

IDENTIFICACIÓN *in silico* DE GENES PUTATIVOS DE PECTATO-LIASAS (E.C.  
4.2.2.2) DE INTERÉS BIOPROSPECTIVO A PARTIR DEL METAGENOMA DE  
BOSQUE ALTOANDINO - PARQUE NACIONAL NATURAL LOS NEVADOS



DANIELA QUIBANO ORDOÑEZ

UNIVERSIDAD DEL CAUCA  
FACULTAD DE CIENCIAS NATURALES, EXACTAS Y DE LA EDUCACIÓN  
PROGRAMA DE BIOLOGÍA  
POPAYÁN  
2019

IDENTIFICACIÓN *in silico* DE GENES DE PECTATO-LIASAS (E.C. 4.2.2.2) DE  
INTERÉS BIOPROSPECTIVO A PARTIR DEL METAGENOMA DE BOSQUE  
ALTO ANDINO - PARQUE NACIONAL NATURAL Los Nevados

DANIELA QUIBANO ORDOÑEZ

Trabajo de Grado para optar al título de

Bióloga

Directora

PATRICIA E. VELEZ V. M.Sc Ph.D

Codirectora

IVON ANDREA BOLAÑOS MARTINEZ M.Sc

UNIVERSIDAD DEL CAUCA

FACULTAD DE CIENCIAS NATURALES, EXACTAS Y DE LA EDUCACIÓN

PROGRAMA DE BIOLOGÍA

POPAYÁN

2019

Nota de aceptación

---

---

---

Directora: \_\_\_\_\_

Patricia Eugenia Velez, M.Sc Ph.D

Codirectora: \_\_\_\_\_

Ivon Andrea Bolaños Martinez, M.Sc

Jurado: \_\_\_\_\_

William Orlando Castillo, Ph.D

Jurado: \_\_\_\_\_

Fabian Tosse, Ph.D

Fecha de sustentación: Popayán, 8 de octubre de 2019

## **AGRADECIMIENTOS**

A mi familia, mi mamá y mi hermano por su amor, su acompañamiento durante todo este proceso y el apoyo que me han brindado desde el principio de éste camino en cada una de las decisiones que he tomado. A Christian Franco enseñarme que es la paciencia y decir siempre las palabras exactas que necesito en mis momentos de desesperación.

A mi codirectora Ivon Andrea Bolaños, por guiarme en este camino tanto académico como personalmente, ser mi tutora, consejera y amiga todo este tiempo y compartir de tu conocimiento conmigo.

A mi directora la Doctora Patricia E. Vélez, por confiar en mi y tenerme en cuenta siempre, por tus historias y por enseñarme que todo se puede lograr si uno insiste y trabaja arduo, además de bríndame todo ese conocimiento que me ha hecho crecer todos estos años.

A Adrián Rodríguez por tener la paciencia de sacarme de dudas y explicarme una y otra vez cuando no entendía, por ser uno de los mejores maestros aunque no fue mi profesor y más por sacar de su tiempo para colaborar conmigo.

Al grupo de Investigación y al Semillero de Investigación de Biología Molecular Ambiental y Cáncer - BIMAC, que me abrió las puertas al mundo de la metagenómica y la bioinformática y del cual he aprendido mucho; a mis compañeros del Semillero, con los que hemos aprendido y crecido académicamente y me han mostrado un gran apoyo.

Al Departamento de Biología (Profesores y Laboratoristas) de la Universidad del Cauca que me ayudaron a forjar la Bióloga que soy ahora.

Finalmente, agradecer a todos los que de una u otra forma hicieron parte de este proceso.

## TABLA DE CONTENIDO

<b>RESUMEN</b> .....	1
<b>1. INTRODUCCIÓN</b> .....	2
<b>2. JUSTIFICACIÓN</b> .....	5
<b>3. MARCO TEÓRICO</b> .....	6
<b>3.1. MICROORGANISMOS DEL SUELO, METAGENÓMICA Y BIOPROSPECCIÓN</b> 6	
<b>3.2. EL BOSQUE ALTOANDINO DEL PARQUE NACIONAL NATURAL LOS</b> <b>NEVADOS</b> .....	7
<b>3.3. PECTATO LIASAS (E.C. 4.2.2.2)</b> .....	9
2.3.1. Distribución de las pectato liasas.....	9
2.3.2. Estructura y clasificación de las pectato liasas.....	10
2.3.3. Funciones de las pectato liasas.....	13
<b>3.4. FILOGENIA MOLECULAR</b> .....	13
3.4.1. Selección de secuencias.....	13
3.4.2. Alineamiento múltiple.....	13
3.4.3. Especificación del modelo evolutivo.....	14
3.4.4. Construcción del árbol.....	14
3.4.5. Análisis del árbol.....	15
<b>4. ANTECEDENTES</b> .....	15
<b>5. OBJETIVOS</b> .....	19
<b>5.1. OBJETIVO GENERAL</b> .....	19
<b>5.2. OBJETIVOS ESPECÍFICOS</b> .....	19
<b>6. METODOLOGÍA</b> .....	19
<b>6.1. CREACIÓN DE UNA BASE DE DATOS LOCAL A PARTIR DE BASES DE</b> <b>DATOS PÚBLICA</b> .....	19
<b>6.2. BUSQUEDA DE HOMOLOGÍAS DE LA DASE DE DATOS LOCAL CON EL</b> <b>METAGENOMA</b> .....	20

<b>6.3.</b>	<b>BÚSQUEDA DE DOMINIOS DE PECTATO LIASAS .....</b>	<b>20</b>
6.3.1.	Relación de secuencias por medio de heatmap .....	20
6.3.2.	Alineamiento y verificación de motivos en la página del NCBI.....	20
<b>6.4.</b>	<b>ESTABLECIMIENTO DE RELACIONES FILOGENÉTICAS.....</b>	<b>21</b>
6.4.1.	Arboles filogenéticos para búsqueda de secuencias ancestrales .....	21
6.4.2.	Verificación de filogenia por taxonomía .....	22
6.4.3.	Arboles filogenéticos incluyendo las secuencias de pectato liasas del metagenoma de bosque altoandino .....	22
<b>7.</b>	<b>RESULTADOS .....</b>	<b>23</b>
7.1.	BASES DE DATOS LOCAL .....	23
7.2.	BÚSQUEDA DE HOMOLOGÍAS DE BASE DE DATOS LOCAL CON EL METAGENOMA .....	23
7.3.	BÚSQUEDA DE DOMINIOS DE PECTATO LIASAS .....	24
7.3.1.	Relación de secuencias por medio de heatmap .....	24
7.3.2.	Alineamiento y verificación de los motivos encontrados en la página del NCBI.	25
7.4.	RELACIONES FILOGENÉTICAS .....	30
7.4.1.	Verificación de filogenia por taxonomía .....	33
7.4.2.	Cladograma de secuencias de pectato liasas y péptidos del metagenoma de bosque altoandino.....	34
<b>8.</b>	<b>DISCUSIÓN.....</b>	<b>36</b>
8.1.	BASE DE DATOS DE ENZIMAS PUTATIVAS DE PECTATO LIASAS Y HEATMAP.....	36
8.2.	DOMINIOS DE PECTATO LIASAS PUTATIVAS .....	37
8.3.	ANÁLISIS CON ÁRBOLES FILOGENÉTICOS.....	38
<b>9.</b>	<b>CONCLUSIONES .....</b>	<b>40</b>
	<b>BIBLIOGRAFÍA .....</b>	<b>41</b>
	<b>ANEXOS.....</b>	<b>52</b>

## LISTA DE FIGURAS

<b>Figura 1</b> Reacción de $\beta$ -eliminación de las pectato liasas.....	9
<b>Figura 2</b> Estructura 3D de pectato liasas .....	12
<b>Figura 3</b> Heatmap de la tabla de similitud arrojada por el BLASTp entre el metagenoma y la base de datos de pectato liasas.. .....	25
<b>Figura 4</b> Logo del alineamiento de las secuencias del clúster 1 .....	26
<b>Figura 5</b> Búsqueda en la herramienta CD-Search de una de las secuencias del grupo C1 de pectato liasas del metagenoma de boque altoandino. ....	27
<b>Figura 6</b> Logo del alineamiento de las secuencias del clúster 2. ....	28
<b>Figura 7</b> Búsqueda en la herramienta CD-Search de una de las secuencias del grupo C1 de pectato liasas del metagenoma de boque altoandino .....	29
<b>Figura 8a</b> Árbol filogenético del grupo C1 .....	32
<b>Figura 8b</b> Árbol filogenético del grupo C2 .....	33
<b>Figura 9a</b> Cladograma de ancestros del grupo C1 y las secuencias del metagenoma de Bosque Alto Andino .....	35
<b>Figura 9b</b> Cladograma de ancestros del grupo C2 y las secuencias del metagenoma de Bosque Alto Andino.....	35
<b>Figura 10</b> Cladograma de C1 y C2.....	36

## LISTA DE DIAGRAMAS

<b>Diagrama 1</b> Diagrama de flujo de la metodología usada. ....	23
---	----

## LISTA DE TABLAS

<b>Tabla 1.</b> Puntajes de e-value de los dominios de cada una de las secuencias de pectato liasas del metagenoma del bosque altoandino .....	30
--	----

