

**IDENTIFICACIÓN POLIFÁSICA DE UNA COLECCIÓN BACTERIANA DEL
GÉNERO *Bacillus* OBTENIDA DEL SUBSUELO PROFUNDO EN ZONAS
CARBONÍFERAS DE COLOMBIA**

CATALINA LUCERO RIVERA GIRALDO

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE BIOLOGÍA
BUCARAMANGA**

2014

**IDENTIFICACIÓN POLIFÁSICA DE UNA COLECCIÓN BACTERIANA DEL
GÉNERO *Bacillus* OBTENIDA DEL SUBSUELO PROFUNDO EN ZONAS
CARBONÍFERAS DE COLOMBIA**

CATALINA LUCERO RIVERA GIRALDO

**Trabajo de Grado para optar al título de
Bióloga**

Director

JORGE LUIS FUENTES LORENZO

Microbiólogo, M.Sc., Ph.D.

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE BIOLOGÍA
BUCARAMANGA**

2014

AGRADECIMIENTOS

Al profesor Jorge Luis Fuentes Lorenzo por permitirme entrar al Laboratorio de Microbiología y Mutagénesis Ambiental y por su colaboración en el desarrollo de este proyecto. A mis amigos y compañeros de estudio Jessica Calderón, Oscar Hernández y Claudia Moreno por su compañía durante toda la carrera. A mi novio Félix Álvarez por su compañía, apoyo y amor incondicional en todo momento. Finalmente a mi familia, especialmente a mis padres Sergio Rivera y Lucero Giraldo, a quienes debo todo lo que soy y por quienes seguiré trabajando para cumplir mis metas.

CONTENIDO

	Pág.
INTRODUCCIÓN.....	12
1. ANTECEDENTES O MARCO TEÓRICO.....	14
2. MATERIALES Y MÉTODOS.....	18
2.1 COLECCIÓN DE AISLADOS.....	18
2.2 DESCRIPTORES MOLECULARES.....	18
2.2.1 Extracción de ADN.....	18
2.2.2 Amplificación y secuenciación del gen ARNr 16S.....	18
2.2.3 Análisis filogenético basado en las secuencias del gen ARNr 16S.....	19
2.2.4 Análisis de Restricción del ADN Ribosomal Amplificado (ARDRA).....	20
2.3 DESCRIPTORES FENOTÍPICOS.....	20
2.4 ANÁLISIS DE CONGLOMERADOS.....	21
2.5 CLAVE INTERACTIVA DE IDENTIFICACIÓN.....	22
3. RESULTADOS.....	23
4. DISCUSIÓN.....	34
5. CONCLUSIONES.....	41
6. RECOMENDACIONES.....	42
BIBLIOGRAFÍA.....	43
ANEXOS.....	56

LISTA DE FIGURAS

	Pág.
Figura 1. Árbol filogenético basado en la secuencia del gen ARNr 16S.	26
Figura 2. Variantes de restricción obtenidas mediante el ARDRA usando las enzimas TaqI, HaeIII, AluI, DdeI y XbaI.	27
Figura 3. Dendrograma basado en descriptores fenotípicos y del ARDRA.	31
Figura 4. Ventana Principal de la clave interactiva de identificación.	33

LISTA DE TABLAS

	Pág.
Tabla 1. Historia de la clasificación de procariotas	14
Tabla 2. Asignación de aislados a especies bacterianas basado en el porcentaje de similitud en la secuencia del gen ARNr 16S.	23
Tabla 3. Variantes de restricción obtenidas con el ARDRA.	28
Tabla 4. Índice de Concordancia Fenotípica y sus correspondientes errores estándar.	29

ANEXOS

	Pág.
Anexo A. Perfiles de bandas obtenidos con la enzima de restricción AluI.	56
Anexo B. Perfiles de bandas obtenidos con la enzima de restricción DdeI.	57
Anexo C. Perfiles de bandas obtenidos con la enzima de restricción TaqI.	58
Anexo D. Perfiles de bandas obtenidos con la enzima de restricción HaeIII.	59
Anexo E. Perfiles de bandas obtenidos con la enzima de restricción XbaI.	60
Anexo F Cálculo del Índice de Concordancia Fenotípica (ICF).	61
Anexo G. Listado de los aislados y la especie identificada.	66
Anexo H. Árbol filogenético basado en secuencias del gen ARNr 16S sin el aislado UC-03.	67
Anexo I. Dendrograma basado en descriptores fenotípicos.	67

RESUMEN

TÍTULO: IDENTIFICACIÓN POLIFÁSICA DE UNA COLECCIÓN BACTERIANA DEL GÉNERO *Bacillus* OBTENIDA DEL SUBSUELO PROFUNDO EN ZONAS CARBONÍFERAS DE COLOMBIA*

AUTOR: CATALINA LUCERO RIVERA GIRALDO**

PALABRAS CLAVE: *Bacillus*, gen ARNr 16S, ARDRA, taxonomía polifásica.

DESCRIPCIÓN:

En este estudio se caracterizaron e identificaron 102 aislados de una colección bacteriana del género *Bacillus* obtenida del subsuelo profundo de zonas carboníferas de Colombia (Guajira y Cucunubá) mediante taxonomía polifásica. Como descriptores genotípicos se usaron el análisis de homología de secuencias del gen ARNr 16S y un Análisis de Restricción del ADN Ribosomal Amplificado (ARDRA) con las enzimas AluI, DdeI, HaeIII, TaqI y XbaI; y como descriptores fenotípicos se utilizaron 12 pruebas morfológicas, fisiológicas y metabólicas. El análisis de homología de secuencias del gen ARNr 16S resultó poco efectivo en la distinción de especies dentro de la colección. Por el contrario, los perfiles de bandas del ARDRA obtenidos con las enzimas AluI y DdeI resultaron los más informativos permitiendo identificar la mayoría de las especies encontradas en la colección; así como revelar la existencia de un posible híbrido entre las especies *B. subtilis* y *B. licheniformis*. De conjunto, los 12 descriptores fenotípicos y los 5 descriptores genotípicos ARDRA se usaron para construir una clave interactiva que permitió identificar de forma rápida y precisa siete especies de *Bacillus* altamente relacionadas: *B. cereus*, *B. thuringiensis*, *B. licheniformis*, *B. subtilis*, *B. safensis*, *B. altitudinis* y *B. megaterium*. En el trabajo se discute la utilidad de combinar métodos fenotípicos y genotípicos para la identificación de especies altamente relacionadas del género *Bacillus* que habitan el subsuelo profundo de zonas carboníferas de Colombia.

* Proyecto de grado

** Facultad de Ciencias. Escuela de Biología. Director: Jorge Luis Fuentes Lorenzo

ABSTRACT

TITLE: POLYPHASIC IDENTIFICATION OF A BACTERIAL COLLECTION OF THE GENUS *Bacillus* OBTAINED FROM SUBSURFACE DEEP IN CARBONIFEROUS ZONES OF COLOMBIA*

AUTHOR: CATALINA LUCERO RIVERA GIRALDO**

KEY WORDS: *Bacillus*, 16S rRNA gene, ARDRA, polyphasic taxonomy.

DESCRIPTION:

In this study we used a polyphasic taxonomy test to characterize and identify 102 isolated from bacterial collection of the genus *Bacillus* obtained from two deep subsoil carboniferous zones in Colombia (Guajira, department and Cucunubá, municipality). We used genotypic and phenotypic descriptors; for the first descriptor we evaluated the 16S rRNA gene sequences homology and an Amplified Ribosomal DNA Restriction Analysis (ARDRA) using the following enzymes AluI, DdeI, HaeIII, TaqI and XbaI. For the second descriptor we used 12 morphologic, physiologic and metabolic tests. We found that the 16S rRNA gene sequences homology was not sufficient in the distinction among the bacteria species from the collection. Nevertheless, the AluI and DdeI ARDRA profile bands allowed us not only a better taxonomic identification but the possible existence of a hybrid between *B. subtilis* and *B. licheniformis*. We also developed an interactive taxonomic key using the 12 phenotypic descriptors and 5 ARDRA genotypic descriptors; this taxonomic key allows precise and easy identification of 7 *Bacillus* species highly related such as: *B. cereus*, *B. thuringiensis*, *B. licheniformis*, *B. subtilis*, *B. safensis*, *B. altitudinis* y *B. megaterium*. Here we discuss the importance and usefulness of implement combine phenotypic and genotypic methods in the identification of species highly related of the *Bacillus* genus that inhabit subsurface deep in carboniferous zones of Colombia.

* Bachelor Thesis

** Science Faculty. Biology program. Director: Jorge Luis Fuentes Lorenzo,

INTRODUCCIÓN

La clasificación e identificación taxonómica de los procariotas, a lo largo de su historia, ha estado determinada por el desarrollo de diferentes métodos microbiológicos y tipos de análisis. Pese a que la taxonomía microbiana se ha nutrido de una gran cantidad de técnicas que estudian los microorganismos considerando su variación al nivel morfológico, fisiológico, bioquímico y genético; ninguna de ellas ha garantizado por sí sola una completa y efectiva clasificación del mundo microbiano. Actualmente, existe un consenso en que una adecuada aproximación taxonómica al mundo microbiano se logra combinando técnicas que miden la variación a diferentes niveles con aproximaciones al fenotipo, el genotipo y a la filogenia de los microorganismos; enfoque conocido como taxonomía polifásica (Vandamme *et al.*, 1996; Janda & Abbott, 2002; Schleifer, 2009; Moore *et al.*, 2010).

Uno de los desafíos actuales de la taxonomía polifásica es encontrar combinaciones efectivas de descriptores fenotípicos y genotípicos que permitan una clasificación microbiana con una clara significación biológica (Cohan, 2002; Achtman & Wagner, 2008; Koeppel *et al.*, 2008). Esto es particularmente cierto para grupos microbianos como el género *Bacillus*, donde varios de los métodos usados para la identificación y clasificación taxonómica han resultado inconsistentes. La limitación más frecuente encontrada con este género es que una alta similitud en la secuencia del gen ARNr 16S hace muy difícil la clasificación de especies estrechamente relacionadas (Fox *et al.*, 1992; Maughan & Van der Auwera, 2011).

El Análisis de Restricción del ADN Ribosomal Amplificado (ARDRA, de sus siglas en inglés) es una técnica alternativa al análisis basado en similitud de secuencias del gen ARNr 16S que supera las limitaciones de este último en cuanto a su capacidad discriminante de especies muy relacionadas (Heyndricks *et al.*, 1996;

Wu *et al.*, 2006). El ARDRA ha resultado efectiva en la identificación de especies del género *Bacillus* presentes en ambientes termófilos (Duffner *et al.*, 1997; Mora *et al.*, 1998; Nicolaus *et al.*, 2000; Cihan *et al.*, 2012), de especies que degradan compuestos fenólicos del carbón (Felföldi *et al.*, 2010) y celulosa (Mohanty *et al.*, 2011) y de estirpes bacterianas de interés industrial (De Clerck *et al.*, 2004; Wu *et al.*, 2006).

Dado que una adecuada selección de las enzimas de restricción usada para el ARDRA es la clave para lograr diferenciar aislados estrechamente relacionados (Moyer *et al.*, 1996); el presente trabajo tiene como hipótesis que la combinación de descriptores ARDRA desarrollados con enzimas adecuadas con otros descriptores fenotípicos (morfológicos, fisiológicos y bioquímicos) y genotípicos (secuenciación del gen ARNr 16S), deben permitir la identificación de estirpes del género *Bacillus* del subsuelo profundo de zonas carboníferas de Colombia. Por el ello, el objetivo general del trabajo es identificar mediante taxonomía polifásica una colección de aislados bacterianos del género *Bacillus* obtenida del subsuelo profundo de dos zonas carboníferas de Colombia (Cerrejón y Cucunubá). Consecuentemente se plantearon cuatro objetivos específicos: 1) Caracterizar molecularmente (secuencia del gen ARNr 16S y ARDRA) los aislados del género *Bacillus* presentes en mantos de carbón. 2) Establecer cuáles son las características fenotípicas más importantes para la identificación de las especies de *Bacillus* presentes en mantos de carbón. 3) Evaluar la utilidad de la taxonomía polifásica con los marcadores descritos en la identificación de especies del género *Bacillus* presentes en mantos de carbón. 4) Construir una clave interactiva para la identificación de especies del género *Bacillus* presentes en mantos de carbón.

1. ANTECEDENTES O MARCO TEÓRICO

La clasificación taxonómica de los microorganismos, a lo largo de su historia, ha estado determinada por el desarrollo de diferentes métodos microbiológicos y tipos de análisis (Tabla 1).

Tabla 1. Historia de la clasificación de procariotas

Tiempo	Clasificación basada en:
Finales siglo 19	Morfología, requisitos de crecimiento, potencial patogénico
1900-1960	Morfología, fisiología, bioquímica
1960-1980	Quimiotaxonomía, taxonomía numérica, hibridación ADN-ADN
1980- Actualidad	Análisis genotípicos, análisis de secuencias multilocus, promedio de identidad de nucleótidos, análisis del genoma completo

Schleifer, 2009

Los primeros métodos usados producían descriptores morfológicos basados mayormente en la forma celular (Maughan & Van der Auwera, 2011); siendo justamente especies del género *Bacillus* (*B. anthracis* y *B. subtilis*) las primeras descritas con tales métodos dando inicio a la clasificación procariota.

A comienzos del siglo 20 nuevos descriptores basados en propiedades fisiológicas y bioquímicas se unieron a aquellos morfológicos para lograr una clasificación más efectiva de los microorganismos (Schleifer, 2009). Justo en el año 1923, se publica la primera versión del “Bergey’s Manual of Determinative Bacteriology”, manual que unificó criterios taxonómicos convirtiéndose en una obra de referencia obligada para la clasificación bacteriana (Roselló-Mora & Amann, 2001).

A finales de la década de los años 50 surge la taxonomía numérica que elimina los inconvenientes de trabajar con grandes tablas de datos, mejorando la identificación fenotípica de las bacterias mediante el cálculo de coeficientes de

similitud entre cepas y especies (Sneath, 1995). En este enfoque taxonómico, todos los caracteres (morfológicos, fisiológicos, bioquímicos, etc.) tienen el mismo peso. El número de características comunes entre dos cepas se considera una medida cuantitativa de la relación taxonómica; pero no de su relación filogenética (Schleifer, 2009). Estos caracteres fenotípicos, aunque útiles para la clasificación, no contienen información genética que permita establecer relaciones evolutivas entre las especies (Maughan & Van der Auwera, 2011).

A principios de la década de los años 70 se desarrolló la técnica de hibridación ADN-ADN (Jhonson, 1973). Dicha técnica evalúa la similitud genética entre especies considerando el grado en que sus genomas hibridan bajo condiciones estándares. Un valor de hibridación $\geq 70\%$ fue considerado para definir especies (Goris *et al.*, 2007). Durante los noventa, la técnica de secuenciación del gen ARNr 16S permitió dar un salto cualitativo en la taxonomía microbiana (Woese *et al.*, 1990). Las características del gen ARNr 16S (secuencia altamente conservada, ubicuidad y funcionalidad constante), permiten por primera vez el análisis filogenético de especies; contribuyendo significativamente a la taxonomía bacteriana en un contexto evolutivo (Olsen & Woese, 1993). Un hallazgo importante derivado del uso de esta técnica, fue que un porcentaje de hibridación ADN-ADN mayor al 70% generalmente coincide con un valor de similitud en la secuencia del gen ARNr 16S mayor al 97% (Stackebrandt & Goebel, 1994). Así, un valor de similitud en la secuencia del gen ARNr 16S $\geq 97\%$, fue considerado como un criterio para definir especies (Roselló-Mora & Amann, 2001; Cohan, 2002). El análisis de similitud en la secuencias del gen ARNr 16S ha sido sin dudas uno de los métodos más útiles en taxonomía microbiana y la disponibilidad en bases de datos gratuitas con las secuencias de las especies reportadas lo han convertido en un método de uso extensivo (Peplies *et al.*, 2008; Yarza *et al.*, 2008). Sin embargo, se ha encontrado que el método puede ser poco resolutivo en diferentes géneros bacterianos cuyas especies muestran claras divergencias fenotípicas. Un buen ejemplo, es el género *Bacillus* donde la divergencia en

similitud del gen ARNr 16S es insuficiente para delimitar especies (Fox *et al.*, 1992; Maughan & Van der Auwera, 2011).

Ninguna de las técnicas hasta aquí mencionadas garantizan por sí sola una completa y efectiva clasificación del mundo microbiano (Janda & Abbott, 2002). Con el fin de superar estas limitaciones, a mediados de los noventa fue propuesto un enfoque taxonómico conocido como taxonomía polifásica (Vandamme *et al.*, 1996). Dicho enfoque plantea que la mejor aproximación taxonómica al mundo microbiano se logra combinando técnicas fenotípicas, genotípicas y de inferencias filogenéticas que midan la variación a diferentes niveles con aproximaciones al metabolismo y la evolución de los microorganismos (Scheliefer, 2009; Moore *et al.*, 2010).

Varias técnicas de genotipado de ADN en combinación con descriptores fenotípicos y filogenéticos, han sido usadas en taxonomía polifásica de microorganismos (Tiedje *et al.*, 1999; Nocker *et al.*, 2007; Ansari & Malik, 2013). Una de estas técnicas, el ARDRA ha sido usada exitosamente en taxonomía polifásica de géneros microbianos como: *Lactobacillus* (Ruiz *et al.*, 2000; Cleenwerck & De Vos, 2008; Samelis *et al.*, 2011), *Pseudomonas* (Peix *et al.*, 2009), *Rhizobium* (Laguerre *et al.*, 1994; Pinto *et al.*, 2007; Suzuki *et al.*, 2008; Wadhwa *et al.*, 2011), *Streptomyces* (Anderson & Wellington, 2001), *Vibrio* (Urakawa *et al.*, 1997) y archeobacterias metanogénicas (Whitby *et al.*, 2004).

