaje con ABySS se utiliza el programa *kmergenie*, el cual implementa histogramas de diferentes *kmeros* para determinar la longitud que maximiza la lectura en el set de datos, sin embargo, puede generar errores de predicción dependiendo del resultado del histograma (Behera *et al.*, 2018).

Al realizar la comparación general de los tres ensamblajes (ver cuadros 9, 10 y 11), se puede observar que *SPAdes* genera con más frecuencia mejores resultados, debido a que la longitud del ensamblaje es mayor, en la mayoría de las muestras se obtiene un menor número de *Scaffolds* y un menor número de ellos para completar el 50% del total (L50). Al observar los cuadros 9, 10 y 11, particularmente en la muestra 2AP9, se puede observar este suceso. En *SPAdes* se obtuvo un número de *scaffolds* de 19, mientras que para *Velvet* y ABySS se obtuvo 27 y 35 respectivamente; adicionalmente, el valor L50 es solamente de dos, y L90 es 8; por lo que solamente se necesitan 8 *Scaffolds* para generar el 90% de todo el ensamblaje. Cabe mencionar que los valores de porcentaje de GC en todos los ensamblajes no fue variable. Antes de proseguir con el análisis bioinformático, se analizó individualmente la calidad de los ensamblajes para cada muestra; a pesar de que en algunos casos ABySS y *Velvet* tenían un menor número de *Scaffolds* que *SPAdes*, se escogió este último ensamblaje para todas las muestras a excepción de la muestra 1AP23, debido a que en todos estos casos el valor N50 era menor, por lo que, a pesar de tener más *scaffolds*, algunos de ellos son de mayor longitud, por lo que en definitiva estaba menos fragmentado.

Se prosiguió con la determinación del genoma de referencia y la anotación en *Prokka* como se encuentra descrito en la metodología. En el cuadro 12 se muestran los resultados del análisis por *GTDB-Tk*. En el cual cabe destacar además de la clasificación final, los valores de similitud calculados por *FastANI* y la cobertura del ensamblaje realizado en base al genoma de referencia obtenido con *Quast*. Se determinó una especie de *Bacillus tropicus*, tres representantes de la especie *Pseudomonas paraglycinae*, dos representantes de *Pseudomonas canadensis*, una especie de *Bacillus subtilis*, una especie de *Acinetobacter baumannii* y dos muestras que no fue posible determinarlas a nivel de especie (*Pseudomonas sp.* y *Acinetobacter sp.*) lo que puede deberse a que no existe información suficiente en la base de datos de referencia.

B. tropicus ha comenzado a ser estudiada como una bacteria promotora del crecimiento vegetal. Wang *et al.*, (2020) aisló y caracterizó bacterias endofíticas del género *Bacillus* de la raíz de la caña de azúcar, en este estudio se determinó que esta especie era capaz de producir sideróforos y de combatir el tizón de la vaina de arroz causada por el hongo patógeno *Rhizoctonia solani*. No obstante, en este estudio *B. tropicus* no se caracterizó por producir fitohormonas o por solubilizar fósforo, lo que puede deberse al nicho ecológico de esta cepa. Cabe mencionar que en otro estudio

llevado a cabo por Efe, (2020), aislaron bacterias de suelo, *B. tropicus* fue reportada por no promover el crecimiento al analizar: producción de IAA, sideróforos, fijación de nitrógeno y solubilización de fósforo. Esto difiere de los resultados obtenidos en este estudio, debido a que se determinó que *B. tropicus* aislada de suelo de cultivo de caña de azúcar en Guatemala, sí presenta varias características para promover el crecimiento (ver *cuadro 6*), esto se puede deber principalmente a la región en donde fueron aisladas, además, en el estudio de Efe, (2020), las bacterias fueron aisladas de suelos ricos en metales, por lo que es posible que la presión selectiva no haya permitido a las *B. tropicus* que tenían estas características proliferar.

Ahora bien, las especies *P. paraglycinae* y *P. canadensis* no han sido tan estudiadas como PGBP. Esto se debe a que existen pocos estudios de caracterización para ambas bacterias, no obstante, varios autores han descrito distintas especies del género *Pseudomonas* como bacterias con potencial promotor del crecimiento vegetal, debido a que muchas presentan la capacidad de fijar nitrógeno, de producir IAA, sideróforos, proteasas, pectinas, quitinasas y en algunos casos combatir a hongos patógenos (Shin *et al.*, 2007; Pandey y Gupta, 2020). Como puede observarse nuevamente en el *cuadro 7.2.1*, las especies aisladas *P. paraglycinae* y *P. canadensis* sí presentan potencial de ser utilizadas para agricultura sostenible, además, son de un particular interés porque a pesar de que se caracterizó como fijadora de nitrógeno, no fue posible identificar el gen *nifH* (ver *figura 16*), por lo que podrían codificar a operones menos convencionales o en su defecto, presentar genes que aún no han sido descritos.

En el caso del género *Acinetobacter*, a pesar de ser estudiado principalmente por sus características patogénicas, las especies de este género aisladas de suelo poseen varias capacidades para promover el crecimiento de plantas (Qadir *et al.*, 2021; Rana *et al.*, 2020). En un estudio llevado a cabo por Singh *et al.*, (2020). se caracterizó particularmente a *Acinetobacter baumannii* aislada de suelo como una PGPB, en este estudio se determinó que tenía la capacidad de solubilizar potasio, solubilizar zinc, producir IAA y también sideróforos. Estos resultados resaltan los hallazgos de este estudio, debido a que, de igual manera, *A. baumannii* 4MO3 presentó una gran capacidad promotora, ya que tenía la capacidad de fijar nitrógeno, solubilizar varios compuestos y producir IAA.

B. subtilis ha sido estudiado más ampliamente por sus características promotoras de crecimiento de plantas, particularmente se caracteriza por ser una rizobacteria con un gran potencial. La revisión realizada por Sivasakthi *et al.*, (2014). mencionan varias características de *B. subtilis* tales como fijación de nitrógeno, síntesis de IAA, producción de sideróforos, solubilización de fosfatos e incluso presenta potencial de utilizarse como un control de patógenos como algunas especies de *Aspergillus*, *Rhizoctonia* y *Fusarium*. La mayoría de las características mencionadas anteriormente se mencionan en el estudio llevado por Hashem *et al.*, (2019). en donde también se menciona la capacidad *B. subtilis* para aumentar la tolerancia a estrés. En el *cuadro 6* se muestra las características de la bacteria *B. subtilis* 2MO12 caracterizada en este estudio, es capaz de fijar nitrógeno, solubilizar fósforo y potasio, producir fitohormonas y sideróforos. Por lo tanto, es de gran interés, particularmente porque existen más estudios que ratifican su uso como PGPRB.

La dinámica general de los genes identificados y analizados se muestra en la *figura 18B*. Esto se ve complementado en los resultados de la *figura 18A*, donde se evidencia que todas tienen los genes *phoR*, *phoP* y *phoB*, los cuales se han estudiado como genes promotores o reguladores para el procesamiento de fosfatos (Liu *et al.*, 2023). Estos genes se regulan al alza como respuesta a bajas condiciones de fósforo para proveer suministros a las bacterias (Hsieh y Wanner, 2010; Liu *et al.*, 2023; Devine, 2018). Asimismo, se identificaron los genes *phoA* y *phoD* los cuales son esenciales para la producción de fosfatasa alcalina para la mineralización de fosfatos (Liang *et al.*, 2020; Fraser *et al.*, 2015). Los genes determinados *phnX* y *phnJ* son importantes para la producción de fosfonoacetaldehido hidrolasa y una C-P liasa respectivamente que participa en la conversión de fosfatos orgánico a inorgánico (Pi) (Kamat y Raushel, 2015; Liang *et al.*, 2020). Por último, el gen *appA* codifica para una fitasa, la cual de igual manera ayuda en la producción de Pi, es de particular importancia ya que aproximadamente el 60% del fósforo inorgánico se encuentra en la

forma de fitato, el cual no está directamente disponible para las plantas (Wang *et al.*, 2013). La determinación de genes reportados para el aumento de la biodisponibilidad de fósforo demostró que las bacterias estudiadas tienen un gran potencial para la solubilización de este nutriente, lo que es un resultado esperado debido a que las cepas analizadas presentaron la capacidad de llevar a cabo estos procesos a través de la determinación de solubilización de fosfatos llevado a cabo por Ochoa, (2022).

De igual manera, se realizó un listado de genes importante para la síntesis de IAA, como se puede ver en la *figura 19A*, la mayoría de las especies cuentan con genes implicados para la producción de esta fitohormona, sin embargo, la bacteria 1AP23 *Pseudomonas sp.* presentó una mayor incidencia de estos, lo cual es esperado debido a que varios estudios han caracterizado distintas especies del género *Pseudomonas* por la producción de esta fitohormona (Tabatabaei *et al.*, 2016; Bharucha *et al.*, 2013). Como se puede observar en la *figura 19B*, los genes *iaaM* e *iaaH* que codifican a la triptófano-2-monooxigenasa y a la indol-acetamida hidrolasa (Vega-Celedón, 2016), representan la ruta más directa para la producción de IAA a partir del L-triptófano. Existen otras rutas directas que han sido caracterizadas para la ruta de biosíntesis de forma directa; particularmente el metabolismo de dos pasos mediado por el L-triptófano oxigenasa y la indol-3-acetaldehído oxidasa (Duca *et al.*, 2020). A pesar de que la mayoría de las especies tienen presente el gen *iaaM* solamente *Pseudomonas sp.* presentó el gen *iaaH*, no obstante, se hipotetiza que solo no fue posible realizar la identificación con *prokka* por la fragmentación del genoma o por la variación que pueda existir en el gen para cada especie.

Se identificaron varios genes implicados en la biosíntesis de distintos sideróforos los cuales se encuentran resumidos en el *cuadro 13*. *Acinetobacter baumannii* presentó los genes codificantes de enterobactina, el cual es un sideróforo encontrado principalmente en *Escherichia* y *salmonella* (Valdebenito *et al.*, 2006) pero se ha caracterizado en algunas cepas de *A. baumannii* (Penwell *et al.*, 2012). La síntesis se lleva a cabo por los genes *entA*, *entB* y *entC*, asimismo, el gen *fepA* es de particular importancia porque codifica a una proteína de membrana para complejos de enterobactina férrica (Crossa & Walsh, 2002; Cao *et al.*, 2000). Cabe mencionar que el operón consta de 14 genes estudiados, de los cuales no fue posible determinar en el genoma la mayoría de los genes *fep* (como *fepE/C/G*) y *fes* (Crossa y Walsh, 2002), probablemente debido a la fragmentación del genoma ensamblado. Por el otro lado, se determinó que *B. tropicus* y *B. subtilis* tenían la presencia de varios genes implicados en la síntesis de la bacilibactina. El estudio de Miethke *et al.*, (2006). evidencia la capacidad de producción de este sideróforo por *B. subtilis*, el cual es capaz de producir este agente quelante para capturar hierro. La biosíntesis de la bacilibactina está regulada por el cluster de genes *dhnA*, *dhbC*, *dhbE*, *dhbB* y *dhbF* (Hertlein *et al.*, 2014), de los cuales todas fueron determinadas para las dos especies del género *Bacillus* en este estudio, lo cual es esperado debido a que varias especies como *B. subtilis*, *B. cereus* y *B. velezensis* (Hertlein *et al.*, 2014; Segond *et al.*, 2014; Xie *et al.*, 2023). Las especies del género *Pseudomonas* en el presente estudio no presentaron todos los genes para ninguna de estos dos sideróforos, sin embargo, sí fue posible identificar algunos genes importantes en la producción del sideróforo *pyochelin*, el cual ha sido reconocido como un sideróforo determinado comúnmente en *Pseudomonas aeruginosa* (Brandel *et al.*, 2012). El estudio de Youard *et al.*, (2010). menciona que el cluster de genes necesario para llevar a cabo la biosíntesis de *pyochelin* depende del cluster comprendido siguientes genes: *pchA*, *pchB*, *pchC*, *pchR*, *pchE* para *P. aeruginosa* de estos, solamente la especie *P. canadensis* presentó la mayoría de estos; específicamente *pchA/R/B*.

En el *cuadro 14* se muestran los resultados de algunos genes que potencialmente confieren la capacidad combatir organismos fitopatógenos, particularmente del género *Fusarium*. Los genes *prnA*, *prnB*, *prnC* y *prnD* son necesarios para llevar a cabo la síntesis de la pirrolnitrina (Garbeva *et al.*, 2004), sin embargo, a pesar de que, en el estudio realizado por Ochoa, (2022). se reportó que fueron capaces de producirlo, los genes no están presentes. Lo que se puede deber a la fragmentación en el ensamblaje del genoma o porque existe una gran variabilidad en el gen. Moghabieb *et al.*, (2021) estudió como la sobreexpresión de *katE* por bacterias mejoraba la resistencia de plantaciones de tomate a *Fusarium oxysporum*. En este estudio fue posible identificar la presencia de los genes *kat*, en específico *katA*, *katB* y *katE*, Los genes *katA* y *katB* codifican para catalasas que

potencialmente pueden utilizarse para el control biológico y para respuestas en contra al estrés oxidativo (Romero-Gutiérrez *et al.*, 2020), no obstante, no existen tantos estudios para los genes *katA* y *katB*. Contrariamente, el gen *katE* se ha estudiado extensivamente para el control biológico llevado a cabo por bacterias.

Se llevó a cabo una identificación de factores de virulencia basados en los genes reportados en la *virulence factors Data Base* VFDB para cada género (Chen *et al.*, 2005). Como puede observarse en el *cuadro 15*, ninguna de las bacterias tiene presentes genes que codifiquen a la producción de exotoxinas, que es el factor al que se le proporcionó una mayor prioridad, debido a su implicación en el desarrollo de enfermedades y daño al hospedero. Por el otro lado, sí fue posible identificar otros factores de virulencia, principalmente para la formación de biofilm (*pgaA/B/C/D*, *alpA/B/C* y *mucA/B/D*) y para modular el sistema inmune humano en caso de una infección (*lpxA/B/C/D*, *waaA*, *capa/B/C/D*). Es relevante mencionar que a pesar de que estos genes codifican a factores de virulencia reportados, no necesariamente implica que serán altamente patogénicas (Malovichko *et al.*, 2019). En el caso de la formación de biofilm, existen varios estudios que reportan que la síntesis de esta capa de polisacáridos es beneficiosa para las PGBP, debido a que permite que las bacterias colonicen las regiones de las raíces y formen una barrera protectora para tener una resistencia ante condiciones adversas que puedan afectar la abundancia de nuestras bacterias de interés (Ajjjah *et al.*, 2023; Liu *et al.*, 2011).

En la *figura 17A* se muestran los resultados de la identificación de genes asociados a la fijación de nitrógeno. En este, se muestra un resumen de la identificación basado en el complejo nitrogenasa con el operón más incidente, el cual está compuesto por los genes *nif*. Se realizó una búsqueda de todos los genes implicados en la formación de las tres distintas nitrogenasas (operón *nif*-Molibdeno, *vnf*-vanadio y *anf*-hierro), no obstante, no fue posible identificar los genes *anf*, asimismo, solo se identificó el gen *vnfA* para el complejo nitrogenasa dependiente de vanadio, lo que podría deberse a que no se pudo identificar los genes con *Prokka* debido a la fragmentación o posible variabilidad de secuencias. La identificación con la anotación del genoma solo permitió identificar los genes *nifS*, *nifU* y *nifA* del cluster necesario para el funcionamiento de la nitrogenasa-molibdeno. Sin embargo, como puede verse en la *figura 17B*, es posible inferir la presencia del gen por la posición del gen *nifL* en la estructura del operón, pero debido a la fragmentación del genoma no fue posible identificarse a través de *Prokka*. De igual manera, la disposición del operón *nif* indica que la presencia de los genes *nifU* y *nifS* sugiere que los demás genes están presentes en el genoma, pero la calidad de ensamblaje no fue suficiente para su identificación. Cabe mencionar que existen complicaciones para realizar la identificación de los genes *nif* debido a una alta variabilidad (Coelho *et al.*, 2009).

Es importante mencionar que se identificó la presencia del gen *nifH* a través de PCR, sin embargo, este mismo no se logró identificar en los genomas ensamblados (ver *figura 16* y la *figura 17A*), de igual manera se atribuye a la variabilidad entre las secuencias de los genes *nif* (Coelho *et al.*, 2009). Cabe mencionar que existen algunos estudios en donde se han caracterizado bacterias fijadoras de nitrógeno aisladas de suelos de cultivo, en donde tampoco fue posible identificar todos los genes para la función de la nitrogenasa al momento de realizar el ensamblaje. El estudio de Franco-Sierra *et al.*, (2020). identificó los genes *nifU*, *nifS*, *nifV* y *nifF*, en *B. subtilis* aislada de suelos, los cuales están involucrados en la capacidad enzimática del complejo nitrogenasa, sin embargo, los demás componentes no fueron identificados. Cabe resaltar que para entender a profundidad el mecanismo de fijación de nitrógeno de la bacteria, es necesario identificar los genes mínimos para el funcionamiento correcto del complejo, debido a que esto puede variar entre especies (Shamseldin, 2013). En la *figura 17B* se muestra la función de cada gen del operón. Los genes *nifHDK* se encuentran en la misma región y codifican a las subunidades necesarias para llevar a cabo la formación del complejo (dinitrogenasa reductasa y dinitrogenasa). Por lo tanto, a pesar de que no fue posible identificar estos genes a través del análisis de genoma completo, es pertinente hipotetizar que sí se encuentran estas secuencias por los resultados de la PCR (*figura 16*).

El análisis de caracterización de genoma completo permitió identificar distintas características promotoras de crecimiento de plantas por parte de las bacterias analizadas. Se escogió

la especie *B. subtilis* para la propuesta del sistema de edición génica para el mejoramiento en la fijación de nitrógeno debido a que fue posible identificar más genes del operón *nif*, asimismo, existe más información reportada en la literatura para poder respaldar las decisiones tomadas para el desarrollo del sistema. Además, *B. subtilis* también se caracterizó por presentar genes para el control biológico, para aumentar biodisponibilidad de fósforo a través de la solubilización de fosfatos, producción del sideróforo bacilibactina y algunos genes para la producción de IAA. En la *figura 20* se muestra un diagrama general de la distribución de características del genoma ensamblado de *B. subtilis*. Se obtuvo un total de 3940 secuencias de ADN codificante (CDS) y 57 secuencias de tRNA, un porcentaje de GC de 43.54 % dentro de un genoma de aproximadamente 4.06 MB.

Se realizó tres propuestas para una transformación bacteriana con el propósito de clonar el gen de regulación positivo *nifA* para poder realizar una comparación de qué plásmido presenta mejores resultados en base al inductor. En la *figura 21* se muestra uno plásmidos inducibles propuestos; este está basado en el constructo planteado por Schuster y Reisch., (2021) el gen reportero y de resistencia permitirán realizar la selección de las bacterias transformantes. Este sistema es inducido por ácido salicílico, un compuesto que no ha sido reportado como un agente perjudicial para las bacterias. El plásmido cuenta con un gen regulador *nahR* y un promotor (Meyer *et al.*, 2019). Se determinó que este plásmido puede llegar a una inducción máxima con una concentración de 10 μ M de ácido salicílico y un aumento de hasta 100 veces la expresión normal (Schuster y Reisch, 2021), por lo que se hipotetiza que su nivel de expresión inducirá un aumento significativo en la tasa de fijación de *B. subtilis*. Cabe mencionar que debido a que esta cepa modificada será utilizada para un bioinóculo, se eliminará el gen de resistencia utilizando enzimas restricción una vez se haya realizado la transformación.

La estructura y distribución del plásmido *pCinOHC14*, es la misma al descrito previamente, lo único que cambia es la secuencia del gen regulador y el promotor asociado para construir el sistema inducible (ver *figura 23*); para este se utilizó el regulador *cinR* y el promotor *Pcin* (Meyer *et al.*, 2019). Según Schuster y Reisch., (2021). la inducción de este plásmido demostró estimular hasta 120 veces la expresión normal del plásmido. Una particularidad de este sistema es que la molécula inductora es la hidroxitetradecanoil-homoserina lactona (OHC14), la cual tiene un rol importante debido a que es una molécula de *Quorum sensing* (Stock *et al.*, 2020), asimismo, este tiene una capacidad autoinductora reportada para los sistemas del promotor *Cin* (Vainana *et al.*, 2022). por lo que se hipotetiza que el ambiente de suelo inducirá la expresión del plásmido sin necesidad de la aplicación del compuesto.