## LISTA DE ANEXOS

<b>Anexo S 1</b> Árbol filogenético de la base de datos local completa para identificar el grupo más divergente.. .....	54
<b>Anexo S 2</b> Tabla de algunos organismos representativos miembros de las diferentes familias de pectato liasas .....	55
<b>Anexo S 3</b> Antecedentes.....	57
<b>Anexo S 4</b> Taxonomía de las bacterias correspondientes al grupo C1 .....	60
<b>Anexo S 5</b> Taxonomía de las bacterias correspondientes al grupo C2.....	66



## RESUMEN

En Colombia el único estudio metagenómico en el suelo de ambientes extremos ha sido del proyecto GeBiX, el cual anteriormente ha descrito la presencia de enzimas bioprospectivas y de biorremediación. Uno de los ecosistemas de estudio es el suelo de bosque altoandino del parque Nacional Natural los Nevados, este suelo tiene temperaturas entre 6 y 14 °C, altas precipitaciones (hasta 4000 mm por año), suelo ácido (pH de 5.2) y una gran acumulación de materia orgánica que permite el crecimiento de microorganismos. En el marco del proyecto GeBiX se realizó una base de datos de pectato liasas, enzimas que cortan enlaces alfa-1,4 glicosídicos del ácido poligalacturónico y se encuentra en la pared celular de plantas, bacterias, hongos, arqueas y algunos animales. Se utilizaron herramientas bioinformáticas como BLASTp, análisis de Heatmap y algoritmos propios para depurar bases de datos públicas y crear una base de datos nueva con enzimas putativas de pectato liasas con interés en aplicaciones biotecnológicas. Además, se realizaron análisis filogenéticos para establecer relaciones entre los genes putativos y determinar si esta enzima sirve para identificación taxonómica.

**Palabras Clave:** metagenómica, bioprospección, bioinformática, análisis filogenéticos.

## 1. INTRODUCCIÓN

En los últimos años ha crecido el interés por el estudio de enzimas de microorganismos extremófilos debido a sus propiedades termoestables, tales como: resistencia a valores extremos de pH y salinidad (Van den Burg, 2003). Estos pueden producir biocatalizadores únicos que soportan condiciones adversas que pueden presentarse en las producciones industriales, aminorar desechos, costos y permitir el desarrollo adicional procesos mas eficientes, además de esto, a nivel ambiental se ha demostrado que algunas enzimas psicrófilas aparte de ser eficientes para la producción de textiles o químicos (Demirjian, Morís-Varas, & Cassidy, 2001) se pueden emplear para la degradación de hidrocarburos y descontaminación de suelos (Eriksson, Ka, & Mohn, 2001).

En la industria textil, existen problemáticas en la parte del biodecudado del algodón en el que normalmente se usan químicos alcalinos para la degradación del material no celulolítico; sin embargo, esta práctica causa una perdida en la calidad de la fibra de algodón y genera un impacto ambiental, debido a que existe un inevitable desecho de los químicos que usan (dialdehido, glioxal y urea-formaldehido o algunas resinas) los cuales se pueden filtrar en el suelo o en fuentes de agua generando un problema ambiental (Slater, 2003; Solbak et al., 2005), estos problemas pueden ir desde contaminar agua potable para humanos o animales que puede generar problemas de salud acelerando mutaciones, hasta la destrucción de las partículas del suelo que se vuelven más finas haciendo así que por medio del agua y el viento se pierdan los nutrientes del suelo y por tanto la fertilidad de los mismo, éstas partículas llevadas por el viento pueden ser también inhaladas por la población y resultar fatal (Slater, 2003).

En otras industrias como en la de procesamiento de frutas para la obtención de jugos claros como de manzana, pera, uva o para con jugos con sobrenadantes como los néctares, jugo de limón o tomate, los pasos son largos, primero es necesario la selección y lavado de la fruta, después los procesos de extracción en el cual se deben separar los fuertes enlaces pectínicos en las células de las frutas

para la posterior separación, aglutinación y posterior decantación y filtrado de los jugos, esto puede ser costoso debido a que tiene un bajo rendimiento, necesita mucho equipo y una gran labor en el proceso de concentración del jugo (Kashyap, Vohra, Chopra, & Tewari, 2001), el rompimiento de estos enlaces son los que principalmente generan altos costos en las industrias de fermentación de café y té, preparación de vino y papel, ya que las alternativas usadas son dañinas para la materia y para el ambiente (Pandey, Negi, & Soccol, 2017).

Por otro lado, en el tratamiento de aguas residuales pécticas provenientes de las industrias de jugos como el de limón, la naranja o la manzana, portan alto contenido de material pectínico que las hace muy difíciles de descomponer; en estos casos se ha intentado su tratamiento con técnicas como deshidratación física, irrigación en spray, coagulación química, tratamiento directo de lodos activados e hidrólisis química seguida de fermentación por metano (Reginatto, Amante, Gerhardy, Kunst, & Duran, 2009; Tanabe, Kobayashi, & Akamatsu, 1986) sin embargo, estos procesos son muy complejos y tienen un alto costo.

Las pectato liasas son enzimas que han sido ampliamente utilizadas en la industria textil (Araújo, Casal, & Cavaco-Paulo, 2008; Wang, Li, Ma, & Song, 2014) industrias alimenticias de fermentación de jugos, té y café (Alkorta, Garbisu, Llama, & Serra, 1998; Pandey *et al.*, 2017), fabricación de papel, tratamiento de aguas residuales, extracción de aceite cítrico, entre otras (Pandey *et al.*, 2017). Sin embargo, aunque las pectato liasas más utilizadas provengan de hongos, las bacterias provenientes de ambientes extremos pueden ser una fuente importante de éstas enzimas dado que presentan adaptaciones especiales como estabilidad a diferentes valores de pH, altas o bajas temperaturas o resistencia a sales, útiles a nivel industrial, biotecnológico o en procesos de restauración ambiental (A. Dubey, Yadav, Kumar, Anand, & Yadav, 2016).

Algunas enzimas pécticas como las pectato liasas se han implementado en algunos procesos industriales, sin embargo existen problemas de resistencia a temperaturas, pH extremos o patogenicidad de los microorganismos usados

(Kashyap *et al.*, 2001) por esto la búsqueda de extremozimas puede ser una alternativa para mejorar el desarrollo y rendimiento de estas industrias.

En Colombia, la investigación de ecosistemas extremos como las que se llevan a cabo en el Bosque Alto Andino del Parque Nacional Natural Los Nevados (PNN Los Nevados) (GEBIX, 2009) pueden ser un potencial biotecnológico y contener fuente de nuevas enzimas de interés bioprospectivo debido a los suelos que por su composición acida sustentan una amplia gama de microorganismos (Bautista Zúñiga, Rivas Solórzano, Durán de Bazúa, & Palacio, 1998).

La búsqueda de pectato liasas en metagenoma altoandino es importante ya que este es un ecosistema que ha sido muy poco explorado a nivel de diversidad de microorganismos y a nivel funcional (Minambiente, 2017; Parques Nacionales Naturales., 2006), aunque la presencia de pectato liasas tiene una baja proporción en el suelo de bosque altoandino, probablemente por las condiciones de pH y temperatura del mismo, resulta interesante la aparición de estas enzimas que pueden tener un gran potencial en la creación de una amplia gama de productos.

Es por esto que lo que busca este trabajo al identificar genes putativos de pectato liasas en el metagenoma de bosque altoandino es ser un punto de partida para posteriores estudios de bioprospección, las características de los resultados que se entregarán son importantes para la continuación del estudio de estas enzimas y su posterior utilización en diferentes industrias.

## 2. JUSTIFICACIÓN

Desde siempre las enzimas han sido utilizadas para el beneficio humano en varios procesos industriales como la fabricación de cerveza, papeles, lácteos, entre otros (Castillo Rosales & Rodríguez Alegría, 2014), muchas industrias usan enzimas sintéticas pero muchas veces la eficacia no es tan buena como, por esto, la búsqueda de enzimas con propiedades de resistencia a ambientes extremos es de gran interés para muchas industrias que necesitan agilizar procesos y reducir costos de producción (Demirjian et al., 2001; Yuan et al., 2011), en los últimos años, se han encontrado algunas enzimas de ambientes extremos que han mejorado procesos industriales de una manera significativa (J. Montaña, 2015; Moya A. & Torres R., 2012; Sasaki, Koide, Kobayashi, Usami, & Horikoshi, 2015). Colombia se caracteriza por ser un país megadiverso, tanto a nivel de ecosistemas, como de flora, fauna y también de microorganismos; uno de los ecosistemas más especiales que tiene Colombia es el bosque altoandino, el cual tiene bajas temperaturas y suelos ácidos que pueden ser un foco interesante para la búsqueda de enzimas extremófilas de importancia bioprospectiva (Bautista Zúñiga et al., 1998).