La reciente re-evaluación del concepto procarionta de especie en un contexto ecológico impone nuevos retos a la taxonomía microbiana (Gevers *et al.*, 2005). La incorporación de la Ecología en la sistemática bacteriana tiene como paradigma la simulación de ecotipos microbianos bajo un modelo de “ecotipo estable”; en el cual nuevos ecotipos son formados solo raramente y cada ecotipo perdura muchos eventos de selección durante su vida (Cohan, 2006; Gevers *et al.*, 2006; Achtman & Wagner, 2008). Bajo tales circunstancias, la agrupación de

secuencias del gen ARNr 16S que se observa en estos sistemas es dominada por la selección periódica, mostrando una estrecha correspondencia entre ecotipos y los grupos de secuencias (Cohan & Perry, 2007; Koepfel *et al.*, 2007). Gevers *et al.* (2005) indicaron que una aproximación taxonómica basada en el Análisis de Secuencias Multilocus (Maiden, 2006) ofrece suficiente flexibilidad para dar cabida a las grandes diferencias biológicas de los procariotas y se han hecho nuevas reconstrucciones del árbol filogenético de procariotas basados en tales tecnologías (Wu *et al.*, 2009; Rinke *et al.*, 2013). Sin embargo, la genómica comparada resulta aún muy costosa para establecerse como una herramienta de uso universal y se requerirá bases de datos genómicas y de diseños de muestreo más cuidadosos de las poblaciones microbianas naturales, para materializar una taxonomía basada en datos genómicos (Konstantinidis & Tiedje, 2005ab, 2007).

El subsuelo profundo juega un papel central en los ciclos biogeoquímicos (Tang *et al.*, 2012), albergando entre el 75 y 94% de los procariotas terrestres. Los mantos de carbón, contenidos en este subsuelo, aportan su materia orgánica e inorgánica como sustratos de los microorganismos que participan en los mencionados ciclos (Whitman *et al.*, 1998). Es por ello, que el estudio de la composición y diversidad de las especies microbianas presentes en mantos de carbón ha tenido creciente atención (Li *et al.*, 2008; Fry *et al.*, 2009; Penner *et al.*, 2010; Tang *et al.*, 2012; Gallagher *et al.*, 2013). El género *Bacillus*, uno de los más representativos en estos ambientes, juega un papel fundamental en la biodegradación de residuos orgánicos, en los ciclos biogeoquímicos del carbono y del nitrógeno y en la producción industrial de vitaminas, proteínas, enzimas, insecticidas, entre otros (Imelda & Raj, 2007; de Maagd *et al.*, 2003; Tang & You 2012). Sin embargo, a pesar de la importancia del género *Bacillus*, su clasificación e identificación taxonómica sigue siendo un tema por resolver (Fox *et al.*, 1992; Maughan & Van der Auwera, 2011).

2. MATERIALES Y MÉTODOS

2.1 COLECCIÓN DE AISLADOS

Se utilizó una colección de 102 aislados bacterianos del género *Bacillus*, previamente conservados en el Laboratorio de Microbiología y Mutagénesis Ambiental (LMMA) de la Escuela de Biología de la Universidad Industrial de Santander (UIS). Dichos aislados fueron obtenidos a partir de muestras de carbón mineral, aguas y lodos de los yacimientos carboníferos del Cerrejón (La Guajira, Colombia) y del Sinclinal Checua-Lenguazaque (Cundinamarca, Colombia). Adicionalmente, se utilizaron 4 estirpes bacterianas certificadas por la Colección Americana de Cultivos Tipo (ATCC, de sus siglas en inglés) previamente descritos para mantos carboníferos: *Bacillus cereus* ATCC®10876™, *Bacillus licheniformis* ATCC®12759™, *Bacillus megaterium* ATCC®9885™, *Bacillus subtilis* ATCC®6633™, obtenidos de la casa comercial OXOID LTD. (Basingstoke, Inglaterra).

2.2 DESCRIPTORES MOLECULARES

2.2.1 Extracción de ADN. La extracción de ADN genómico de bacterias se realizó usando básicamente el procedimiento descrito por Maniatis et al., (1989) con las modificaciones propuestas por Liu (2009). La concentración del ADN obtenido se cuantificó mediante espectrofotometría a una longitud de onda de 260nm. Para comprobar el estado de pureza del ADN obtenido, se midió la absorbancia del ADN a 280nm y se calculó la relación de DO260/280, considerándose como criterio de pureza valores DO260/280 entre 1.8 y 2.0.

2.2.2 Amplificación y secuenciación del gen ARNr 16S. La amplificación del gen ARNr 16S de las estirpes bacterianas se desarrolló mediante la Reacción en

Cadena de la Polimerasa (PCR, de sus siglas en inglés) usando los cebadores universales 530F (5'-GTCCCAGCMGCCGCGG-3') y 1490R (5'-GGTTACCTTGTTACGACTT-3') descritos por Wani *et al.*, (2006). Para la mezcla de reacción (22 µl) se usaron 3µl de buffer 10x, 4µl de dNTPs (2mM), 0.4µl de primer forward, 0.4µl de primer reverse (100 µM), 0.2µl de DreamTaq™ DNA Polymerase (Fermentas, USA), 12µl de agua grado molecular y 2µl del ADN molde (5ng/µL). La amplificación se llevó a cabo en un Termociclador MasterCycler® Pro-Realplex4 (Eppendorf, Alemania), usando el siguiente programa: 94°C/3min seguido de 35 ciclos de 94°C/45s, 55°C/1min, 72°C/1min y una fase final de extensión de 72°C/5min. Los fragmentos amplificados se visualizaron mediante electroforesis en gel de agarosa al 0.8% usando el colorante fluorescente EZVISION (Amresco, USA) y fueron documentados usando un Sistema para Análisis de Geles (Syngene, USA). Los productos amplificados fueron purificados y secuenciados usando el servicio SSiGmol del Instituto de Genética de la Universidad Nacional de Colombia (Bogotá, Colombia).

2.2.3 Análisis filogenético basado en las secuencias del gen ARNr 16S. Las secuencias obtenidas de los fragmentos amplificados del gen ARNr 16S de cada aislado, se compararon con las disponibles en la base de datos NCBI (www.ncbi.nlm.nih.gov) y RDP (Maidak *et al.*, 1994) usando los algoritmos BLAST (Altschul *et al.*, 1990) y SeqMatch (Cole *et al.*, 2014) respectivamente; esto con el fin de identificar las secuencias depuradas (Yarza *et al.*, 2010; Munoz *et al.*, 2011) más próximas. Para la comparación fueron además usadas, secuencias correspondientes a estirpes bacterianas certificadas (ATCC) disponibles en las mencionadas bases de datos y a las aquí estudiadas. Se consideró un valor de similitud de secuencia superior al 97% para definir especie. Las secuencias asignadas a especie, serán depositadas en la mencionada base de datos NCBI. Todas las anteriormente mencionadas fueron alineadas usando el algoritmo CLUSTAL W (Thompson *et al.*, 1994) y posteriormente, se construyó un árbol filogenético basado en el modelo evolutivo de Tamura-Nei (Tamura y Nei, 1993) y

usando el método de Neighbor-Joining (Saitou y Nei, 1987). El nivel de confianza de los nodos del árbol se computo mediante un análisis de re-muestreo de 1000 permutaciones usando el programa MEGA 6 (Tamura *et al.*, 2013).

2.2.4 Análisis de Restricción del ADN Ribosomal Amplificado (ARDRA).

Usando el paquete estadístico BioEdit (Hall, 1999), se determinaron las enzimas útiles para revelar el polimorfismo de nucleótidos existente en el gen ARNr 16S del género *Bacillus* y sus correspondientes sitios de restricción. Se utilizaron un total de seis enzimas de restricción: AluI, DdeI, HaeIII, TaqI, XbaI y EcoRI. Se realizaron digestiones individuales en 30µl de reacción conteniendo 1µl (1U) de enzima FastDigest® (Fermentas, USA), 2µl de Buffer FastDigest® 10X, 17µl de agua grado molecular y 10µl (0,2 – 0,5ug) de ADN molde. Se incubó la mezcla de reacción a 37°C (64°C para la enzima TaqI) durante 5-20min según las recomendaciones del fabricante. Los fragmentos de ADN obtenidos de la digestión, fueron separados electroforéticamente en geles de agarosa de alta resolución al 2%, usando Buffer de corrida TBE (0.5 X) y un voltaje de 100V durante 2.5 - 3h. Los fragmentos separados fueron visualizados usando el colorante fluorescente EZVISION (Amresco, USA). En todas las corridas electroforéticas se utilizó un patrón de peso molecular: GeneRuler 100kb plus DNA Ladder (Fermentas, USA) o Ready Ladder™ 100bp (AMRESCO, USA). A partir de los patrones de bandas de las estirpes bacterianos, se generó una matriz de datos binarios.

2.3 DESCRIPTORES FENOTÍPICOS

Se utilizó una base de datos fenotípicos de la colección de aislados descrita en el acápite 2.1 previamente desarrollados por Mayorga (2012) y por Barrios (2014). Dicha base de datos contiene 26 descriptores morfológicos, fisiológicos y bioquímicos informativos para la identificación de especies del género *Bacillus* según el Bergey's Manual of Systematic Bacteriology (De Vos *et al.*, 2009). Estos

descriptores son los siguientes: 1) respuesta a tinciones diferenciales (Gram y tinción de esporas), 2) forma bacteriana, 3) motilidad, 4) relación con el oxígeno, 5) actividad catalasa y oxidasa, 6) uso de citrato como fuente de carbono, 7) producción de gas y/o ácido por vía fermentativa de diferentes fuentes de carbono (Glicerol, Glucosa, Lactosa, Maltosa, Manitol, Sacarosa), 8) producción de 2,3-butanediol a partir de glucosa (prueba Voges Proskauer), 9) utilización de urea como fuente de carbono, 10) tolerancia a NaCl, 11) hidrólisis enzimática almidón y gelatina 12) crecimiento a temperatura 50°C.

Dado que este tipo de descriptores pueden producir resultados no conclusivos (+/-) con diferentes aislados de una misma especie, se estimó un índice de concordancia fenotípica (ICF) para especies del género *Bacillus*, con el fin de seleccionar descriptores confiables para su identificación. Dicho índice se calculó como sigue:

$$\text{Índice de Concordancia Fenotípica (ICF)} = \frac{F1 + F2 + F3 + \dots + Fn}{n}$$

Donde F es la frecuencia del resultado esperado (correcto) con la prueba 1, 2, 3... n; y n es el número total de pruebas fenotípicas. Se consideraron aquellas pruebas fenotípicas que producen $ICF \geq 0.8$; es decir pruebas que producen resultados esperados, ya sean positivos o negativos, en más del 80% de las cepas de referencias de especies de género *Bacillus* publicadas en Bergey's Manual of Systematic Bacteriology (De Vos *et al.*, 2009). Solo en el caso de la especie *B. safensis* se usó la información publicada por Satomi *et al.* (2006).

2.4 ANÁLISIS DE CONGLOMERADOS

Para los análisis fenéticos se usaron los descriptores fenotípicos con $ICF \geq 0.8$ y los ARDRA. Con dichos descriptores fenotípicos se construyó una matriz binaria

considerando los resultados positivos como unos (1) y los negativos como ceros (0). En el caso del ARDRA se consideró la presencia de variantes genotípicas como unos (1) y las ausencias como ceros (0). El nivel de similitud entre unidades taxonómicas se estimó mediante el índice de Dice (S_D); el cual se consideró como una medida de su proximidad fenotípica y genotípica. La matriz de similitud se usó para construir un árbol usando el método de la media aritmética no ponderada (UPGMA). Los análisis fenéticos se realizaron usando el programa Xlstat versión 54.6MB (Addinsoft SARL, Francia).

2.5 CLAVE INTERACTIVA DE IDENTIFICACIÓN

Con base en los descriptores fenotípicos con ICF ≥ 0.8 y los perfiles de bandas del ARDRA, se construyó una clave interactiva de identificación mediante el programa Open-DELTA (Dallwitz, 1980; Dallwitz *et al.*, 1993). Los nombres de las especies bacterianas, los caracteres, el peso de los caracteres y las fotos de los perfiles de bandas del ARDRA fueron importados en el Editor del programa DELTA (Dallwitz *et al.*, 1999) y el programa Open-DELTA intkey fue utilizado para crear la clave (Dallwitz *et al.*, 1995).

3. RESULTADOS

En la Tabla 2 se resumen los resultados obtenidos al comparar las secuencias de la colección de aislados con las bases de datos NCBI y RDP.

Tabla 2. Asignación de aislados a especies bacterianas basado en el porcentaje de similitud en la secuencia del gen ARNr 16S.

Especie(s) relacionada	No. de Aislados	% de similitud
Secuencias asignadas a una especie:		
<i>B. licheniformis</i>	21	98 - 99%
<i>B. safensis</i>	6	99%
<i>B. altitudinis</i>	1	99%
<i>B. megaterium</i>	1	97%
Secuencias asignadas a más de una especie:		
<i>B. subtilis</i> / <i>B. vallismortis</i>	51	98 - 99%
<i>B. cereus</i> / <i>B. thuringiensis</i>	19	97 - 99%
<i>B. licheniformis</i> / <i>B. subtilis</i>	1	98%
Secuencias no asignadas a especies:		
<i>B. licheniformis</i> / <i>B. safensis</i> / <i>B. altitudinis</i> / <i>B. subtilis</i> / <i>B. vallismortis</i>	6	92-95%

Todas las secuencias analizadas presentaron "valor e" igual a 0.

Durante las comparaciones, se consideró un porcentajes de similitud $\geq 97\%$ del gen ARNr 16S para la asignación de especie (Roselló-Mora & Amann, 2001; Cohan, 2002). Las cuatro cepas de referencia usadas en el estudio (*B. cereus* ATCC®10876™, *B. licheniformis* ATCC®12759™, *B. megaterium* ATCC®9885™, *B. subtilis* ATCC®6633™) mostraron similitud de secuencias $\geq 97\%$ con secuencias depuradas de las correspondientes especies en las bases de datos (Yarza *et al.*, 2010; Muñoz *et al.*, 2011); indicando su idoneidad como material de referencia de la colección de aislados. De los aislados estudiados, 29 fueron

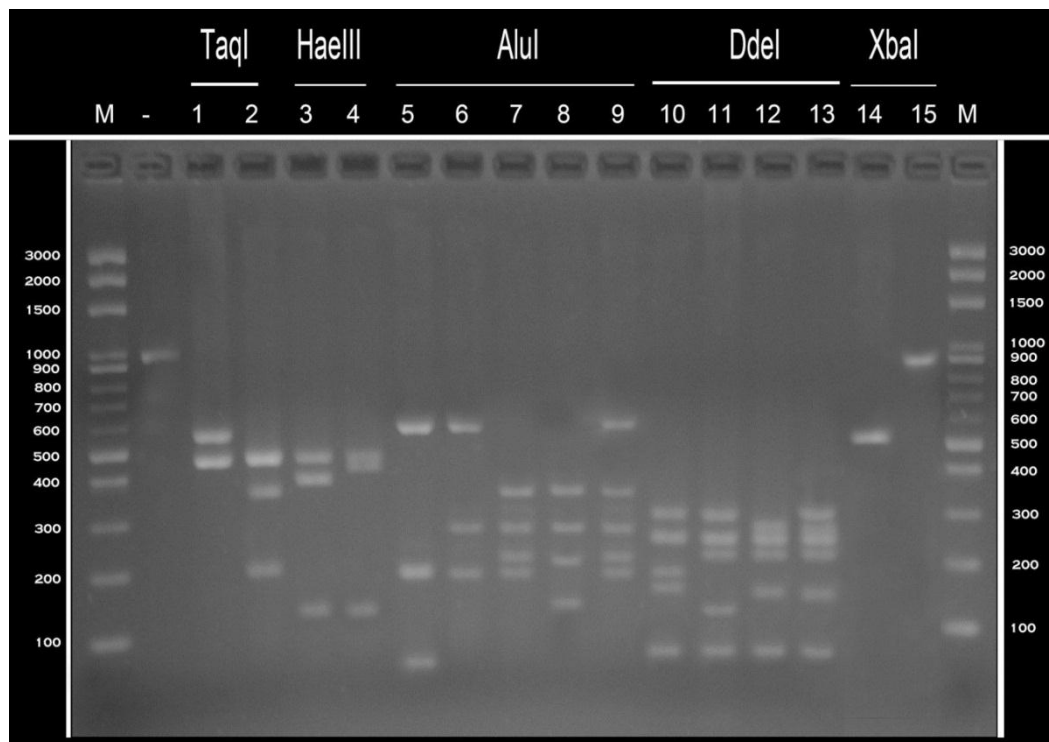
asignados a una única especie con porcentaje de similitud $\geq 97\%$ como sigue: *B. licheniformis* (21), *B. safensis* (6), *B. altitudinis* (1) y *B. megaterium* (1). Otros 71 aislados fueron asignados a más de una especie con porcentaje de similitud $\geq 97\%$ como sigue: *B. subtilis* / *B. vallismortis* (51), *B. cereus* / *B. thuringiensis* (19) y *B. licheniformis* / *B. subtilis* (1). Por último, 6 aislados mostraron porcentaje de similitud $< 97\%$ con accesiones de la base de datos correspondientes a las especies *B. altitudinis*, *B. licheniformis*, *B. safensis*, *B. subtilis* y *B. vallismortis*. Estos resultados claramente ratifican la inconveniencia del análisis de similitud de secuencias para la asignación de especie en el género *Bacillus*.