De igual manera, la estructura y distribución del plásmido *pLuxOC6* solo difiere en las secuencias del gen regulador y el promotor que conforman el sistema inducible (ver *figura 22*); para este se utilizó el regulador *luxR* y el promotor (Meyer *et al.*, 2019). Según Schuster y Reisch., (2021). la inducción de este plásmido demostró estimular hasta 120 veces la expresión normal del plásmido. Una particularidad de este sistema es que la molécula inductora es la N-(3-oxohexanoil) homoserina lactona (OC6), la cual tiene un rol importante en la comunicación bacteriana debido a que se ha reportado como una molécula de *Quorum sensing* (Coquant *et al.*, 2020). No obstante, la diferencia entre el sistema *LuxR/pLuxB* y *CinR/Pcin* radica en que el OC6 tiene una mayor incidencia en ambientes donde proliferan bacterias que forman relaciones simbióticas para la fijación de nitrógeno con las raíces de las plantas (Nazari *et al.*, 2020; Yajima *et al.*, 2008; Schripsema *et al.*, 1996). Por lo tanto, este sistema tiene un gran potencial debido a que el aumento de fijación de nitrógeno de *B. subtilis*. La mejora en la tasa de fijación de nitrógeno, y la diversidad de bacterias fijadoras, podrían llevar a un aumento en la concentración de OC6 por parte de las bacterias, y por ende estimularía indefinidamente el sistema inducible. Es por esto, que se hipotetiza que el sistema de inducción más favorable para la formulación es a través del plásmido *pLuxOC6*.

Por último, se realizó una propuesta para llevar a cabo un *knock-out* (KO) del gen *nifL*, el cual regula negativamente la activación del operón *nif* al irrumpir la transcripción del gen *nifA* (Martínez-Argudo *et al.*, 2004). En la *figura 24* se muestra el sistema general que se propone. Debido a que la bacteria ya habrá sido transformada, es posible que la edición utilizando un plásmido

CRISPR pueda llevar a una disminución en el fitness de la bacteria, o en su defecto, complicar el proceso de KO por el tamaño del constructo (Diaz & Hernández, 2000; Smith y Bidochka, 1998). Por esto, se propone realizar un complejo de ribonucleoproteína (RNP), el cual estará compuesto por la endonucleasa SpCas9, el tracrRNA y el gRNA (Arroyo-Olarte *et al.*, 2021) debido a que una disminución en el tamaño del sistema, se puede obtener una mayor tasa de éxito (Zhang *et al.*, 2021). El gRNA se presenta en la *figura 24B*, en esta se muestran gRNA con el mejor puntaje (5'-CCACAGCGTCTGATACACCT-3'), como puede observarse el valor *on-target* es de 73.4% lo que sugiere que el sistema tiene una alta eficacia teórica, asimismo, se obtuvo un valor *off-target* de 100%, esto indica que no habrá cortes en regiones inespecíficas del genoma de *B. subtilis*, lo que podría aumentar la tasa de eficacia de la edición genética (Doench *et al.*, 2016). Una de las complicaciones con este sistema, es que debido a que el sistema no presenta un gen para realizar la selección, se debe de realizar un *screening* a través de técnicas moleculares. Se han utilizado cebadores para flanquear la región de KO, de esta manera se puede evaluar por PCR y gel de agarosa o por secuenciación para determinar si se ha realizado la edición génica correctamente.

Posterior a la realización de la transformación y el CRISPR-KO se deberá realizar una evaluación de qué sistema fomenta una mejora significativa en la fijación de nitrógeno a través de una prueba cuantitativa. De igual manera, deberá estudiarse la interacción que tiene la cepa *B. subtilis* modificada con las demás cepas potenciales para la bioformulación, por último, deberá estudiarse la capacidad de mejora de nitrógeno al ser inoculada en el suelo. Los hallazgos de este estudio presentan, no solamente información de genes implicados en capacidades promotoras del crecimiento de plantas, sino una propuesta potencial para la formulación de un biofertilizante con el potencial de poder suministrar altas concentraciones de nitrógeno a los cultivos. Por lo tanto, se podría disminuir significativamente el uso de fertilizantes químicos, mejorando la salud del suelo y proporcionando alternativas para el diseño de sistemas agrícolas más sostenibles.

Conclusiones

1. Se seleccionaron doce cepas bacterianas candidatas con capacidades promotoras del crecimiento bacteriano con base en la caracterización bioquímica tomando como prioridad la fijación de nitrógeno para la secuenciación de genoma completo.
2. Se completaron exitosamente diez ensamblajes de genoma completo de las bacterias candidatas, con base en sus resultados de calidad, se determinó que el algoritmo más apropiado para la secuenciación de genoma fue SPAdes, debido a que resultó en los ensamblajes menos fragmentados.
3. Se identificaron genes implicados en la ruta de fijación de nitrógeno, particularmente regulada por el operón *nif*, sin embargo, no fue posible determinar todos los genes necesarios para el completo funcionamiento del complejo nitrogenasa, posiblemente por la fragmentación en el ensamblaje y por la variabilidad que existe en estos genes.
4. Se identificaron exitosamente genes implicados en la producción de IAA, aumento en la biodisponibilidad de fósforo, producción de sideróforos y con un potencial para el control biológico de *Fusarium*.
5. Las cepas estudiadas no presentaron tener un potencial de patogenicidad a humanos debido a que se determinó pocos genes que, a pesar de ser factores de virulencia, no son intrínsecamente patogénicos. Asimismo, no se determinó genes para la producción de endotoxinas. Se identificaron pocos genes que codifican a factores de virulencia asociados a los géneros determinados en este estudio, sin embargo, ninguna cepa fue reportada como productora de endotoxinas.
6. Se planteo exitosamente un sistema de edición génica para el mejoramiento en la fijación de nitrógeno en *B. subtilis* a través de una transformación con un plásmido inducible para clonar el gen nativo *nifA* y una edición génica para realizar un KO del gen *nifL*.

Primero se recomienda realizar una identificación de una mayor cantidad de genes que componen el operón *nif* a través del diseño de primers degenerados para llevar a cabo una PCR con el propósito de poder ratificar que las cepas estudiadas sí se tienen la presencia de ellos. De igual manera, se recomienda realizar una curva de nitrógeno para poder cuantificar la capacidad fijadora de las cepas estudiadas, debido a que la caracterización está limitada a la identificación de la capacidad, más no que tan efectivo es. Esto a su vez es importante, debido a que es necesario contar con un método para evaluar que el sistema de edición génica propuesto sí aumente de forma significativa la capacidad fijadora de nitrógeno.

Como siguientes pasos en la investigación se recomienda previamente a realizar la edición génica, llevar a cabo una secuenciación de genoma completo con una mayor profundidad, con el propósito de ensamblar un genoma menos fragmentado. De esta manera, se podrá obtener una mejor calidad en la anotación de genes y la caracterización de las rutas de interés será más completa. Asimismo, se recomienda llevar a cabo una secuenciación del gen *nifL* para poder contar con la secuencia cruda del gen de *B. subtilis* para cerciorar la exactitud del diseño de CRISPR para llevar a cabo el KO.

Se recomienda de igual manera complementar los resultados obtenidos en este estudio con los análisis de antagonismo de las cepas bacterianas, principalmente para evaluar que la cepa seleccionada para llevar a cabo la edición génica sí pueda ser utilizada en la formulación del biofertilizante.

Ahemad, M., Kibret, M. (2014). Mechanisms and applications of plant growth promoting rhizobacteria: Current perspective. [Mecanismos y aplicaciones de rizobacterias promotoras del crecimiento vegetal: Perspectiva actual]. *Journal of King Saud University*. ISSN: 10183647. DOI: 10.1016/j.jksus.2013.05.001

Ahemad, M. (2015). Phosphate-solubilizing bacteria-assisted phytoremediation of metalliferous soils: a review. [Fitorremediación asistida por bacterias solubilizadoras de fosfatos de suelos metalíferos: una revisión]. *3 Biotech*. ISSN: 2190-572X. DOI: 10.1007/s13205-014-0206-0

Ahmed, E., & Holmström, S. J. (2014). Siderophores in environmental research: roles and applications. [Sideróforos en la investigación ambiental: roles y aplicaciones]. *Microbial biotechnology*, 7(3), 196-208. DOI: 10.1111/1751-7915.12117

Ajjjah, N., Fiodor, A., Pandey, A. K., Rana, A., & Pranaw, K. (2023). Plant Growth-Promoting Bacteria (PGPB) with biofilm-forming ability: A multifaceted agent for sustainable agriculture. [Bacterias promotoras del crecimiento vegetal (PGPB) con capacidad de formación de biofilm: un agente multifacético para la agricultura sostenible]. *Diversity*, 15(1), 112. DOI: 10.3390/d15010112

Arroyo-Olarte, R. D., Bravo Rodríguez, R., & Morales-Ríos, E. (2021). Gene editing in bacteria via SpCas9/sgRNA ribonucleoprotein complexes. [Edición genética en bacterias a través de complejos ribonucleoproteicos SpCas9/sgRNA]. *bioRxiv*, 2021-06. DOI: 10.1101/2021.06.08.447524

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A., Dvorkin, M., Kulikov, A., Lesin, V., Nikolenko, S., Pham, S., Prijbelski, A., Pyshkin, A., Sirotkin, A., Vyahni, N., Tesler, G., Alekseyev, M., Pevzner, P. A. (2012). SPAdes: A New Genome Assembly Algorithm and Its Application to Single-Cell Sequencing. [SPAdes: Un nuevo algoritmo de ensamblaje de genomas y su aplicación a la secuenciación de células individuales]. *Journal of Computational Biology*. ISSN: 1066-5277. DOI: 10.1089/cmb.2012.0021

Barros, E., Grimaldi, M., Sarrazin, M., Chauvel, A., Mitja, D., Desjardins, T., Lavelle, P. (2004). Soil physical degradation and changes in macrofaunal communities in Central Amazon. [Degradación física del suelo y cambios en las comunidades de macrofauna en la Amazonia central]. *Applied Soil Ecology*. ISSN: 0929-1393. DOI: 10.1016/j.apsoil.2003.10.012

Barker, A. (2018). Fertilizers. [Fertilizantes]. DOI: 10.1016/B978-0-12-409547-2.00142-6

Banerjee, S., van der Heijden, M.G.A. (2023). Soil microbiomes and one health. [Microbiomas del suelo y una salud]. *Nat Rev Microbiol* 21, 6–20. DOI: 10.1038/s41579-022-00779-w

Beneduzi, A., Moreira, F., Costa, P. B., Vargas, L. K., Lisboa, B. B., Favreto, R., . . . Passaglia, L. M. P. (2013). Diversity and plant growth promoting evaluation abilities of bacteria isolated from sugarcane cultivated in the South of Brazil. [Diversidad y evaluación de capacidades de promoción del crecimiento vegetal de bacterias aisladas de caña de azúcar cultivada en el sur de Brasil]. *Applied Soil Ecology*, 63, 94–104. DOI: 10.1016/j.apsoil.2012.08.010

Behera, S., Gayen, S., Deogun, J. S., & Vinodchandran, N. V. (2018, August). KmerEstimate: A streaming algorithm for estimating k-mer counts with optimal space usage. [KmerEstimate: Un algoritmo de transmisión para estimar recuentos de k-mer con uso óptimo del espacio]. In *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics* (pp. 438-447). DOI: 10.1145/3233547.3233573

Billah, M., Khan, M., Bano, A., Hassan, T., Munir, A., Gurmani, A. (2019). Phosphorus and phosphate solubilizing bacteria: Keys for sustainable agriculture. [Fósforo y bacterias solubilizadoras de fosfatos: Claves para una agricultura sostenible]. *Geomicrobiology Journal*. ISSN: 0149-0451. DOI: 10.1080/01490451.2019.1654043

Bijay-Singh., & Craswell, E. (2021). Fertilizers and nitrate pollution of surface and ground water: an increasingly pervasive global problem. [Fertilizantes y contaminación por nitratos del agua superficial y subterránea: un problema global cada vez más extendido]. *SN Applied Sciences*. ISSN: 2523-3963. DOI: 10.1007/s42452-021-04521-8

Bindraban, P., Van der Velde, M., Ye, L., Van der Berg, M., Materechera, S., Kiba, D., Tamene, L., Ragnarsdottir, K., Jongschaap, R., Hoogmoed, M., Van Beek, C., Ban Lynden, G. (2012). Assessing the impact of soil degradation on food production. [Evaluación del impacto de la degradación del suelo en la producción de alimentos]. *Current Opinion in Environmental Sustainability*. ISSN: 1877-3435. DOI: 10.1016/j.cosust.2012.09.015

Bharucha, U., Patel, K., & Trivedi, U. B. (2013). Optimization of indole acetic acid production by *Pseudomonas putida* UB1 and its effect as plant growth-promoting rhizobacteria on mustard (*Brassica nigra*). [Optimización de la producción de ácido indolacético por *Pseudomonas putida* UB1 y su efecto como rizobacteria promotora del crecimiento vegetal en mostaza (*Brassica nigra*)]. *Agricultural Research*, 2, 215-221. DOI: 10.1007/s40003-013-0065-7

Brandt, K., Nethery, M., O’Flaherty, S., Barrangou, R. (2020). Genomic Characterization of *Lactobacillus fermentum* DSM. [Caracterización genómica de *Lactobacillus fermentum* DSM]. *BMC Genomics*. ISSN: 1471-2164. DOI: 10.1186/s12864-020-6740-8

Brown, J., Pirrung, M., McCue, L. (2017). FQC Dashboard: Integrated FastQC results into a web-based, interactive, and extensive FASTQ quality control tool. [FQC Dashboard: Resultados de FastQC integrados en una herramienta de control de calidad FASTQ interactiva, extensiva y basada en la web]. *Bioinformatics*. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btx373

Bolger, A., Lohse, M., Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. [Trimmomatic: un recortador flexible para datos de secuencias de Illumina]. *Bioinformatics*. ISSN: 1367-4811. DOI: 10.1093/bioinformatics/btu170

Canisares, L., Cherubin, M., Silva, L., Franco, A., Cooper, M., Mooney, S., Cerri, C. (2020). Soil microstructure alterations induced by land use change for sugarcane expansion in Brazil. [Alteraciones en la microestructura del suelo inducidas por el cambio de uso de la tierra para la expansión de la caña de azúcar en Brasil]. *Geoderma*. ISSN: 0016-7061. DOI: 10.1016/j.geoderma.2020.114511

Carini, P., Marsden, P., Leff, J., Morgan, E., Strickland, M., Fierer, N. (2016). Relic DNA is

abundant in soil and obscures estimates of soil microbial diversity. [El ADN relicto es abundante en el suelo y oscurece las estimaciones de la diversidad microbiana del suelo]. *Nature Microbiology*. ISSN: 2058-5276. DOI: 10.1038/nmicrobiol.2016.242

Checker, V., Kushwaha, H., Kumari, P., Yadav, S. (2018). Role of phytohormones in Plant Defense: Signaling and cross talk. [Papel de las fitohormonas en la defensa de las plantas: Señalización y comunicación cruzada]. *Molecular Aspect of Plant-Pathogen Interaction*. DOI: 10.1007/978-981-10-7371-7_7

Chen, C., Khaleel, S., Huang, H., Wu, C. (2014). Software for pre-processing Illumina next-generation sequencing short read sequences. [Software para el preprocesamiento de secuencias de lectura corta de secuenciación de próxima generación Illumina]. *Source Code for Biology and Medicine*. ISSN: 1751-0473. DOI: 10.1186/1751-0473-9-8

Chen, S., Huang, T., Zhou, Y., Han, Y., Xu, M., & Gu, J. (2017). AfterQC: automatic filtering, trimming, error removing and quality control for fastq data. [AfterQC: filtrado automático, recorte, eliminación de errores y control de calidad para datos fastq]. *BMC Bioinformatics*, 18(3), 91-100.

Chen, L., Yang, J., Yu, J., Yao, Z., Sun, L., Shen, Y., & Jin, Q. (2005). VFDB: a reference database for bacterial virulence factors. [VFDB: una base de datos de referencia para factores de virulencia bacteriana]. *Nucleic Acids Research*, 33(suppl_1), D325-D328. DOI: 10.1093/nar/gki008

Christou, P. (2013). Plant genetic engineering and agricultural biotechnology 1983-2013. [Ingeniería genética de plantas y biotecnología agrícola 1983-2013]. *Trends in Biotechnology*. ISSN: 01677799. DOI: 10.1016/j.tibtech.2013.01.006

Chanderban, M., Hill, C. A., Dhamad, A. E., & Lessner, D. J. (2023). Expression of V-nitrogenase and Fe-nitrogenase in *Methanosarcina acetivorans* is controlled by molybdenum, fixed nitrogen, and the expression of Mo-nitrogenase. [La expresión de V-nitrogenasa y Fe-nitrogenasa en *Methanosarcina acetivorans* está controlada por molibdeno, nitrógeno fijado y la expresión de Mo-nitrogenasa]. *Applied and Environmental Microbiology*, 89(9), e01033-23. DOI: 10.1128/aem.01033-23

Cock, P., Fields, C., Goto, N., Heuer, M., Rice, P. (2010). The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. [El formato de archivo FASTQ de Sanger para secuencias con puntajes de calidad y las variantes FASTQ de Solexa/Illumina]. *Nucleic Acid Research*. ISSN: 0305-1048. DOI: 10.1093/nar/gkp1137

Coelho, M. R. R., Marriel, I. E., Jenkins, S. N., Lanyon, C. V., Seldin, L., & O'Donnell, A. G. (2009). Molecular detection and quantification of *nifH* gene sequences in the rhizosphere of sorghum (*Sorghum bicolor*) sown with two levels of nitrogen fertilizer. [Detección molecular y cuantificación de secuencias del gen *nifH* en la rizosfera de sorgo (*Sorghum bicolor*) sembrado con dos niveles de fertilizante nitrogenado]. *Applied Soil Ecology*, 42(1), 48–53. DOI: 10.1016/j.apsoil.2009.01.010

Coquant, G., Grill, J. P., & Seksik, P. (2020). Impact of N-acyl-homoserine lactones, quorum sensing molecules, on gut immunity. [Impacto de las N-acil-homoserina lactonas, moléculas de detección de quórum, en la inmunidad intestinal]. *Frontiers in Immunology*, 11, 1827. DOI: 10.3389/fimmu.2020.01827

Company, S., Samad, A., Faist, H., Sessitsch, A. (2019). *Journal of Advanced Research*. [Revista de Investigación Avanzada]. ISSN: 20901232. DOI: 10.1016/j.jare.2019.03.004

Crossa, J. H., & Walsh, C. T. (2002). Genetics and Assembly Line Enzymology of Siderophore Biosynthesis in Bacteria. [Genética y enzimología de línea de ensamblaje de la biosíntesis de sideróforos en bacterias]. *Microbiology and Molecular Biology Reviews*, 66(2), 223–249. DOI: 10.1128/mmbr.66.2.223-249.2002

Curatti, L., & Rubio, L. M. (2014). Challenges to develop nitrogen-fixing cereals by direct nif-gene transfer. [Desafíos para desarrollar cereales fijadores de nitrógeno mediante la transferencia directa del gen nif]. *Plant Science*, 225, 130-137. DOI: 10.1016/j.plantsci.2014.05.011

Dastogeer, K., Tumpa, F., Sultana, A., Akter, M., Chakraborty, A. Plant microbiome-an account of the factors that shape community composition and diversity. [Microbioma de plantas: una explicación de los factores que moldean la composición y diversidad de la comunidad]. *Current Plant Biology*. ISSN: 22146628. DOI: 10.1016/j.cpb.2020.100161

De Sena Brandine, G., & Smith, A. (2021). Falco: Highspeed FASTQC emulation for quality control of sequencing data. [Falco: emulación rápida de FASTQC para el control de calidad de datos de secuenciación]. *F1000Research*. ISSN: 2046-1402. DOI: 10.12688/f1000research.21142.2

Delgado-Baquerizo, M., & Eldridge, D. (2019). Cross-Biome Drivers of Soil Bacterial Alpha Diversity on a Worldwide Scale. [Factores impulsores de la diversidad alfa bacteriana del suelo en diferentes biomas a escala mundial]. *Ecosystems*. ISSN: 1432-9840. DOI: 10.1007/s10021-018-0333-2