Entre las enzimas que han destacado en el uso de procesos industriales están las pectato liasas, que rompen fuertes enlaces en las paredes celulares de las plantas y han sido muy útiles en industrias de alimentos, jugos, té, café, fabricación de papel, tratamiento de aguas residuales, entre otras (Alkorta et al., 1998; Madhavan, Sindhu, Parameswaran, Sukumaran, & Pandey, 2017). Aunque ya se han hecho varias investigaciones sobre estas enzimas en metagenomas a nivel internacional, el conocimiento de pectato liasas en Colombia es muy limitado, ya que el estudio de microorganismos de ambientes extremos en nuestro país tiene un corto recorrido. El proyecto GeBiX (Centro Colombiano de Genómica y Bioinformática de Ambientes Extremos) ha sido la iniciativa en Colombia con respecto a la metagenómica de ambientes extremos (GEBIX, 2009) dentro de los trabajos realizados en este proyecto, se encuentra búsqueda de enzimas de nivel bioprospectivo como esterasas, lipasas o lacasas (Bolaños, 2016 (unpublished

work); Borda-Molina, Montaña, Zambrano, & Baena, 2017; Jiménez, Montaña, Álvarez, & Baena, 2012; López-López, Cerdán, & González Siso, 2014). Enmarcado en este mismo proyecto, este estudio pretende hacer un análisis *in silico* de pectato liasas putativas que tengan resistencia a ambientes extremos y que puedan ser de importancia industrial o de biorremediación.

### 3. MARCO TEÓRICO

#### 3.1. MICROORGANISMOS DEL SUELO, METAGENÓMICA Y BIOPROSPECCIÓN

La biodiversidad de los microorganismos en el suelo es la más alta comparada con cualquier otra comunidad ya que éste resulta ser heterogéneo con respecto a las condiciones de crecimiento y distribución de sustancias de la matriz, lo cual genera nichos ampliamente variados para los microorganismos (Lee & Lee, 2013) la abundancia de los diferentes filos depende del tipo del suelo, así, suelo de bosques tienen mayor presencia de *Acidobacteria* mientras que en cultivos de suelos supresores (aquellos suelos en los que los cultivos sufren menos patógenos por acción de microorganismos del suelo) se encuentra mayor presencia de  $\lambda$  y  $\beta$ -Proteobacteria (Lee & Lee, 2013; Lim et al., 2005; Mendes et al., 2011). La mayoría de éstos microorganismos no son cultivables, dejando un gran vacío en el conocimiento de la diversidad microbiana en general, sin embargo, el desarrollo de nuevas tecnologías ómicas han permitido la amplificación y secuenciación de ADN a partir de la unidad ribosomal 16S y la identificación de muchos de estos microorganismos que no son cultivables (Lorenz & Schleper, 2002; Torsvik, 2002).

La metagenómica es una herramienta que hace parte de las ciencias ómicas, y se refiere al análisis de los genomas de microorganismos por medio del estudio de ADN directamente extraído, ésta herramienta ha ayudado a identificar microorganismos no cultivables, encontrar nuevos genes e interacciones entre los microorganismos (Schloss & Handelsman, 2003). Así, la metagenómica se ha

convertido en una herramienta para acceder a los genomas de los organismos y por lo tanto es uno de los mejores enfoques para encontrar enzimas a partir de muestras ambientales (Eriksson *et al.*, 2001; Ferrer, Martínez-Abarca, & Golyshin, 2005; Kennedy, Marchesi, & Dobson, 2008).

Finalmente la bioprospección busca aislar, identificar o seleccionar componentes de la diversidad biológica para mejorar procesos biotecnológicos, de biorremediación o para fines comerciales (Torres & Velho, 2009), es por esto que además de la identificación de los microorganismos del suelo, el aislamiento de ADN y la creación de librerías de clones facilita el estudio de características como identificación de genes, obtención de información genética de filogenias, extraer enzimas o biocatalizadores y proteínas con importancia bioprospectiva; igualmente permite encontrar interacciones entre comunidades microbianas, descubrir genes de resistencia a antibióticos o producción de vitaminas con ayuda de métodos computacionales (Daniel, 2005; Lorenz & Schleper, 2002; National Research Council (US) Committee on Metagenomics: Challenges and Functional Applications. & Council, 2007; Streit, Daniel, & Jaeger, 2004; Timmermans, 2001; Uchiyama & Miyazaki, 2009), no solo en ambientes normales, sino también en ambientes de temperaturas o valores de pH extremos.

### **3.2. EL BOSQUE ALTOANDINO DEL PARQUE NACIONAL NATURAL LOS NEVADOS**

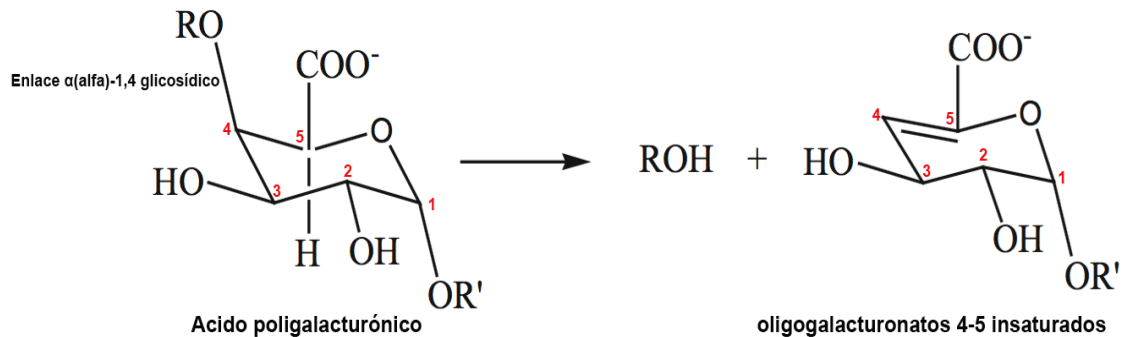
El Parque Nacional Natural Los Nevados tiene un área aproximada de 58.300 hectáreas distribuidas en la cordillera central de Colombia entre las vertientes oriental y occidental, presenta alturas entre 2600 y 5321 msnm. (Minambiente, 2017). Según Hernández Camacho & Sánchez Páez (1992) las alturas entre 3000 y 4000 msnm corresponden a un orobioma de bosque andino y entre 4000 a 4700 msnm corresponden a páramo, el cual a su vez se encuentra dividido entre el suborobioma de subpáramo y superpáramo; la altura de 4700 a 5100 msnm. corresponde a orobioma nival, teniendo así el Parque Nacional Natural Los Nevados una riqueza natural característica de biomas zonales de montaña.

Según el Plan de Manejo del Parque Nacional Natural Los Nevados (2006-2010) el porcentaje de Bosque Altoandino del parque es de 6.5%, siendo el segundo más alto después de páramo con 85.9%, ocupando aproximadamente 3790 hectáreas. Sus temperaturas varían entre los 6 °C a 12 °C con una precipitación entre los 500 y 4000 mm al año. Sus suelos pertenecen al orden andisol, caracterizados por ser ácidos (pH 5.2), contener altas cantidades de agua y favorecer la acumulación de materia orgánica la cual puede llegar a tener más de 40 cm de profundidad (Bautista Zúñiga *et al.*, 1998) razón por la cual este tipo de suelo puede ser fuente de una alta diversidad de microorganismos.



### 3.3. PECTATO LIASAS (E.C. 4.2.2.2)

Las pectato liasas (E.C. 4.2.2.2) son enzimas que cortan enlaces  $\alpha$ (alfa)-1,4 glicosídicos de ácido poligalacturónico por eliminación de una reacción  $\beta$ (beta) que produce un enlace  $\Delta$ 4, 5 insaturado en el extremo no reductor del polisacárido y genera oligogalacturonatos 4,5-insaturados (Figura 1) (A. K. Dubey *et al.*, 2010). Si la enzima ataca aleatoriamente dentro de la cadena de poligalacturonato generando múltiples productos se dice que es **endo-pectato liasa**, por el contrario, si ataca solo en la parte final del poligalacturonato se le denomina **exopectato liasa** (Hugouvieux-Cotte-Pattat, Condemine, & Shevchik, 2014). Las pectato liasas juegan un papel importante en las paredes celulares al remodelar y reciclar los polisacáridos de pectina que son insolubles, de esta manera, estas enzimas forman uno de los enlaces más fuertes en la naturaleza (Ali, Søndergaard, Teixeira, & Pickersgill, 2015).



**Figura 1** Reacción de  $\beta$ -eliminación de las pectato liasas la cual rompe enlaces  $\alpha$ (alfa)-1,4 glicosídicos de ácido poligalacturónico y produce un enlace  $\Delta$ 4, 5 insaturado en el extremo no reductor del polisacárido generando oligogalacturonatos 4,5-insaturados. (Tomado de Ali *et al.*, 2015; modificado por Daniela Quibano).

**2.3.1. Distribución de las pectato liasas:** se encuentran distribuidas en todas las ramas del árbol de la vida, en nematodos parásitos, plantas, bacterias y hongos fitopatógenos, este último es la fuente más predominante con *Aspergillus* (P. Micheli, 1729), *Fusarium* (Link, 1809) y *Penicillium* (Link, 1809) (Payasi, Sanwal, & Sanwal, 2009), sin embargo han sido reportadas numerosas pectato liasas en

géneros de bacterias como, *Erwinia* (Winslow, Broadhurst, Buchanan, Krumwiede, Rogers and Smith, 1920), *Bacillus* (Cohn, 1872), *Pseudomonas* (Migula, 1894), *Xanthomonas* (Dowson, 1939), *Sinorhizobium* (Chen *et al.*, 1988) entre otras (A. K. Dubey *et al.*, 2010; A. Dubey *et al.*, 2016).(Anexo S 2).