En la Figura 1, se muestra la relación filogenética entre aislados de la colección, cepas de referencias (ATCC) y accesiones depuradas de las bases de datos. Este análisis indicó la existencia de tres grupos consistentemente robustos. Un primer grupo (Grupo 1) formado con un valor de confianza del nodo del 100% que contiene: la cepa de referencia *B. cereus* ATCC®10876™, 2 accesiones de las bases de datos (*B. cereus* 877AEO16 y *B. thuringiensis* D16281) y 18 aislados de la colección. El segundo grupo (Grupo 2) formado con un valor de confianza del 97% contiene la cepa de referencia *B. megaterium* ATCC®9885™ y una accesión relacionada de la base de datos (*B. megaterium* D16273). Por último, el Grupo 3 resultó menos robusto (88% de confianza) y contiene las cepas de referencia *B. licheniformis* ATCC®12759™ y *B. subtilis* ATCC®6633™, 5 accesiones de las bases de datos (*B. altitudinis* AJ831842, *B. licheniformis* CP000002, *B. safensis* AF234854, *B. subtilis* AJ276351 y *B. vallismortis* AB021198) y 84 aislados de la colección. Con mucha menor robustez (62% de confianza), se pueden identificar dos subgrupos dentro del Grupo 3. Un subgrupo A que contiene las accesiones *B. safensis* AF234854 y *B. altitudinis* AJ831842 y 7 aislados de la colección; y un subgrupo B que contiene las accesiones *B. licheniformis* CP000002, *B. subtilis* AJ276351 y *B. vallismortis* AB021198 y 76 aislados de la colección. Cinco de los aislados de la colección que mostraron porcentaje de similitud menor al 97% con las accesiones de la base de datos (CB-07, MGS-23, MGS-39, WG-20 y WG-23),

quedaron incluidos en el subgrupo B. Solo el aislado WG-02 no agrupa con los subgrupos y se sitúa independiente dentro del Grupo 3. En general, si consideramos el agrupamiento de subgrupos, el análisis filogenético confirma los resultados observados durante la comparación de las secuencias de los aislados de la colección con las cepas de referencia (ATCC) y accesiones depuradas de las bases de datos.

En la Figura 2 son presentadas las variantes de restricción obtenidos mediante el ARDRA. De las 6 enzimas identificadas *in silico* por presentar sitios de restricción en la secuencia del gen ARNr 16S (AluI, DdeI, HaeIII, TaqI, XbaI y EcoRI), 5 resultaron informativas. Solo la enzima EcoRI, aunque cortó el ADN, no resultó polimórfica en la colección estudiada. La efectividad de las enzimas en términos del número de patrones o variantes de restricción obtenidas es como sigue: AluI (5) > DdeI (4) > TaqI y HaeIII (2) > XbaI (1).

Figura 2. Variantes de restricción obtenidas mediante el ARDRA usando las enzimas TaqI, HaeIII, AluI, DdeI y XbaI.



De izquierda a derecha: M) Marcador de peso molecular 100pb, (-) Control negativo (fragmento amplificado sin digerir), Las líneas de la 1 a la 15 corresponden los patrones de restricción obtenidos con las enzimas TaqI, HaeIII, AluI, DdeI y XbaI (ver detalles en los anexos A-E).

La mayoría de las variantes de restricción encontradas en el estudio (12 de 14 variantes) están representadas en las cepas de referencias usadas (*B. cereus* ATCC®10876™, *B. licheniformis* ATCC®12759™, *B. megaterium* ATCC®9885™,

B. subtilis ATCC®6633™), permitiendo así su diferenciación mediante los patrones ARDRA (Tabla 3). El resto de los aislados de la colección, mostraron alguna de las 12 variantes de restricción encontradas en las cepas de referencia (ver anexos A-E); excepto el aislado UC-03. Dicho aislado mostró además dos patrones infrecuentes (V9 y V13) ausentes en el resto de la colección.

Tabla 3. Variantes de restricción obtenidas con el ARDRA.

ATCC / Aislados	Variantes o Patrones de Restricción													
	TaqI		HaeIII		AluI					DdeI				XbaI
	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14
<i>B. cereus</i> ATCC®10876™	■		■		■					■				
<i>B. licheniformis</i> CC®12759™		■		■		■					■			
<i>B. megaterium</i> ATCC®9885™	■			■				■		■				■
<i>B. subtilis</i> ATCC®6633™		■					■					■		
Aislado UC-03								■					■	

Las celdas sombreadas significan presencia del patrón de restricción indicado en la Figura 2.

En correspondencia con el análisis filogenético presentado en la Figura 1, los 18 aislados incluidos en el Grupo 1, presentan alta similitud en sus variantes de restricción con la cepa de referencia *B. cereus* ATCC®10876™ (ver anexos A-E). El análisis *in silico* de los sitios de restricción de las enzimas AluI, DdeI, HaeIII y TaqI en la secuencia del gen ARNr 16S de las accesiones *B. cereus* 877AEO16 y *B. thuringiensis* D16281 muestra que ambas accesiones contienen los sitios de restricción sin modificación; lo que no ayuda a resolver la asignación de los mencionados aislados a alguna de las dos especies mencionadas. Por otra parte, los aislados contenidos en el subgrupo A del Grupo 3 según el análisis filogenético (MGS-01, MGS-03, MGS-07, WG-12, WG-16, WGS-04 y WGS-23), mostraron un patrón de restricción similar a la cepa de referencia *B. subtilis* ATCC®6633™ (ver anexos A-E). El análisis *in silico* de los sitios de restricción de las enzimas AluI, DdeI, HaeIII y TaqI en la secuencia del gen ARNr 16S de las accesiones *B.*

safensis AF234854 y *B. altitudinis* AJ831842 tampoco mostró modificación en los sitios de restricción; limitando la asignación de dichos aislados a algunas de las mencionadas especies.

En la Tabla 4 son presentados los valores de ICF calculados para especies del género *Bacillus* previamente reportadas para zonas carboníferas (Meslé *et al.*, 2013) que usamos en este estudio.

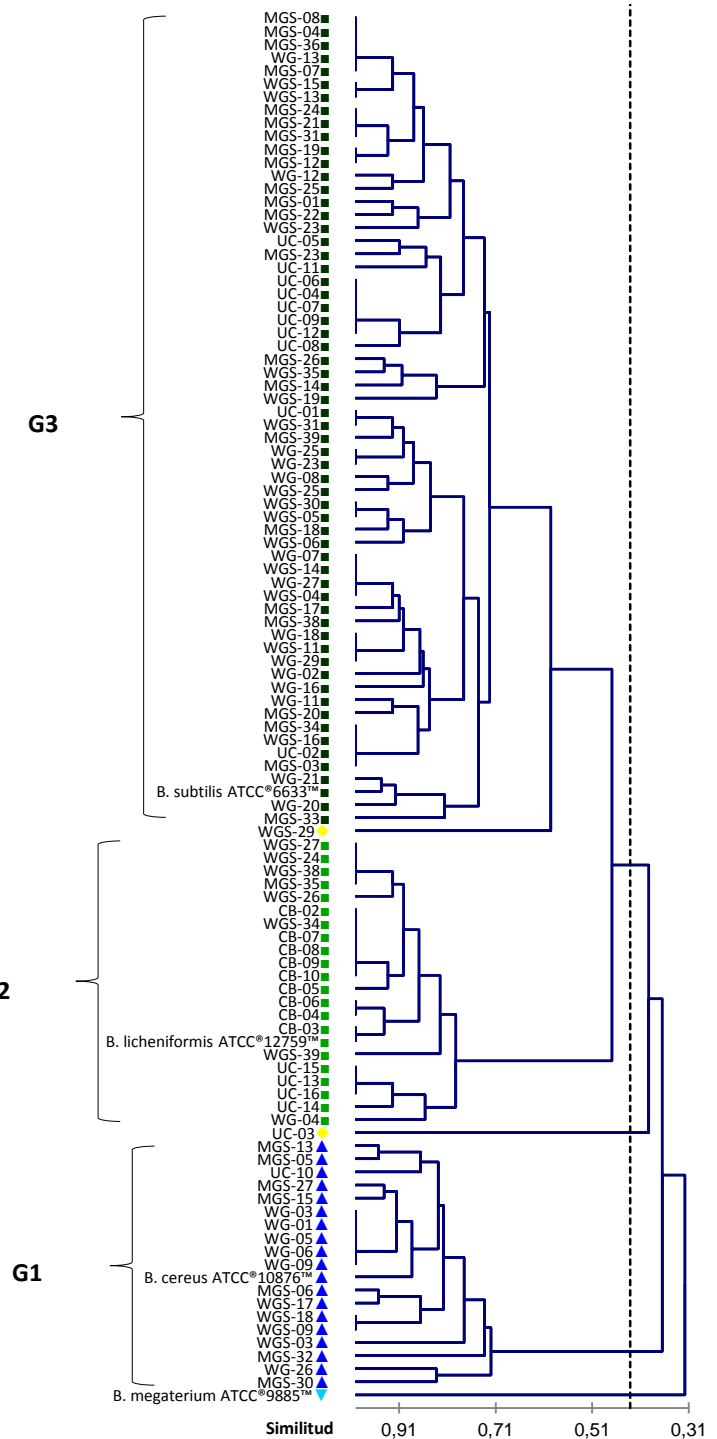
Tabla 4. Índice de Concordancia Fenotípica y sus correspondientes errores estándar.

Especies	ICF ± EE
<i>B. licheniformis</i>	0,872 ± 0,083
<i>B. cereus</i> / <i>B. thuringiensis</i>	0,864 ± 0,064
<i>B. safensis</i>	0,816 ± 0,140
<i>B. subtilis</i> / <i>B. vallismortis</i>	0,811 ± 0,110

En el anexo F son presentadas las frecuencias de resultados esperados, ya sean positivos o negativos, para los 26 descriptores fenotípicos usados para el cálculo del ICF. De esos 26 descriptores, solo 21 mostraron frecuencias de resultados esperados ≥ 0.8 en las diferentes especies analizadas. Estos son los siguientes: 1) respuesta a tinciones diferenciales (Gram y tinción de esporas), 2) forma bacteriana, 3) motilidad, 4) relación con el oxígeno, 5) actividad catalasa y oxidasa, 6) producción de gas y/o ácido por vía fermentativa de diferentes fuentes de carbono (Lactosa, Sacarosa, Glicerol, Maltosa, Manitol y Glucosa), 7) producción de 2,3-butanediol a partir de glucosa (prueba Voges Proskauer), 8) utilización de urea como fuente de carbono, 9) tolerancia a NaCl, 10) hidrólisis enzimática de almidón y 11) crecimiento 50°C. A pesar de ser descriptores robustos, la respuesta a tinción de Gram y de esporas, la forma bacteriana y la producción de gas por vía fermentativa de Glicerol y Lactosa no resultaron polimórficos en la colección de aislados, por lo que estos descriptores fueron descartados en futuros análisis.

Las variables fenotípicas descritas arriba y los datos del ARDRA fueron usadas para estudiar la similitud (Índice de Dice) entre los aislados estudiados y la diversidad fue representada mediante un análisis de conglomerados (Figura 3). Considerando un nivel de corte de 0.5, este análisis mostró tres grupos fenotípicos principales (1, 2 y 3). Cada grupo incluyó una cepa de referencia (ATCC) diferente (1: *B. cereus* ATCC®10876™, 2: *B. licheniformis* ATCC®12759™, 3: *B. subtilis* ATCC®6633™); indicando la proximidad fenotípica de los aislados de cada grupo con la cepa de referencia correspondiente. Dos aislados de la colección (UC-03 y WGS-29) y una cepa de referencia (*B. megaterium* ATCC®9885™) se ubicaron independientemente de los mencionados grupos. En general, el análisis fenético basado en descriptores fenotípicos y de ARDRA mostró concordancia parcial con el análisis filogenético; aunque el fenético resultó más resolutivo. Esto sugiere que el uso de estos tipos de descriptores puede ser esencial para la identificación de especies del género *Bacillus*.

Figura 3. Dendrograma basado en descriptores fenotípicos y del ARDRA.



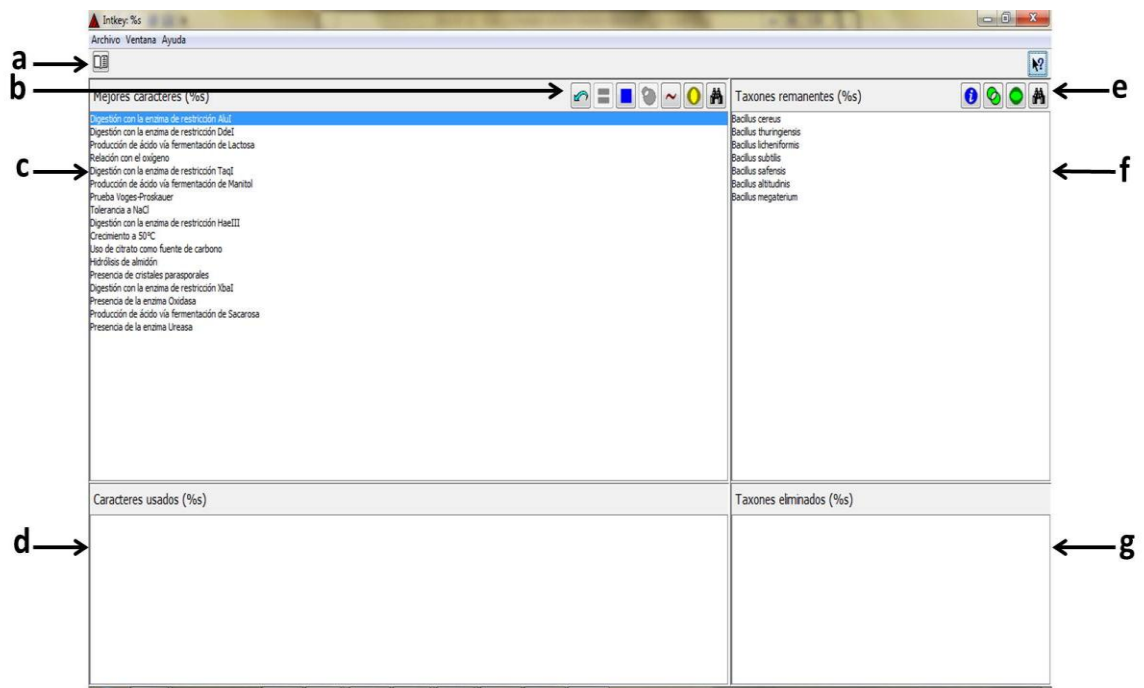
El árbol se construyó con el método UPGMA, basado en el índice de similaridad de Dice. Los colores y figuras geométricas indican la especie con la que se relacionó cada uno de los aislados: ■, *B. licheniformis*; ■, *B. subtilis*; ▲, *B. cereus*; ▼, *B. megaterium* y ◆, Aislados de la colección no relacionados.

Las variables fenotípicas y los datos del ARDRA también fueron usados para construir la clave interactiva de identificación (<https://www.dropbox.com/s/nmtrmp7pnppan31/Clave%20Bacillus.rar>). En total, se usaron 12 descriptores fenotípicos (morfológicos, fisiológicos y metabólicos) y 5 genotípicos (Perfiles de bandas del ARDRA) para definir siete especies (*B. cereus*, *B. thuringiensis*, *B. licheniformis*, *B. subtilis*, *B. safensis*, *B. altitudinis* y *B. megaterium*). En el diseño de la clave, los descriptores más excluyentes como: la digestión con las enzimas Alul y Ddel, la producción de ácido vía fermentación de la Lactosa y la relación con el oxígeno, se presentan como las variables de mayor peso para iniciar la identificación. A medida que van quedando las especies altamente relacionadas, se presentan descriptores como: la presencia de cristales proteínicos en la especie *B. thuringiensis*, la incapacidad de hidrolizar almidón en la especie *B. safensis*, la no utilización de citrato como fuente de carbono en la especie *B. altitudinis*, y el corte del gen ARNr 16S con la enzima XbaI de la especie *B. megaterium*. Dichos descriptores fueron muy útiles para diferenciar especies altamente relacionadas. El uso de la mencionada clave de identificación, permitió identificar los aislados de la colección. Las especies contenidas en la colección son las siguientes: *Bacillus altitudinis*, *Bacillus cereus*, *Bacillus licheniformis*, *Bacillus safensis*, *Bacillus subtilis* y *Bacillus thuringiensis*. Los detalles pueden ser consultados en el anexo G.