Devine, K. M. (2018). Activation of the PhoPR-mediated response to phosphate limitation is regulated by wall teichoic acid metabolism in *Bacillus subtilis*. [La activación de la respuesta mediada por PhoPR a la limitación de fosfato está regulada por el metabolismo del ácido teicoico de la pared en *Bacillus subtilis*]. *Frontiers in Microbiology*, 9, 2678. DOI: 10.3389/fmicb.2018.02678

Diaz Ricci, J. C., & Hernández, M. E. (2000). Plasmid effects on *Escherichia coli* metabolism. [Efectos del plásmido en el metabolismo de *Escherichia coli*]. *Critical Reviews in Biotechnology*, 20(2), 79-108. DOI: 10.1080/07388550008984167

Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., . . . Root, D. E. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. [Diseño optimizado de sgRNA para maximizar la actividad y minimizar los efectos fuera del objetivo de CRISPR-Cas9]. *Nature Biotechnology*, 34(2), 184–191. DOI: 10.1038/nbt.3437

Duan, F., Liu, J., March, J. (2015). Engineered Commensal Bacteria Reprogram Intestinal Cells Into Glucose-Responsive Insulin-Secreting Cells for the Treatment of Diabetes. [Bacterias comensales diseñadas reprograman células intestinales en células secretoras de insulina sensibles a la glucosa para el tratamiento de la diabetes]. *Diabetes*. ISSN: 0012-1797. DOI: 10.2337/db14-0635

Dubey, A., Malla, M., Khan, F., Chowdhary, K., Yadav, S., Kumar, A., Sharma, S., Khare, P., Khan, M. (2019). Soil microbiome: A key player for conservation of soil health under changing climate. [Microbioma del suelo: Un actor clave para la conservación de la salud del suelo bajo un clima cambiante]. ISSN: 0960-3115. DOI: 10.1007/s10531-019-01760-5

Duca, D. R., & Glick, B. R. (2020). Indole-3-acetic acid biosynthesis and its regulation in plant-associated bacteria. [Biosíntesis del ácido indol-3-acético y su regulación en bacterias asociadas a plantas]. *Applied Microbiology and Biotechnology*. doi:10.1007/s00253-020-10869-5

Eisen-Dorman, A., Lawrence, D., Bolívar, V. (2009). Cautionary insights on knockout mouse Studies: The gene or not the gene? [Perspectivas precautorias sobre estudios de ratones knockout: ¿El gen o no el gen?]. *Brain, behavior, and immunity*. ISSN: 08891591. DOI: 10.1016/j.bbi.2008.09.001

Ehrlich, P., & Harte, J. (2015). Opinion: To feed the world in 2050 wil require a global revolution. [Opinión: Alimentar al mundo en 2050 requerirá una revolución global]. *Proceedings of the National Academy of Science*. ISSN: 00278424

Efe, D. (2020). Potential Plant Growth-Promoting Bacteria with Heavy Metal Resistance. [Bacterias potencialmente promotoras del crecimiento vegetal con resistencia a metales pesados]. *Current Microbiology*, 77(12), 3861–3868. doi:10.1007/s00284-020-02208-8

Fraser, T. D., Lynch, D. H., Bent, E., Entz, M. H., & Dunfield, K. E. (2015). Soil bacterial phoD

gene abundance and expression in response to applied phosphorus and long-term management. [Abundancia y expresión del gen bacteriano *phoD* del suelo en respuesta al fósforo aplicado y al manejo a largo plazo]. *Soil Biology and Biochemistry*, 88, 137-147.

FAO. (2011). The State of the World's Land and water resources of foos and agriculture (SOLAW) managing systems at risk. [El estado de los recursos mundiales de tierra y agua para la alimentación y la agricultura (SOLAW) gestionando sistemas en riesgo].

Ferreti, G., Galamini, G., Deltedesco, E., Gorfer, M., Fritz, J., Faccini, B., Mentler, A., Zechmeister-Boltenstern, S., Coltorti, M., Keiblinger, K. (2021). Gross Ammonification and Nitrification Rates in soil Amended with Natural and NH₄-Enriched chabazite zeolite and nitrification inhibitor DMPP. [Tasas brutas de amonificación y nitrificación en suelo enmendado con zeolita chabazita natural y enriquecida con NH₄ e inhibidor de nitrificación DMPP]. *Applied Sciences*. ISSN: 2079-3417. DOI: 10.3390/app11062605

Fierer, N. (2017): Embracing the Unknown: Disentangling the complexities of the soil microbiome. [Abrazando lo desconocido: Desentrañando las complejidades del microbioma del suelo]. *Nature Reviews Microbiology*. ISSN: 1740-1526. DOI: 10.1038/nrmicro.2017.87

Fhad, S., Hussain, S., Matloon, A., Khan, F., Khaliq, A., Saud, S., Hassan, S., Shan, D., Khan, F., Ullah, N., Faiq, M., Khan, M., Tareen, A., Khan, A., Ullah, A., Ullah, N., Huang, J. (2015). Plant Growth Regulation. [Regulación del crecimiento vegetal]. ISSN: 0167-6903. 10.1007/s10725-014-0013-y

Fuqua, C., Winans, S. C., & Greenberg, E. P. (1996). CENSUS AND CONSENSUS IN BACTERIAL ECOSYSTEMS: The LuxR-LuxI Family of Quorum-Sensing Transcriptional Regulators. [CENSO Y CONSENSO EN ECOSISTEMAS BACTERIANOS: La familia LuxR-LuxI de reguladores transcripcionales de detección de quórum]. *Annual Review of Microbiology*, 50(1), 727-751. doi:10.1146/annurev.micro.50.1.7

Gamalero, E., & Glick, B. (2011). Mechanisms Used by Plant Growth-Promoting Bacteria. [Mecanismos utilizados por las bacterias promotoras del crecimiento vegetal]. *Bacteria in Agrobiolgy: Plant Nutrient Management*. DOI: 10.1007/978-3-642-21061-7_2

Garbeva, P., Voesenek, K., & Elsas, J. D. va. (2004). Quantitative detection and diversity of the pyrrolnitrin biosynthetic locus in soil under different treatments. [Detección cuantitativa y diversidad del locus biosintético de pirrolnitrina en suelo bajo diferentes tratamientos]. *Soil Biology and Biochemistry*, 36(9), 1453-1463. doi:10.1016/j.soilbio.2004.03.009

Goulding, K. (2016). Soil acidification and the importance of liming agricultural soils with particular reference to the United Kingdom. [Acidificación del suelo y la importancia del encalado de suelos agrícolas con referencia particular al Reino Unido]. ISSN: 02660032. DOI: 10.1111/sum.12270

Gollnick, P., Babitzke, P., Antson, A., & Yanofsky, C. (2005). Complexity in regulation of tryptophan biosynthesis in *Bacillus subtilis*. [Complejidad en la regulación de la biosíntesis de triptófano en *Bacillus subtilis*]. *Annu. Rev. Genet.*, 39, 47-68.

Gutiérrez-Santana, A., Carrillo-Cerda, H., Rodríguez-Campos, J., Kirchmayr, M., Contretas-Ramos, S., Velásquez-Fernández, J. (2020). Volatile emission compounds from plant growth-promoting bacteria are responsible for the antifungal activity against *F. solani*. [Los compuestos de emisión volátil de bacterias promotoras del crecimiento vegetal son responsables de la actividad antifúngica contra *F. solani*]. *Biotech*. ISSN: 2190-572X. DOI: 10.1007/s13205-020-02290-6

Gulati, A., Vyas, P., Rahi, P., & Kasana, R. C. (2009). Plant growth-promoting and rhizosphere-competent *Acinetobacter rhizosphaerae* strain BIHB 723 from the cold deserts of the Himalayas. [*Acinetobacter rhizosphaerae* cepa BIHB 723 promotora del crecimiento vegetal y

competente en la rizosfera de los desiertos fríos del Himalaya]. *Current Microbiology*, 58, 371-377.

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUASt: quality assessment tool for genome assemblies. [QUASt: herramienta de evaluación de calidad para ensamblajes genómicos]. *Bioinformatics*, 29(8), 1072-1075.

Hasan, N., Young, B., Minard-Smith, A., Saeed, K., Li, H., Heizer, E., McMillan, N., Isom, R., Abdullah, A., Bornmad, D., Faith, S., Choi, S., Dickens, M., Cebula, T., Colwell, R. (2014). Microbial community profiling of Human Saliva Using Shotgun Metagenomic Sequencing. [Perfilado de la comunidad microbiana de la saliva humana utilizando secuenciación metagenómica de escopeta]. *PLoS ONE*. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0097699

Hashem, A., Tabassum, B., & Abd_Allah, E. F. (2019). *Bacillus subtilis*: A plant-growth promoting rhizobacterium that also impacts biotic stress. [Bacillus subtilis: Una rizobacteria promotora del crecimiento vegetal que también impacta el estrés biótico]. *Saudi journal of biological sciences*, 26(6), 1291-1297

Heather, J., & Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. [La secuencia de secuenciadores: La historia de la secuenciación del ADN]. ISSN: 08887543. DOI: 10.1016/j.ygeno.2015.11.003. DOI: 10.1007/s00005-016-0427-5

Henson, J., Tischler, G., Ning, Z. (2012). Next-generation sequencing and large genome assemblies. [Secuenciación de próxima generación y ensamblajes de genomas grandes]. *Pharmacogenomics*. ISSN: 1462-2416. DOI: 10.2217/pgs.12.72

Hsieh, Y. J., & Wanner, B. L. (2010). Global regulation by the seven-component Pi signaling system. [Regulación global por el sistema de señalización Pi de siete componentes]. *Current opinion in microbiology*, 13(2), 198-203.

Heyhorowicz, M., Lipiński, d., Zeyland, J., Słomski, R. (2017). *Archivum immunologiae et Therapiae Experimentalis*. [Archivo de Inmunología y Terapia Experimental]. ISSN: 0004-069X. DOI: 10.1016/j.bbi.2008.09.001

Hertlein, G., Müller, S., Garcia-Gonzalez, E., Poppinga, L., Süßmuth, R. D., & Genersch, E. (2014). Production of the catechol type siderophore bacillibactin by the honey bee pathogen *Paenibacillus larvae*. [Producción del sideróforo tipo catecol bacilibactina por el patógeno de las abejas melíferas *Paenibacillus larvae*]. *PLoS One*, 9(9), e108272.

Hinsiger, P. (2001). Bioavailability of soil inorganic P in the rhizosphere as affected by root-induced chemical changes: A review. [Biodisponibilidad del P inorgánico del suelo en la rizosfera afectada por cambios químicos inducidos por las raíces: Una revisión]. *Plant and soil*. ISSN: 0032079X. DOI: 10.1023/A:1013351617532

Holkar, S., Balasubramaniam, P., Kumar, A., Kardivel, N., Shingote, P., Chhabra, M., Kumar, S., Kumar, P., Viswanathan R., Jain, R., Pathak, A. (2020). Present status and future management Strategies for sugarcane yellow leaf virus: A major constraint to the global sugarcane production. [Estado actual y estrategias futuras de manejo para el virus de la hoja amarilla de la caña de azúcar: Una restricción importante para la producción mundial de caña de azúcar]. *The plant pathology journal*. ISSN: 1598-2254. DOI: 10.5423/PPJ.RW.09.2020.0183

Hyatt D, et al. Prodigal: prokaryotic gene recognition and translation initiation site identification. [Prodigal: reconocimiento de genes procariotas e identificación de sitios de inicio de traducción]. *BMC Bioinformatics*, 2010, vol. 11 pg. 119

IFIA (2000). *Fertilizers and their Use*. [Los fertilizantes y su uso]. International Fertilizer industry Association.

Jackman, S., Vandervalk, B., Mohamadi, H., Chu, J., Yeo, S., Hammond, S., Jahesh, G., Khan, H.,

Coombe, L., Warren, R., Biron, I. (2017). ABySS 2.0: Resource-efficient assembly of large genomes using Bloom filter. [ABYSS 2.0: Ensamblaje eficiente en recursos de grandes genomas utilizando filtro de Bloom]. *Genome Research*. ISSN: 1088-9051. DOI: 10.1101/gr.214346.116

Jasim, B., Jimtha John, C., Shimil, V., Jyothis, M., & Radhakrishnan, E. K. (2014). Studies on the factors modulating indole-3-acetic acid production in endophytic bacterial isolates from *Piper nigrum* and molecular analysis of ipdc gene. [Estudios sobre los factores que modulan la producción de ácido indol-3-acético en aislados bacterianos endofíticos de *Piper nigrum* y análisis molecular del gen ipdc]. *Journal of applied microbiology*, 117(3), 786-799.

Jetten, M. (2008). The microbial nitrogen cycle. [El ciclo microbiano del nitrógeno]. *Environmental Microbiology*. ISSN: 14622912. DOI: 10.1111/j.1462-2920.2008.01786.x

Kamat, S. S., & Raushel, F. M. (2015). PhnJ–A novel radical SAM enzyme from the C–P lyase complex. [PhnJ–Una nueva enzima SAM radical del complejo liasa C–P]. *Perspectives in Science*, 4, 32-37.

Karthika, S., Midhun, S., Jisha, M. (2020). A potential antifungal and growth-promoting bacterium *Bacillus* sp. KTMA4 from tomato rhizosphere. [Una bacteria potencialmente antifúngica y promotora del crecimiento *Bacillus* sp. KTMA4 de la rizosfera del tomate]. *Microbial Pathogenesis*. ISSN: 08824010. DOI: 10.1016/j.micpath.2020.104049

Kershaw, M., Westwood, J., Darcy, P. (2013). Gene-engineered T cells for cancer therapy. [Células T modificadas genéticamente para la terapia del cáncer]. *Nature Reviews Cancer*. ISSN: 1474-175X. DOI: 10.1038/nrc3565

Koonin, E., Makarova, K., Zhang, F. (2017). Diversity, Classification and Evolution of CRISPR-Cas systems. [Diversidad, clasificación y evolución de los sistemas CRISPR-Cas]. *Current Opinion in Microbiology*. ISSN: 13695274. DOI: 10.1016/j.mib.2017.05.008

Koonin, E., & Makarova, K. (2019). Origin and evolution of CRISPR-Cas Systems. [Origen y evolución de los sistemas CRISPR-Cas]. *Philosophical transactions of the Royal Society B: Biological Sciences*. ISSN: 0962-8436. DOI: 10.1098/rstb.2018.0087

Kolbe DL, Eddy SR. Fast filtering for RNA homology search. [Filtrado rápido para búsqueda de homología de ARN]. *Bioinformatics*, 2011, vol. 27 (pg. 3102-3109)

Kuman, S., Diksha., Sindhu, S., Kumar, R. (2022). Biofertilizers: An ecofriendly technology for nutrient recycling and environmental sustainability. [Biofertilizantes: Una tecnología ecológica para el reciclaje de nutrientes y la sostenibilidad ambiental]. *Current Research in Microbial Sciences*. ISSN: 26665174. DOI: 10.1016/j.crmicr.2021.100094

Lanigan, T., Hopera, H., Saunders, T. (2020). Principles of genetic Engineering. [Principios de ingeniería genética]. *Genes*. ISSN: 2073-4425. DOI: 10.3390/genes11030291

Lagesen K, et al. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. [RNAmmer: anotación consistente y rápida de genes de ARN ribosomal]. *Nucleic Acids Res.*, 2007, vol. 35 (pg. 3100-31008)

Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. [ARAGORN, un programa para detectar genes de ARNt y genes de ARNtm en secuencias de nucleótidos]. *Nucleic Acids Res.*, 2004, vol. 32 (pg. 11-16)

Lehman, R., Cambardella, C., Stott, D., Acosta-Martinez, V., Manter, D., Buyen, J., Maul, J., Smith, J., Collins, H., Halvorson, J., Kremer, R., Lundgren, J., Ducey, T., Jin, V., Karlen, D. (2015). Understanding and enhancing soil Biological health: The solution or reversing soil Degradation. [Comprendiendo y mejorando la salud biológica del suelo: La solución para revertir la degradación del suelo]. *Sustainability*. ISSN: 2071-1050. DOI: 10.3390/su7010988

Leitão, J. H. (2020). Microbial virulence factors. [Factores de virulencia microbiana]. *International journal of molecular sciences*, 21(15), 5320

Lischer, H., & Shimizu, K. (2017). Reference-guided de novo assembly approach improves genome reconstruction for related species. [El enfoque de ensamblaje de novo guiado por referencia mejora la reconstrucción del genoma para especies relacionadas]. *BMC Bioinformatics*. ISSN: 1471-2105. DOI: 10.1186/s12859-017-1911-6

Liu, L., Gao, Z., Yang, Y., Gao, Y., Mahmood, M., Jiao, H., ... & Liu, J. (2023). Long-term high-P fertilizer input shifts soil P cycle genes and microorganism communities in dryland wheat production systems. [La entrada a largo plazo de fertilizantes altos en P cambia los genes del ciclo P del suelo y las comunidades de microorganismos en los sistemas de producción de trigo de secano]. *Agriculture, Ecosystems & Environment*, 342, 108226.

Liu, Y., Wang, H., Sun, X., Yang, H., Wang, Y., & Song, W. (2011). Study on mechanisms of colonization of nitrogen-fixing PGPB, *Klebsiella pneumoniae* NG14 on the root surface of rice and the formation of biofilm. [Estudio sobre los mecanismos de colonización de PGPB fijadoras de nitrógeno, *Klebsiella pneumoniae* NG14 en la superficie de la raíz del arroz y la formación de biopelícula]. *Current microbiology*, 62, 1113-1122.

Liang, J. L., Liu, J., Jia, P., Yang, T. T., Zeng, Q. W., Zhang, S. C., ... & Li, J. T. (2020). Novel phosphate-solubilizing bacteria enhance soil phosphorus cycling following ecological restoration of land degraded by mining. [Nuevas bacterias solubilizadoras de fosfato mejoran el ciclo del fósforo del suelo después de la restauración ecológica de tierras degradadas por la minería]. *The ISME journal*, 14(6), 1600-1613.

Lu, G., Wang, Z., Xu, F., Pan, Y., Grisham, M., Xu, L. (2021). Sugarcane Mosaic Disease: Characteristics, identification and Control. [Enfermedad del mosaico de la caña de azúcar: Características, identificación y control]. *Microorganisms*. ISSN: 2076-2607. DOI: 10.3390/microorganisms9091984

Lyu, N., Feng, Y., Pan, Y., Huang, Y., Liu, H., Xue, C., Zhu, B., Hu, Y. (2021). Genomic Characterization of *Salmonella enterica* isolates from Retail Meat in Beijing, China. [Caracterización genómica de aislados de *Salmonella enterica* de carne minorista en Beijing, China]. *Frontiers in Microbiology*. ISSN: 1664-302X. DOI: 10.3389/fmicb.2021.636332

Maçik, M., Gryta, A., Fraç, M. (2020). Biofertilizers in agriculture. An overview on concepts, strategies and effects on soil microorganisms. [Biofertilizantes en agricultura. Una visión general sobre conceptos, estrategias y efectos en los microorganismos del suelo]. DOI: 10.1016/bs.agron.2020.02.001

Malovichko, Y. V., Nizhnikov, A. A., & Antonets, K. S. (2019). Repertoire of the *Bacillus thuringiensis* virulence factors unrelated to major classes of protein toxins and its role in specificity of host-pathogen interactions. [Repertorio de factores de virulencia de *Bacillus thuringiensis* no relacionados con las principales clases de toxinas proteicas y su papel en la especificidad de las interacciones huésped-patógeno]. *Toxins*, 11(6), 347.

Martini, A., Valani, G., Boschmi, R., Bovi, R., Simoes da Silva, L., Cooper, M. (2020): Is soil quality a concern in sugarcane cultivation? A bibliometric review. [¿Es la calidad del suelo una preocupación en el cultivo de caña de azúcar? Una revisión bibliométrica].

Martinez-Argudo, I., Little, R., Shearer, N., Johnson, P., & Dixon, R. (2004). The NifL-NifA system: a multidomain transcriptional regulatory complex that integrates environmental signals. [El sistema NifL-NifA: un complejo regulador transcripcional multidominio que integra señales ambientales]. *Journal of bacteriology*, 186(3), 601-610.