**2.3.2. Estructura y clasificación de las pectato liasas:** en general presentan una estructura de un dominio de hélice- $\beta$  formado por hebras paralelas dobladas en una hélice derecha grande y una región de bucle principal (A. K. Dubey *et al.*, 2010) tienen generalmente un pH básico (entre 8 y 10) y necesitan un catión divalente que en la mayoría de los casos es  $\text{Ca}^{2+}$ . Su mecanismo de  $\beta$ -eliminación consta de un residuo de Arg, Lys o His (que depende de la familia) la cuál extrae el protón de la posición de C-5 del residuo de galacturonato (Hugouvieux-Cotte-Pattat *et al.*, 2014). Las pectato liasas pertenecen a las familias de polisacárido liasas denominadas PL1, PL2, PL3, PL9, PL10 (Soriano Lasheras, 2004).

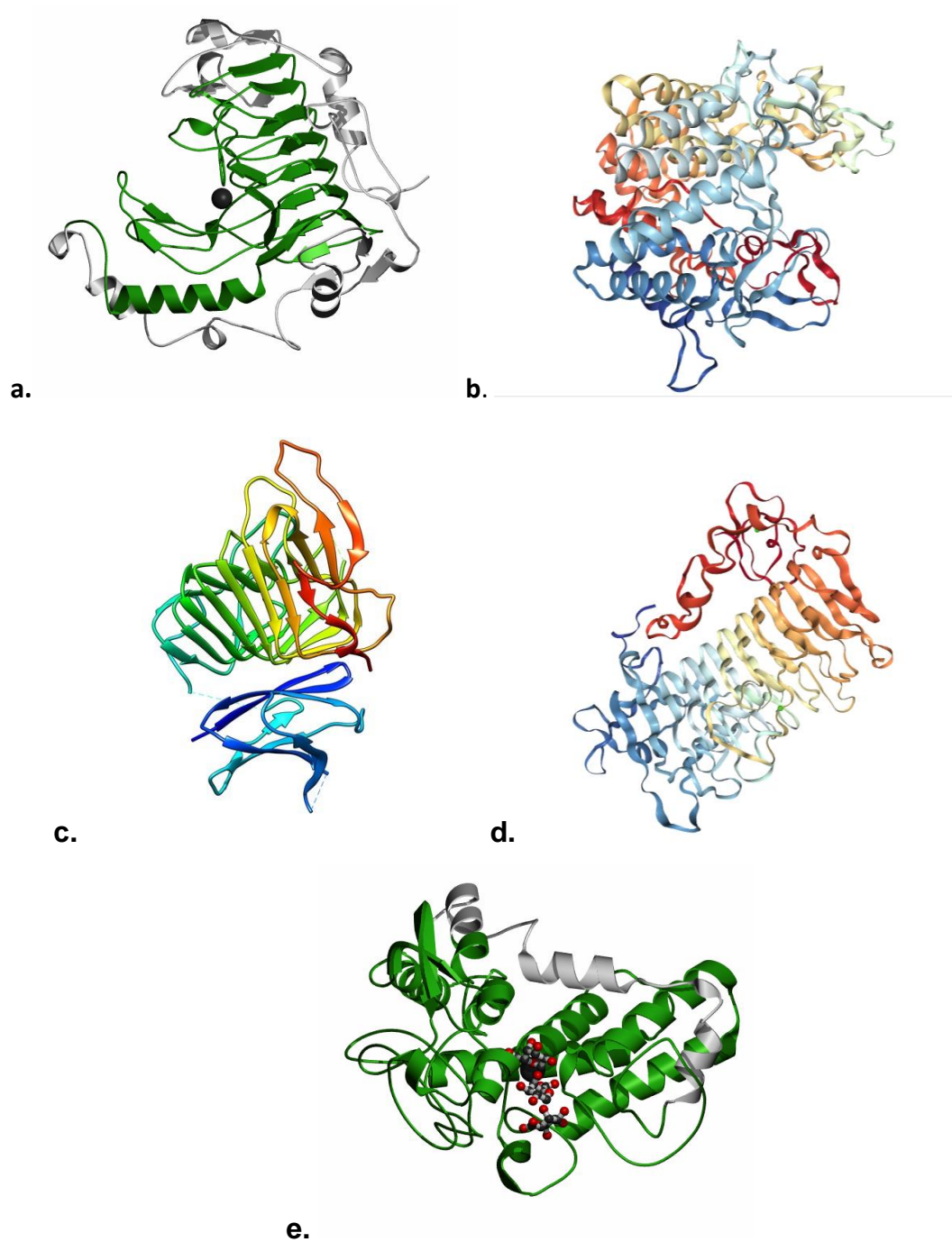
- **PL1:** presentan una estructura de hebras  $\beta$  plegadas en una superhélice derecha, normalmente difieren de la conformación y el tamaño de los loops sobresalientes de las hélices  $\beta$ . El catión más representativo en sus sitios activos es  $\text{Ca}^{2+}$  y su tamaño difiere entre 300 a 400 residuos (Hugouvieux-Cotte-Pattat *et al.*, 2014). Presentan las secuencias consenso **vWIDH** que puede estar involucrada en el plegamiento de la proteína y transporte de membrana, **AxDIKxxxxxVTxS** y **VxxRxPxxRxGxxHxxxN** que presenta el residuo de arginina (R) catalítico y se han clasificado en dos subfamilias PelADE y PelBC (Soriano Lasheras, 2004).
- **PL2:** son pectato liasas exclusivamente intracelulares que se pliegan en barras  $(\alpha/\alpha)_7$ , pueden ser endo o exo-pectato liasas y usan otros cationes diferentes de  $\text{Ca}^{2+}$  como  $\text{Co}^{2+}$ ,  $\text{Mn}^{2+}$ , and  $\text{Ni}^{2+}$ , su tamaño difiere entre 300 y 400 residuos (Hugouvieux-Cotte-Pattat *et al.*, 2014).
- **PL3:** son endo-pectato liasas altamente dependientes de  $\text{Ca}^{2+}$ . Adoptan una topología de hélices  $\beta$  paralelas y son las más cortas entre las pectato liasas con aproximadamente 200 residuos (Hugouvieux-Cotte-Pattat *et al.*,

2014). Éstas contienen 4 regiones de residuos conservados **VxxxDxCEDAxT, DKixQxN, DYGKLxRxCGxC y LxGxNxNYGDxA.**

- **PL9:** presenta una topología de hélices  $\beta$  paralelas más larga que aquellas de PL1 y PL3. Contiene endo y exo-pectato liasas. En su sitio catalítico pueden tener cationes divalentes como  $Fe^{2+}$  y en otros casos estar reemplazados por residuos aromáticos que los hace más específicos a sustratos metilados o sustratos cargados (Hugouvieux-Cotte-Pattat *et al.*, 2014; Mayans *et al.*, 1997).
- **PL10:** éstas pectato liasas son encontradas exclusivamente en bacterias de plantas, suelo o intestino, dependientes de  $Ca^{2+}$ , atacan sustratos de pectina con poca metilación y a pesar de que adoptan una topología de barra  $(\alpha/\alpha)_3$  diferentes a PL1, PL3 y PL9 presentan una disposición conservadas de los sitios activos (Charnock, Brown, Turkenburg, Black, & Davies, 2002; Hugouvieux-Cotte-Pattat *et al.*, 2014).

El dominio correspondiente a esta familia se encuentra en la base de datos de dominios conservados de la NCBI (<https://www.ncbi.nlm.nih.gov/cdd>) con PSS-ID pfam09492.

En la Figura 2 se pueden ver algunos ejemplos de las estructuras de las diferentes familias de polisacárido liasas.



**Figura 2** Estructura 3D de pectato liasas **a.** Pectato liasa de la familia PL1 de *Bacillus subtilis* ((Ehrenberg, 1835) Cohn, 1872) **b.** Pectato liasa de la familia PL2 de *Yersinia enterocolitica* ((Schleifstein and Coleman, 1939) Frederiksen, 1964) **c.** Pectato liasa de la familia PL3 de *Pectobacterium carotovorum* ((Jones, 1901) Waldee, 1945 emend. Hauben et al., 1998 emend.

Gardan et al., 2003) **d.** Pectato liasas de la familia PL9 de *Erwinia chrysanthemi* (Burkholder, McFadden & Dimock, 1953) **e.** Pectato liasa de la familia PL10 de mutante D189A de *Cellvibrio cellulosa* (Humphry et al. 2003) en complejo con el ácido trigalacturónico. (Tomado de <https://pfam.xfam.org>)

**2.3.3. Funciones de las pectato liasas:** algunos de los diversos roles que tienen las pectato liasas son: a) Ayuda a la enfermedad de pudrición blanda (A. Dubey et al., 2016; Hassan, Shevchik, Robert, & Hugouvieux-Cotte-Pattat, 2013; Hugouvieux-Cotte-Pattat et al., 2014; Sasaki et al., 2015), b) Degradación de pectina como fuente de carbono para crecimiento bacteriano (Hugouvieux-Cotte-Pattat et al., 2014), c) Degradación de productos de pectina como fuente de carbono para zoopatógenos (Hugouvieux-Cotte-Pattat et al., 2014), d) Señalamiento de moléculas en plantas y bacterias por los productos de la degradación de pectina (Hugouvieux-Cotte-Pattat et al., 2014), e) Simbiosis en plantas y bacterias (Hugouvieux-Cotte-Pattat et al., 2014), f) Digestión de material vegetal en animales (Hugouvieux-Cotte-Pattat et al., 2014), g) procesos de desarrollo de plantas (A. Dubey et al., 2016).