Para utilizar la clave interactiva de identificación se debe descargar e instalar el programa open-delta (<https://code.google.com/p/open-delta/downloads/list>), una vez realizado esto se podrá abrir el fichero Intkey de la carpeta Clave Bacillus. En la ventana principal (Figura 3), se puede observar a la derecha la lista de caracteres en el orden mencionado arriba (c) y a la izquierda la lista de especies (f). Al seleccionar un carácter aparecerá una nueva ventana con las opciones de respuesta a dicho carácter, en el caso de los descriptores ARDRA aparecerá una imagen con los perfiles de bandas y las respectivas opciones de respuesta. A medida que se van seleccionando los resultados obtenidos con el aislado a

identificar, los caracteres usados se irán mostrando en la parte inferior izquierda de la ventana (d) y las especies eliminadas en la parte inferior derecha (g); finalmente, se llegará a la identificación cuando solo quede una especie en la lista de la parte superior derecha (f). Además de la identificación mediante la selección de carácter por carácter, la clave interactiva ofrece otras aplicaciones que se podrán realizar a través de los iconos de las barras de herramientas de caracteres (b) y especies (e), estas aplicaciones nos permiten comparar especies, conocer los caracteres exclusivos de cada especie, reducir el número de caracteres y/o especies, cambiar el orden de los caracteres, entre otros. Por último, En el icono “Introduction and referencias” ubicado en la parte superior derecha de la ventana (a), se podrá encontrar una breve descripción de la clave y las especies estudiadas para construirla. Para más detalles sobre el uso de la clave interactiva se recomienda descargar el tutorial en: <http://delta-intkey.com/www/programs.htm> opción DELTA intkey tutorial PDF.

Figura 4. Ventana Principal de la clave interactiva de identificación.



El significado de las letras (A-G) esta descrito en el texto.

4. DISCUSIÓN

El análisis de similitud de secuencias del gen ARNr 16S relacionó los 102 aislados de la colección con siete especies de la base de datos NCBI. 21 de los aislados fueron identificados como *B. licheniformis* (similitud $\geq 98\%$), 6 como *B. safensis* (similitud 99%) y 1 como *B. altitudinis* (similitud 99%). Los demás aislados, no pudieron ser identificados debido a que presentaron porcentajes de similitudes $< 97\%$ o a que se relacionaron con más de una especie de la base de datos. Estos resultados confirman las limitaciones del análisis de homología de secuencias del gen ARNr 16S para distinguir especies estrechamente relacionadas del género *Bacillus*, como ha sido previamente indicado por otros autores (Earl *et al.*, 2007; Maughan & Van der Auwera, 2011).

Dentro de las especies que encontramos como estrechamente relacionadas se encuentran *B. cereus* y *B. thuringiensis*. La secuencia del gen ARNr 16S de estas dos especies fue prácticamente idéntica y 19 de los aislados (incluyendo la cepa *B. cereus* ATCC®10876™) presentaron similitudes de secuencias $\geq 98\%$ con ambas especies. Por su alta similitud, deferentes autores han recomendado considerar estas dos especies como ecotipos de la misma especie (Carlson *et al.*, 1994; Helgason *et al.*, 2000; Priest *et al.*, 2004; Rasko *et al.*, 2005; Tourasse *et al.*, 2006). Maughan & Van der Auwera (2011) plantean que el problema de identificación de las especies *B. cereus* y *B. thuringiensis* tiene raíces históricas y radica en que esta se ha definido con base en caracteres fenotípicos erróneamente utilizados para seguir diferencias filogenéticas. La mayor diferencia fenotípica entre estas especies es la presencia de cristales proteínicos con toxinas insecticidas en *B. thuringiensis*. Sin embargo, los genes que codifican para tales toxinas se encuentran en plásmidos que pueden recombinar con los presentes en la especie *B. cereus* (Carlson *et al.*, 1994). Adicionalmente, las cepas de *B. thuringiensis* que pierden estos plásmidos se vuelven indistinguibles de la especie *B. cereus* (Maughan & Van der Auwera, 2011). Estos eventos de recombinación

hacen básicamente indistinguible la especie *B. cereus* de *B. thuringiensis*. Por ello, los estudios dirigidos a desarrollar nuevos descriptores fenotípicos que permitan identificar dichas especies son de alta importancia taxonómica.

Así mismo, 51 aislados (incluyendo la cepa *B. subtilis* ATCC®6633™) presentaron similitudes $\geq 98\%$ con las especies *B. subtilis* y *B. vallismortis*, indicando las limitaciones del análisis de secuencias de gen ARNr 16S con fines de identificación en correspondencia con un estudio previo (Earl *et al.*, 2007). Estas dos especies están estrechamente relacionadas tanto a nivel fenotípico como genotípico (Roberts *et al.*, 1996). En este caso, ellas solo pueden distinguirse mediante el porcentaje de hibridación ADN-ADN y perfiles de restricción de genes específicos (Roberts *et al.*, 1996).

El análisis filogenético basado en secuencias de gen ARNr 16S (Figura 1) nos permitió ver las relaciones entre los aislados y/o especies. Los aislados pertenecientes a las especies *B. cereus/B. thuringiensis* (triángulos) se agruparon en un grupo robusto (100% de confianza) que se distingue genéticamente de los demás grupos. Por su parte, los aislados aquí asignados a las especies *B. safensis* y *B. altitudinis* (círculos) también formaron un grupo distinguible pero estrechamente relacionado con los aislados de las especies *B. licheniformis* y *B. subtilis/B. vallismortis* (cuadrados). Estas últimas especies, se observan dentro de la misma rama en el árbol filogenético a pesar de que existen diferencias entre sus secuencias (Tabla 1). Esto podría deberse a limitantes de las herramientas matemáticas y computacionales usadas para la construcción del árbol. En tal sentido, la presencia o ausencia de una sola entrada (secuencia) puede reordenar completamente las ramas del árbol. Por ejemplo, aquí encontramos que al eliminar la secuencia del aislado UC-03, que presentó un 98% de similitud con las especies *B. licheniformis* y *B. subtilis*, las ramas del árbol se reorganizan y se pueden observar los 20 aislados de la especie *B. licheniformis* en una rama independiente soportada con un 86% de confianza del nodo (Anexo H). Sin embargo, no se

puede olvidar que, aún con las diferencias identificadas (Tabla 1), se trata de especies que son estrechamente relacionadas. Por último, los aislados CB-07, MGS-23 Y WG-02, se observaron en ramas independientes de gran longitud posiblemente por sus bajos porcentajes de similitud (94%, 92% y 95%, respectivamente) con las secuencias de especies conocidas usadas para el análisis filogenético; sugiriendo que podría tratarse de nuevas especies.

Las relaciones filogenéticas encontradas aquí para especies del género *Bacillus* son congruentes con las reportadas previamente en la literatura que muestran una estrecha relación entre las especies *B. safensis*, *B. subtilis*, *B. vallismortis* y *B. licheniformis*; mientras que difieren ligeramente de las indicadas para las especies *B. cereus*, *B. thuringiensis* y *B. megaterium* (Xu & Côté, 2003; Huang *et al.*, 2005; Shivaji *et al.*, 2006; Yarza *et al.*, 2008, 2010; Munoz *et al.*, 2011; Maughan & Van der Auwera, 2011). Maughan & Van der Auwera (2011) han estudiado las relaciones taxonómicas de especies de *Bacillus* desde el punto de vista genotípico y fenotípico indicando que cuando las cepas y/o especies muestran alta homología de secuencia del gen ARNr 16S su fenotipo probablemente también es muy similar; sin embargo, cuando las cepas y/o especies tienen fenotipos similares no necesariamente tienen alta homología del gen ARNr 16S. Los resultados aquí presentados fueron congruentes con tales argumentos. Algunos de los aislados relacionados en el árbol basado en el gen ARNr 16S también están relacionados en el árbol basado en descriptores fenotípicos, pero aislados que se encuentran relacionados en el árbol basado en fenotipo no lo están totalmente en el árbol basado en homología de secuencia del gen ARNr 16S (Anexos H, I). En tal sentido, se debe tener en cuenta que el uso de caracteres fenotípicos para la identificación o clasificación puede resultar inconsistente. De acuerdo con el Bergey's Manual of Systematic Bacteriology (De Vos *et al.*, 2009) una prueba fenotípica es considerada positiva para una especie cuando más del 85% de las cepas y/o aislados producen resultados positivos; mientras que es considerada negativa cuando menos del 15% dan resultado positivo. Sin embargo, pueden

existir pruebas donde entre el 16 y el 84% de los aislados dan resultado positivo. Es por ello, que los descriptores fenotípicos deben ser usados como complemento de la información genética, la cual resulta más confiable en la mayoría de los casos.

En tal sentido, el ARDRA es uno de los métodos moleculares más confiables en la identificación de las especies. Mediante este método, se logró identificar los aislados basándonos en la comparación del perfil de bandas obtenidos entre estos y las cepas de referencia (*B. cereus* ATCC®10876™, *B. licheniformis* ATCC®12759™, *B. megaterium* ATCC®9885™, *B. subtilis* ATCC®6633™). En total, con cinco enzimas se observaron 18 aislados que presentaron el mismo perfil de bandas de la cepa de referencia *B. cereus* ATCC®10876™, 21 de *B. licheniformis* ATCC®12759™ y 62 de *B. subtilis* ATCC®6633™. Uno de los aislados (UC-03), presentó un perfil de bandas único (Anexos A - E).

La enzima AluI fue la más informativa permitiendo distinguir por sí sola las cuatro cepas de referencia y el aislado UC-03 mostrando resultados congruentes con estudios previos (Wu *et al.*, 2006; Mohanty *et al.*, 2011; Cihan *et al.*, 2012). Aunque con probada utilidad para distinguir especies del género *Bacillus* (Wu *et al.*, 2006, Cihan *et al.*, 2012), la enzima TaqI no resultó muy útil en este estudio mostrando solo dos perfiles de bandas en la colección de aislados y cepas de referencia estudiadas. Por su parte, la enzima DdeI, que no había sido previamente usada en la identificación de especies de *Bacillus* mediante ARDRA mostró, como la enzima AluI, sitios de restricción polimórficos en la secuencia del gen ARNr 16S de la colección de aislados y cepas de referencia; lo que la hace muy informativa. Por el contrario, la otra enzima no usada previamente XbaI mostró alta especificidad por la cepa *B. megaterium* ATCC®9885™.

En general, la combinación de las enzimas usadas en este trabajo mostró ser efectivas para la identificación de todas las especies del género *Bacillus* en

estudio. Al igual que el análisis de homología de secuencias, los ARDRA demostraron que la especie *B. cereus* es la más distante genéticamente. Su identificación fue posible con solo una de las enzimas siguientes: HaeIII, TaqI, AluI y DdeI; mientras que para identificar especies más relacionadas como *B. licheniformis* y *B. subtilis* se requirió al menos dos enzimas, AluI y DdeI. Por otro lado, los 8 aislados anotados aquí como *B. safensis* y *B. altitudinis* mostraron el mismo perfil de bandas para todas las enzimas utilizadas que la cepa de referencia *B. subtilis* ATCC®6633™. Así mismo, los aislados CB-07, MGS-23 y WG-02 presentaron mayormente perfiles de bandas similares a las cepas de referencia *B. licheniformis* ATCC®12759™ y *B. subtilis* ATCC®6633™. Este hallazgo sugiere que dichos aislados posiblemente corresponden a alguna de estas especies. Sin embargo, se recomienda realizar más análisis moleculares con estos tres aislados, que incluyan regiones del genoma diferentes al gen ARNr 16S, para descartar totalmente la posibilidad de que sean especies nuevas.

Por último, el aislado UC-03 presentó perfiles de bandas únicos con las enzimas AluI y DdeI; pero que comparten bandas con los perfiles de las cepas de referencia *B. subtilis* ATCC®6633™ y *B. licheniformis* ATCC®12759™. Dichos resultados son congruentes con el análisis de homología de secuencias en el cual este aislado presentó 98% de similitud con ambas especies, indicando que probablemente se trata de un híbrido entre ambas especies como se ha mostrado previamente por otros autores (Duncan *et al.*, 1989). Sin embargo, se recomienda estudiar el genoma completo del aislado UC-03 para confirmar si es un híbrido o una especie nueva.

Con los resultados obtenidos mediante el ARDRA y los resultados de los caracteres fenotípicos se generó un dendrograma que distribuyó los aislados de la colección en tres grupos (Figura 3), este dendrograma fue parcialmente congruente con el análisis filogenético aunque más resolutivo. Además de los tres principales grupos observados (1: 18 aislados y *B. cereus* ATCC®10876™, 2: 21

aislados y *B. licheniformis* ATCC®12759™ y 3: 61 aislados y *B. subtilis* ATCC®6633™) el aislado UC-03 se observa en una rama independiente debido a sus diferencias genotípicas ya mencionadas (posible híbrido), la especie *B. megaterium* ATCC®9885™ se observa en la base del árbol sin ningún aislado relacionado a ella y el aislado WGS-29 se observa en una rama independiente del grupo 3, este aislado aunque no mostró diferencias genotípicas (análisis de similitud de secuencias del gen ARNr 16S y ARDRA) se encuentra separado del grupo por sus diferencias fenotípicas, por lo que probablemente podría considerarse como un ecotipo de la especie *B. subtilis*.

La identificación final de las especies fue posible gracias a la combinación de las tres técnicas utilizadas, el análisis de similitud de secuencias dio una primera aproximación de las posibles especies a las que pertenecían los aislados y logró identificar 27 de ellos; los caracteres fenotípicos permitieron conocer la capacidad metabólica de cada uno de los aislados y diferenciar entre las especies *B. cereus* y *B. thuringiensis*. Por último, mediante el ARDRA se pudo asignar a la especie *B. subtilis* aislados de la colección relacionados con dos posibles especies *B. subtilis* y *B. vallismortis*. Adicionalmente, con el ARDRA se pudo asignar a una especie aquellos aislados con homología de secuencias del gen ARNr 16S < 95% con las entradas de la bases de datos.

Los problemas de identificación encontrados en el presente estudio evidencian claramente la necesidad de consensos entre taxónomos para redefinir el concepto de especie microbiana. Por un lado, la utilización de genes altamente conservados pueden estar subestimando el número de especies ecológicamente distintas; y por el otro, las características metabólicas/ecológicas de las especies pueden no ser congruentes con la reconstrucción filogenética (Maughan & Van der Auwera, 2011). Se requerirá desarrollar técnicas que permitan medir la variación a diferentes niveles con aproximaciones al fenotipo, el genotipo y a la filogenia de los microorganismos como ya ha sido sugerido (Vandamme *et al.*, 1996; Janda &

Abbott, 2002; Schleifer, 2009; Moore *et al.*, 2010). Aproximar el concepto de especie en un contexto ecológico y evolutivo (Cohan 2002, 2006; Koepfel *et al.*, 2007), podría ser el modo más realista de identificar los microorganismos. Sin embargo, hasta tanto esto ocurra la taxonomía polifásica seguirá siendo la mejor aproximación para la identificación de especies.

5. CONCLUSIONES

El análisis de homología de secuencias del gen ARNr 16S no es suficiente para distinguir especies estrechamente relacionadas dentro del género *Bacillus*. Aunque esta técnica es más informativa, los descriptores fenotípicos son indispensables para la identificación de especies estrechamente relacionadas dentro de este género.

La identificación polifásica basada en los descriptores genotípicos y fenotípicos aquí usados constituye una estrategia útil para la identificación de especies del género *Bacillus* que habitan el subsuelo profundo de minas de carbón colombianas. Mientras no se llegue a un consenso en la definición de especie bacteriana o al desarrollo de un método universal que permita la identificación de todas las especies de *Bacillus* conocidas hasta hoy, la identificación polifásica será la mejor opción para la identificación de especies de este género.

La clave interactiva aquí desarrollada proporciona una forma rápida y sencilla para identificar las especies de *Bacillus* que habitan el subsuelo profundo de zonas carboníferas de Colombia.

6. RECOMENDACIONES

Se recomienda realizar más estudios con los aislados CB-07, MGS-23, WG-02 Y UC-03 que incluyan genes diferentes al ARNr 16S, para confirmar o descartar lo expuesto aquí. Dado su utilidad, se recomienda ampliar la clave de identificación con la incorporación de otros descriptores y especies de diferentes zonas y ambientes del país.

BIBLIOGRAFÍA

- Achtman, M., & Wagner, M. (2008). Microbial diversity and the genetic nature of microbial species. *Nature Reviews Microbiology*, 6(6), 431-440.
- Altschul S.F., Gish W., Miller W., Myers E.W. & Lipman D.J. (1990) Basic Local Alignment Search Tool. *Journal of Molecular Biology*, 215, 403-410.
- Anderson, A. S., & Wellington, E. M. (2001). The taxonomy of *Streptomyces* and related genera. *International Journal of Systematic and Evolutionary Microbiology*, 51(3), 797-814.
- Ansari M.I. & Malik A. (2013) Recent Development in the Methods of Studying Microbial Diversity. In: *Management of Microbial Resources in the Environment*. Malik A. et al. (eds.) Springer. Cap. 2.
- Barrios S.X. (2014). Caracterización metabólica de microorganismos aislados de mantos de carbón en la mina cerrejón (Guajira, Colombia). Tesis de grado. Escuela de Biología, Universidad Industrial de Santander (UIS).
- Carlson, C. R., Caugant, D. A., & Kolstø, A. B. (1994). Genotypic diversity among *Bacillus cereus* and *Bacillus thuringiensis* strains. *Applied and environmental microbiology*, 60(6), 1719-1725.
- Cihan, A. C., Tekin, N., Ozcan, B., & Cokmus, C. (2012). The genetic diversity of genus *Bacillus* and the related genera revealed by 16S ARNrgene sequences and ARDRA analyses isolated from geothermal regions of turkey. *Brazilian Journal of Microbiology*, 43(1), 309-324.

Cleenwerck I. & De Vos P. (2008) Polyphasic taxonomy of acetic acid bacteria: An overview of the currently applied methodology. *International Journal of Food Microbiology*. 125, 2-14.

Cohan F.M. (2002) What are Bacterial Species?. *Annual Review of Microbiology*, 56, 457-487.

Cohan, F. M. (2006). Towards a conceptual and operational union of bacterial systematics, ecology, and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1475), 1985-1996.