Mahanty, T., Bhattacharjee, S., Goswami, M., Bhattacharyya, P., Das, B., Ghost, A., Tribedi, P. (2017). Biofertilizers: A potential approach for sustainable agriculture development. [Biofertilizantes:

Un enfoque potencial para el desarrollo de la agricultura sostenible]. *Environmental Science and Pollution Research*. ISSN: 0944-1344. DOI: 10.1007/s11356-016-8104-0

Medina-Cordoba, L., Chande, A., Rishishwar, L., Mayer, L., Valderrama-Aguirre, L., Gaby, J., Kostka, J., Jordan, I. (2021). *Scientific Reports*. [Informes Científicos]. ISSN: 2045-2322. DOI: 10.1038/s41598-021-88380-8

Meena, R. S., Kumar, S., Datta, R., Lal, R., Vijayakumar, V., Brtnicky, M., ... & Marfo, T. D. (2020). Impact of agrochemicals on soil microbiota and management: A review. [Impacto de los agroquímicos en la microbiota del suelo y su manejo: Una revisión]. *Land*, 9(2), 34.

Meyer, A. J., Segall-Shapiro, T. H., Glassey, E., Zhang, J., & Voigt, C. A. (2019). *Escherichia coli* "Marionette" strains with 12 highly optimized small-molecule sensors. [Cepas "Marioneta" de *Escherichia coli* con 12 sensores de moléculas pequeñas altamente optimizados]. *Nature chemical biology*, 15(2), 196-204.

Miethke, M., Klotz, O., Linne, U., May, J. J., Beckering, C. L., & Marahiel, M. A. (2006). Ferri-bacillibactin uptake and hydrolysis in *Bacillus subtilis*. [Absorción e hidrólisis de ferri-bacillobactina en *Bacillus subtilis*]. *Molecular Microbiology*, 61(6), 1413-1427. doi:10.1111/j.1365-2958.2006.05321.x

Mir, A., Edraki, A., Lee, J., Sontheimer, E. (2018). Type II-C CRISPR-Cas9 Biology, Mechanism, and Application. [Biología, mecanismo y aplicación del CRISPR-Cas9 tipo II-C]. *ACS Chemical Biology*. ISSN: 1554-8929. DOI: 10.1021/acscchembio.7b00855

Mishra, U., & Pabbi, S. (2004). Cyanobacteria : A potential biofertilizer for rice. [Cianobacterias: Un biofertilizante potencial para el arroz]. *Resonance*. ISSN: 0971-8044. DOI: 10.1007/BF02839213

Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D., & Gurevich, A. (2018). Versatile genome assembly evaluation with QUAST-LG. [Evaluación versátil del ensamblaje del genoma con QUAST-LG]. *Bioinformatics*, 34(13), i142-i150.

Moghaieb, R. E., Ahmed, D. S., Gaber, A., & Abdelhadi, A. A. (2021). Overexpression of bacterial *katE* gene improves the resistance of modified tomato plant against *Fusarium oxysporum* f. sp. *lycopersici*. [La sobreexpresión del gen bacteriano *katE* mejora la resistencia de la planta de tomate modificada contra *Fusarium oxysporum* f. sp. *lycopersici*]. *GM crops & food*, 12(1), 315-327.

Muraro, G. B., Carvalho-Estrada, P., Pasetti, M. H., Santos, M. C., & Nussio, L. G. (2021). Bacterial dynamics of sugarcane silage in the tropics. [Dinámica bacteriana del ensilaje de caña de azúcar en los trópicos]. *Environmental Microbiology*. doi:10.1111/1462-2920.15428

Mus, F., Alleman, A. B., Pence, N., Seefeldt, L. C., & Peters, J. W. (2018). Exploring the alternatives of biological nitrogen fixation. [Explorando las alternativas de la fijación biológica del nitrógeno]. *Metallomics*, 10(4), 523-538.

Nazari, M., & Smith, D. L. (2020). A PGPR-produced bacteriocin for sustainable agriculture: a review of thuricin 17 characteristics and applications. [Una bacteriocina producida por PGPR para la agricultura sostenible: una revisión de las características y aplicaciones de la turicina 17]. *Frontiers in plant science*, 11, 916.

Naveed, M., Qureshi, M. A., Zahir, Z. A., Hussain, M. B., Sessitsch, A., & Mitter, B. (2015). L-Tryptophan-dependent biosynthesis of indole-3-acetic acid (IAA) improves plant growth and colonization of maize by *Burkholderia phytofirmans* PsJN. [La biosíntesis de ácido indol-3-acético (AIA) dependiente de L-triptófano mejora el crecimiento de las plantas y la colonización del maíz por *Burkholderia phytofirmans* PsJN]. *Annals of microbiology*, 65, 1381-1389. DOI: 10.1007/s13213-015-1027-4

Nicoll, D. (2002). An introduction to genetic Engineering (Studies in Biology). [Una introducción

a la ingeniería genética (Estudios en Biología)]. 1-304. ISBN: 978-0-521-00727-1

Pandey, S., & Gupta, S. (2020). Evaluation of *Pseudomonas* sp. for its multifarious plant growth promoting potential and its ability to alleviate biotic and abiotic stress in tomato (*Solanum lycopersicum*) plants. [Evaluación de *Pseudomonas* sp. por su potencial multifacético para promover el crecimiento de las plantas y su capacidad para aliviar el estrés biótico y abiótico en plantas de tomate (*Solanum lycopersicum*)]. *Scientific Reports*, 10(1), 20951. DOI: 10.1038/s41598-020-77840-3

Pervez, M., Hasnain, M., Abbas, S., Moustafa, M., Aslam., Shah, S. (2022). A comprehensive Review of Performance of Next-Generation Sequencing Platforms. [Una revisión exhaustiva del rendimiento de las plataformas de secuenciación de próxima generación]. *BioMed Research International*. ISSN: 2314-6141. DOI: 10.1155/2022/3457806

Penwell, W. F., Arivett, B. A., & Actis, L. A. (2012). The *Acinetobacter baumannii* entA gene located outside the acinetobactin cluster is critical for siderophore production, iron acquisition and virulence. [El gen entA de *Acinetobacter baumannii* ubicado fuera del grupo de acinetobactina es crítico para la producción de sideróforos, la adquisición de hierro y la virulencia]. *PLoS one*, 7(5), e36493. DOI: 10.1371/journal.pone.0036493

Petersen TN, et al. SignalP 4.0: discriminating signal peptides from transmembrane regions. [SignalP 4.0: discriminando péptidos señal de regiones transmembrana]. *Nat. Methods*, 2011, vol. 8 (pg. 785-786) DOI: 10.1038/nmeth.1701

Phillips, P. L., & Schultz, G. S. (2012). Molecular mechanisms of biofilm infection: biofilm virulence factors. [Mecanismos moleculares de la infección por biopelículas: factores de virulencia de biopelículas]. *Advances in Wound Care*, 1(3), 109-114. DOI: 10.1089/wound.2011.0301

Prjibelski, A., Antipov, D., Meeleshko, D., Lapidus, A., Korobeynikov, A. (2020). Using SPAdes De Novo Assembler. [Usando el ensamblador de novo SPAdes]. *Current protocols in Bioinformatics*. ISSN: 1934-3396. DOI: 10.1002/cpbi.102

Qadir, M., Hussain, A., Hamayun, M., Shah, M., Iqbal, A., Irshad, M., ... & Lee, I. J. (2021). Phytohormones producing *Acinetobacter bouvetii* P1 mitigates chromate stress in sunflower by provoking host antioxidant response. [Acinetobacter bouvetii P1 productora de fitohormonas mitiga el estrés por cromato en girasol provocando la respuesta antioxidante del huésped]. *Antioxidants*, 10(12), 1868. DOI: 10.3390/antiox10121868

Ramakrishna, W., Yadav, R., Li, K. (2019). Plant growth promoting bacteria in agriculture: Two sides of a coin. [Bacterias promotoras del crecimiento vegetal en la agricultura: Dos caras de una moneda]. *Applied Soil Ecology*. ISSN: 09291393. DOI: 10.1016/j.apsoil.2019.02.019

Rana, K. L., Kour, D., Kaur, T., Sheikh, I., Yadav, A. N., Kumar, V., ... & Dhaliwal, H. S. (2020). Endophytic microbes from diverse wheat genotypes and their potential biotechnological applications in plant growth promotion and nutrient uptake. [Microbios endófitos de diversos genotipos de trigo y sus potenciales aplicaciones biotecnológicas en la promoción del crecimiento de las plantas y la absorción de nutrientes]. *Proceedings of the national academy of sciences, India section B: biological sciences*, 90, 969-979. DOI: 10.1007/s40011-020-01211-6

Raymond, K. N., Dertz, E. A., & Kim, S. S. (2003). Enterobactin: an archetype for microbial iron transport. [Enterobactina: un arquetipo para el transporte microbiano de hierro]. *Proceedings of the national academy of sciences*, 100(7), 3584-3588. DOI: 10.1073/pnas.0630018100

Redman, M., King, A., Watson, C., King, D. (2016). What is CRISPR/Cas9?. [¿Qué es CRISPR/Cas9?]. *Archives of disease in childhood – Education & practice edition*. ISSN: 1743-0585. DOI: 10.1136/archdischild-2016-310459

Rezazade Bazaz, M., & Dehghani, H. (2022). From DNA break repair pathways to

CRISPR/Cas-Mediated gene Knock-in Methods. [De las vías de reparación de roturas de ADN a los métodos de inserción de genes mediados por CRISPR/Cas]. ISSN: 00243205. DOI: 10.1016/j.lfs.2022.120409

Richmond, N. (2015). Chemical Soil degradation as a result of contamination: A review. [Degradación química del suelo como resultado de la contaminación: Una revisión]. *Journal of soil Sciences and Environmental Management*. ISSN: 2141-2391. DOI: 10.5897/JSSEM15.0499

Rieber, N., Zapatka, M., Lasitschka, B., Jones, D., Northcott, P., Hutter, B., ... & Eils, R. (2013). Coverage bias and sensitivity of variant calling for four whole-genome sequencing technologies. [Sesgo de cobertura y sensibilidad de la llamada de variantes para cuatro tecnologías de secuenciación del genoma completo]. *PloS one*, 8(6), e66621. DOI: 10.1371/journal.pone.0066621

Romero-Gutiérrez, K. J., Dourado, M. N., Garrido, L. M., Olchanheski, L. R., Mano, E. T., Dini-Andreote, F., ... & Araújo, W. L. (2020). Phenotypic traits of *Burkholderia* spp. associated with ecological adaptation and plant-host interaction. [Rasgos fenotípicos de *Burkholderia* spp. asociados con la adaptación ecológica y la interacción planta-huésped]. *Microbiological research*, 236, 126451. DOI: 10.1016/j.micres.2020.126451

Rubio, L. M., & Ludden, P. W. (2008). Biosynthesis of the Iron-Molybdenum Cofactor of Nitrogenase. [Biosíntesis del cofactor hierro-molibdeno de la nitrogenasa]. *Annual Review of Microbiology*, 62(1), 93–111. doi:10.1146/annurev.micro.62.083107.153215

Sánchez-Cañizares, C., Jorin, B., Poole, P., Tkacz, A. (2017). Current Option in Microbiology. [Opción actual en microbiología]. ISSN: 13695274. DOI: 10.1016/j.mib.2017.07.001

Santoyo, G., Orozco-Mosqueda, M. D. C., & Govindappa, M. (2012). Mechanisms of biocontrol and plant growth-promoting activity in soil bacterial species of *Bacillus* and *Pseudomonas*: a review. [Mecanismos de biocontrol y actividad promotora del crecimiento vegetal en especies bacterianas del suelo de *Bacillus* y *Pseudomonas*: una revisión]. *Biocontrol Science and Technology*, 22(8), 855-872. DOI: 10.1080/09583157.2012.694413

Sansinenea, E. (2019). *Bacillus* spp.: As plant growth-promoting bacteria. [Bacillus spp.: Como bacterias promotoras del crecimiento vegetal]. *Secondary metabolites of plant growth promoting rhizomicroorganisms: Discovery and applications*, 225-237. DOI: 10.1007/978-3-030-18933-4_10

Savci, S. (2012). Investigation of Effect of Chemical Fertilizers on Environment. [Investigación del efecto de los fertilizantes químicos en el medio ambiente]. *APCBEE*. ISSN: 22126708. DOI: 10.1016/j.apcbee.2012.03.047

Sastalla, I., Monack, D. M., & Kubatzky, K. F. (2016). Bacterial exotoxins: How bacteria fight the immune system. [Exotoxinas bacterianas: Cómo las bacterias luchan contra el sistema inmunológico]. *Frontiers in immunology*, 7, 300.

Sahu, P., & Brahma Prakash, G. (2016). Formulations of biofertilizers-Approaches and Advances. [Formulaciones de biofertilizantes: Enfoques y avances]. *Microbial inoculants in sustainable Agricultural Productivity*. DOI: 10.1007/978-81-322-2644-4_12

Sato, M. P., Ogura, Y., Nakamura, K., Nishida, R., Gotoh, Y., Hayashi, M., ... & Hayashi, T. (2019). Comparison of the sequencing bias of currently available library preparation kits for Illumina sequencing of bacterial genomes and metagenomes. [Comparación del sesgo de secuenciación de los kits de preparación de bibliotecas actualmente disponibles para la secuenciación Illumina de genomas bacterianos y metagenomas]. *DNA Research*, 26(5), 391-398. DOI: 10.1093/dnares/dsz014

Schütz, L., Gattinger, A., Meier, M., Müller, A., Boller, T., Mäder, P., Mathimaran, N. (2018). Improving Crop Yield and Nutrient Use Efficiency via Biofertilization -A global Meta-analysis. [Mejora del rendimiento de los cultivos y la eficiencia en el uso de nutrientes mediante biofertilización

- Un metaanálisis global]. *Frontiers in Plant Science*. ISSN: 1664-462X. DOI: 10.3389/fpls.2017.02204

Schuster, L. A., & Reisch, C. R. (2021). A plasmid toolbox for controlled gene expression across the Proteobacteria. [Una caja de herramientas de plásmidos para la expresión génica controlada en las Proteobacterias]. *Nucleic Acids Research*, 49(12), 7189-7202. DOI: 10.1093/nar/gkab524

Schenk zu Schweinsberg-Michan, M., & Müller, T. (2009). Impact of effective microorganisms and other biofertilizers on soil microbial characteristics, organic-matter decomposition, and plant growth. [Impacto de microorganismos efectivos y otros biofertilizantes en las características microbianas del suelo, la descomposición de materia orgánica y el crecimiento de las plantas]. *Journal of plant nutrition and soil science*. ISSN: 1436-8730. DOI: 10.1002/jpln.200800021

Schirmer, M., Ijaz, U. Z., D'Amore, R., Hall, N., Sloan, W. T., & Quince, C. (2015). Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. [Perspectiva sobre sesgos y errores de secuenciación para la secuenciación de amplicones con la plataforma Illumina MiSeq]. *Nucleic acids research*, 43(6), e37-e37. DOI: 10.1093/nar/gku1341

Schripsema, J., De Rudder, K. E., Van Vliet, T. B., Lankhorst, P. P., De Vroom, E., Kijne, J. W., & Van Brussel, A. A. (1996). Bacteriocin small of *Rhizobium leguminosarum* belongs to the class of N-acyl-L-homoserine lactone molecules, known as autoinducers and as quorum sensing co-transcription factors. [La bacteriocina pequeña de *Rhizobium leguminosarum* pertenece a la clase de moléculas de N-acil-L-homoserina lactona, conocidas como autoinductores y como factores de co-transcripción de detección de quórum]. *Journal of bacteriology*, 178(2), 366-371.

Schwartz, S. L., & Farman, M. L. (2010). Systematic overrepresentation of DNA termini and underrepresentation of subterminal regions among sequencing templates prepared from hydrodynamically sheared linear DNA molecules. [Sobrerrepresentación sistemática de los extremos del ADN y subrepresentación de regiones subterminales entre las plantillas de secuenciación preparadas a partir de moléculas de ADN lineal cortadas hidrodinamicamente]. *BMC genomics*, 11(1), 1-12.

Sewe, S., Silva, G., Sicat, P., Seal, S., Visendi, P. (2022). Trimming and validation of Illumina Short Reads Using Trimmomatic, Trinity Assembly, and Assessment of RNA-Seq Data. [Recorte y validación de lecturas cortas de Illumina usando Trimmomatic, ensamblaje Trinity y evaluación de datos de RNA-Seq]. DOI: 10.1007/978-1-0716-2067-0_11

Sergeeva, E., Hirkala, D. L., & Nelson, L. M. (2007). Production of indole-3-acetic acid, aromatic amino acid aminotransferase activities and plant growth promotion by *Pantoea agglomerans* rhizosphere isolates. [Producción de ácido indol-3-acético, actividades de aminotransferasa de aminoácidos aromáticos y promoción del crecimiento de plantas por aislados de la rizosfera de *Pantoea agglomerans*]. *Plant and Soil*, 297, 1-13.

Segond, D., Abi Khalil, E., Buisson, C., Daou, N., Kallassy, M., Lereclus, D., ... & Nielsen Le Roux, C. (2014). Iron acquisition in *Bacillus cereus*: the roles of IIsA and bacillibactin in exogenous ferritin iron mobilization. [Adquisición de hierro en *Bacillus cereus*: los roles de IIsA y bacilibactina en la movilización de hierro de ferritina exógena]. *PLoS Pathogens*, 10(2), e1003935. DOI: 10.1371/journal.ppat.1003935

Torsten Seemann, Prokka: rapid prokaryotic genome annotation. [Prokka: anotación rápida del genoma procariota]. *Bioinformatics*, Volume 30, Issue 14, July 2014, Pages 2068–2069, <https://doi.org/10.1093/bioinformatics/btu153>

Simpson, J., Jackman, S., Schein, J., Jones, S., Birol, I. (2009). ABySS: A parallel assembler for short read sequence data. [ABySS: Un ensamblador paralelo para datos de secuencia de lectura corta]. *Genome Research*. ISSN: 1088-9051. DOI: 10.1101/gr.089532.108

Silva, G., Dutilh, B., Matthews, T., Elkins, K., Schmieder, R., Dinsdale, E., Edwards, R. (2013).

Combining de novo and reference-guided assembly with scaffold_builder. [Combinando el ensamblaje de novo y guiado por referencia con scaffold_builder]. Source Code for biology and Medicine. ISSN: 1751-0473. DOI: 10.1186/1751-0473-8-23

Smith, M. A., & Bidochka, M. J. (1998). Bacterial fitness and plasmid loss: the importance of culture conditions and plasmid size. [Aptitud bacteriana y pérdida de plásmidos: la importancia de las condiciones de cultivo y el tamaño del plásmido]. *Canadian journal of microbiology*, 44(4), 351-355.