### **3.4. FILOGENIA MOLECULAR**

La filogenia molecular es una rama de la filogenética que combina técnicas estadísticas y moleculares para inferir las relaciones evolutivas entre genes u organismos, ésta estudia relaciones evolutivas por medio de ADN, ARN o aminoácidos (Dowell, 2008; Pevsner, 2009). El análisis filogenético de secuencias de aminoácidos consta de 5 pasos esenciales, 1) Selección de las secuencias 2) Alineamiento múltiple, 3) Especificación del modelo evolutivo, 4) construcción del árbol 5) análisis del árbol (Pevsner, 2009).

**3.4.1. Selección de secuencias:** las secuencias de ADN, ARN o proteínas, pueden provenir de datos tomados por el investigador o de bases de datos públicas como NCBI, Pfam para proteínas o Rfam para RNA. Las secuencias deben tener un nivel de homología (Pevsner, 2009).

**3.4.2. Alineamiento múltiple:** herramienta en la que se toman varias secuencias y se organizan en una matriz para identificar regiones de homología (Dowell,

2008) hay varios algoritmos que permiten hacer esto, los que presenta el programa MEGA 7 son ClustalW y Muscle.

- **ClustalW:** usa alineamiento progresivo en el que se toman las secuencias más similares y se alinean, después se van agrupando más secuencias hasta tener un alineamiento global (Thompson, Higgins, & Gibson, 1994).
- **Muscle:** es un método iterativo el cual funciona parecido al progresivo pero puede realinear las primeras secuencias mientras agrega unas nuevas al alineamiento, este método ha resultado ser más exacto comparado con T-Coffee y ClustalW (Edgar, 2004).

**3.4.3. Especificación del modelo evolutivo:** según Abascal et al., 2014 “son descripciones matemáticas de la evolución de las secuencias y constituyen el engranaje que nos permite conectar los datos (alineamientos) con los métodos de reconstrucción filogenética”. Para métodos de distancia como Neighbor-Joining se utilizan modelos estadísticos que estiman el número de cambios de DNA o aminoácidos en comparaciones entre dos secuencias, éstos métodos suponen información del puntaje de distancia entre cada OTU (unidad taxonomica operacional) en el alineamiento múltiple obteniendo puntajes por parejas de secuencias de todas las proteínas alineadas y obteniendo con esto el árbol filogenético (Pevsner, 2009), para métodos probabilísticos basados en caracteres como Maximun Likelihood o el de inferencia bayesiana se usan modelos estadísticos que se aplican a los residuos (caracteres individuales) y tienen en cuenta los cambios en cada posición del alineamiento (Abascal et al., 2014; Pevsner, 2009).

**3.4.4. Construcción del árbol:** pueden escogerse entre los métodos de distancia como Neighbor-Joining (NJ), los métodos estadísticos como el de inferencia bayesiana o Maximun Likelihood (ML) y el método de máxima parsimonia (MP).

- **Métodos de distancia – Neighbor-Joining:** creado por Saitou & Nei (1987) Analizan primero un par de secuencias y posteriormente se buscan las siguientes más parecidas, así usan esas distancias para inferir la relación entre los taxones (Peña, 2011; Pevsner, 2009).

- **Métodos estadísticos – Maximun Likelihood e inferencia bayesiana:** el primero determina la topología del árbol de acuerdo a la mayor probabilidad de producir un grupo de datos, el segundo calcula la probabilidad de observar algunos datos teniendo en cuenta el modelo que se esté usando (Peña, 2011; Pevsner, 2009).
- **Método de Máxima parsimonia:** descrito por Hennig (1966), y Eck y Dayhoff (1966), éste método prioriza la menor cantidad de cambios en un grupo de secuencias buscando así la explicación más parsimoniosa de los datos (Pevsner, 2009).

**3.4.5. Análisis del árbol:** uno de los métodos más comunes para el análisis de la robustez del árbol filogenético es el bootstrap, en este método se realiza el análisis varias veces con las secuencias tomadas aleatoriamente generando otro árbol con dichos datos, al final se comparan todos los arboles generados. Al final se obtiene la frecuencia con la cual se puede observar cada clado en el árbol original.

#### 4. ANTECEDENTES

En 1986 Tanabe, Kobayashi, & Akamatsu usaron una endo-pectato liasa aislada de la bacteria *Erwina carotovora* (Jones, 1901) para tratar aguas residuales de la industria del procesamiento de cítricos encontrando que el tratamiento con este tipo de enzima degrada la mayoría de las sustancias pécticas en aproximadamente 24 horas. Roy y colaboradores en 1999 analizaron el modo de acción de 5 endo-pectato liasas de *Erwinia chrysanthemi* (Burkholder, McFadden & Dimock, 1953) (bacteria fitopatógena), en sustratos poliméricos y oligoméricos. Benen, Kester, Pařenicová, & Visser, en 2000 clonaron un gen de pectato liasa de *Aspergillus niger* (van Tieghem, 1867) encontrando dos formas de la enzima con actividades específicas. En 2001, Tamaru & Doi, 2001 encontraron el gen de una pectato liasa en *Clostridium cellulovorans* (Sleat *et al.*, 1985) el cual codifica para un complejo enzimático llamado celulosoma que presentaba regiones bastantes conservadas y fue el primer reporte para este género, ese mismo año Brown, Mallen, Charnock, Davies, & Black, en 2001 identificaron la pectato liasa 10A de

*Pseudomonas cellulosa* que tenían sitios parecidos a los módulos de clivaje de carbohidratos y sitios afines con celulosa cristalina, observaron también una actividad a un pH de 10 y una temperatura de 62 °C. En el 2002, Yutaka Tamaru y colaboradores realizaron extracción de protoplastos por medio de cultivos de *Clostridium cellulovorans* (Sleat *et al.*, 1985) debido a sus genes de celulasa y pectato liasas, demostraron que el aislamiento de protoplastos por medio de esta enzima en cultivo era más eficiente que el comercial.

Los trabajos con bibliotecas de genes y metagenómica empezaron a volverse más comunes, en 2003, Voget y colaboradores identificaron varios genes en muestras de suelo colectadas en un campo cerca del Departamento de Microbiología de la Universidad de Göttingen, dentro de éstos genes se incluyeron dos pectato liasas, PelA la cual es muy similar a la pectato liasa de *Pseudoalteromonas haloplanktis* (ZoBell & Upham, 1944) y la segunda (Uvs119) similar a una enzima putativa de *Bacillus halodurans* ((ex Boyer, 1973) Nielsen *et al.*, 1995) y que posiblemente hidrolizan goma gellan, la cual es ampliamente usada como gelificante y en industrias como la cosmetología y la alimentación se usa como emulsificador, espesante y estabilizador (Dea, 1989). En 2005 Solbak y colaboradores descubrieron más de 40 pectato liasas nuevas en una librería de ADN de varias muestras ambientales ricas en materia vegetal, realizaron métodos de evolución directa para mejorar el rendimiento de una enzima candidata a altas temperaturas y hallaron 19 combinaciones para generar termorresistencia, de esta manera estas enzimas podrían ser utilizadas en el biodecudado del algodón, generando así la posibilidad de reemplazar el lavado tradicional con químicos. Kikuchi, Shibuya, Aikawa, & Jones en 2006 clonaron y analizaron dos pectato liasas de un nematodo de madera de pino que sugieren que estas enzimas juegan un papel importante en las interacciones nematodo-planta. En 2007, Gummadi & Kumar analizaron los efectos del cambio de pH, agitación y maceración para la máxima producción de pectin y pectato liasas de la levadura *Debaryomyces nepalensis* (Goto & Sugiy, 1968) en un biorreactor.