Cohan, F. M., & Perry, E. B. (2007). A systematics for discovering the fundamental units of bacterial diversity. *Current Biology*, 17(10), R373-R386.

Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., Brown, T., Porras-Alfaro, A., Kuske, C. R. & Tiedje, J. M. (2014). Ribosomal Database Project: data and tools for high throughput ARNr analysis. *Nucleic acids research*, 42, Database issue D633–D642.

Dallwitz, M. J., Paine, T. A. & Zurcher, E. J. (1999). User's guide to the DELTA Editor.

Dallwitz, M. J., Paine, T. A. & Zurcher, E. J. (1993). User's guide to the delta system: a general system for processing taxonomic descriptions.

Dallwitz, M. J., Paine, T. A. & Zurcher, E. J. (1995). User's guide to intkey: a program for interactive identification and information retrieval.

Dallwitz, M.J. (1980) A general system for coding taxonomic descriptions. *Taxon*, 29, 41-46.

De Clerck, E., Rodriguez-Diaz, M., Forsyth, G., Lebbe, L., Logan, N. A., & DeVos, P. (2004). Polyphasic Characterization of *Bacillus coagulans* Strains, Illustrating Heterogeneity within this Species, and Emended Description of the Species. *Systematic and applied microbiology*, 27(1), 50-60.

De Maagd, R. A., Bravo, A., Berry, C., Crickmore, N., & Schnepf, H. E. (2003). Structure, diversity, and evolution of protein toxins from spore-forming entomopathogenic bacteria. *Annual review of genetics*, 37(1), 409-433.

De Vos, P., Garrity, G. M., Jones, D., Krieg, N. R., Ludwig, W., Rainey, F. A., Schleifer, K. H. & Whitman, W. B. (2009). Bergey's Manual of Systematic Bacteriology. 2nd edn, Vol. 3: The *Firmicutes*. Springer, New York, pp. 21-128.

Duffner, F. M., Reinscheid, U. M., Bauer, M. P., Mutzel, A., & Müller, R. (1997). Strain differentiation and taxonomic characterisation of a thermophilic group of phenol-degrading bacilli. *Systematic and applied microbiology*, 20(4), 602-611.

Duncan, K. E., Istock, C. A., Graham, J. B., & Ferguson, N. (1989). Genetic exchange between *Bacillus subtilis* and *Bacillus licheniformis*: variable hybrid stability and the nature of bacterial species. *Evolution*, 1585-1609.

Earl, A. M., Losick, R., & Kolter, R. (2007). *Bacillus subtilis* genome diversity. *Journal of bacteriology*, 189(3), 1163-1170.

Felföldi, T., Székely, A. J., Gorál, R., Barkács, K., Scheirich, G., András, J., Rácz A. & Márialigeti, K. (2010). Polyphasic bacterial community analysis of an aerobic activated sludge removing phenols and thiocyanate from coke plant effluent. *Bioresource technology*, 101(10), 3406-3414.

Fox, G. E., Wisotzkey, J. D., & Jurtshuk, P. (1992). How close is close: 16S ARNrsequence identity may not be sufficient to guarantee species identity. *International Journal of Systematic Bacteriology*, 42(1), 166-170.

Fry J.C., Horsfield B., Sykes R., Cragg B.A., Heywood C., Kim G.T., Mangelsdorf K., Mildenhall D.C., Rinna J. & Vieth A. (2009). Prokaryotic populations and activities in an interbedded coal deposit, including a previously deeply buried section (1.6–2.3 km) above~ 150 Ma basement rock. *Geomicrobiology Journal* 26, 163-178.

Gallagher L.K., Glossner A.W., Landkamer L.L., Figueroa L.A., Mandernack K.W. & Munakata-Marr J. (2013). The effect of coal oxidation on methane production and microbial community structure in Powder River Basin coal. *International Journal of Coal Geology*. 115, 71-78

Gevers D., Cohan F.M., Lawrence J.G., Spratt B.G., Coenye T., Feil E.J., Stackebrandt E., Vand de Peer Y., Vandamme P., Thompson F.L. & Swings J. (2005) Re-evaluating prokaryotic species. *Nature Reviews Microbiology*, 3, 733-739.

Gevers D., Dawyndt P, Vandamme P, Willems A, Vancanneyt M, Swings J. & De Vos P. (2006) Stepping Stones Towards a New Procaryotic Taxonomy. *Philosophical Transactions: Biological Sciences*. 361, 1911-1916.

Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., & Tiedje, J. M. (2007). DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *International Journal of Systematic and Evolutionary Microbiology*, 57(1), 81-91.

Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95-98.

Helgason, E., Økstad, O. A., Caugant, D. A., Johansen, H. A., Fouet, A., Mock, M., Hegna, I. & Kolstø, A. B. (2000). *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*—one species on the basis of genetic evidence. *Applied and environmental microbiology*, 66(6), 2627-2630.

Heyndrickx, M., Vauterin, L., Vandamme, P., Kersters, K., & De Vos, P. (1996). Applicability of combined amplified ribosomal DNA restriction analysis (ARDRA) patterns in bacterial phylogeny and taxonomy. *Journal of Microbiological Methods*, 26(3), 247-259.

Huang, X. W., Niu, Q. H., Zhou, W., & Zhang, K. Q. (2005). *Bacillus nematocida* sp. nov., a novel bacterial strain with nematotoxic activity isolated from soil in Yunnan, China. *Systematic and applied microbiology*, 28(4), 323-327.

Imelda, J., & Raj, R. P (2007). Isolation and characterization of phytase producing *Bacillus* strains from mangrove ecosystem. *Journal of the Marine Biological Association of India*, 49(2), 177-182.

Janda, J. M., & Abbott, S. L. (2002). Bacterial identification for publication: when is enough enough?. *Journal of clinical microbiology*, 40(6), 1887-1891.

Johnson J.L. (1973) Use of nucleic-acid homologies in the taxonomy of anaerobic bacteria. *International Journal of Sistematic Bacteriology*. 23, 308-315.

Koeppel, A., Perry, E. B., Sikorski, J., Krizanc, D., Warner, A., Ward, D. M., & Cohan, F. M. (2007). Identifying the fundamental units of bacterial diversity: a

paradigm shift to incorporate ecology into bacterial systematics. *Proceedings of the National Academy of Sciences*, 105(7), 2504-2509.

Konstantinidis K. T. & Tiedje J. M. (2005a). Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences*. 102(7), 2567-2572.

Konstantinidis K.T. & Tiedje J M. (2007) Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead. *Current Opinion in Microbiology* 10(5), 504-509.

Konstantinidis, K. T., & Tiedje, J. M. (2005b). Towards a genome-based taxonomy for prokaryotes. *Journal of bacteriology*, 187(18), 6258-6264.

Laguerre, G., Allard, M., Revoy, F. & Amarger, N. (1994). Rapid Identification of Rhizobia by Restriction Fragment Length Polymorphism Analysis of PCR-Amplified 16s ARNrGenes. *Applied and Environmental Microbiology*. 61, 56-63.

Liu D. (2009). Purification of Nucleic Acids from Bacteria. In: *Handbook of Nucleic Acid Purification*. Liu D. (Ed.), Taylor & Francis CRC Press, Cap. 5.

Li D., Hendry P. & Faiz M., (2008). A survey of the microbial populations in some Australian coalbed methane reservoirs. *International Journal of Coal Geology* 76, 14-24.

Maidak, B. L., Larsen, N., McCaughey, M. J., Overbeek, R., Olsen, G. J., Fogel, K., Blandy, J. & Woese, C. R. (1994). The ribosomal database project. *Nucleic acids research*, 22(17), 3485-3487.

- Maiden, M. C. (2006). Multilocus sequence typing of bacteria. *Annual Reviews Microbiology*, 60, 561-588.
- Maniatis T., Fritsch E.F., Sambrook J. (1989) Molecular cloning. *Cold Spring Harbor Laboratory Press, Cold Spring Harbor , New York*.
- Maughan, H., & Van der Auwera, G. (2011). *Bacillus* taxonomy in the genomic era finds phenotypes to be essential though often misleading. *Infection, Genetics and Evolution*, 11(5), 789-797.
- Mayorga C.M. (2012). Caracterización microbiológica de consorcios microbianos asociados a la formación guaduas (cundinamarca, colombia). Tesis de grado. Escuela de Biología, Universidad Industrial de Santander (UIS).
- Meslé, M., Dromart, G., & Oger, P. (2013). Microbial methanogenesis in subsurface oil and coal. *Research in microbiology*, 164(9), 959-972.
- Mohanty, S., Choudhury P.K., Dash, A., Samanta M. & Maiti, N. K. (2011). Genotypic and phenotypic diversity of *Bacillus* spp. isolated from Freshwater Ecosystems. *Journal of Aquaculture Research & Development*, 2(2), 1-4.
- Moore, E. R., Mihaylova, S. A., Vandamme, P., Krichevsky, M. I., & Dijkshoorn, L. (2010). Microbial systematics and taxonomy: relevance for a microbial commons. *Research in microbiology*, 161(6), 430-438.
- Mora, D., Fortina, M. G., Nicastro, G., Parini, C., & Manachini, P. L. (1998). Genotypic characterization of thermophilic bacilli: a study on new soil isolates and several reference strains. *Research in Microbiology*, 149(10), 711-722.

Moyer, C. L., Tiedje, J. M., Dobbs, F. C., & Karl, D. M. (1996). A computer-simulated restriction fragment length polymorphism analysis of bacterial small-subunit ARNr genes: efficacy of selected tetrameric restriction enzymes for studies of microbial diversity in nature. *Applied and Environmental Microbiology*, 62(7), 2501-2507.

Munoz, R., Yarza, P., Ludwig, W., Euzéby, J., Amann, R., Schleifer, K. H., Glöckner, F. O. & Rosselló-Móra, R. (2011). Release LTPs104 of the all-species living tree. *Systematic and applied microbiology*, 34(3), 169-170.

Nicolaus, B., Panico, A., Manca, M. C., Lama, L., Gambacorta, A., Maugeri, T., Gugliandolo C. & Caccamo, D. (2000). A Thermophilic *Bacillus* Isolated From an Eolian Shallow Hydrothermal Vent Able to Produce Exopolysaccharides. *Systematic and applied microbiology*, 23(3), 426-432.

Nocker, A., Burr, M., & Camper, A. K. (2007). Genotypic microbial community profiling: a critical technical review. *Microbial Ecology*, 54(2), 276-289.

Olsen G. J. & Woese C. R. (1993) Ribosomal RNA: a key to phylogeny. *The Journal of the Federation of American Societies for Experimental Biology*. 7, 113-123.

Peix, A., Ramírez-Bahena, M. H., & Velázquez, E. (2009). Historical evolution and current status of the taxonomy of genus *Pseudomonas*. *Infection, Genetics and Evolution*, 9(6), 1132-1147.

Penner T.J., Foght J.M. & Budwill K. (2010). Microbial diversity of western Canadian subsurface coal beds and methanogenic coal enrichment cultures. *International Journal of Coal Geology* 82, 81-93.

Peplies, J., Kottmann, R., Ludwig, W., & Glöckner, F. O. (2008). A standard operating procedure for phylogenetic inference (SOPPI) using (rRNA) marker genes. *Systematic and applied microbiology*, 31(4), 251-257.

Pinto, F. G. S., Hungria, M., & Martins Mercante, F. (2007). Polyphasic characterization of Brazilian *Rhizobium tropici* strains effective in fixing N₂ with common bean (*Phaseolus vulgaris* L.). *Soil Biology and Biochemistry*, 39(8), 1851-1864.

Priest, F. G., Barker, M., Baillie, L. W., Holmes, E. C., & Maiden, M. C. (2004). Population structure and evolution of the *Bacillus cereus* group. *Journal of Bacteriology*, 186(23), 7959-7970.

Rasko, D. A., Altherr, M. R., Han, C. S., & Ravel, J. (2005). Genomics of the *Bacillus cereus* group of organisms*. *FEMS microbiology reviews*, 29(2), 303-329.

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N. N., Anderson, I. J., Cheng, J. F. & Woyke, T. (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature*. 1-7.

Roberts, M. S., Nakamura, L. K., & Cohan, F. M. (1996). *Bacillus vallismortis* sp. nov., a close relative of *Bacillus subtilis*, isolated from soil in Death Valley, California. *International journal of systematic bacteriology*, 46(2), 470-475.

Roselló-Mora R. & Amann R. (2001) The species concept for prokaryotes. *Federation of European Microbiological Societies, Microbiology Reviews*. 25,39-67.

Ruiz A., Poblet M., Mas A. & Guillamón J.M. (2000) Identification of acetic acid bacteria by RFLP of PCR-amplified 16S rDNA 16S-23S rDNA intergenic spacer. *International Journal of Systematic and Evolutionary Microbiology*. 50, 1981-1987.

Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution*, 4(4), 406-425.

Samelis, J., Bleicher, A., Delbès-Paus, C., Kakouri, A., Neuhaus, K., & Montel, M. C. (2011). FTIR-based polyphasic identification of lactic acid bacteria isolated from traditional Greek Graviera cheese. *Food microbiology*, 28(1), 76-83.

Satomi, M., La Duc, M. T., & Venkateswaran, K. (2006). *Bacillus safensis* sp. nov., isolated from spacecraft and assembly-facility surfaces. *International journal of systematic and evolutionary microbiology*, 56(8), 1735-1740.

Schleifer K.H. (2009) Classification of Bacteria and Archae: Past, present and future. *Systematic and Applied Microbiology*, 32, 533-542.

Shivaji, S., Chaturvedi, P., Suresh, K., Reddy, G. S. N., Dutt, C. B. S., Wainwright, M., Narlikar, J. V. & Bhargava, P. M. (2006). *Bacillus aerius* sp. nov., *Bacillus aerophilus* sp. nov., *Bacillus stratosphericus* sp. nov. and *Bacillus altitudinis* sp. nov., isolated from cryogenic tubes used for collecting air samples from high altitudes. *International Journal of Systematic and Evolutionary Microbiology*, 56(7), 1465-1473.

Sneath P.H.A. (1995) Thirty Years of Numerical taxonomy. *Systematic Biology*, 44(3), 281-298.

Stackebrandt E. & Goebel B.M. (1994) Taxonomic note: a place for DNA-DNA reassociation and 16S ARNrsequence analysis in the present species definition in bacteriology. *International Journal of Sistematic Bacteriology*. 44, 846-849.

Suzuki, K., Oguro, H., Yamakawa, T., Yamamoto, A., Akao, S., & Saeki, Y. (2008). Diversity and distribution of indigenous soybean-nodulating rhizobia in the Okinawa islands, Japan. *Soil Science & Plant Nutrition*, 54(2), 237-246.

- Tamura, K., & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular biology and evolution*, 10(3), 512-526.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular biology and evolution*, 30(12), 2725-2729.
- Tang, M., & You, M. (2012). Isolation, identification and characterization of a novel triazophos-degrading *Bacillus* sp.(TAP-1). *Microbiological Research*, 167(5), 299-305.
- Tang Y. Q., Ji P., Lai G. L., Chi C. Q., Liu Z.S. & Wu X. L. (2012). Diverse microbial community from the coalbeds of the Ordos Basin, China. *International Journal of Coal Geology*, 90, 21-33.
- Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research*, 22(22), 4673-4680.
- Tiedje, J. M., Asuming-Brempong, S., Nüsslein, K., Marsh, T. L., & Flynn, S. J. (1999). Opening the black box of soil microbial diversity. *Applied Soil Ecology*, 13(2), 109-122.
- Tourasse, N. J., Helgason, E., Økstad, O. A., Hegna, I. K., & KolstØ, A. B. (2006). The *Bacillus cereus* group: novel aspects of population structure and genome dynamics. *Journal of applied microbiology*, 101(3), 579-593.

Urakawa, H., Kita-Tsukamoto, K., & Ohwada, K. (1997). 16S rDNA genotyping using PCR/RFLP (Restriction Fragment Length Polymorphism) analysis among the family Vibrionaceae. *FEMS microbiology letters*, 152(1), 125-132.

Vandamme P., Pot B., Gillis M., De Vos P., Kersters K. & Swings J. (1996) Polyphasic Taxonomy, a Consensus Approach to Bacterial Systematics. *Microbiological Reviews*, 60(2), 407-438.

Wadhwa, K., Dudeja, S. S., & Yadav, R. K. (2011). Molecular diversity of native rhizobia trapped by five field pea genotypes in Indian soils. *Journal of basic microbiology*, 51(1), 89-97.

Wani, A.A., Surakasi, V.P., Siddharth, J., Raghavan, R.G., Patole, M.S., Ranade, D. & Shouche, Y.S. (2006) Molecular analyses of microbial diversity associated with the Lonar soda lake in India: an impact crater in a basalt area. *Research in Microbiology*, 57, 928–937.

Whitby, C., Earl, J., Lanyon, C., Gray, S., Robinson, J., Meadows, J., & Edwards, C. (2004). The molecular diversity of the methanogenic community in a hypereutrophic freshwater lake determined by PCR-RFLP. *Journal of applied microbiology*, 97(5), 973-984.

Whitman W.B., Coleman D.C. & Wiebe W.J. (1998). Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences* 95, 6578-6583.

Woese, C.R., Kandler, O. & Wheelis, M.L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences*. 87, 4576–4579.

Wu, D., Hugenholtz, P., Mavromatis, K., Pukall, R., Dalin, E., Ivanova, N. N., Kunin, V. ... & Eisen, J. A. (2009). A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature*, 462(7276), 1056-1060.