Strachan S., Bhuiyan, S., Thompson, N., Nguyen, N., Ford, R., Shiddiky, M. (2022). Latent potential of current plant diagnostics for detection of sugarcane diseases. [Potencial latente de los diagnósticos actuales de plantas para la detección de enfermedades de la caña de azúcar]. *Current Research in biotechnology*. ISSN: 25902628. DOI: 10.1016/j.crbiot.2022.10.002

Stock, F., Bilcke, G., De Decker, S., Osuna-Cruz, C. M., Van den Berge, K., Vancaester, E., ... & Vyverman, W. (2020). Distinctive growth and transcriptional changes of the diatom *Seminavis robusta* in response to quorum sensing related compounds. [Crecimiento distintivo y cambios transcripcionales de la diatomea *Seminavis robusta* en respuesta a compuestos relacionados con la detección de quórum]. *Frontiers in Microbiology*, 11, 1240. DOI: 10.3389/fmicb.2020.01240

Sontheimer, E., & Barrangou, R. (2015). The bacterial Origin of the CRISPR Genome-Editing. [El origen bacteriano de la edición del genoma CRISPR]. *Human Gene Therapy*. ISSN: 1043-0342. DOI: 10.1089/hum.2015.091

Sohn, J., & Nam, J. (2016). The Present and future of de Novo Whole-genome assembly. [El presente y futuro del ensamblaje de novo del genoma completo]. *Briefings in Bioinformatics*. ISSN: 1467-5463. DOI: 10.1093/bib/bbw096

Souza, E., Chubatsu, L., Huergo, L., Monteiro, R., Camilios-Neto, D., Wassem, R., De Oliveira Pedrosa, F. (2014). Use of Nitrogen-fixing bacteria to improve agricultural productivity. [Uso de bacterias fijadoras de nitrógeno para mejorar la productividad agrícola]. *BMC Proceedings*. ISSN: 1753-6561. DOI: 10.1186/1753-6561-8-S4-O23

Sokol, N., Slessarev, E., Marschmann, G., Nicolas, A., Blazewicz, S., Brodie, E., Firestone, M., Foley, M., Hestrin, R., Hungate, B., Koch, B., Stone, B., Sullivan, M., Zablocki, O., Trubl, G., McFarlane, K., Stuart, R., Nuccio, E., Weber, P., Jiao, Y., Zavarin, M., Kimbrel, J., Morrison, K., Adhikari, D., Bhattacharaya, A., Nico, P., Tang, J., Didonato, N., Paša-Tolić, L., Greenlon, A., Sieradzki, E., Dijkstra, P., Schwartz, E., Sachdeva, R., Banfield, J., Pett-Ridge J. (2022). Life and death in the soil microbiome: How ecological processes influence biogeochemistry. [Vida y muerte en el microbioma del suelo: Cómo los procesos ecológicos influyen en la biogeoquímica]. *Nature Reviews Microbiology*. ISSN: 1740-1526. DOI: 10.1038/s41579-022-00695-z

Sivasakthi, S., Usharani, G., & Saranraj, P. (2014). Biocontrol potentiality of plant growth promoting bacteria (PGPR)-*Pseudomonas fluorescens* and *Bacillus subtilis*: A review. [Potencial de biocontrol de las bacterias promotoras del crecimiento vegetal (PGPR)-*Pseudomonas fluorescens* y *Bacillus subtilis*: Una revisión]. *African journal of agricultural research*, 9(16), 1265-1277.

Singh, T. B., Sahai, V., Goyal, D., Prasad, M., Yadav, A., Shrivastav, P., ... Dantu, P. K. (2020). Identification, Characterization and Evaluation of Multifaceted Traits of Plant Growth Promoting Rhizobacteria from Soil for Sustainable Approach to Agriculture. [Identificación, caracterización y evaluación de rasgos multifacéticos de rizobacterias promotoras del crecimiento vegetal del suelo para un enfoque sostenible de la agricultura]. *Current Microbiology*, 77(11), 3633-3642. doi:10.1007/s00284-020-02165-2

SK, V. (2013). Pokkah Boeng: An Emergign disease of sugarcane. [Pokkah Boeng: Una enfermedad emergente de la caña de azúcar]. *Journal of plant pathology & microbiology*. ISSN: 21577471. DOI: 10.4172/2157-7471.1000170

Sparacino-Watkins, C., Stolz, J., Basu, P. (2014). Nitrate and periplasmic nitrate reductases. [Nitrato y nitrato reductasas periplásmicas]. *Chen. Soc. Rev.* ISSN: 0306-0012. DOI: 10.1039/C3CS60249D

Srivastav, A. (2020). Chemical fertilizers and pesticides: Role in groundwater contamination. [Fertilizantes químicos y pesticidas: Papel en la contaminación de las aguas subterráneas]. *Agrochemical Detection, Treatment and Remediation*. DOI: 10.1016/B978-0-08-103017-2.00006-4

Stüeken, E, Kipp, M., Koehler, M., Buick, R. (2016). The evolution of Earth's biogeochemical nitrogen cycle. [La evolución del ciclo biogeoquímico del nitrógeno de la Tierra]. *Earth-Science Reviews*. ISSN: 00128252. DOI: 10.1016/j.earscirev.2016.07.007

Tabatabaei, S., Ehsanzadeh, P., Etesami, H., Alikhani, H. A., & Glick, B. R. (2016). Indole-3-acetic acid (IAA) producing *Pseudomonas* isolates inhibit seed germination and α -amylase activity in durum wheat (*Triticum turgidum* L.). [Los aislados de *Pseudomonas* productores de ácido indol-3-acético (AIA) inhiben la germinación de semillas y la actividad de la α -amilasa en el trigo duro (*Triticum turgidum* L.)]. *Spanish Journal of Agricultural Research*, 14(1), 15.

Tahat, M., Alananbeh, K., Othman, A., Leskovar, D. Sustainability. [Sostenibilidad]. ISSN: 2071-1050. DOI: 10.3390/su12124859

Thomason L., Sawitzke, J., Li, X., Constantino, N., Court, D. (2014). Recombineering: Genetic Engineering in Bacteria Using Homologous Recombination. [Recombineering: Ingeniería genética en bacterias utilizando recombinación homóloga]. *Current Protocols in Molecular Biology*. ISSN: 1934-3639. DOI: 10.1002/0471142727.mb0116s106

Tripathi, S., Srivastava, P., Devi, R., Bhadouria, R. (2020). Influence of synthetic fertilizers and pesticides on soil health and soil microbiology. [Influencia de los fertilizantes sintéticos y pesticidas en la salud del suelo y la microbiología del suelo]. DOI: 10.1016/B978-0-08-103017-2.00002-7

Trüper, G., & Krämer, J. (1981). Principles of Characterization and identification of Prokaryotes. [Principios de caracterización e identificación de procariotas]. *The Prokaryotes*. DOI: 10.1007/978-3-662-13187-9_6

Umrit, G., Cheong, R., Gillabel, J., Merckx, R. (2014). Effect of conventional versus mechanized sugarcane cropping systems on soil organic carbon stocks and labile carbon pools in Mauritius as revealed by ^{13}C natural abundance. [Efecto de los sistemas de cultivo de caña de azúcar convencionales frente a los mecanizados en las reservas de carbono orgánico del suelo y los pools de carbono lábil en Mauricio revelado por la abundancia natural de ^{13}C]. *Plant and Soil*. ISSN: 0032-079X. DOI: 10.1007/s11104-014-2053-5

Vaiana, C. A., Kim, H., Cottet, J., Oai, K., Ge, Z., Conforti, K., ... & Buie, C. R. (2022). Characterizing chemical signaling between engineered "microbial sentinels" in porous microplates. [Caracterización de la señalización química entre centinelas microbianas diseñadas en microplacas porosas]. *Molecular Systems Biology*, 18(3), e10785.

Van Dijk, E. L., Jaszczyszyn, Y., & Thermes, C. (2014). Library preparation methods for next-generation sequencing: tone down the bias. [Métodos de preparación de bibliotecas para la secuenciación de próxima generación: reducir el sesgo]. *Experimental cell research*, 322(1), 12-20.

Van Nuland, M., Ware, I., Schadt, C., Yang, Z., Bailey, J., Schweitzer, J. (2021). Natural soil microbiome Variation affects spring foliar phenology with consequences for plant productivity and climate-driven range shifts. [La variación natural del microbioma del suelo afecta la fenología foliar de primavera con consecuencias para la productividad de las plantas y los cambios de rango impulsados por el clima]. *New phytologist*. ISSN: 0028-646X. DOI: 10.1111/nph.17599

Vega-Celedón, P., Canchignia Martínez, H., González, M., & Seeger, M. (2016). Biosynthesis of

indole-3-acetic acid and plant growth promoting by bacteria. [Biosíntesis de ácido indol-3-acético y promoción del crecimiento vegetal por bacterias]. *Cultiv. Trop*, 37, 33-39.

Wang, Y., Ye, X., Ding, G., & Xu, F. (2013). Overexpression of *phyA* and *appA* genes improves soil organic phosphorus utilisation and seed phytase activity in *Brassica napus*. [La sobreexpresión de los genes *phyA* y *appA* mejora la utilización del fósforo orgánico del suelo y la actividad de la fitasa de las semillas en *Brassica napus*]. *PloS one*, 8(4), e60801.

Wang, Z., Yu, Z., Solanki, M. K., Yang, L., Xing, Y., Dong, D., & Li, Y. (2020). Diversity of sugarcane root-associated endophytic *Bacillus* and their activities in enhancing plant growth. [Diversidad de *Bacillus* endofíticos asociados a la raíz de la caña de azúcar y sus actividades en la mejora del crecimiento de las plantas]. *Journal of Applied Microbiology*. doi:10.1111/jam.14512

Wei, Z., Friman, Y. Kowalchuk, G., Xu, Y., Shen, Q., Jousset, A. (2019). Initial soil microbiome composition and functioning predetermine future plant health. [La composición y el funcionamiento inicial del microbioma del suelo predeterminan la salud futura de las plantas]. *Science advances*. ISSN: 2375-2548. DOI: 10.1126/sciadv.aaw0759

Wu, H. J., Wang, A. H., & Jennings, M. P. (2008). Discovery of virulence factors of pathogenic bacteria. [Descubrimiento de factores de virulencia de bacterias patógenas]. *Current opinion in chemical biology*, 12(1), 93-101.

Xie, J., Sun, X., Xia, R., Zhang, R., Shen, Q., Xu, Z., & Kovács, Á. T. (2023). *Burkholderia* Contaminans Induces the Bacillibactin Production of *Bacillus velezensis* and Initiates a Positive Metabolic Cycle in the Tri-Species Biofilms. [Burkholderia Contaminans induce la producción de bacillibactina de *Bacillus velezensis* e inicia un ciclo metabólico positivo en las biopelículas de tres especies]. In 16th Symposium on Bacterial Genetics and Ecology.

Xu, Q., Ling, N., Chen, H., Duan, Y., Wang, S., Shen, Q., Vandenkoornhuyse, P. (2020). Long-Term chemical-only Fertilization induces a diversity Decline and deep selection on the soil Bacteria. [La fertilización a largo plazo solo con productos químicos induce una disminución de la diversidad y una selección profunda en las bacterias del suelo]. *mSystems*. ISSN: 2379-5077. DOI: 10.1128/mSystems.00337-20

Yajima, A., van Brussel, A. A., Schripsema, J., Nukada, T., & Yabuta, G. (2008). Synthesis and stereochemistry-activity relationship of small bacteriocin, an autoinducer of the symbiotic nitrogen-fixing bacterium *Rhizobium leguminosarum*. [Síntesis y relación estereoquímica-actividad de la pequeña bacteriocina, un autoinductor de la bacteria simbiótica fijadora de nitrógeno *Rhizobium leguminosarum*]. *Organic letters*, 10(10), 2047-2050.

Yao, Z., Zou, C., Peng, C., Zhu, Y., Bao, Y., Zhou, Q., Wu, Q., Chen, B., Zhang, M. (2020). Virome Identification and characterization of *Fusarium sacchari* and *F. andiyazi*: Causative Agents of Pokkah Boeng Disease in sugarcane. [Identificación y caracterización del viroma de *Fusarium sacchari* y *F. andiyazi*: Agentes causales de la enfermedad Pokkah Boeng en la caña de azúcar]. *Frontiers in Microbiology*. ISSN: 1664-302X. DOI: 10.3389/fmicb.2020.00240

Youard, Z. A., Wenner, N., & Reimann, C. (2010). Iron acquisition with the natural siderophore enantiomers pyochelin and enantio-pyochelin in *Pseudomonas* species. [Adquisición de hierro con los enantiómeros naturales de sideróforos pioquelina y enantio-pioquelina en especies de *Pseudomonas*]. *BioMetals*, 24(3), 513-522. doi:10.1007/s10534-010-9399-9

Zaidi, A., Khan, M., Ahemad, M., Oves, M. (2009). Plant Growth promotion by phosphate solubilizing bacteria. [Promoción del crecimiento vegetal por bacterias solubilizadoras de fosfato]. *Acta Microbiologica et immunologica hungarica*. ISSN: 1217-8950. DOI: 10.1556/AMicr.56.2009.3.6

Zhang, S., Shen, J., Li, D., & Cheng, Y. (2021). Strategies in the delivery of Cas9 ribonucleoprotein for CRISPR/Cas9 genome editing. [Estrategias en la entrega de ribonucleoproteína

Cas9 para la edición del genoma CRISPR/Cas9]. *Theranostics*, 11(2), 614.

Zhang, W., Chen, Y., Huang, K., Wang, F., & Mei, Z. (2023). Molecular Mechanism and Agricultural Application of the NifA–NifL System for Nitrogen Fixation. [Mecanismo molecular y aplicación agrícola del sistema NifA-NifL para la fijación de nitrógeno]. *International Journal of Molecular Sciences*, 24(2), 907. <https://doi.org/10.3390/ijms24020907>

Anexo A Análisis bioinformático

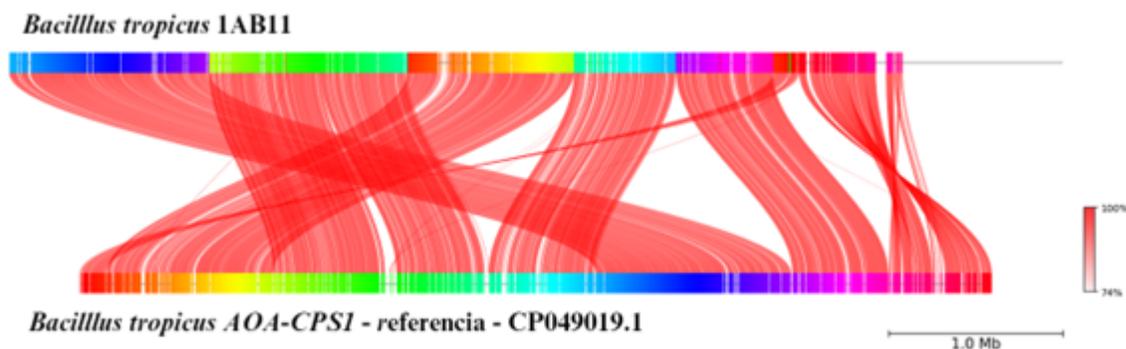
La sección de Anexos A proporciona información específica de identificación de cada una de las especies secuenciadas a través de NGS. Los valores fueron determinados a través de *GTDB-Tk*, en los cuales se ve el genoma de referencia más cercano, se muestra el algoritmo de inferencia y los valores de proximidad, Asimismo, se muestran los diagramas de alineamiento entre cada una de las muestras con su respectivo genoma de referencia a través del algoritmo *FastANI*.

Cuadro 16
Identificación con GTDB-Tk de la muestra 1AB11

Muestra	Clasificación	Referencia de fastANI	Radio de referencia de fastANI	Taxonomía de fastANI	fastANI_anl	fastANI_of	método de clasificación	Nota	Otras referencias relacionadas	porcentaje MSA	valor REE	alete	Cobertura
1AB11	d_Bacetiop_Baciltoz c_Bacillo_Baciltoef Bacillaceae_Gg_Bac Btu_As_Bacillu_A tropicus	GCF_001114035	93.3838	d_Bacetiop_Baciltoz Bacillo_Baciltoef Bacillaceae_Gg_Bacillu As_Bacillu_A tropicus	96.05	0.914	ani screen	classification based on ANI only	GCF_001113995.1 Btu_Bacillu_A paratropic	N/A	N/A	N/A	

Nota. Elaboración propia

Figura 25
Alineamiento de B. tropicus 1AB11 con genoma de referencia por FastANI



Nota. Elaboración propia

Cuadro 17
Identificación con GTDB-Tk de la muestra 1AP23

Muestra	Clasificación	Referencia de fastANI	Ratio de referencia de fastANI	Taxonomía de fastANI	fastANI_ani	fastANI_af	método de clasificación	Nota	Otras referencias relacionadas	porcentaje MS	valor RED	alerta	Cobertura
1AP23	d_Bacteria_p_Pseudomonadota; c_Gammaproteobacteria_o_Pseudomonadales_f_Pseudomonadales_g_Pseudomonas_Es_Pseudomonas_E_sp003445295	GCF_007677	725.1	95	d_Bacteria_p_Pseudomonadota; c_Gammaproteobacteria_o_Pseudomonadales_f_Pseudomonadales_g_Pseudomonas_Es_Pseudomonas_E_sp003445295	98.7	0.94	ani_screen	classification based on ANI	N/A	N/A	N/A	N/A

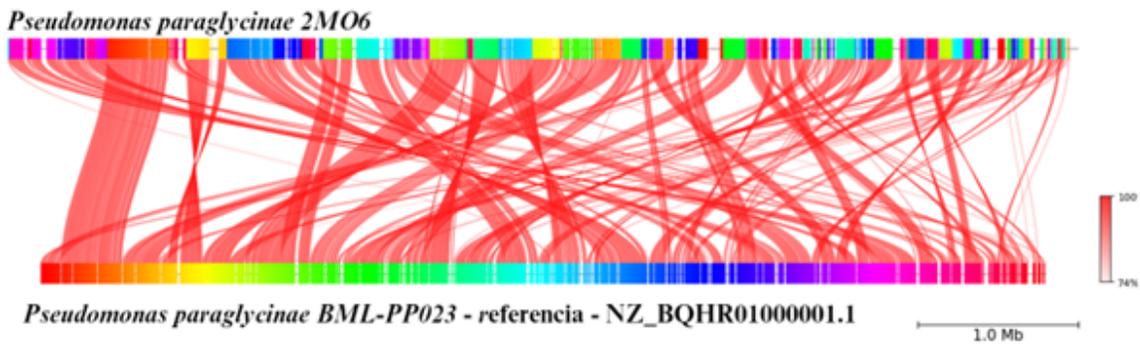
Nota. Elaboración propia

Cuadro 18
Identificación con GTDB-Tk de la muestra 2MO6

Muestra	Clasificación	Referencia de fastANI	Ratio de referencia de fastANI	Taxonomía de fastANI	fastANI_ani	fastANI_af	método de clasificación	Nota	Otras referencias relacionadas	porcentaje MS	valor RED	alerta	Cobertura
2MO6	d_Bacteria_p_Pseudomonadota; c_Gammaproteobacteria_o_Pseudomonadales_f_Pseudomonadales_g_Pseudomonas_Es_Pseudomonas_E_sp021601	GCF_021601	1625.1	95	d_Bacteria_p_Pseudomonadota; c_Gammaproteobacteria_o_Pseudomonadales_f_Pseudomonadales_g_Pseudomonas_Es_Pseudomonas_E_sp021601	95.77	0.917	ani_screen	classification based on ANI only	N/A	N/A	N/A	N/A

Nota. Elaboración propia

Figura 26
Alineamiento de *P. paraglycinae* 2MO6 con genoma de referencia por FastANI



Nota. Elaboración propia

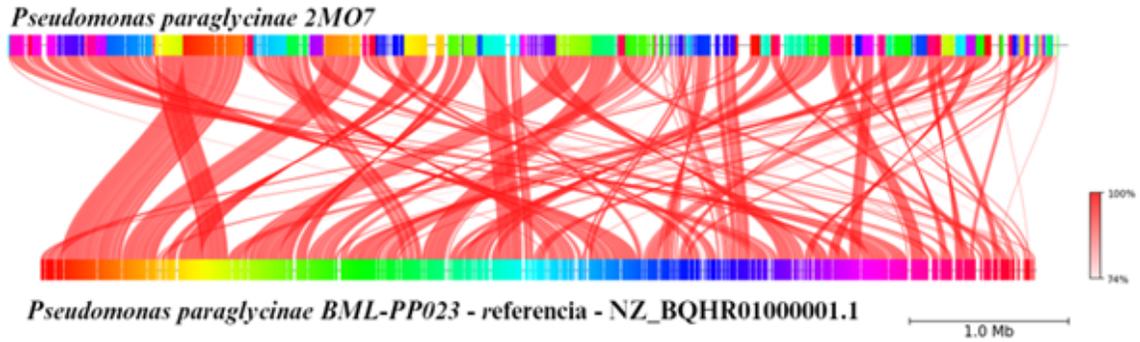
Cuadro 19
Identificación con GTDB-Tk de la muestra 2MO7

Muestra	Clasificación	Referencia de fastANI	Ratio de referencia de fastANI	Taxonomía de fastANI	fastANI_ani	fastANI_af	método de clasificación	Nota	Otras referencias relacionadas	porcentaje MS	valor RED	alerta	Cobertura
2MO7	d_Bacteria_p_Pseudomonadota; c_Gammaproteobacteria_o_Pseudomonadales_f_Pseudomonadales_g_Pseudomonas_Es_Pseudomonas_E_sp021601	GCF_021601	625.1	95	d_Bacteria_p_Pseudomonadota; c_Gammaproteobacteria_o_Pseudomonadales_f_Pseudomonadales_g_Pseudomonas_Es_Pseudomonas_E_sp021601	95.79	0.918	ani_screen	classification based on ANI only	5,95.0	N/A	N/A	78.92