En la última década, los trabajos sobre pectato liasas se han centrado en la caracterización y análisis de nuevas enzimas. En el 2010 Ferreira, Da Silva, Silva, & Gomes reportaron las propiedades fisicoquímicas y la producción de una pectato liasa obtenida de *Penicillium viridicatum* (Westling, R., 1911) RFC3 por fermentación en estado sólido y fermentación sumergida en una mezcla de salvado de trigo y bagazo de naranja como fuente de carbón; ese mismo año Dubey y colaboradores hicieron un análisis *in silico* de 121 secuencias de pectato liasas descargadas de GenBank y analizaron homologías entre ellas, realizaron alineamiento múltiple, construcción de árbol filogenético y análisis de motivos conservados. En 2011 Yuan y colaboradores clonaron, expresaron y caracterizaron una pectato liasa de *Klebsiella* sp. Y1 y la compararon con novozimas comerciales de la industria textil mostrando que la pectato liasa clonada tenía una eficacia similar a la comercial; al año siguiente el mismo equipo recuperó 103 fragmentos de pectinasas a partir de rumen de oveja, posteriormente clonaron y expresaron 2 pectato liasas que servirían para el uso industrial a niveles fisiológicos. En el 2013 Hassan y colaboradores caracterizaron 8 endopectato liasas de *Dickeya dadanti* (Samson et al., 2005) (bacteria fitopatógena) y encontraron una nueva pectato liasa que juega un rol importante en la infección de las plantas. Sasaki y colaboradores en 2015 purificaron una nueva enzima de la bacteria de suelo submarino *Georgenia muralis* (P. Altenburger, P. Kämpfer, P. Schumann, D. Vybiral, W. Lubitz & H.J. Busse, 2002) la cual mostró tener buena actividad a pH de 10 y una temperatura de 50 °C; ese mismo año, Ali y colaboradores produjeron una pectato liasa de *Bacillus subtilis* ((Ehrenberg, 1835) Cohn, 1872) ya descrita y demostraron la pérdida de actividad de la enzima a pH bajo. En 2016 Peng y colaboradores clonaron y caracterizaron 4 cDNAs de pectato liasas de un nematodo de la hoja de soya (*Heterodera glycines* Ichinohe, 1952) y encontraron que juegan un rol importante su ciclo de vida y lograron reducir el número de parásitos de la planta al reprimir la expresión de una de las pectato liasas aisladas.

En Colombia se han realizado estudios en el suelo de bosque altoandino dentro del Proyecto GeBiX (“Conformación de una Plataforma en Metagenómica y Bioinformática para la caracterización y el aprovechamiento de recursos genéticos de ambientes extremos”), el cual tiene como objetivo conocer y potencializar la biodiversidad microbiana tanto a nivel de ambiente como en utilidad de enzimas de metagenomas. Entre los productos generados dentro de éste proyecto está el trabajo de Montaña y colaboradores en el 2012, en el cual se realizó un análisis de composición funcional y taxonómica encontrando que la mayoría de los microorganismos presentes se clasificaban como *Bacteria*, dentro de éstos se encontraban los filos *Actinobacteria*, *Proteobacteria*, y *Acidobacteria*; para el análisis funcional se encontraron secuencias asociadas al metabolismo de carbohidratos, proteínas, lípidos y funciones catalíticas como fosfatasas, glicosiltransferas, metiltransferasa, deshidratasa y epóxido hidrolasas; ese mismo año Jiménez y colaboradores, identificaron una nueva esterasa cuyo gen fue denominado “*estGX1*” la cual puede pertenecer al filo de las actinobacterias, a una nueva familia de lipasas, ésta codifica a un polipéptido de 210 aminoácidos con una masa de 22.4 kDa y mantiene el 30% de su actividad a 10°C lo cual la hace termoestable; Moya A. & Torres R., ese mismo año evaluaron la actividad hidrolítica de sobrenadantes producidas por hongos aislados del bosque altoandino del PNN Los Nevados en tusas de palma de aceite obtenidos de UNIPALMA, y encontraron que la hidrólisis de tusa a glucosa fue más eficiente en 400% en comparación con la hidrólisis normal, demostrando un gran potencial en la descomposición de tusas de palma, esto permite una reducción de problemas sanitarios y ambientales.

En 2014 López y colaboradores obtuvieron a partir de muestras de suelo de bosque altoandino una nueva esterasa de una proteobacteria termoacidofílica *Acidicaldus*, la cual es un nuevo miembro de la familia IV de las esterasa/lipasas de bacterias, esta enzima resulta ser la primera esterasa termoalcaloestable derivada de una bacteria Gram-negativa, en 2017 Borda-Molina, Montaña, Zambrano, & Baena, encontraron en ADN metagenómico de bosque altoandino un

gen que codificaba a una lipasa de una especie de *Ralstonia*, dicha enzima presentaba estabilidad a altas temperaturas y condiciones alcalinas (80°C y pH 9.0), entre otros estudios desarrollados por el Centro. (Un resumen de los antecedentes se puede observar en Anexo S 3).

## **5. OBJETIVOS**

### **5.1. OBJETIVO GENERAL**

Determinar genes putativos de pectato liasas en el metagenoma de suelo de bosque altoandino del Parque Nacional Natural Los Nevados.

### **5.2. OBJETIVOS ESPECÍFICOS**

- Crear una base de datos de genes putativos de pectato liasas del metagenoma de bosque altoandino a partir de regiones conservadas.
- Identificar regiones conservadas de las pectato liasas del metagenoma de bosque altoandino del Parque Nacional Natural Los Nevados.
- Establecer relaciones filogenéticas entre los genes putativos de pectato liasas encontradas en el metagenoma de bosque altoandino.

## **6. METODOLOGÍA**

### **6.1. CREACIÓN DE UNA BASE DE DATOS LOCAL A PARTIR DE BASES DE DATOS PÚBLICA**

- Se descargaron las secuencias de pectato liasas de bacterias de la base de datos RefSeq del NCBI.
- Se realizaron posteriores filtros a la base de datos dejando solo secuencias con “pectate lyase” en el título (eliminando hypothetical o partial), número de accesoión WP\_ (las cuales son secuencias de proteínas no redundantes).
- Se usó el programa cd-hit versión 4.7 para agrupar las secuencias por similitud y reducir el número de secuencias, usando un umbral de identidad de secuencias de 0.41. Estas se redujeron de 16753 a 211.

## **6.2. BUSQUEDA DE HOMOLOGÍAS DE LA DASE DE DATOS LOCAL CON EL METAGENOMA**

- Se usó una predicción de proteínas del metagenoma de bosque altoandino que fue realizada con el programa Metagenmark ([http://exon.gatech.edu/meta\\_gmhmp.cgi](http://exon.gatech.edu/meta_gmhmp.cgi)). Dicha predicción se encuentra almacenada en el servidor de GeBiX y presenta 619288 péptidos.
- Se hizo un BLASTp entre la base de datos local con los péptidos predichos del metagenoma de bosque altoandino, para tener altas probabilidades de tener secuencias parecidas con alta similaridad se disminuyó el parámetro de e-value a  $1 \times 10^{-20}$ .

## **6.3. BÚSQUEDA DE DOMINIOS DE PECTATO LIASAS**

- Se creó un script por el cual se generaron tres archivos, 1) una tabla con los porcentajes de similaridad de las coincidencias del BLASTp, 2) un archivo fasta con las pectato liasas encontradas en el metagenoma y 3) un archivo fasta con las coincidencias entre la base de datos local y el metagenoma de bosque altoandino. Estos archivos sirvieron para hallar los dominios y realizar los árboles filogenéticos.

**6.3.1. Relación de secuencias por medio de heatmap:** usando el programa online heatmapper (<http://www.heatmapper.ca/>) se realizó un análisis de correlación multivariable con representación mediante heatmap de expresión con la tabla de porcentaje de similaridad anteriormente generada, éste muestra un análisis de correlación en el que se puede observar la organización las secuencias dependiendo de su valor de similaridad. Los parámetros de visualización para el agrupamiento fueron: “complete linkage” para el método de agrupamiento y “manhattan” para el método de medición de distancia.

**6.3.2. Alineamiento y verificación de motivos en la página del NCBI:** el heatmap muestra con intensidad de color las secuencias que mejor se correlacionan entre la base de datos local y las secuencias del metagenoma. Se

tomaron dos grupos o clústeres de las secuencias con alta similaridad o correlacion y con estas secuencias se realizó un alineamiento múltiple en el programa MEGA 7.2 con el algoritmo de Muscle con parámetros por defecto (método de agrupamiento = UPGMB; gap open = -2.9; gap extend = 0), éste muestra las regiones conservadas en las secuencias y por tanto los motivos presentes.

El alineamiento se editó en BioEdit 7.2.6 y posteriormente se realizó un Web Logo en la página <https://weblogo.berkeley.edu/logo.cgi> para una visualización más sencilla de los motivos.

Una vez encontrados los motivos de cada grupo y para verificar si pertenecían a algún dominio reportado se usó la herramienta “Conserved Domains Search” (CD-Search) adjunta a la página del NCBI (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>), en la cual se suben la secuencias y se buscan los dominios en todas las bases de datos adjuntas las cuales incluyen secuencias curadas de Pfam, SMART, COG, PRK, and TIGRFAM.

#### **6.4. ESTABLECIMIENTO DE RELACIONES FILOGENÉTICAS**

**6.4.1. Árboles filogenéticos para búsqueda de secuencias ancestrales:** se tomaron las secuencias de la base de datos local de cada clúster y se realizó un árbol filogenético con el método de Maximum Likelihood (ML) con parámetros por defecto (500 bootstrap; modelo JTT) en el programa MEGA 7.2 para poder seleccionar las secuencias ancestrales de dichos grupos. El alineamiento para dicho árbol se realizó con el algoritmo Muscle con parámetros por defecto (arriba descritos). El outgroup usado para los realizar los árboles filogenéticos fue de una secuencia proteica de pectin metilesterasa (WP\_107142492.1).