Wu, X. Y., Walker, M. J., Hornitzky, M., & Chin, J. (2006). Development of a group-specific PCR combined with ARDRA for the identification of *Bacillus* species of environmental significance. *Journal of microbiological methods*, 64(1), 107-119.

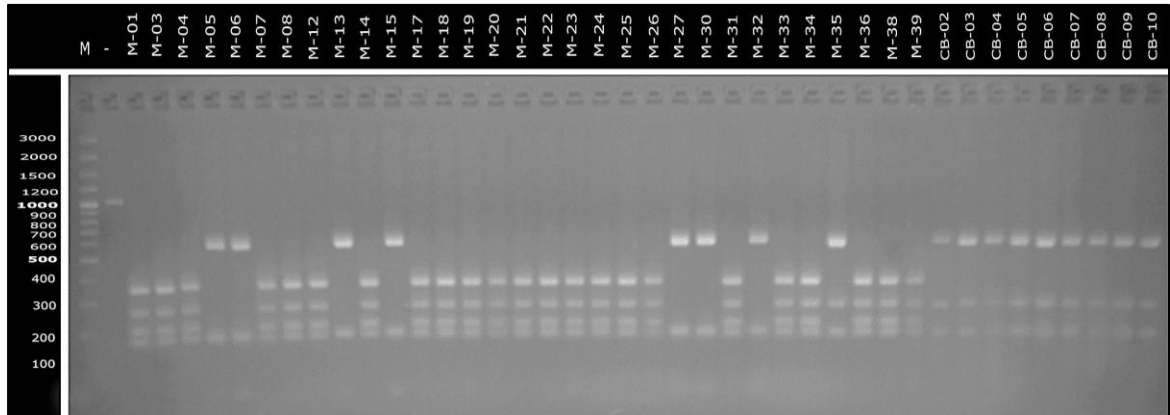
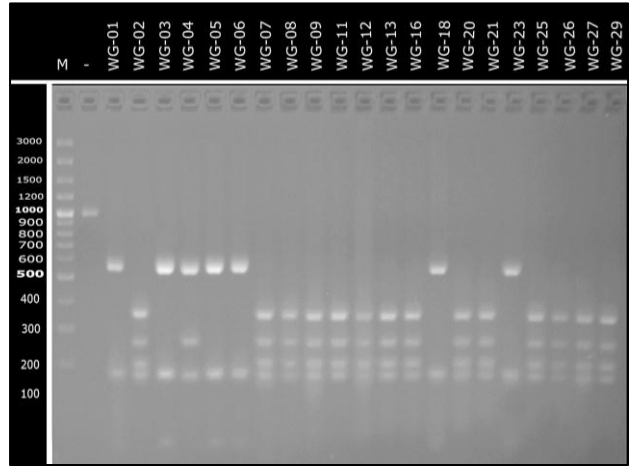
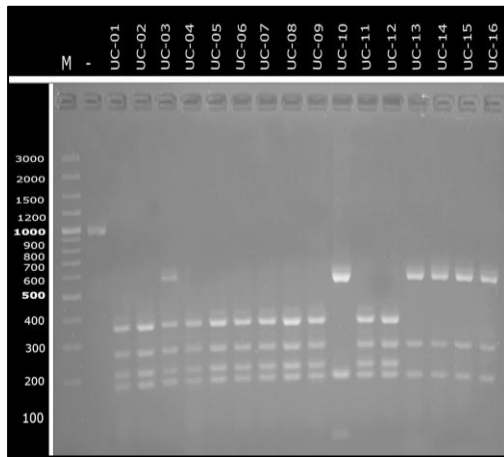
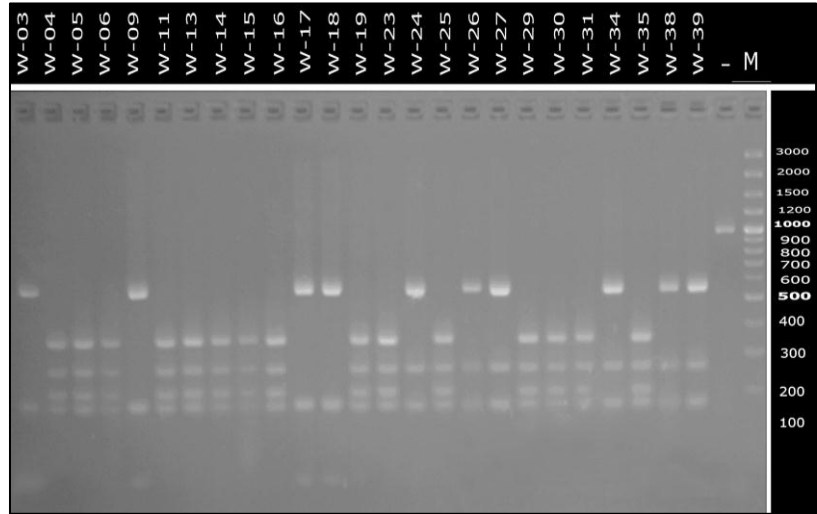
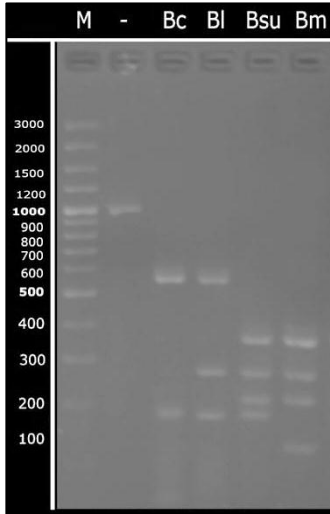
Xu, D., & Côté, J. C. (2003). Phylogenetic relationships between *Bacillus* species and related genera inferred from comparison of 3' end 16S rDNA and 5' end 16S–23S ITS nucleotide sequences. *International Journal of Systematic and Evolutionary Microbiology*, 53(3), 695-704.

Yarza, P., Ludwig, W., Euzéby, J., Amann, R., Schleifer, K. H., Glöckner, F. O., & Rosselló-Móra, R. (2010). Update of the All-Species Living Tree Project based on 16S and 23S ARNrsequence analyses. *Systematic and applied microbiology*, 33(6), 291-299.

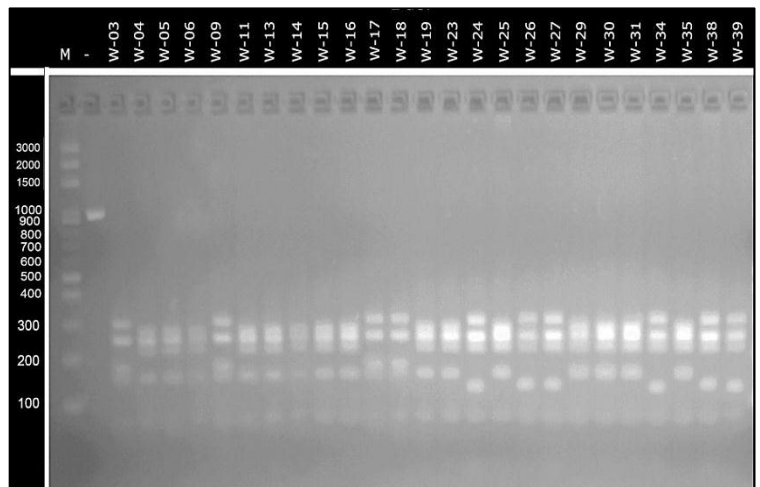
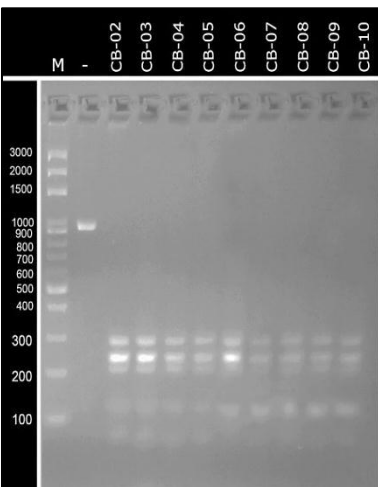
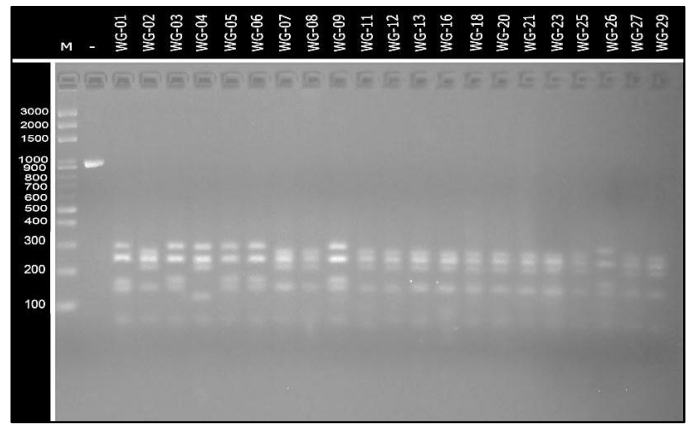
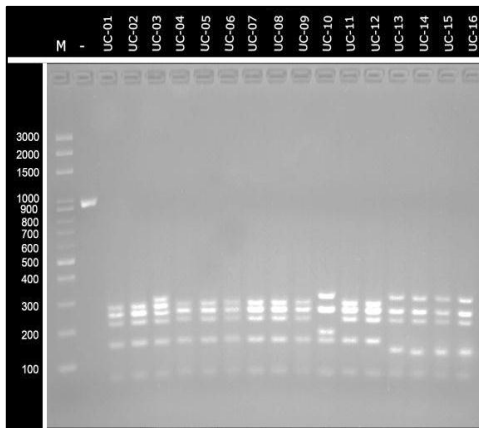
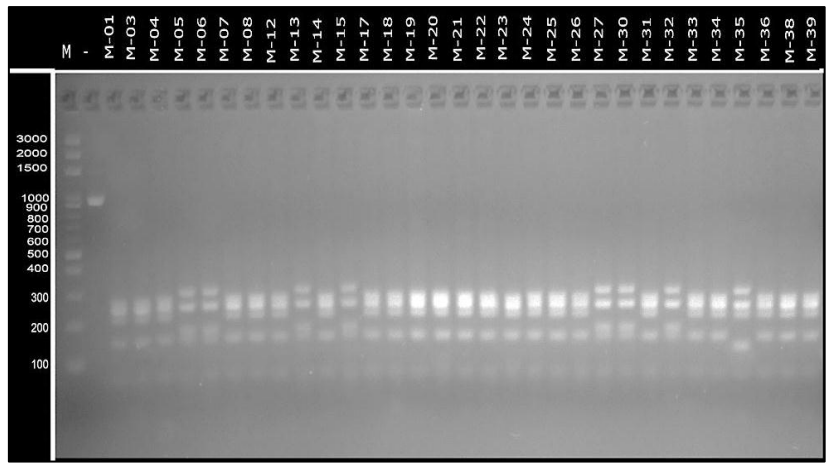
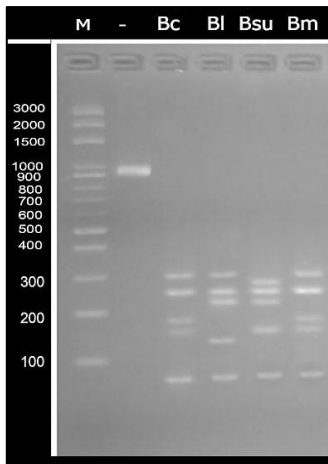
Yarza, P., Richter, M., Peplies, J., Euzéby, J., Amann, R., Schleifer, K. H., Ludwig, W., Glöckner, F. O. & Rosselló-Móra, R. (2008). The All-Species Living Tree project: A 16S rRNA-based phylogenetic tree of all sequenced type strains. *Systematic and applied microbiology*, 31(4), 241-250.

ANEXOS

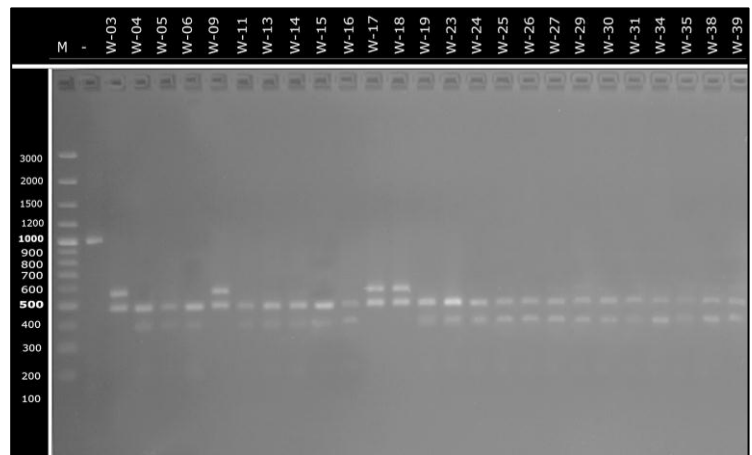
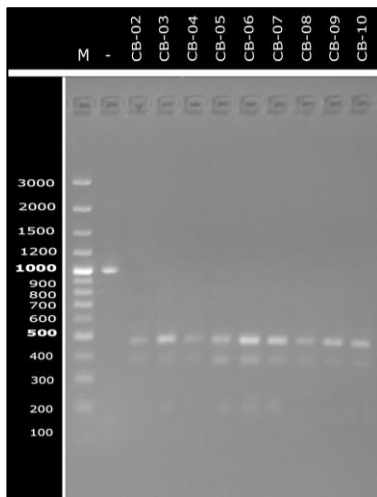
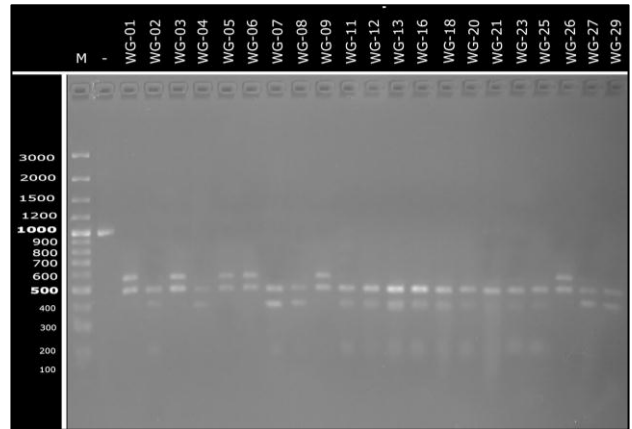
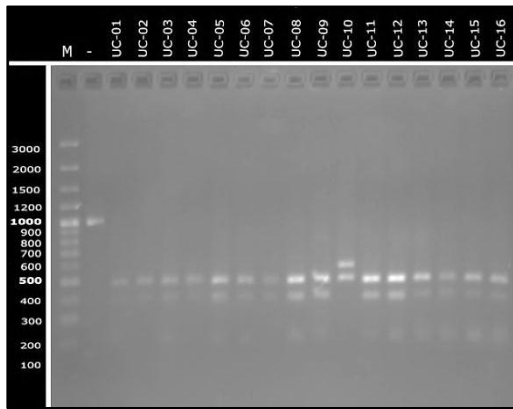
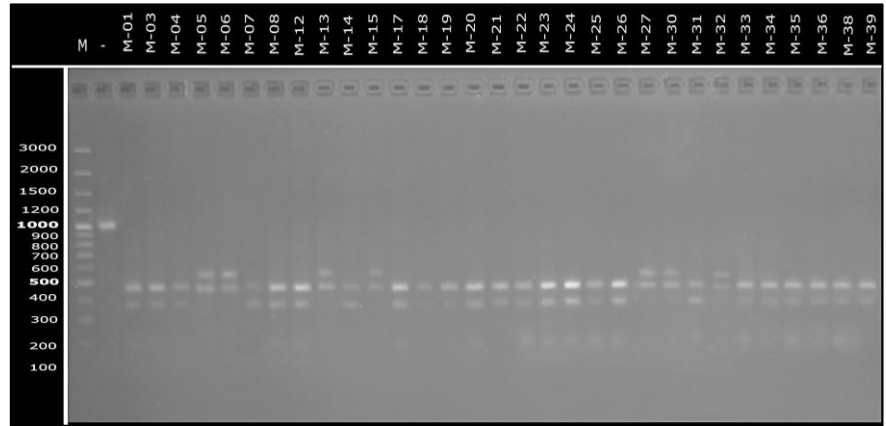
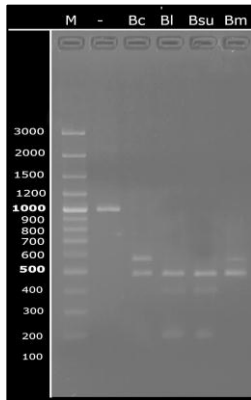
Anexo A. Perfiles de bandas obtenidos con la enzima de restricción AluI.