Nota. Elaboración propia

Figura 27

Alineamiento de *P. paraglycinae* 2MO7 con genoma de referencia por FastANI



Nota. Elaboración propia

Cuadro 20

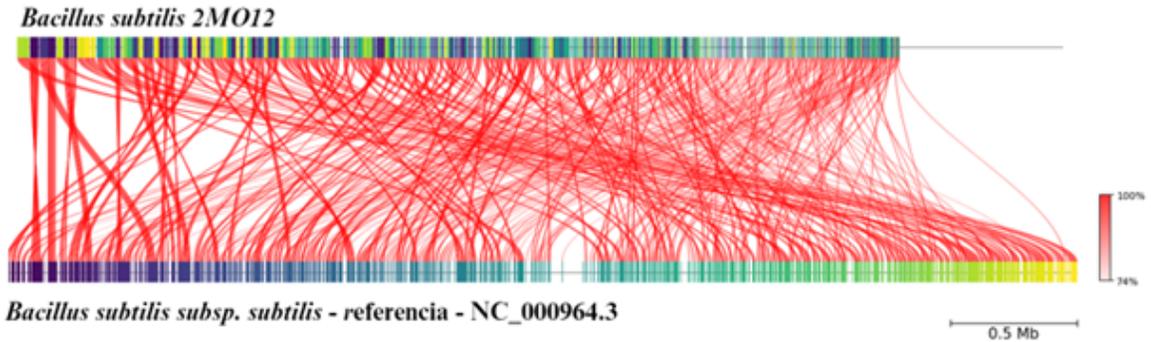
Identificación con GTDB-Tk de la muestra 2MO12

Muestra	Clasificación	Referencia de fastANI	Radio de referencia de fastANI	Taxonomía de fastANI	fastANI_ani	fastANI_af	método de clasificación	Nota	Otras referencias relacionadas	porcentaje MS	valor RED	alerta	Cobertura
2MO12	d__Bacteria; p__Bacillota; c__Bacilli; o__Bacillales; f__Bacillales; g__Bacillus; s__Bacillus subtilis	GCF_000009045.1	95	d__Bacteria; p__Bacillota; c__Bacilli; o__Bacillales; f__Bacillales; g__Bacillus; s__Bacillus subtilis	99.55	0.959	ani_screen	classification based on ANI only	N/A	N/A	N/A	N/A	

Nota. Elaboración propia

Figura 28

Alineamiento de *B. subtilis* 2MO12 con genoma de referencia por FastANI



Nota. Elaboración propia

Cuadro 21

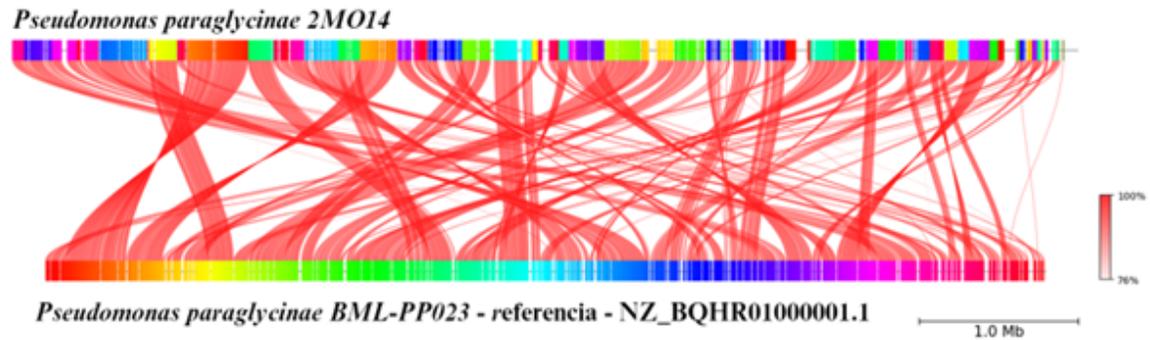
Identificación con GTDB-Tk de la muestra 2MO14

Muestra	Clasificación	Referencia de fastANI	Radio de referencia de fastANI	Taxonomía de fastANI	fastANI_ani	fastANI_af	método de clasificación	Nota	Otras referencias relacionadas	porcentaje MS	valor RED	alerta	Cobertura
2MO14	d__Bacteria; p__Pseudomonadota; c__X__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadales; g__Pseudomonas; s__Pseudomonas_E__paraglycinae	GCF_021909325.1	95	d__Bacteria; p__Pseudomonadota; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadales; g__Pseudomonas; s__Pseudomonas_E__paraglycinae	95.83	0.911	ani_screen	only	GCF_021918765.1, GCF_021918765.1, classification based on ANI	95.0-95.03-0.887	N/A	N/A	N/A

Nota. Elaboración propia

Figura 29

Alineamiento de *P. paraglycinae* 2MO14 con genoma de referencia por FastANI



Nota. Elaboración propia

Cuadro 22

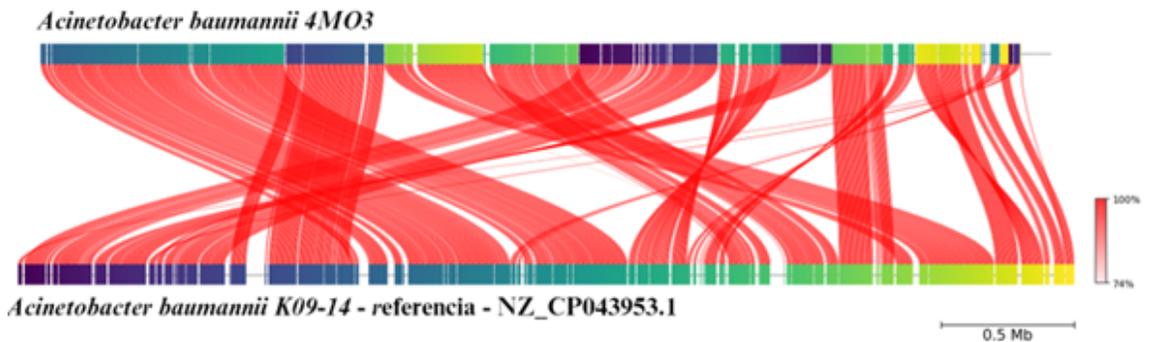
Identificación con GTDB-Tk de la muestra 4MO3

Muestra	Clasificación	Referencia de fastANI	Ratio de referencia de fastANI	Taxonomía de fastANI	fastANI_ani	fastANI_af	método de clasificación	Nota	Otras referencias relacionadas	porcentaje	MS	valor RED	alerta	Cobertura
4MO3	d__Bacteria;_Pseudomonadota;_Gammaproteobacteria;_Pseudomonadales;_Moraxellaceae;_Acinetobacter;_Acinetobacter baumannii	GCF_00075_0655.1	95	d__Bacteria;_Pseudomonadota;_Gammaproteobacteria;_Pseudomonadales;_Moraxellaceae;_Acinetobacter;_Acinetobacter baumannii	97.96	0.934	ani_screen	classification based on ANI only	N/A	N/A	N/A	N/A	N/A	N/A

Nota. Elaboración propia

Figura 30

Alineamiento de *A. baumannii* 4MO3 con genoma de referencia por FastANI



Nota. Elaboración propia

Cuadro 23

Identificación con GTDB-Tk de la muestra 4MO3

Muestra	Clasificación	Referencia de fastANI	Ratio de referencia de fastANI	Taxonomía de fastANI	fastANI_ani	fastANI_af	método de clasificación	Nota	Otras referencias relacionadas	porcentaje	MS	valor RED	alerta	Cobertura
2AP13	d__Bacteria;_Pseudomonadota;_Gammaproteobacteria;_Pseudomonadales;_Pseudomonadaceae;_Pseudomonas;_Escherichia;_Pseudomonas_Escherichia coli	GCF_000503_215.1	95	d__Bacteria;_Pseudomonadota;_Gammaproteobacteria;_Pseudomonadales;_Pseudomonadaceae;_Pseudomonas;_Escherichia;_Pseudomonas_Escherichia coli	99.94	0.963	ani_screen	classification based on ANI only	N/A	N/A	N/A	N/A	N/A	N/A

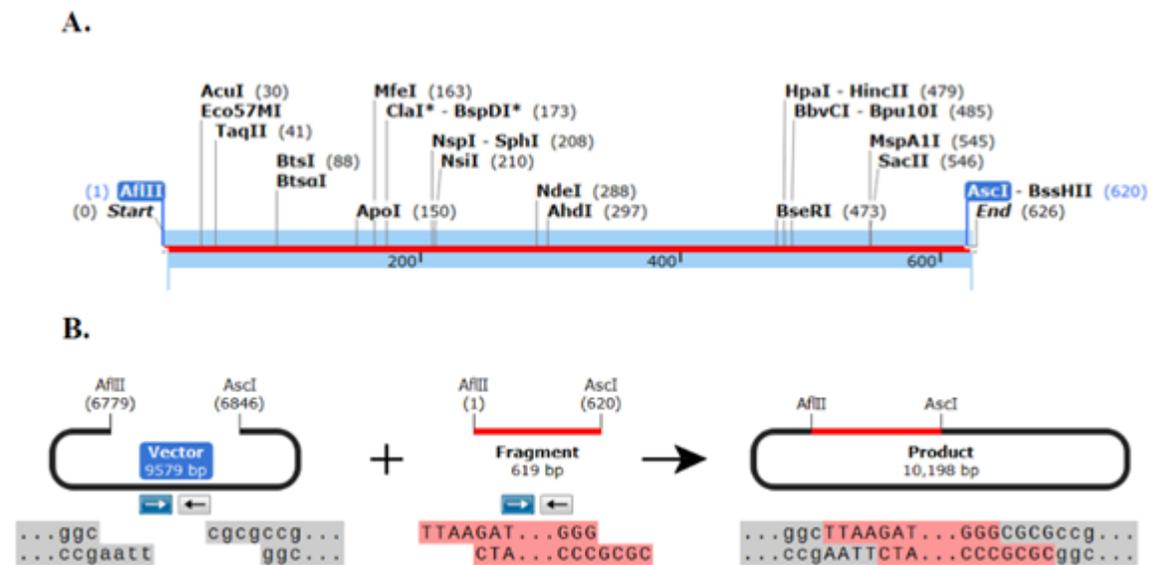
Nota. Elaboración propia

Anexo B Sistema de clonación

En el anexo B, se muestra cómo se realizará la clonación del gen regulador transcripcional, *nifA*, en cada uno de los plásmidos inducibles para la transformación bacteriana. Se muestran las regiones de corte, así como también las enzimas de restricción propuestas y las regiones de ligación. El diagrama fue generado a través del software *Snappgene*.

Figura 33

Propuesta de clonación del gen nifA en los plásmidos inducibles



Nota. Elaboración propia

Anexo C Metadatos de secuenciación

En el anexo C se muestran los valores crudos de cada una de las muestras extraídas de ADN de las bacterias candidatas. Se muestra el OD600 estándar para cada una de las extracciones, la concentración de ADN genómico medido en el *Nanodrop*, la pureza en relación A260/280 y A260/230, por último se muestran también la concentración de ADN genómico medida en el fluorómetro *Quantus*.

Se muestran los valores de estandarización para las bibliotecas genómicas y cuanto espacio del cartucho de secuenciación representa cada muestra. Asimismo, se muestran los gels de calidad de las bibliotecas para la evaluación de la molaridad de cada una de las muestras.

Cuadro 26*Pureza y concentración para las bibliotecas genómicas*

ID extracción	Cepa	OD 600	Nanodrop [ng/uL]	A260/280	A260/230	Quantus [ng/uL]
1	1AB11	NA	7,30	1,77	1,58	0,706
2	1AB21	NA	7,70	2,09	1,96	0,661
3	1AP23	NA	12,40	1,88	2,01	4,37
4	2MO6	NA	21,60	1,93	1,99	6,5
5	2MO7	NA	20,70	1,94	2,13	6,2
6	2MO12	NA	17,30	1,51	0,71	1,71
7	2MO14	NA	19,10	2,04	1,12	5,5
8	4MO3	NA	7,00	1,65	1,29	3,79
9	2AP13	NA	15,50	2,02	2,27	7
10	2AP16	NA	15,80	1,85	1,51	7

ID extracción	Cepa	OD 600	Nanodrop [ng/uL]	A260/280	A260/230	Quantus [ng/uL]
11	1AB11	2.61	172.5	2.24	2	9,6
12	1AB21	2.7	282.3	2.26	2.21	6
13	1AP23	0.85	87.6	2.15	2.43	9,9
15	2MO7	1.13	188.5	2.01	2.29	38
16	2MO12	0.54	103.2	1.53	0.62	11
17	2MO14	1.13	179.4	1.98	2.2	57
18	4MO3	1.76	256.3	1.78	1.71	38
19	2AP13	1.49	124.2	1.92	1.37	53
20	2AP16	1.77	101	1.99	2.06	48

Nota. Elaboración propia

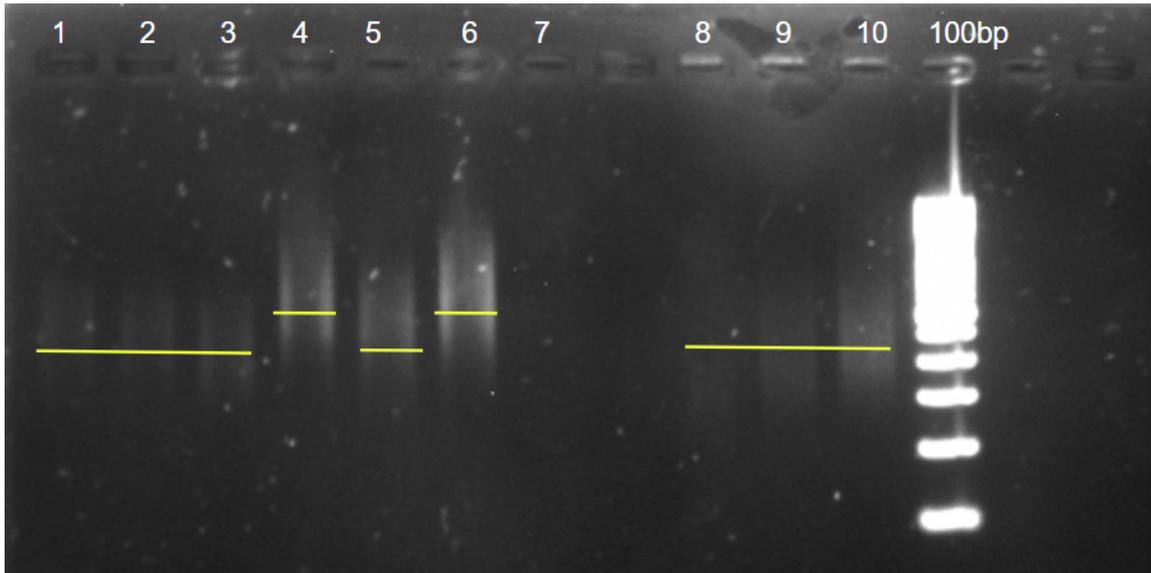
Cuadro 27*Preparación de bibliotecas genómicas*

ID extracción	Concentración en 30uL	Volumen mx para obtener 100ng	Volumen H2O para llegar a 30ul	Bibliotecas ng/ul	Tamaño Mpb
1	21,18	141,6	-111,6	1,030	6
2	19,83	151,3	-121,3	0,843	
3	131,1	22,9	7,1	1,430	
4	195	15,4	14,6	3,88	4
5	186	16,1	13,9	2,92	4
6	51,3	29,2	0,8	6,6	
7	165	18,2	11,8	0,747	
8	113,7	26,4	3,6	0,800	
9	210	14,3	15,7	1,710	
10	210	14,3	15,7	3,180	
ID extracción	Concentración en 30uL	Volumen mx para obtener 100ng	Volumen H2O para llegar a 30ul	Bibliotecas ng/ul	Tamaño Mpb
11	288	20,8	9,2	0,221	
12	180	30,0	0,0	4,17	6
13	297	20,2	9,8	2,72	6
15	1140	5,3	24,7	0,249	
16	330	18,2	11,8	0,316	4
17	1710	3,5	26,5	2,7	9
18	1140	5,3	24,7	1,39	4
19	1590	3,8	26,2	2,88	3
20	1440	4,2	25,8	1,69	3

Nota. Elaboración propia

Figura 34

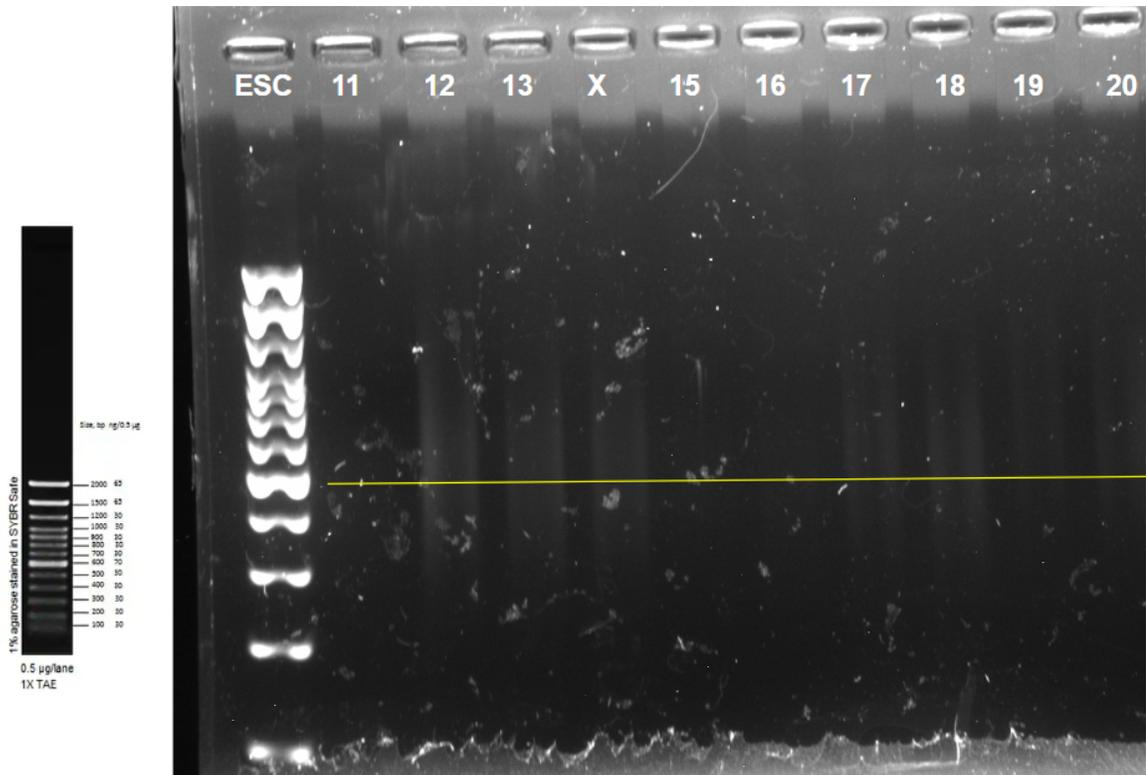
Calidad de bibliotecas de las muestras 1-10



Nota. Elaboración propia

Figura 35

Calidad de bibliotecas de las muestras 11-20



Nota. Elaboración propia

Anexo D Protocolos

Procedimiento Normado de Operación	SOP 19 Proyecto Suelos Ingenio Magdalena
	Páginas Totales:
Extracción de ADN bacteriano con Qiaamp(r) DNA mini and blood Mini Handbook	Autores: Andrés Say-Agosto, Mellissa Caballeros, Isabella García Caffaro.

1. Propósito

Llevar a cabo una extracción de ADN genómico de bacterias en suspensión para análisis superiores como PCR y secuenciación.

2. Aplicación

La extracción de ADN tiene diversas aplicaciones, esta permite llevar a cabo estudios moleculares para la identificación de marcadores específicos a través de PCR. Se puede utilizar los resultados para evaluar la presencia de un gen, en ADN comunal o ambiental puede amplificarse un gen conservado para llevar a cabo *DNA barcoding* o metagenómica. Asimismo, se utiliza para llevar a cabo secuenciación de genoma completo de bacterias aisladas.