**6.4.1.1. Selección del outgroup:** para seleccionar el outgroup se realizó un árbol filogenético sin raíz que permitió agrupar las secuencias para saber cual de ellas es la más ancestral. El árbol se hizo con toda la base de datos local usando el método de Maximum Likelihood (ML) con parámetros por defecto (500 bootstrap) en el programa MEGA 7.2 (Anexo S 1), del árbol resultante, la secuencia más divergente fue de *Cellvibrio japonicus*

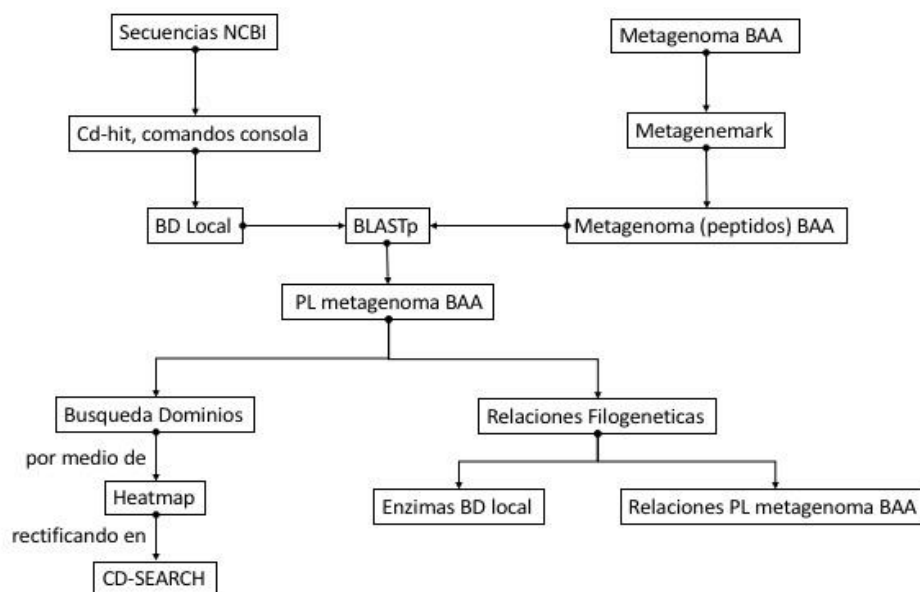
(WP\_012488698.1) y con esta se realizó un BLASTp en la página del NCBI con e-value de  $1.0e^{-10}$  en la que escogió la secuencia de otra familia con mayor bitscore (mayor identidad entre secuencias) (WP\_107142492.1 pectin metilesterasa).

**6.4.2. Verificación de filogenia por taxonomía:** para verificar si las pectato liasas pueden ser un referente taxonómico o si solamente se podían comparar por cambio en las secuencias, se compararon las especies de bacterias en cada uno de los clústeres con los arboles realizados, para esto se usaron las bases de datos de NCBI taxonomy para aquellas secuencias que aún no estuvieran validadas por el Manual de Sistemática de Arqueas y Bacterias de Bergey o que fueran extraídas de metagenomas e ITIS (<http://www.itis.gov>) (Integrated Taxonomic Information System), la cual es una base de datos formada por agencias federales que permite la búsqueda de información taxonómica exacta y nomenclatura estandarizada.

**6.4.3. Arboles filogenéticos incluyendo las secuencias de pectato liasas del metagenoma de bosque altoandino:** tomando las secuencias ancestrales halladas en el punto 5.4.1. se realizaron arboles filogenéticos usando el método ML con parámetros por defecto en el programa MEGA 7.2. incluyendo los péptidos encontrados en el BLAST para cada clúster, el alineamiento para dichos árboles se realizó con el algoritmo Muscle con parámetro por defecto y el outgroup usado para los árboles fue de una secuencia proteica de pectin metilesterasa (WP\_107142492.1).

Se realizó un árbol filogenético usando el método ML con parámetros por defecto en el programa MEGA 7.2. agrupando ambos clústeres, el alineamiento para dicho árbol se realizó con el algoritmo de Muscle con parámetros por defecto y el outgroup usado para el árbol fue de una secuencia proteica de pectin metilesterasa (WP\_107142492.1).

La metodología resumida se puede ver en el diagrama 1.



**Diagrama 1** Diagrama de flujo de la metodología usada.

## 7. RESULTADOS

### 7.1. BASES DE DATOS LOCAL

Se descargaron un total de 20109 secuencias de bacterias de la base de datos RefSeq del NCBI, al dejar las secuencias de proteínas no redundantes se obtuvieron 8000 secuencias. Finalmente, tras el cd-hit se obtuvo un total de 211 clústeres con una secuencia representativa de cada uno siendo esta **la base de datos local** que se usó para la búsqueda en el metagenoma.

### 7.2. BÚSQUEDA DE HOMOLOGÍAS DE BASE DE DATOS LOCAL CON EL METAGENOMA

Al realizar el BLASTp entre la predicción realizada con el Metagenemark de 619288 péptidos y la base de datos local de pectato liasas con 211 secuencias se encontraron 60 péptidos de pectato liasas.

Por lo tanto, de los 619.288 péptidos el BLASTp reporta 60 péptidos con relación a candidatos putativos de pectato liasas

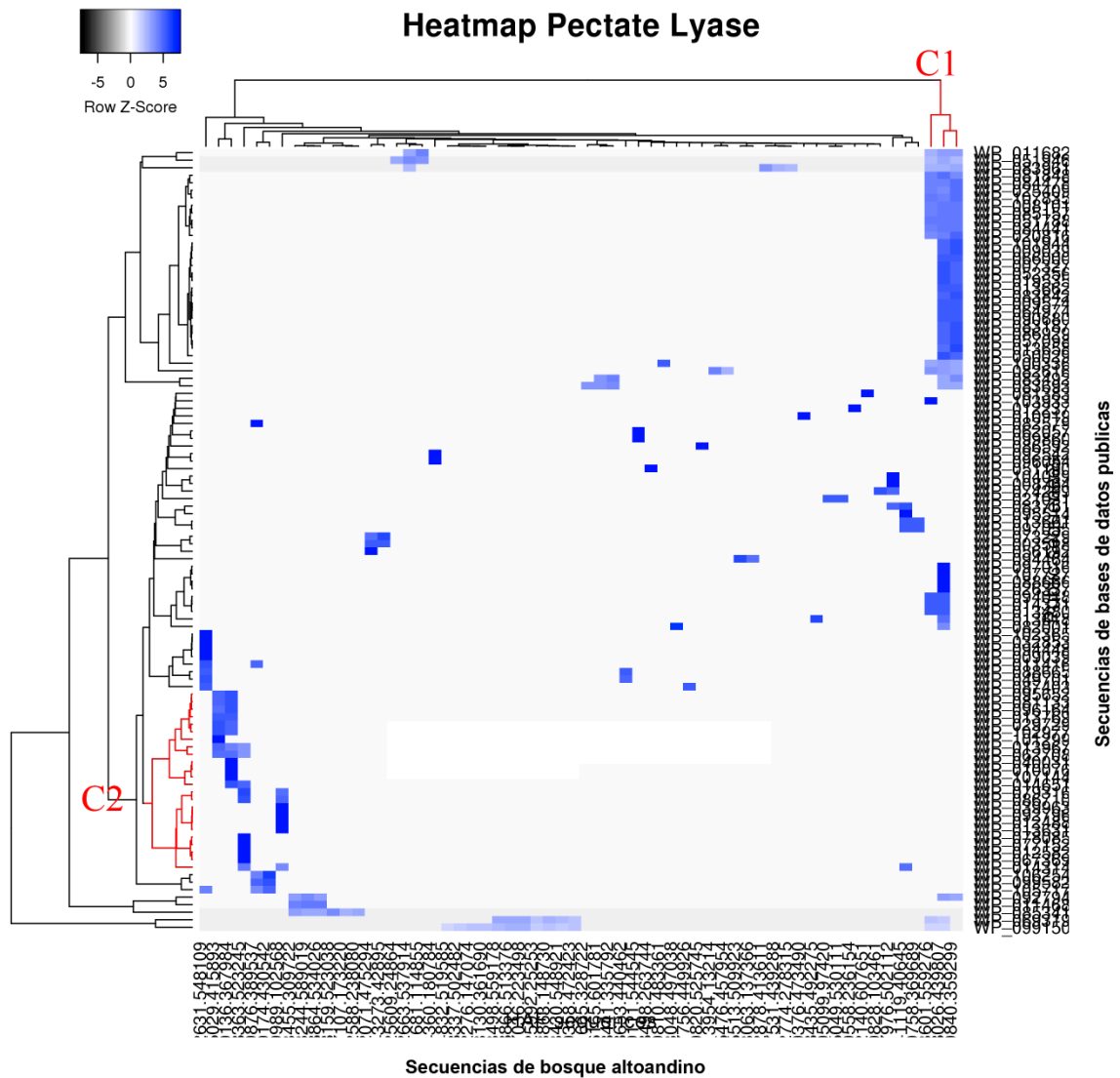
- **Frecuencia relativa PL en metagenoma:**  $\frac{60}{619288} = 9,7 \times 10^{-5}$
- **Porcentaje de PL en metagenoma:**  $9,7 \times 10^{-5} \times 100\% = 0,0097\%$

### 7.3. BÚSQUEDA DE DOMINIOS DE PECTATO LIASAS

**7.3.1. Relación de secuencias por medio de heatmap:** el BLASTp arroja porcentajes de similaridad por cada par de secuencias analizadas, con éstos datos se realizó un heatmap, el cual correlaciona las secuencias y muestra de manera gráfica, por intensidad de color, cuales son más similares. En el heatmap (Figura 3) se observan 2 agrupaciones denominadas C1 (por clúster 1) y C2 (por clúster 2), que son las que se tomaron para realizar la búsqueda de dominios y análisis filogenéticos. Al lado izquierdo de la figura se encuentran las secuencias de pectato liasas de la base de datos local y al lado derecho el dendograma de las mismas, abajo se encuentran los péptidos de pectato liasas del metagenoma y en la parte de arriba su dendograma. En el dendograma se observan en rojo las agrupaciones de las secuencias más parecidas denominadas C1 Y C2.