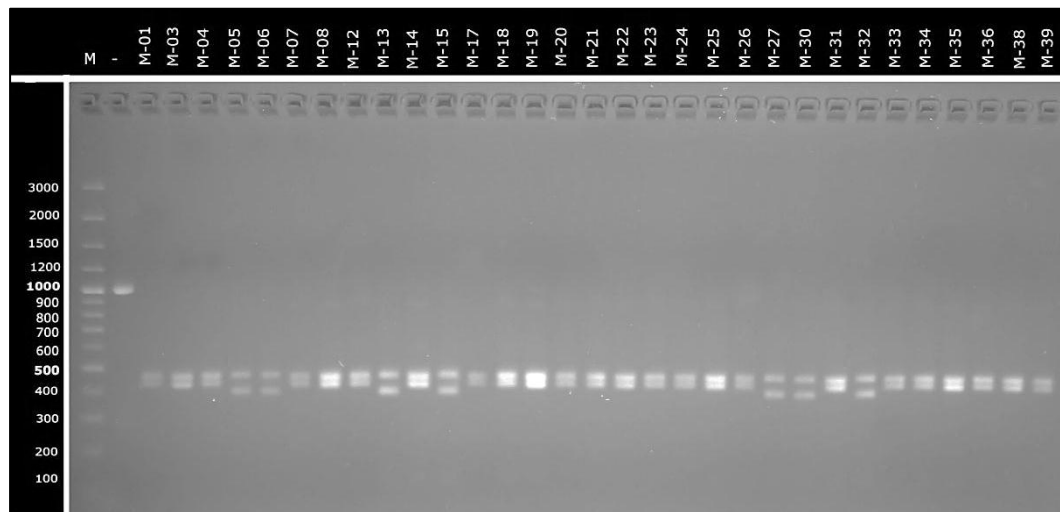
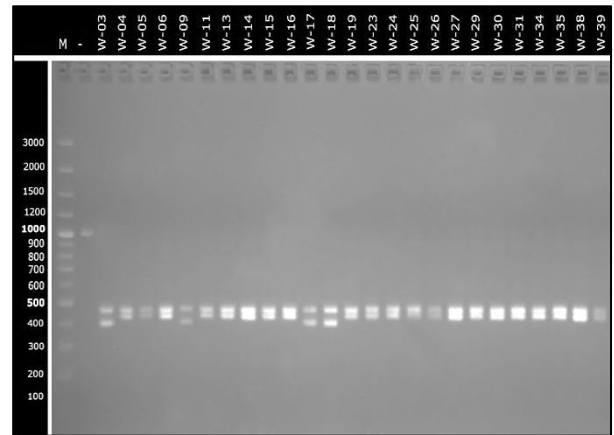
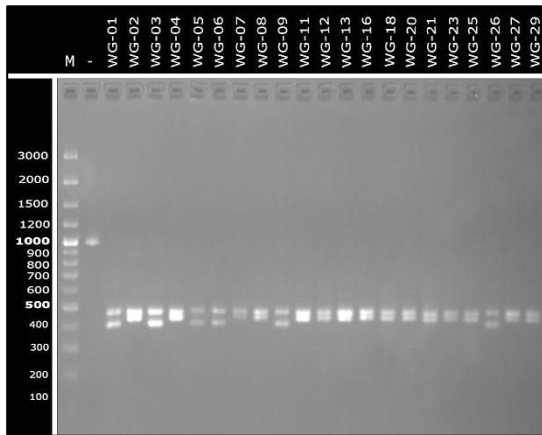
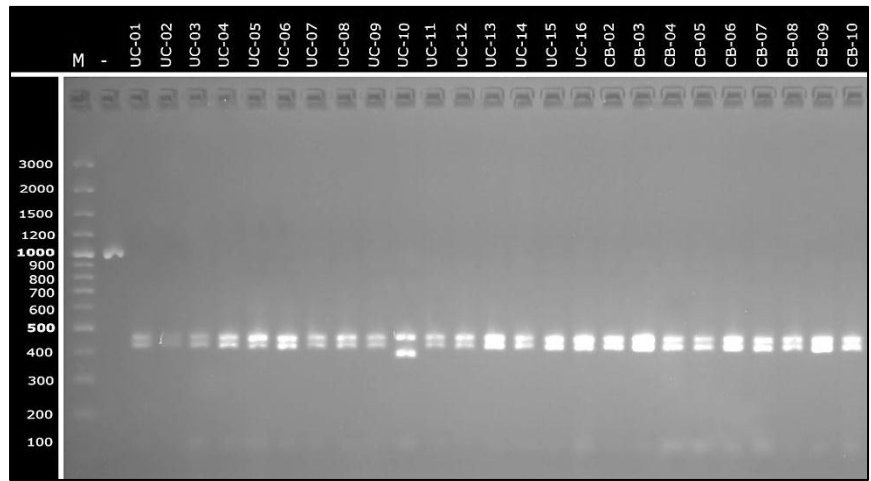
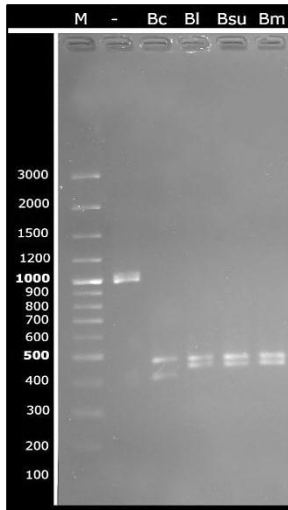
Anexo B. Perfiles de bandas obtenidos con la enzima de restricción DdeI.



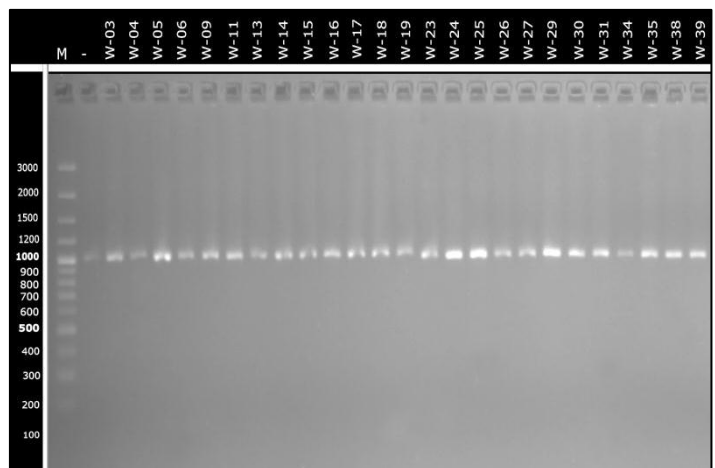
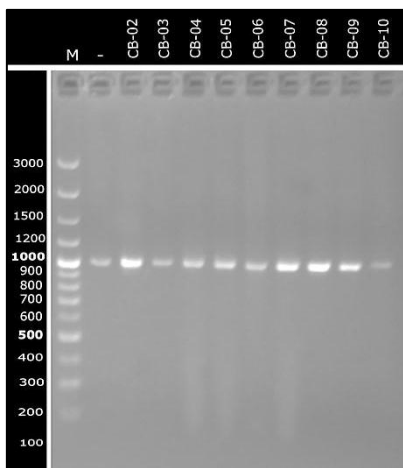
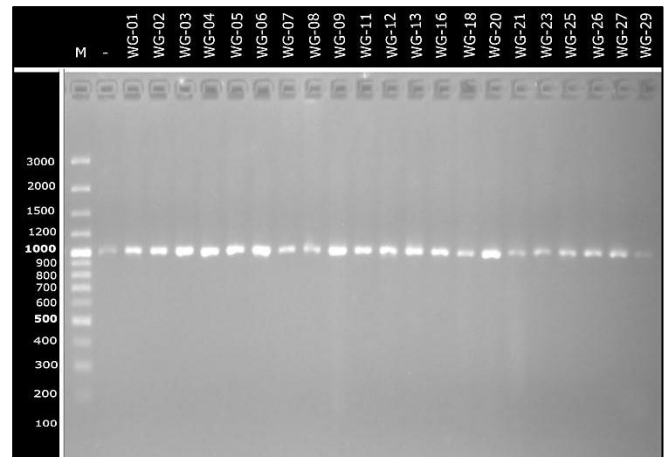
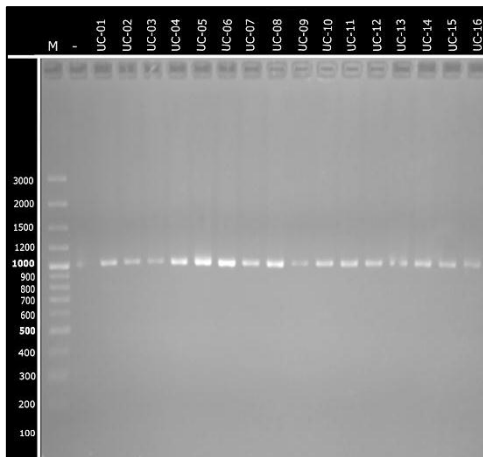
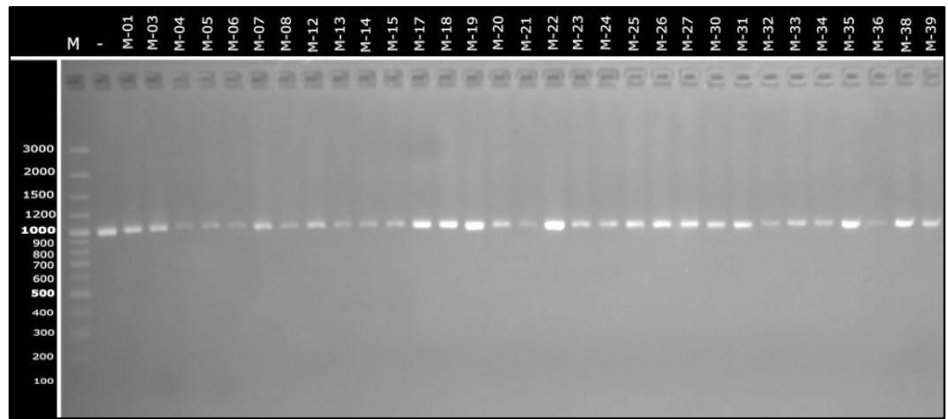
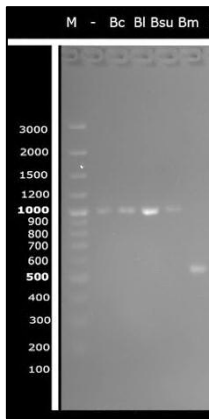
Anexo C. Perfiles de bandas obtenidos con la enzima de restricción TaqI.



Anexo D. Perfiles de bandas obtenidos con la enzima de restricción HaeIII.



Anexo E. Perfiles de bandas obtenidos con la enzima de restricción XbaI.



Anexo F Cálculo del Índice de Concordancia Fenotípica (ICF).

Las columnas sombreadas indican las pruebas bioquímicas no consideradas.

CÓDIGO	PRUEBAS BIOQUÍMICAS Y METABOLISMO																										
	Gr	B	Sf	lac g	gly g	mo	oxr	cat	oxi	lac a	mal g	suc g	suc a	glu g	glu a	man g	vp	ure	ms1	alm	50°C	cit	gly a	mal a	man a	gel	
<i>B. licheniformis</i> (Bergey's, 2009)	1	1	1	0	0	1	0	1	1-0	1	0	0	1	0	1	0	1	1-0	1	1	1	1	1	1	1	1	1
<i>B. licheniformis</i> ATCC®12759™	1	1	1	0	0	1	0	1	1	1	1	0	1	0	0	1	0	1	0	1	0	1	0	0	0	0	1
WGS-24	1	1	1	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	0	0	1	0	0	0
WGS-26	1	1	1	0	0	1	0	1	1	0	0	0	1	0	1	1	1	0	1	0	1	0	1	0	1	0	0
WGS-27	1	1	1	0	0	1	0	1	0	0	0	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	0
WGS-34	1	1	1	0	0	1	0	1	0	0	0	0	1	0	1	0	1	0	1	1	1	1	0	0	1	0	1
WGS-38	1	1	1	0	0	1	0	1	0	0	0	0	1	0	0	0	1	0	1	0	1	0	1	0	1	0	0
WGS-39	1	1	1	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	1	0	1	0	1	0	0	1	1
MGS-35	1	1	1	0	0	1	0	0	0	0	0	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	0
UC-13	1	1	1	0	0	1	0	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	1	0	0
UC-14	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0
UC-15	1	1	1	0	0	1	0	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	1	0	0
UC-16	1	1	1	0	0	1	0	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	1	0	0
WG-04	1	1	1	0	0	0	0	1	1	0	0	0	1	0	1	0	1	1	1	1	1	1	0	0	1	1	0
CB-02	1	1	1	0	0	1	0	1	0	0	0	0	1	0	1	0	1	0	1	1	1	1	0	0	0	0	0
CB-03	1	1	1	0	0	0	0	1	1	1	0	0	1	0	1	0	1	0	1	1	1	1	0	0	1	1	1
CB-04	1	1	1	0	0	1	0	1	0	1	0	0	1	0	1	0	1	0	1	1	1	1	0	0	0	0	1
CB-05	1	1	1	0	0	1	1	1	0	0	0	0	1	0	1	0	1	0	1	1	1	1	0	0	0	0	0
CB-06	1	1	1	0	0	1	0	1	0	1	0	0	1	0	1	0	1	0	1	1	1	1	0	0	1	1	1
CB-07	1	1	1	0	0	1	0	1	0	0	0	0	1	0	1	0	1	0	1	1	1	1	0	1	1	0	0
CB-08	1	1	1	0	0	1	0	1	0	0	0	0	1	0	1	0	1	0	1	1	1	1	0	1	1	0	1
CB-09	1	1	1	0	0	1	0	1	0	0	0	0	1	0	1	0	1	0	1	1	1	1	0	1	1	0	0
CB-10	1	1	1	0	0	1	0	1	0	0	0	0	1	0	1	0	1	0	1	1	1	1	0	1	1	1	1
F. R. Esperado	1,00	1,00	1,00	1,00	1,00	0,91	0,95	0,86	0,82	0,18	0,95	0,95	0,73	0,95	0,64	0,95	0,95	0,77	1,00	0,68	1,00	0,00	0,18	0,73	0,18	0,64	
ICF	0,87																										
ERROR ESTANDAR	0,08																										

CÓDIGO	PRUEBAS BIOQUÍMICAS Y METABOLISMO																												
	Gr	B	sf	lac g	gly g	mo	oxr	cat	oxi	lac a	mal g	suc g	suc a	glu g	glu a	man g	vp	ure	ms1	alm	50°C	cit	gly a	mal a	man a	gel			
<i>B. thuringiensis</i> (Bergey's, 2005)	1	1	1	0	0	1	0	1	0	0	0	1-0	0	1	0	1	1-0	1	1	0	1	1	1	0	1	1	1	0	1
<i>B. cereus</i> (Bergey's, 2005)	1	1	1	0	0	1	0	1	0	0	0	1-0	0	1	0	1	0	1	1	0	1	1	0	1	1	1	0	1	
<i>B. cereus</i> ATCC® 10876™	1	1	1	0	0	1	0	1	1	1	0	1	0	1	0	0	1	1	1	1	0	1	0	0	0	0	1		
WGS-03	1	1	1	0	0	1	0	1	0	1	0	0	0	1	0	0	0	0	1	1	0	0	0	1	0	0	0		
WGS-09	1	1	1	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0		
WGS-17	1	1	1	0	0	1	0	1	1	1	0	0	1	0	1	0	1	0	1	1	1	0	0	1	0	1	1		
WGS-18	1	1	1	0	0	1	0	1	0	0	0	1	0	1	0	0	0	0	1	1	1	0	0	1	0	0	0		
MGS-05	1	1	1	0	0	1	1	1	1	0	0	0	0	0	0	1	0	1	1	1	0	0	0	0	0	1	1		
MGS-06	1	1	1	0	0	1	0	1	0	1	0	1	0	1	0	1	0	1	1	1	1	0	0	0	0	1	1		
MGS-13	1	1	1	0	0	1	1	1	1	1	0	0	0	1	0	1	0	1	1	1	1	1	1	0	1	0	0		
MGS-15	1	1	1	0	0	1	0	1	1	1	0	0	0	1	0	1	0	1	1	0	1	1	0	1	0	1	0		
MGS-27	1	1	1	0	0	1	0	0	1	0	0	0	0	1	0	1	0	1	1	0	1	1	0	0	0	0	0		
MGS-30	1	1	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0		
MGS-32	1	1	1	0	0	1	0	1	0	0	0	0	0	1	0	1	1	1	0	1	1	0	0	1	0	0	0		
UC-10	1	1	1	0	0	1	0	0	1	0	0	0	0	1	0	1	0	0	1	1	0	1	0	0	0	0	0		
WG-01	1	1	1	0	0	0	0	1	1	0	0	1	0	1	0	1	0	1	1	0	0	0	1	0	0	0	0		
WG-03	1	1	1	0	0	1	0	1	1	0	0	0	1	0	1	0	1	0	1	1	0	0	0	1	0	0	0		
WG-05	1	1	1	0	0	1	0	1	1	0	0	1	0	1	0	1	0	1	1	0	0	0	1	0	0	0	0		
WG-06	1	1	1	0	0	1	0	1	1	0	0	1	0	1	0	1	0	1	1	0	0	0	1	0	0	0	0		
WG-09	1	1	1	0	0	1	0	1	1	0	0	1	0	1	0	1	0	1	1	0	0	0	1	0	0	0	0		
WG-26	1	1	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
F R. Esperado	1,00	1,00	1,00	1,00	1,00	0,95	0,89	0,79	0,74	0,68	1,00	0,95	0,56	0,95	0,74	1,00	0,74	0,89	0,89	0,84	0,53	0,16	0,00	0,50	1,00	0,21			
ICF	0,86																												
ERROR ESTANDAR	0,06																												