3. Referencias

N/A

4. Terminología

- 1.1. WB1: *Wash buffer 1*
- 1.2. WB2: *Wash buffer 2*
- 1.3.

5. Documentos asociados

- 1.4. Qiagen. (s.f.). QIAamp® DNA mini and blood Mini Handbook.
<https://www.qiagen.com/us/resources/resourcedetail?id=62a200d6-faf4-469b-b50f-2b59cf738962&lang=en>
- 1.5. Quantifluor(R) dsDNA System. <https://worldwide.promega.com/es-es/products/rna-analysis/dna-and-rna-quantitation/quantifluor-dsdna-system/?catNum=E2671>

6. Seguridad**1.6. Debe de tomarse condiciones especiales en caso de:**

- 6.1.1. Formación de aerosoles.
- 6.1.2. Trabajo con grandes cantidades o concentraciones de microorganismos. Concentraciones elevadas se necesita trabajar en campana de flujo laminar para evitar propagación de aerosoles.
- 6.1.3. Exceso de material o personal.

1.7. Aplicar prácticas estandarizadas en laboratorio

- 6.1.4. Lavado frecuente de manos, especialmente después de quitarse los guantes y antes de salir del laboratorio.
- 6.1.5. Mantener una puerta cerrada mientras se trabaja.
- 6.1.6. Limitar el acceso al laboratorio.
- 6.1.7. No fumar, comer, beber, o almacenar alimentos en el laboratorio.
- 6.1.8. Colocar especial atención en condiciones en las que se puedan crear salpicaduras y aerosoles .
- 6.1.9. Descontaminar superficies de trabajo después de uso y después de cualquier derrame.
- 6.1.10. Descontaminar desechos del laboratorio o mandarlos a planta de tratamiento.
- 6.1.11. Utilizar pipetas mecánicas (no usar la pipeta por medio de succión oral).
- 6.1.12. Ser precavido al usar objetos punzantes.
- 6.1.13. Mantenimiento de un programa de control de insectos y roedores.

- 6.1.14. Uso de equipo de protección personal (bata de laboratorio, guantes de látex y protección para los ojos o máscaras para el rostro).

7. Equipo, materiales y reactivos

1.8. Equipo

- 7.1.1. Fluorómetro Quantus™ Fluorometer Promega
 7.1.2. Centrifugadora
 7.1.3. Microcentrifuga
 7.1.4. Vórtex
 7.1.5. Baño maría

1.9. Materiales/ Cristalería

- 7.1.6. Micropipetas p10
 7.1.7. Micropipetas p100
 7.1.8. Micropipetas p1000
 7.1.9. Puntas de micropipetas
 7.1.10. Microtubos de 0.2 mL
 7.1.11. Microtubos de 0.6 mL
 7.1.12. Microtubos de 1.5 mL

1.10. Reactivos

- 7.1.13. ATL
 7.1.14. AL
 7.1.15. WB1
 7.1.16. WB2
 7.1.17. Etanol absoluto
 7.1.18. Agarosa
 7.1.19. Buffer TBE 1X
 7.1.20. *TrackIt™ 100 bp DNA Ladder*
 7.1.21. *GelRed Nucleic Acid Stain 1000x DMSO*

1.11. Medicamentos

- 7.1.22. N/A

1.12. Material biológico

- 7.1.23. ADN extraído de cepas bacterianas

8. Procedimiento

- 1.13. Preparación del cultivo en medio líquido
 1.13.1. Inocular una colonia aislada en agua peptonada y dejar incubar ON a 37°C con 1,5000 rpm de agitación
- 1.14. Extracción de ADN
 1.14.1. Agregar 1 mL de la suspensión bacteriana ON a un microtubo de 1.5 mL y centrifugar a 7,500 rpm/5000g
 1.14.2. Agregar 180 uL de Buffer ATL
 1.14.3. Agregar 20 uL de proteinasa K y vortexear. Incubar a 56°C durante 3h mezclando por vortex (mínimo) tres veces cada hora
 1.14.4. Realizar un spin-down de los microtubos de 1.5 mL
 1.14.5. Dependiendo si se quiere ADN genómico libre de ARN realizar lo siguiente:
 1.14.5.1. LIBRE de ARN: Agregar 4 uL de RNasa A 100 mg/mL y vortexear por 15s. Posteriormente, incubar a temperatura ambiente por 2 minutos. Realizar un spindown y agregar 200 uL de buffer AL. Mezclar por vortex e incubar a 70°C por 10 minutos. Realizar un spin down
 1.14.5.2. NO libre de ARN: Agregar 200 uL de buffer AL, vortexear por 15s e incubar a 70°C por 10 minutos. Realizar un spin down.
 1.14.6. Agregar 200 uL de etanol absoluto y vortexear por 15s. realizar un spin down
 1.14.7. Agregar todo el contenido de los microtubos a una columna de extracción y centrifugar a 8,000 rpm/6000g por 1 minuto. Descartar el tubo de colección y colocar la columna en un tubo nuevo
 1.14.8. Agregar 400 uL de Buffer AW1 (evitar la formación de espuma) y centrifugar a 8,000 rpm/6,000g por 1 minuto. Descartar el tubo de colección y colocar la columna en un nuevo tubo

- 1.14.9. Agregar 500 uL de Buffer AW2 y centrifugar a 14,000 rpm/20,000g por 3 minutos
- 1.14.10. Descartar el tubo de colección y colocar la columna en un nuevo tubo y centrifugar a velocidad máxima por 1 minuto
- 1.14.11. Colocar la columna en un microtubo de 1.5 mL y agregar 50 uL de Buffer AE, posteriormente, incubar por 1 minutos a temperatura ambiente y centrifugar a 8,000 rpm/6,000g 12. Repetir el paso anterior
- 1.15. Medir la concentración utilizando Quantus
 - 1.15.1. Temperar los reactivos de Quantifluor(R) dsDNA system
 - 1.15.2. Calibrar el equipo preparando dos microtubos de 0.6 Quantifluor con lo siguiente:
 - 1.15.2.1. Para el blanco, 200 uL del reactivo Quantifluor dsDNA System
 - 1.15.2.2. Para la referencia, 198 uL del reactivo Quantifluor dsDNA System y 2 uL del reactivo lambda
 - 1.15.3. Realizar un vortex a las muestras y un *Spin-Down*
 - 1.15.4. Incubar las muestras por 10 minutos a oscuridad
 - 1.15.5. Realizar un vortex a las muestras y un *Spin-Down*
 - 1.15.6. Realizar la medición del blanco y la referencia
 - 1.15.7. Preparar las muestras de ADN agregando en un microtubo de 0.6 Quantifluor 198 uL del reactivo Quantifluor dsDNA System y 2 uL de la muestra de ADN
 - 1.15.8. Vortex y *Spin-Down*
 - 1.15.9. Incubar por 10 minutos a oscuridad
 - 1.15.10. Vortex y *Spin-Down*
 - 1.15.11. Realizar las mediciones

Procedimiento Normado de Operación	SOP 18 Proyecto Suelos Ingenio Magdalena
	Páginas Totales:
Amplificación del gen 16S	Autores: Andrés Say-Agosto, Mellissa Caballeros, Isabella García Caffaro.

1. Propósito

Realizar una PCR para la amplificación del gen ARN ribosomal *16S* de bacterias para su posterior identificación a través de secuenciación.

2. Aplicación

La reacción en cadena de la polimerasa (PCR) es una tecnología empleada para amplificar una determinada secuencia de ADN, lo que facilita el aislamiento, la secuenciación o la clonación del segmento amplificado. La PCR encuentra aplicaciones en diversos estudios, especialmente en microbiología, abarcando la clonación de ADN, Southern blotting, secuenciación de ADN e innovaciones en la tecnología del ADN recombinante. También se utiliza en microbiología clínica para identificar bacterias patógenas y en estudios forenses para la identificación de criminales.

Las bacterias pueden identificarse mediante PCR para amplificar las regiones 16S y/o 18S, mientras que los hongos filamentosos y las levaduras se identifican a través de la región ITS. Esta técnica ampliamente utilizada ayuda a la identificación de cepas, normalmente a nivel de género. El proceso de identificación se basa en el análisis de una o dos regiones del ADN ribosómico. Por ejemplo, en la secuenciación 16S, se secuencian los primeros 1.000 nucleótidos del gen, incluidas las zonas hipervariables con la mayoría de las diferencias nucleotídicas entre especies. Al comparar las secuencias obtenidas con las de las bases de datos públicas, las cepas pueden asignarse con seguridad a una especie específica si el porcentaje de homología es del 99% o superior.

3. Referencias

N/A

4. Terminología

1.1. PCR: Reacción en cadena de la Polimerasa

1.2. Primer: Secuencia corta de nucleótidos complementarios a la región de ADN que se desea amplificar.

5. Documentos asociados

1.3. Promega (s.f.). GoTaq(r) G2 Hot Start Colorless Master MX Protocol.

https://www.promega.com/-/media/files/resources/protocols/product-information-sheets/g/gotaq-g2-hot-start-colorless-master-mix-protocol.pdf?rev=cc4ed6e21f5246e08a9f344f2cf4abcb&sc_lang=en

6. Seguridad

1.4. **Debe de tomarse condiciones especiales en caso de:**

6.1.1. Formación de aerosoles.

6.1.2. Trabajo con grandes cantidades o concentraciones de microorganismos. Concentraciones elevadas se necesita trabajar en campana de flujo laminar para evitar propagación de aerosoles.

6.1.3. Exceso de material o personal.

1.5. **Aplicar prácticas estandarizadas en laboratorio**

6.1.4. Lavado frecuente de manos, especialmente después de quitarse los guantes y antes de salir del laboratorio.

6.1.5. Mantener una puerta cerrada mientras se trabaja.

6.1.6. Limitar el acceso al laboratorio.

6.1.7. No fumar, comer, beber, o almacenar alimentos en el laboratorio.

6.1.8. Colocar especial atención en condiciones en las que se puedan crear salpicaduras y aerosoles .

- 6.1.9. Descontaminar superficies de trabajo después de uso y después de cualquier derrame.
- 6.1.10. Descontaminar desechos del laboratorio o mandarlos a planta de tratamiento.
- 6.1.11. Utilizar pipetas mecánicas (no usar la pipeta por medio de succión oral).
- 6.1.12. Ser precavido al usar objetos punzantes.
- 6.1.13. Mantenimiento de un programa de control de insectos y roedores.
- 6.1.14. Uso de equipo de protección personal (bata de laboratorio, guantes de látex y protección para los ojos o máscaras para el rostro).

7. Equipo, materiales y reactivos

1.6. Equipo

- 7.1.1. Termociclador

1.7. Materiales/ Cristalería

- 7.1.2. Micropipetas p10
- 7.1.3. Micropipetas p100
- 7.1.4. Micropipetas p1000
- 7.1.5. Puntas de micropipetas
- 7.1.6. Microtubos de 0.2 mL

1.8. Reactivos

- 7.1.7. Hot Start Master Mix
- 7.1.8. Agua libre de nucleasas
- 7.1.9. Agarosa
- 7.1.10. Buffer TBE 1X
- 7.1.11. *TrackIt™ 100 bp DNA Ladder*
- 7.1.12. *GelRed Nucleic Acid Stain 1000x DMSO*

1.9. Medicamentos

- 7.1.13. N/A

1.10. Material biológico

- 7.1.14. ADN extraído de cepas bacterianas

8. Procedimiento

1.11. PCR de colonia

- 1.11.1. Inocular una colonia en un microtubo de PCR.
- 1.11.2. Incubar las muestras en un termociclador a 95°C por diez minutos
- 1.11.3. Realizar alícuotas de los primers 27F y 1942R para que estén a 10 uM
- 1.11.4. Realizar un master Mix multiplicando los siguientes volúmenes por el número de muestras. Como se presenta en el siguiente cuadro:

Componentes	Concentración inicial	Concentración final	Volumen agregado (µL)
Buffer 2x <i>GoTaq®</i>	2x	1x	7.5
Primer 27f	10 µM	1.5 µM	1
Primer 1492r	10 µM	1.5 µM	1
<i>H₂O</i>	-	-	5.5

- 1.11.5. Posteriormente, realizar la amplificación del gen *16S* Siguiendo la siguiente configuración para el termociclador:

Etapa	Temperatura (°C)	Tiempo (min)	Ciclos
Desnaturalización inicial	95	5:00	1
Desnaturalización	95	0:4	30
Alineamiento	56	0:4	
Extensión	72	1:5	

Extensión final	72	15:00	1
Almacenamiento	4	∞	1

- 1.11.6. Evaluar los productos mediante un gel de electroforesis de agarosa al 1%.
 1.12. PCR 16S De ADN extraído
- 1.12.1. Medir la concentración de ADN en Quantus o en el NanoDrop
- 1.12.2. Realizar una dilución de tal manera que todas las muestras tengan una normalización de 50 ng/uL
- 1.12.3. Seguir la siguiente receta de PCR

Componentes	Concentración inicial	Concentración final	Volumen agregado (µL)
Buffer 2x GoTaq®	2x	1x	7.5
Primer 27f	10 µM	1.5 µM	1
Primer 1492r	10 µM	1.5 µM	1
ADN	50 ng/mL	100 ng	2
H₂O	-	-	3.5

- 1.12.4. Posteriormente, seguir la configuración de amplificación descrita previamente en el inciso 1.11.5 y evaluar los productos de PCR a través de una electroforesis en gel de agarosa al 1%

Procedimiento Normado de Operación	SOP 2 Proyecto Suelos Ingenio Magdalena
	Páginas Totales: 6
Identificación fenotípica de microorganismos fijadores de nitrógeno	Autores: Isabella García Caffaro, Aouda Sanchez, Valeria Ochoa

1. Propósito

Entre los organismos de la microbiota en una muestra de suelo degradado por la actividad agrícola y aplicación de agroquímicos, identificar e aislar aquellos capaces de fijar nitrógeno para aprovechamiento por otros organismos, es decir, de transformar el nitrógeno atmosférico (N_2) en formas orgánicas accesibles, como amoníaco (NH_4^+) o nitratos (NO_3^-).

2. Aplicación

El ciclo del nitrógeno es uno de los ciclos bioquímicos esenciales para la supervivencia y desarrollo de un ecosistema, en el cual el nitrógeno atmosférico se hace biodisponible para otros, especialmente vegetales, que lo utilizan para fabricar aminoácidos y ácidos nucleicos (Audeserik, Audeserik, & Byers, 2012). Esta transformación es mayoritariamente realizada en la naturaleza por bacterias, mediante la fijación de nitrógeno (Mendes dos Santos, Escobar, Bentes, & Cid, 2020) estos organismos son denominados diazotrofos (Baldani, Reis, Videira, Boddey, & Divan, 2014). El cultivo y adición de organismos fijadores de nitrógeno a manera de biofertilizantes a los suelos agrícolas es una alternativa a la aplicación de fertilizantes químicos, previniendo así la erosión y disminución de diversidad y resiliencia del suelo (Vega, 2017).

3. Referencias

- 3.1. Angel, R., Nepel, M., Panhölzl, C., Schmidt, H., Herbold, C. W., Eichorst, S. A., & Woebken, D. 2018. <<Evaluation of primers targeting the diazotroph functional gene and development of NifMAP - A bioinformatics pipeline for analyzing nifH amplicon data>>. *Frontiers in Microbiology*,9(APR):1-15. <https://doi.org/10.3389/fmicb.2018.00703>
- 3.2. Marusina, A. I., Boulygina, E. S., Kuznetsov, B. B., Tourova, T. P., Kravchenko, I. K., & Gal'chenko, V. F. 2001. <<A system of oligonucleotide primers for the amplification of nifH genes of different taxonomic groups of prokaryotes>>. *Microbiology*,70(1):73-78. <https://doi.org/10.1023/A:1004849022417>
- 3.3. Mendes dos Santos, R., Escobar, P., Bentes, L., & Cid, E. 2020. << Use of Plant Growth-Promoting Rhizobacteria in Maize and Sugarcane: Characteristics and Applications>>. *Frontiers in Sustainable Food Systems*, 4. doi:10.3389/fsufs.2020.00136 .
- 3.4.

4. Terminología

- 4.1. Diazotrófico: Nombre otorgado a las bacterias capaces de fijar nitrógeno, es decir, transformar mediante su metabolismo el nitrógeno atmosférico en formas biodisponibles (Baldani, Reis, Videira, Boddey, & Divan, 2014).
- 4.2. Medio selectivo: Medio nutritivo para el cultivo de microorganismos que inhiben el desarrollo de algunos organismos, dando lugar o promoviendo el de otros de interés (Rodríguez & Zhurnenko, 2018).

5. Documentos asociados

- 5.1. Audeserik, T., Audeserik, G., & Byers, B. 2012. << Biología: La vida en la Tierra con Fisiología>>. México : Pearson Educación de México.
- 5.2. Alves, A., Forzani, M., de Souza, R., Sibov, S. & Goncalves, J. 2016. <<Isolation and selection of plant promoting bacteria associated with sugarcane>>. *Pesquisa Agropecuária Tropical*, 46(2). pp. 149-158.

- 5.3. Boldani, J., Reis, V., Videira, S., Boddey, L. & Divan, V. 2014. <<The art of isolating nitrogen-fixing bacteria from non-leguminous plants using N-free semi-solid media: a practical guide for microbiologists>>. *Plant Soil*. DOI: 10.1007/s11104-014-2186-6
- 5.4. Gómez, B., Ramírez, J., Veloz, R., Gasca, J. & Herrera, I. 2014. <<Aislamiento y caracterización de bacterias de la rizósfera de guayaba con capacidad promotora de crecimiento de las plantas>>. *Departamento de Ingeniería Agroindustrial de la Universidad de Guanajuato*.
- 5.5. Rodríguez, C. & Zhurbenko, R. 2018. <<Manual BioCen de Medios de Cultivo>>. Cuarta Edición. Cuba: *Centro Nacional de Biopreparados*.
- 5.6. Tsegaya, Z., Yimam, M., Bekele, D., Chaniyalew, S. & Assefa, F. 2019. <<Characterization and Identification of Native Plant Growth-Promoting Bacteria Colonizing Tef (*Eragrostis Tef*) Rhizosphere During the Flowering Stage for A Production of Bio Inoculants. *Biomedical Journal of Scientific & Technical Research*.
- 5.7. Vega, C. 2017. <<Trabajo de fin de grado: Problemas ambientales y de salud derivados del uso de fertilizantes nitrogenados>>. *Universidad Complutense*:
<http://147.96.70.122/Web/TEG/TEG/Memoria/CRISTINA%20VEGA%20OLIVA.pdf>

Seguridad

6.1. Debe de tomarse condiciones especiales en caso de:

- 6.1.1. Formación de aerosoles.
- 6.1.2. Trabajo con grandes cantidades o concentraciones de microorganismos. Con concentraciones elevadas se necesita trabajar en campana de flujo laminar para evitar propagación de aerosoles.
- 6.1.3. Exceso de material o personal.

6.2. Aplicar prácticas estandarizadas en laboratorio

- 6.2.1. Lavado frecuente de manos, especialmente después de quitarse los guantes y antes de salir del laboratorio.
- 6.2.2. Mantener una puerta cerrada mientras se trabaja.
- 6.2.3. Limitar el acceso al laboratorio.
- 6.2.4. No fumar, comer, beber, o almacenar alimentos en el laboratorio.
- 6.2.5. Colocar especial atención en condiciones en las que se puedan crear salpicaduras y aerosoles.
- 6.2.6. Descontaminar superficies de trabajo después de uso y después de cualquier derrame.
- 6.2.7. Descontaminar desechos del laboratorio o mandarlos a planta de tratamiento.
- 6.2.8. Utilizar pipetas mecánicas (no usar la pipeta por medio de succión oral).
- 6.2.9. Ser precavido al usar objetos punzantes.
- 6.2.10. Mantenimiento de un programa de control de insectos y roedores.
- 6.2.11. Uso de equipo de protección personal (bata de laboratorio, guantes de látex y protección para los ojos o máscaras para el rostro).