CÓDIGO	PRUEBAS BIOQUÍMICAS Y METABOLISMO																										
	Gr	B	sf	lac g	gly g	mo	oxr	cat	oxi	lac a	mal g	suc g	suc a	glu g	glu a	man g	vp	ure	ms1	alm	50°C	cit	gly a	mal a	man a	gel	
<i>B. vallismortis</i> (Bergey's, 2005)	1	1	1	0	0	1	1	1	1	0	0	0	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1
<i>B. subtilis</i> (Bergey's, 2005)	1	1	1	0	0	1	1	1	1-0	1-0	0	0	1	0	1	0	1	0	1	1	1	1-0	1	1	1	1	1
<i>B. subtilis</i> ATCC®6633	1	1	1	0	0	1	1	1	1	0	0	0	0	0	0	0	1	0	1	1	1	0	0	0	0	0	1
WGS-05	1	1	1	0	0	1	0	1	0	1	0	0	1	0	0	0	1	0	1	0	1	0	0	0	0	0	1
WGS-06	1	1	1	0	0	1	0	1	0	1	0	0	1	0	0	0	0	0	1	0	1	0	1	1	0	0	0
WGS-11	1	1	1	0	0	1	1	1	0	0	0	0	0	0	1	0	1	0	1	1	1	0	0	0	0	0	0
WGS-13	1	1	1	0	0	1	0	1	1	0	0	0	0	0	1	0	0	0	1	1	1	0	0	1	0	0	0
WGS-14	1	1	1	0	0	1	1	1	0	0	0	0	0	0	1	0	1	0	1	0	1	0	1	0	0	0	0
WGS-15	1	1	1	0	0	1	0	1	1	0	0	0	0	0	1	0	0	0	1	1	1	0	0	1	0	0	0
WGS-16	1	1	1	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	1	0	1	0	1	0	0	0	1
WGS-19	1	1	1	0	0	1	1	1	1	1	0	0	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0
WGS-25	1	1	1	0	0	1	1	1	0	0	0	0	1	0	0	0	1	0	1	0	1	0	1	0	0	0	1
WGS-29	1	1	1	0	0	1	1	1	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0
WGS-30	1	1	1	0	0	1	0	1	0	1	0	0	1	0	0	0	1	0	1	0	1	0	1	0	0	0	0
WGS-31	1	1	1	0	0	1	0	1	0	0	0	0	1	0	0	0	1	0	1	0	1	0	1	1	0	0	1
WGS-35	1	1	1	0	0	1	0	1	1	1	0	0	0	0	1	0	1	1	1	0	1	0	0	0	0	0	0
MGS-04	1	1	1	0	0	1	0	1	1	0	0	0	0	0	1	0	1	0	1	1	1	1	0	0	0	0	1
MGS-08	1	1	1	0	0	1	0	1	1	0	0	0	0	0	0	0	1	0	1	1	1	0	0	0	0	0	1
MGS-12	1	1	1	0	0	1	0	1	1	1	0	0	0	0	1	0	1	0	1	0	1	0	1	0	1	0	0
MGS-14	1	1	1	0	0	1	0	1	1	1	0	0	0	0	1	0	0	1	1	0	1	0	1	0	0	0	0
MGS-17	1	1	1	0	0	1	1	1	1	0	0	0	0	0	1	0	1	0	1	0	1	0	1	0	0	0	0
MGS-18	1	1	1	0	0	1	0	1	1	1	0	0	1	0	1	0	1	0	1	0	1	0	1	0	0	0	0
MGS-19	1	1	1	0	0	1	0	1	1	1	0	0	0	0	1	0	1	0	1	0	1	0	1	0	0	0	0
MGS-20	1	1	1	0	0	1	0	1	0	1	0	0	0	0	1	0	1	0	1	0	1	0	1	0	0	0	0
MGS-21	1	1	1	0	0	1	0	1	1	1	0	0	0	0	0	0	1	0	1	1	1	0	0	0	0	0	0
MGS-22	1	1	1	0	0	1	1	1	1	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0
MGS-23	1	1	1	0	0	1	0	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
MGS-24	1	1	1	0	0	1	0	1	1	1	0	0	0	0	1	0	1	0	1	1	1	1	0	0	0	0	0
MGS-25	1	1	1	0	0	1	0	1	1	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0	0	0
MGS-26	1	1	1	0	0	1	0	1	1	1	0	0	1	0	1	0	1	1	1	0	1	0	0	1	0	0	0
MGS-31	1	1	1	0	0	1	0	1	1	1	0	0	0	0	1	0	1	0	1	1	1	0	0	1	0	0	0
MGS-33	1	1	1	0	0	1	0	1	0	0	0	0	1	0	1	0	1	1	0	1	1	0	0	1	0	0	0
MGS-34	1	1	1	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0

MGS-36	1	1	1	0	0	1	0	1	1	0	0	0	0	0	0	1	0	1	1	1	0	0	0	0	0	
MGS-38	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0	0	
MGS-39	1	1	1	0	0	1	0	1	0	0	0	0	1	0	1	0	1	1	1	1	0	0	0	0	1	
UC-01	1	1	1	0	0	1	0	1	0	0	0	0	1	0	0	1	0	1	0	1	0	0	0	0	0	
UC-02	1	1	1	0	0	1	0	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0		
UC-03	1	1	1	0	0	1	0	1	0	1	0	0	1	0	1	0	1	0	0	0	1	0	1	0	0	
UC-04	1	1	1	0	0	1	0	1	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	
UC-05	1	1	1	0	0	1	0	1	0	1	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	
UC-06	1	1	1	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	
UC-07	1	1	1	0	0	1	0	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	
UC-08	1	1	1	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	
UC-09	1	1	1	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	
UC-11	1	1	1	0	0	1	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	
UC-12	1	1	1	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	
WG-07	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	1	1	
WG-08	1	1	1	0	0	1	1	1	1	0	0	0	1	0	1	0	1	0	1	0	1	0	1	1	1	
WG-11	1	1	1	0	0	1	1	1	0	1	0	0	0	0	0	1	0	1	0	1	0	0	0	1	1	
WG-13	1	1	1	0	0	1	0	1	1	0	0	0	0	0	1	0	1	0	1	1	1	0	0	0	1	
WG-18	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	1	0	1	1	1	0	0	0	0	0	
WG-20	1	1	1	0	0	1	1	1	0	0	0	0	1	0	1	0	1	1	1	1	1	0	1	1	1	
WG-21	1	1	1	0	0	1	1	1	1	0	0	0	1	0	1	0	1	1	1	1	1	0	1	1	1	
WG-23	1	1	1	0	0	1	0	1	1	0	0	0	1	0	1	0	1	0	1	0	1	0	1	1	1	
WG-25	1	1	1	0	0	1	0	1	1	0	0	0	1	0	1	0	1	0	1	0	1	0	1	1	1	
WG-27	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	
WG-29	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	1	0	1	1	1	0	0	0	0	0	
F R. Esperado	1,00	1,00	1,00	1,00	1,00	1,00	0,30	0,89	0,57	0,64	1,00	1,00	0,30	1,00	0,50	1,00	0,91	0,84	0,79	0,34	0,95	0,04	0,09	0,23	0,09	0,29
ICF	0,81																									
ERROR ESTANDAR	0,11																									

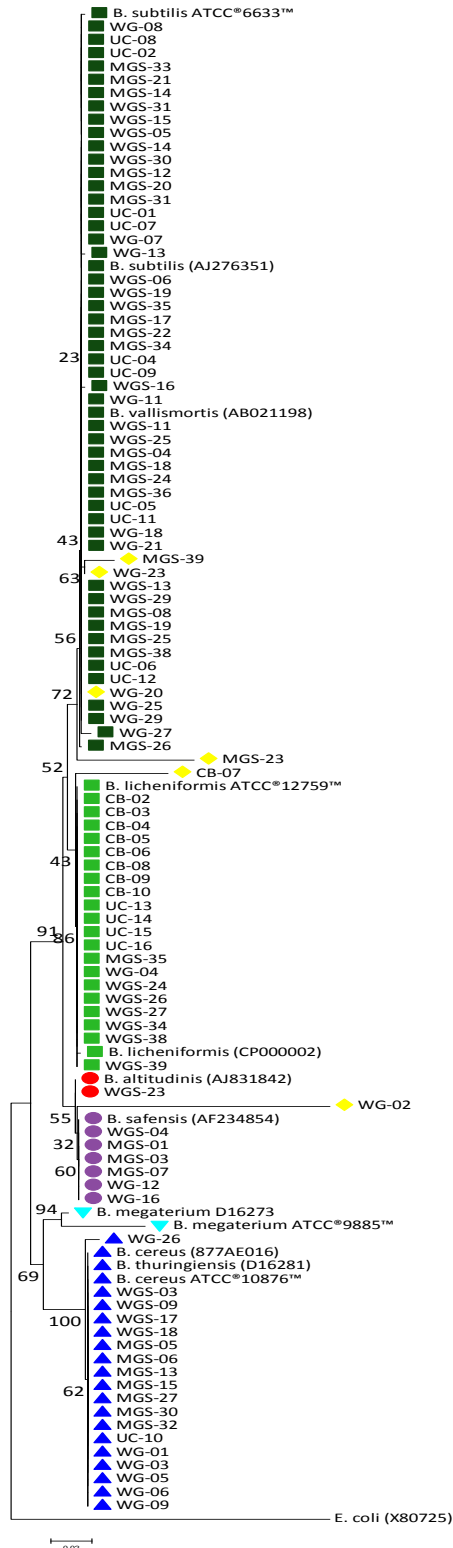
CÓDIGO	PRUEBAS BIOQUÍMICAS Y METABOLISMO																									
	Gr	B	sf	lac g	gly g	mo	oxr	cat	oxi	lac a	mal g	suc g	suc a	glu g	glu a	man g	vp	ure	ms1	alm	50°C	cit	gly a	mal a	man a	gel
B. safensis (Satomy et al., 2006)	1	1	1	-	-	1	1	1	1	1-0	0	-	1	0	1	-	1	0	1	0	1	1	1	1	1	1
WGS-04	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0
MGS-01	1	1	1	0	0	1	0	1	1	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	1
MGS-03	1	1	1	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	1
MGS-07	1	1	1	0	0	1	0	1	1	0	0	0	0	0	0	0	1	0	1	1	1	0	0	0	0	1
WG-02	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	1
WG-12	1	1	1	0	0	1	0	1	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	1
WG-16	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
F. R. Esperado	1,00	1,00	1,00	1,00	1,00	1,00	0,43	1,00	0,43	1,00	1,00	1,00	0,00	1,00	0,00	1,00	0,86	1,00	1,00	0,71	0,71	0,00	0,00	0,00	0,00	0,71
ICF	0,82																									
ERROR ESTANDAR	0,14																									

El código del aislado está relacionado con el origen del aislado (manto muestreado): (Cerrejón: **WGS**: agua del surtidero, **MGS**: lodo, **UC**: carbón y Cucunubá: **WG**: agua, **CB**: carbón); **(1)**: resultados positivos; **(0)**: resultados negativos. **Gr**: tinción de Gram; **B**: formas bacterianas (**1**: bacilo; **0**: coco); **SF**: método tinción de esporas Schaeffer-Fulton; **Lac G**: fermentación de lactosa medido por producción de gas en medio Lactosa caldo OXOID; **Glyc G**: fermentación de glicerol medido por producción de gas en medio Glicerol rojo de fenol caldo; **Mo**: motilidad; **OxR**: relación oxígeno bacteriano (**1**: aerobio; **0**: anaerobio facultativo); **Cat**: presencia enzima catalasa; **Oxi**: presencia enzima oxidasa; **Lac a**: fermentación de lactosa medido por acidificación en medio Lactosa caldo OXOID; **Mal g**: fermentación de maltosa medido por producción de gas en medio Maltosa rojo de fenol caldo; **Suc g**: fermentación de sacarosa medido por producción de gas en medio Sacarosa rojo de fenol caldo; **Suc a**: fermentación de sacarosa medido por acidificación en medio Sacarosa rojo de fenol caldo; **Glu g**: fermentación de glucosa medido por producción de gas en medio Glucosa rojo de fenol caldo; **Glu a**: fermentación de glucosa medido por acidificación en medio Glucosa rojo de fenol caldo; **Man g**: fermentación de manitol medido por producción de gas en medio Manitol rojo de fenol caldo; **VP**: producción de 2,3-butanediol a partir de glucosa medido en medio MR-VP; **Ure**: presencia de enzima ureasa como fuente de carbono medido en medio BBLTM Urea agar base; **MS1**: tolerancia de sal medido como crecimiento en medio Manitol Salado agar OXOID; **Alm**: presencia de almidón medido en medio Almidón agar; **50°C**: crecimiento a temperatura 50°C; **Cit**: uso de citrato como fuente de carbono en medio Citrato Simmons agar MERCK; **Glyc a**: fermentación de glicerol medido por producción de ácido Glicerol rojo de fenol caldo; **Mal a**: fermentación de maltosa medido por acidificación en medio Maltosa rojo de fenol caldo; **Man a**: fermentación de manitol medido por acidificación en medio Manitol rojo de fenol caldo; **Gel**: licuefacción de gelatina en medio Gelatina agar MERCK.

Anexo G. Listado de los aislados y la especie identificada.

AISLADO	ESPECIE	AISLADO	ESPECIE	AISLADO	ESPECIE
WGS-03	<i>B. cereus</i>	MGS-14	<i>B. subtilis</i>	UC-13	<i>B. licheniformis</i>
WGS-04	<i>B. safensis</i>	MGS-15	<i>B. cereus</i>	UC-14	<i>B. licheniformis</i>
WGS-05	<i>B. subtilis</i>	MGS-17	<i>B. subtilis</i>	UC-15	<i>B. licheniformis</i>
WGS-06	<i>B. subtilis</i>	MGS-18	<i>B. subtilis</i>	UC-16	<i>B. licheniformis</i>
WGS-09	<i>B. cereus</i>	MGS-19	<i>B. subtilis</i>	WG-01	<i>B. cereus</i>
WGS-11	<i>B. subtilis</i>	MGS-20	<i>B. subtilis</i>	WG-02	<i>B. safensis</i>
WGS-13	<i>B. subtilis</i>	MGS-21	<i>B. subtilis</i>	WG-03	<i>B. thuringiensis</i>
WGS-14	<i>B. subtilis</i>	MGS-22	<i>B. subtilis</i>	WG-04	<i>B. licheniformis</i>
WGS-15	<i>B. subtilis</i>	MGS-23	<i>B. subtilis</i>	WG-05	<i>B. cereus</i>
WGS-16	<i>B. subtilis</i>	MGS-24	<i>B. subtilis</i>	WG-06	<i>B. thuringiensis</i>
WGS-17	<i>B. cereus</i>	MGS-25	<i>B. subtilis</i>	WG-07	<i>B. subtilis</i>
WGS-18	<i>B. cereus</i>	MGS-26	<i>B. subtilis</i>	WG-08	<i>B. subtilis</i>
WGS-19	<i>B. subtilis</i>	MGS-27	<i>B. cereus</i>	WG-09	<i>B. cereus</i>
WGS-23	<i>B. altitudinis</i>	MGS-30	<i>B. cereus</i>	WG-11	<i>B. subtilis</i>
WGS-24	<i>B. licheniformis</i>	MGS-31	<i>B. subtilis</i>	WG-12	<i>B. safensis</i>
WGS-25	<i>B. subtilis</i>	MGS-32	<i>B. cereus</i>	WG-13	<i>B. subtilis</i>
WGS-26	<i>B. licheniformis</i>	MGS-33	<i>B. subtilis</i>	WG-16	<i>B. safensis</i>
WGS-27	<i>B. licheniformis</i>	MGS-34	<i>B. subtilis</i>	WG-18	<i>B. subtilis</i>
WGS-29	<i>B. subtilis</i>	MGS-35	<i>B. licheniformis</i>	WG-20	<i>B. subtilis</i>
WGS-30	<i>B. subtilis</i>	MGS-36	<i>B. subtilis</i>	WG-21	<i>B. subtilis</i>
WGS-31	<i>B. subtilis</i>	MGS-38	<i>B. subtilis</i>	WG-23	<i>B. subtilis</i>
WGS-34	<i>B. licheniformis</i>	MGS-39	<i>B. subtilis</i>	WG-25	<i>B. subtilis</i>
WGS-35	<i>B. subtilis</i>	UC-01	<i>B. subtilis</i>	WG-26	<i>B. cereus</i>
WGS-38	<i>B. licheniformis</i>	UC-02	<i>B. subtilis</i>	WG-27	<i>B. subtilis</i>
WGS-39	<i>B. licheniformis</i>	UC-03	Posible híbrido	WG-29	<i>B. subtilis</i>
MGS-01	<i>B. safensis</i>	UC-04	<i>B. subtilis</i>	CB-02	<i>B. licheniformis</i>
MGS-03	<i>B. safensis</i>	UC-05	<i>B. subtilis</i>	CB-03	<i>B. licheniformis</i>
MGS-04	<i>B. subtilis</i>	UC-06	<i>B. subtilis</i>	CB-04	<i>B. licheniformis</i>
MGS-05	<i>B. cereus</i>	UC-07	<i>B. subtilis</i>	CB-05	<i>B. licheniformis</i>
MGS-06	<i>B. cereus</i>	UC-08	<i>B. subtilis</i>	CB-06	<i>B. licheniformis</i>
MGS-07	<i>B. safensis</i>	UC-09	<i>B. subtilis</i>	CB-07	<i>B. licheniformis</i>
MGS-08	<i>B. subtilis</i>	UC-10	<i>B. thuringiensis</i>	CB-08	<i>B. licheniformis</i>
MGS-12	<i>B. subtilis</i>	UC-11	<i>B. subtilis</i>	CB-09	<i>B. licheniformis</i>
MGS-13	<i>B. cereus</i>	UC-12	<i>B. subtilis</i>	CB-10	<i>B. licheniformis</i>

Anexo H. Árbol filogenético basado en secuencias del gen ARNr 16S sin el aislado UC-03.



Anexo I. Dendrograma basado en descriptores fenotípicos.