Equipo, materiales y reactivos

7.1. Equipo

- 7.1.1. Balanza semianalítica
- 7.1.2. Estufa con agitación magnética
- 7.1.3. Agitador magnético
- 7.1.4. Probeta de 1~L
- 7.1.5. Beaker de 80mL.
- 7.1.6. Autoclave

7.2. Materiales/ Cristalería

- 7.2.1. Tubos de ensayo o viales de 10mL.
- 7.2.2. Asas microbiológicas en punta.
- 7.2.3. Erlenmeyer 1L
- 7.2.4. Papel encerado
- 7.2.5. Tubos endorff 0.2ml
- 7.2.6. Puntas de micropipeta
- 7.2.7. Micropipetas

7.3. Reactivos

- 7.3.1. Ácido málico
- 7.3.2. K_2HPO_4
- 7.3.3. $MgSO_4 \cdot 7H_2O$
- 7.3.4. NaCl
- 7.3.5. $CaCl_2$
- 7.3.6. Azul de bromotimol
- 7.3.7. KOH
- 7.3.8. Agar microbiológico
- 7.3.9. $Na_2MoO_4 \cdot 2H_2O$
- 7.3.10. $MnSO_4 \cdot H_2O$
- 7.3.11. H_3BO_3
- 7.3.12. $CuSO_4 \cdot 5H_2O$
- 7.3.13. $ZnSO_4 \cdot 7H_2O$
- 7.3.14. Biotina
- 7.3.15. Piridoxina
- 7.3.16. $FeSO_4 \cdot 7H_2O$
- 7.3.17. NaEDTA
- 7.3.18. Etanol
- 7.3.19. Agua destilada
- 7.3.20. ADN
- 7.3.21. Gotaq Mastermix
- 7.3.22. hot start Colorless master mix
- 7.3.23. oligo Forward (Marusina et al., 2001)
 - nifH_F (5'- TGY GAY CCI AAI GCI GA -3')
posición (115 - 131)
- 7.3.24. oligo Reverse (Marusina et al., 2001)
 - nifH_R (5'- GCC ATC ATY TCI CCI GA -3')
posición (457 - 473)
- 7.3.25. Agua libre de nucleasas
- 7.3.26. Agarosa
- 7.3.27. Escalera 1000pb
- 7.3.28. Bromuro de Etidio 10mg/ml
- 7.3.29. Loading dye
- 7.3.30. Buffer TBE

7.4. Medicamentos

N/A

7.5. Material biológico

- 7.5.1. Colonias aisladas de agar nutritivo de muestras de suelo

8. Procedimiento

- 8.1. Preparación de soluciones necesarias para el medio semisólido libre de nitrógeno NFb
 - 8.1.1. Preparación de solución de elementos traza ($V_f=50mL$)
 - 0.004g de $CuSO_4 \cdot 5H_2O$
 - 0.0012g de $ZnSO_4 \cdot 7H_2O$
 - 0.014g de H_3BO_3
 - 0.01g de $Na_2MoO_4 \cdot 2H_2O$
 - 0.01175g de $MnSO_4 \cdot H_2O$
 - 8.1.2. Preparación de solución vitaminada ($V_f=50mL$)
Se disuelven en 50mL de agua destilada en calor 0.0005g de Biotina y 0.001g de piridoxina. Esta solución no se esteriliza en el autoclave, sino que por filtración a través de un filtro de 0.2 μ m.
 - 8.1.3. Preparación de solución FeEDTA ($V_f=100mL$)
 - Hervir agua destilada.
 - Agregar 0.069g de $FeSO_4 \cdot 7H_2O$ y 0.093g de NaEDTA en 8 mL del agua en ebullición.
 - Inmediatamente, aforar a 100mL con agua destilada fría
- 8.2. Preparación del medio de cultivo ($V_f=1000mL$)

- 8.2.1. Se agregaron a 1 L de agua destilada los siguientes componentes, con agitación continua y calor. (Exactamente con el mismo orden)
- 5g de ácido málico
 - 0.5g de K_2HPO_4
 - 0.1g de NaCl
 - 2mL de la mezcla de elementos traza
 - 2mL de azul de bromotimol al 5% (disuelto en etanol)
 - 4mL de la solución de FeEDTA al 1.64%
 - 4g de KOH
 - 1.75g de agar bacteriológico.
- 8.2.2. Llevar a ebullición
- 8.2.3. Ajustar el pH a 6.5
- 8.2.4. Esterilizar el medio
- 8.2.5. Dejar enfriar la solución, cuando ésta haya alcanzado los $\sim 40^\circ C$, agregar:
- 1 mL de la solución vitaminada.
 - Solución $CaCl_2$ filtrada (0.02g en 5ml agua)
 - Solución $MgSO_4 \cdot 7H_2O$ filtrada (0.2g en 5ml agua)
- 8.2.6. Verter ~ 5 mL por tubo en tubos de ensayo. Agitar cada vez que se vierta para asegurar homogeneidad.
- 8.3. Inoculación e incubación
- 8.3.1. Inocular el medio con asa de aguja pinchando en el medio.
- 8.3.2. Seguir el procedimiento de SOP 1 "Muestreo de Suelos, diluciones seriadas y aislamiento de microorganismos en el suelo" en el apartado de "Inoculación en agares selectivos y diferenciales"
- 8.3.3. **Incubar durante 7-10 días sin agitación a $30^\circ C$** y observar la formación de películas bacterianas. Las primeras podrán ser observadas tras 2-3 días de inoculación y el conteo se puede efectuar tras 5-7 días.
- Las bacterias capaces de fijar nitrógeno crecerán formando una película bajo la superficie del medio. Las bacterias incapaces de fijar nitrógeno no crecerán en el medio.
- 8.3.4. Re-aislar las colonias en un medio sólido utilizando un asa microbiológica en anillo.
- 8.3.5. Este procedimiento debe ser repetido 5 veces con las colonias "positivas" para asegurarse de que el crecimiento no se debe a restos de nitrógeno orgánico presente en la solución original.

Luego... Llene un cuadro con la siguiente información.

- Código de colonia
- Prueba realizada
- Reacción positiva o negativa y cuantificación si se obtuvo
- Observaciones
- Fotografía

8.4. PCR de confirmación del gen *nifH*

8.4.1. Colocar en un tubo eppendorf libre de nucleasas:

Reactivo	Volumen final 1rxn
Go Taq master mix	7.5 μ l
oligo <i>nifH_F</i> (10uM)	0.75 μ l
oligo <i>nifH_R</i> (10uM)	0.75 μ l
ADN o Colonia	1 μ l
Agua libre de nucleasas	5 μ l

TOTAL	15µl
--------------	-------------

8.4.2. Realizar un control negativo sin ADN template

8.4.3. Utilizar el siguiente programa

Paso	Temperatura °C	Tiempo	# Ciclos
Iniciación de desnaturalización	94	3min.	1
Anillamiento	50	3 min	
Extensión	72	3 min	
Desnaturalización	94	30seg	5
Anillamiento	50	2 min	
Extensión	72	30seg	
Desnaturalización	94	30seg	30
Anillamiento	40	30seg	
Extensión	72	30seg	
Extensión final	72	7min.	1
	4	Indefinido	1

(Marusina et al., 2001)

Paso	Temperatura °C	Tiempo	# Ciclos
Iniciación de desnaturalización	94	5min.	1
Desnaturalización	94	30 seg	35
Anillamiento	52	45 seg	
Extensión	72	30 seg	
Extensión final	72	10 min	1
	4	Indefinido	1

(Angel et al., 2018)

8.2.1 Electroforesis

8.2.1.1 Preparar buffer TBE 1x

8.2.1.2 Para visualizar el resultado del PCR realice una electroforesis

- 8.2.1.3 Realizar un gel de agarosa de 1.5% con 0.45g de agarosa y 30ml de buffer TBE 1X, luego que se haya disuelto la agarosa en microondas (30 seg * 30 seg y agitar)
- 8.2.1.3.1 Agregar bromuro de etidio 100X (15 ml de gel - 1ul de bromuro de etidio y 30ml de gel - 2ul de bromuro de etidio)
- 8.2.1.4 Dejar que el gel se solidifique
- 8.2.1.5 Cargar el gel (pozos más cerca del lado negativo)
- 8.2.1.6 Llenar la cámara de electroforesis con buffer TBE 1X
- 8.2.1.7 Cargar el gel (7ul de muestra con loading dye y escalera de 100bp)
 - 8.2.1.7.1 Tamaño esperado de producto 370bp
- 8.2.1.8 Correr el gel por (30min) a 100voltios.
- 8.2.1.9 Visualizar utilizando transiluminador

9. Esterilización de material

Es necesario esterilizar el medio de cultivo y contar con puntas estériles.

10. Cálculos

No aplica

Procedimiento Normado de Operación	SOP 21 Proyecto Suelos Ingenio Magdalena
	Páginas Totales:
Secuenciación de genoma completo de bacterias	Autores: Andrés Say-Agosto, Isabella García Caffaro, Josué Castañón.

1. Propósito

Realizar una secuenciación de genoma completo a través de secuenciación de nueva generación con el kit *Illumina DNA Prep*.

2. Aplicación

La secuenciación de genoma completo permite obtener lecturas de secuencias de ADN de una muestra para posteriormente realizar un ensamblaje y realizar análisis a través de genómica funcional y genómica comparativa. El kit de *Illumina DNA Prep* ofrece un flujo de trabajo rápido y fácil de aplicar la preparación de bibliotecas genómicas, amplicones e incluso plásmidos.

3. Referencias

N/A

4. Terminología

- 4.1. WGS: Secuenciación de genoma completo
- 4.2. NGS: Secuenciación de nueva generación
- 4.3. BLT: *Bead-Linked Transposomes*
- 4.4. TWB: *Tagment wash Buffer*
- 4.5. TB1: *Tagmentation Buffer 1*
- 4.6. TSB: *Tagment Stop Buffer*
- 4.7. EPM: *Enhanced PCR Mix*
- 4.8. SPB: Sample purification Beads
- 4.9. RSB: Resuspension Buffer

5. Documentos asociados

- 5.1. Illumina. (s.f.). Illumina DNA Prep Reference Guide. <https://www.illumina.com/products/by-type/sequencing-kits/library-prep-kits/illumina-dna-prep.html>
- 5.2. Promega. (s.f.). Quantifluor(R) dsDNA System. <https://worldwide.promega.com/es-es/products/rna-analysis/dna-and-rna-quantitation/quantifluor-dsdna-system/?catNum=E2671>

6. Seguridad

6.1. Debe de tomarse condiciones especiales en caso de:

- 6.1.1. Formación de aerosoles.
- 6.1.2. Trabajo con grandes cantidades o concentraciones de microorganismos. Concentraciones elevadas se necesita trabajar en campana de flujo laminar para evitar propagación de aerosoles.
- 6.1.3. Exceso de material o personal.

6.2. Aplicar prácticas estandarizadas en laboratorio

- 6.2.1. Lavado frecuente de manos, especialmente después de quitarse los guantes y antes de salir del laboratorio.
- 6.2.2. Mantener una puerta cerrada mientras se trabaja.
- 6.2.3. Limitar el acceso al laboratorio.
- 6.2.4. No fumar, comer, beber, o almacenar alimentos en el laboratorio.
- 6.2.5. Colocar especial atención en condiciones en las que se puedan crear salpicaduras y aerosoles .

- 6.2.6. Descontaminar superficies de trabajo después de uso y después de cualquier derrame.
- 6.2.7. Descontaminar desechos del laboratorio o mandarlos a planta de tratamiento.
- 6.2.8. Utilizar pipetas mecánicas (no usar la pipeta por medio de succión oral).
- 6.2.9. Ser precavido al usar objetos punzantes.
- 6.2.10. Mantenimiento de un programa de control de insectos y roedores.
- 6.2.11. Uso de equipo de protección personal (bata de laboratorio, guantes de látex y protección para los ojos o máscaras para el rostro).

7. Equipo, materiales y reactivos

7.1. Equipo

- 7.1.1. Fluorómetro Quantus TMFluorometer Promega
- 7.1.2. Secuenciador iSeq TM100
- 7.1.3. Termociclador
- 7.1.4. Microcentrífuga
- 7.1.5. Vórtex

7.2. Materiales/ Cristalería

- 7.2.1. Micropipetas p10
- 7.2.2. Micropipetas p100
- 7.2.3. Micropipetas p1000
- 7.2.4. Puntas de micropipetas
- 7.2.5. Microtubos de 0.2 mL
- 7.2.6. Microtubos de 0.6 mL
- 7.2.7. Microtubos de 1.5 mL
- 7.2.8. Magneto

7.3. Reactivos

- 7.3.1. BLT
- 7.3.2. TWB
- 7.3.3. TB1
- 7.3.4. TSB
- 7.3.5. SPB
- 7.3.6. RSB
- 7.3.7. Etanol absoluto
- 7.3.8. Agarosa
- 7.3.9. Buffer TBE 1X
- 7.3.10. *TrackItTM 100 bp DNA Ladder*
- 7.3.11. *GelRed Nucleic Acid Stain 1000x DMSO*

7.4. Medicamentos

- 7.4.1. N/A

7.5. Material biológico

- 7.5.1. ADN extraído de cepas bacterianas

8. Procedimiento

8.1. Determinación de la concentración de ADN

- 8.1.1. Temperar los reactivos de Quantifluor(R) dsDNA system
- 8.1.2. Calibrar el equipo preprando dos microtubos de 0.6 Quantifluor con lo siguiente:
 - 8.1.2.1. Para el blanco, 200 uL del reactivo Quantifluor dsDNA System
 - 8.1.2.2. Para la referencia, 198 uL del reactivo Quantifluor dsDNA System y 2 uL del reactivo lambda
- 8.1.3. Realizar un vortex a las muestras y un *Spin-Down*
- 8.1.4. Incubar las muestras por 10 minutos a oscuridad
- 8.1.5. Realizar un vortex a las muestras y un *Spin-Down*
- 8.1.6. Realizar la medición del blanco y la referencia
- 8.1.7. Preparar las muestras de ADN agregando en un microtubo de 0.6 Quantifluor 198 uL del reactivo Quantifluor dsDNA System y 2 uL de la muestra de ADN
- 8.1.8. Vortex y *Spin-Down*
- 8.1.9. Incubar por 10 minutos a oscuridad
- 8.1.10. Vortex y *Spin-Down*
- 8.1.11. Realizar las mediciones

8.2. Tagmentación de ADN genómico Genomic DNA

- 8.2.1. Añadir de 2 a 30 μL de ADN en tubos de 1.5mL procurando que el ADN este en el rango de 100 a 500ng
- 8.2.2. Para el caso de los volúmenes de ADN sean menores a 30 añadir agua libre de nucleasas hasta aforar a 30 μL .
- 8.2.3. Mezclar de forma vigorosa con vortex el BLT durante 10 segundos hasta resuspender. Repetir si es necesario.
- 8.2.4. Combinar los siguientes volúmenes para preparar el master mix de tagment. Multiplicando cada volumen por el número de muestras procesadas
- BLT (11 μL)
 - TB1 (11 μL)
- 8.2.5. Con ayuda de un vortex mezclar el master mix hasta que se resuspenda.
- 8.2.6. Transferir 20 μL del master mix a cada uno de los tubos que posee una muestra.
- 8.2.7. Mezclar por pipeteo 10 veces cada una de las muestras para resuspender la muestra.
- 8.2.8. Cerrar los tubos y llevarlo al termociclador con el siguiente programa TAG:

Etapas	Temperatura (°C)	Tiempo (min)
Pre-calentamiento	100	-
Incubación	55	15
Almacenamiento	10	-

8.3. Limpieza post-tagmentación

- 8.3.1. Agregar 10 μL de TSB a la reacción de tagmentación
- 8.3.2. Resuspender lentamente por pipeteo 10 veces las perlas
- 8.3.3. Cerrar los tubos y llevarlo al termociclador con el siguiente programa PTC:

Etapas	Temperatura (°C)	Tiempo (min)
Pre-calentamiento	100	-
Incubación	37°C	15
Almacenamiento	10	-

- 8.3.4. Colocar los tubos en la gradilla con el magneto y esperar a que la solución se vea clara y libre de las perlas
- 8.3.5. Descartar el sobrenadante
- 8.3.6. Lavar las perlas dos veces con el siguiente procedimiento:
- 8.3.6.1. Agregar lentamente 100 μL de TWB directamente a las perlas
- 8.3.6.2. Colocar los tubos nuevamente en el magneto y esperar a que la solución se vea clara
- 8.3.6.3. Descartar el sobrenadante
- 8.3.7. Agregar 100 μL de TWB a las perlas y resuspender por pipeteo lentamente
- 8.3.8. Colocar en el magneto las bibliotecas

8.4. Amplificación de ADN tagmentado

- 8.4.1. Preparar el master mix de PCR multiplicando por el número de muestras los siguientes volúmenes:
- 8.4.1.1. 22 μL de EPM
- 8.4.1.2. 22 μL de Agua libre de nucleasas
- 8.4.2. Mezclar por vortex y realizar un *Spin-Down*
- 8.4.3. Remover el sobrenadante de las bibliotecas en el magneto (paso 8.3.8)
- 8.4.4. Remover los tubos del magneto y agregar 40 μL de máster mix de PCR directamente a las perlas. Mezclar por pipeteo para resuspender las perlas.
- 8.4.5. Realizar un *Spin-Down* corto
- 8.4.6. Agregar los adaptadores apropiados para las muestras
- 8.4.6.1. 5 μL de adaptadores i7
- 8.4.6.2. 5 μL de adaptadores i5

8.4.7. Colocar la mezcla en el siguiente programa de PCR (BLT-PCR):

Etapa	Temperatura (°C)	Tiempo
Pre-calentamiento	100	-
1	68	15 min
2	98	-
30 ciclos de la etapa 3	98	45s
	62	30s
	68	2 min
4	68	1 min
Almacenamiento	10	-

8.5. Limpieza de bibliotecas

- 8.5.1. Realizar un *Spin-down* de las muestras
- 8.5.2. Colocar las muestras en el magneto hasta que la solución sea transparente
- 8.5.3. Transferir 45 uL del sobrenadante a un microtubo de 0.2 mL
- 8.5.4. Mezclar por vortex y resuspender el SPB
- 8.5.5. Agregar 40 uL de agua libre de nucleadas
- 8.5.6. Agregar 45uL de SPB
- 8.5.7. Mezclar por pipeteo al menos diez veces
- 8.5.8. Cerrar los tubos e incubar por 5 minutos a temperatura ambiente
- 8.5.9. Colocar en el magneto
- 8.5.10. Descartar el sobrenadante
- 8.5.11. Lavar las perlas dos veces de la siguiente manera:
 - 8.5.11.1. Agregar 200 uL de etanol 80% NO mezclar
 - 8.5.11.2. Incubar por 30s a temperatura ambiente
 - 8.5.11.3. Descartar el sobrenadante
- 8.5.12. Secar por 5 minutos
- 8.5.13. Agregar 32 uL de RSB a las perlas
- 8.5.14. Resuspender mediante pipeteo
- 8.5.15. Incubar a temperatura por 2 minutos
- 8.5.16. Colocar en el magneto por 2 minutos
- 8.5.17. Transferir 30 uL de un microtubo de 0.2 mL

8.6. Calidad de bibliotecas

- 8.6.1. Realizar una medición en la concentración en el Quantus como descrito en el paso (8.1)
- 8.6.2. Realizar un gel de agarosa a 1% con buffer TBE y correr las muestras para evaluar el tamaño del barrido

8.7. Preparación del pool de bibliotecas

- 8.7.1. Preparar las bibliotecas de tal manera que estén a 1-5 uM
- 8.7.2. Tomar 5 uL de cada biblioteca y mezclarlo en un nuevo tubo
- 8.7.3. Calcular para que la concentración tenga exactamente 100 ng
- 8.7.4. Agregarle al tubo PhiX para tener una concentración de 1%
- 8.7.5. Cargar 25 uL en el cartucho y realizar las lecturas

9. Esterilización de material

No aplica

10. Cálculos

No aplica

11. Control de calidad

- 11.1. Verificar con un check list que se llevó al lugar los implementos necesarios

- 11.2. Verificar que el almacenamiento de muestras y reactivos sea el adecuado 4°C, -20°C o -80°C.

12. Conflict solution policy

Anexo E Información suplementaria

Los archivos resultado del análisis bioinformático dividido por cepa, así como el archivo R para realizar los gráficos con base en el .VISUAL por fastANI, la lista de comandos y la secuencia de los plásmidos se encuentra en la siguiente carpeta de Drive:

https://drive.google.com/drive/folders/1Cp3U9JEw_2eWa57RXa-GDIyRm7mxYO96