C1 presentó un total de 32 secuencias de la base de datos local y coincidió con un total de 3 péptidos de pectato liasas del metagenoma: **HAF|543601|559276, HAF|428026|439807, HAF|350840|359299**; por otro lado, C2 presentó un total de 24 secuencias de la base de datos local que coincidió con un total de 4 péptidos de pectato liasas del metagenoma: **HAF|405029|415893, HAF|359136|367984, HAF|508363|523245, HAF|99989|102568.**



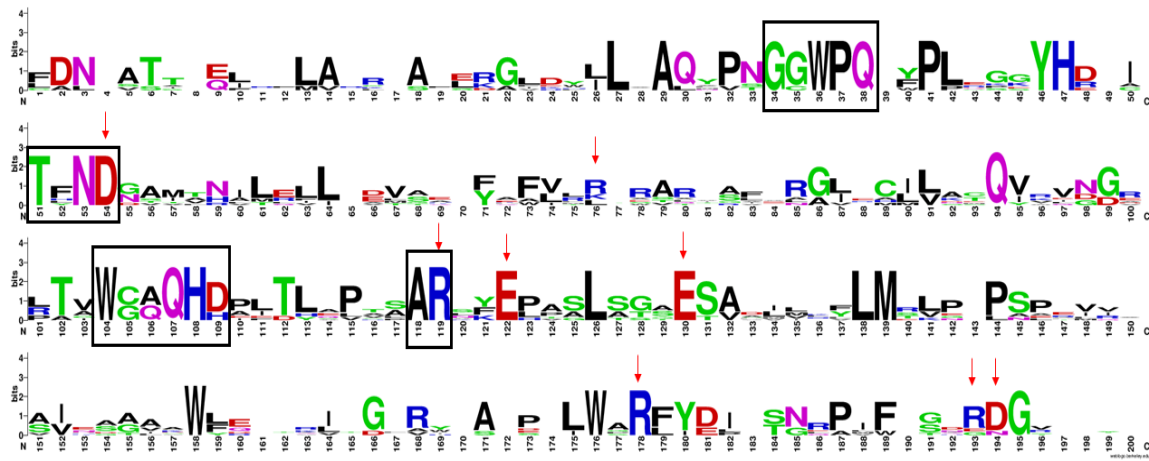


**Figura 3** Heatmap de la tabla de similitud arrojada por el BLASTp entre el metagenoma y la base de datos de pectato liasas. Los cuadros azul oscuro representan mayor grado de similitud entre las secuencias y los dendogramas muestran las agrupaciones de dichas secuencias.

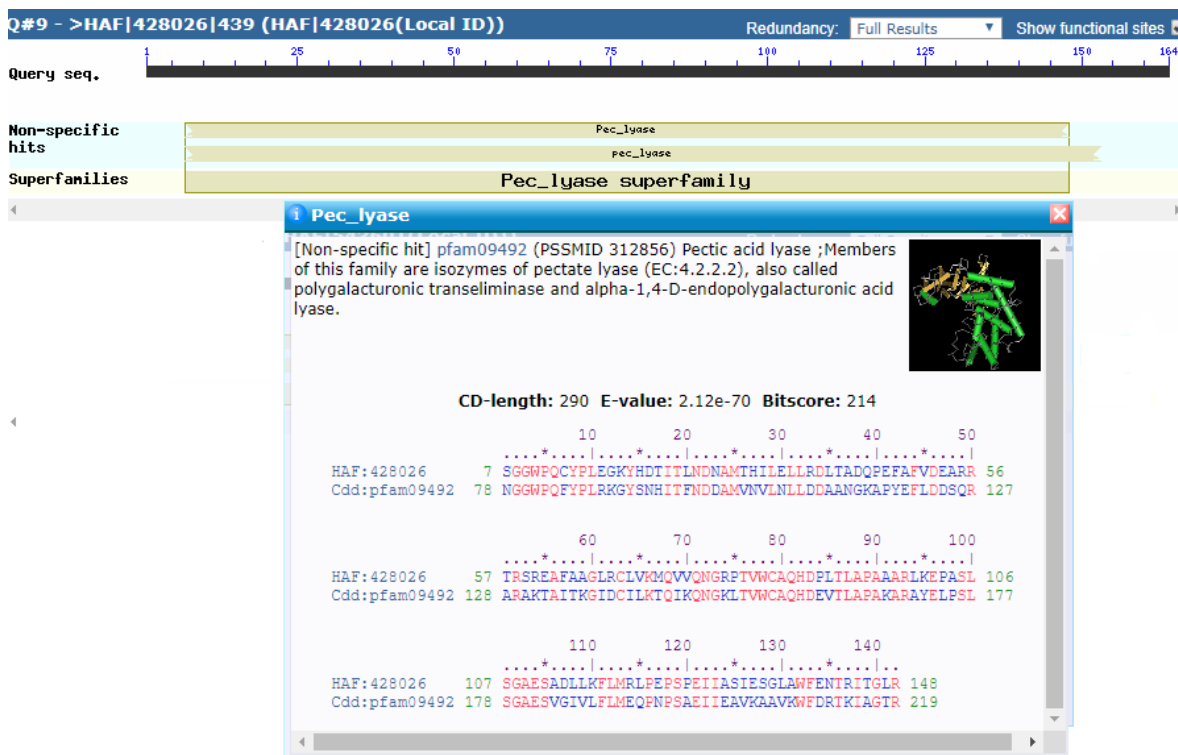
**7.3.2. Alineamiento y verificación de los motivos encontrados en la página del NCBI:** al alinear las secuencias de pectato liasas del metagenoma del bosque altoandino del clúster C1 (Figura 4) y verificar las secuencias en la página del NCBI Conserved Domains Database (<https://www.ncbi.nlm.nih.gov/cdd>) éstas

mostraron corresponder al dominio **pfam09492** (Figura 5) perteneciente a pectato liasas de la superfamilia PL10.

Uno de los parámetros que se tienen en cuenta para identificar la similaridad entre las secuencias es el e-value, entre menor sea, mayor similaridad entre secuencias habrá. Para las secuencias HAF|350840|359299, HAF|428026|439807 y HAF|543601|559276 se obtuvo un e-value de coincidencia con el dominio de pfam09492 de  $3.02e^{-58}$ ,  $2.12e^{-70}$  y  $5.09e^{-46}$  respectivamente (Tabla 1).



**Figura 4** Logo del alineamiento de las secuencias del clúster 1. La flecha roja indica los residuos conservados de arginina (R), ácido aspártico (D) y ácido glutámico (E). Los recuadros muestran los aminoácidos conservados en la mayoría de las secuencias.

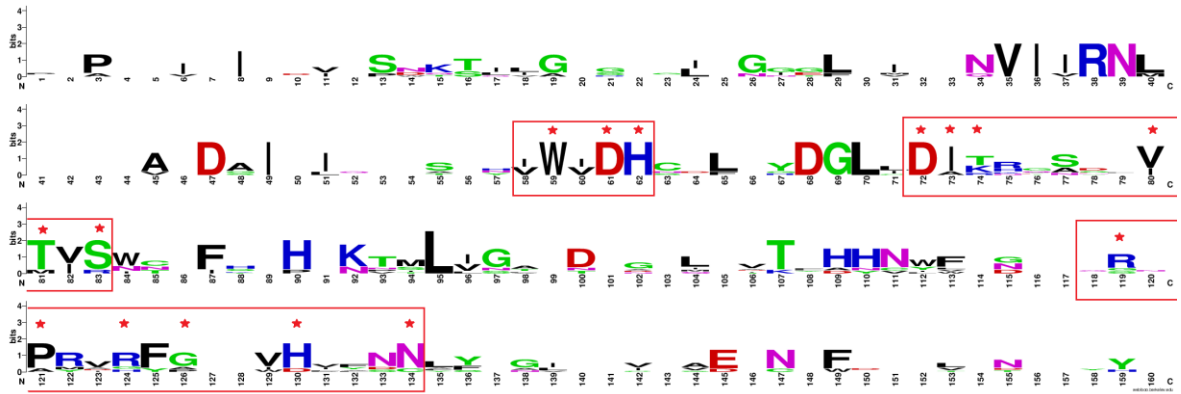


**Figura 5** Búsqueda en la herramienta CD-Search de una de las secuencias del grupo C1 de pectato liasas del metagenoma de boque altoandino (HAF428026) que muestra la coincidencia (en rojo) de una de las secuencias del metagenoma con el dominio pfam09492 correspondiente a la familia PL10.

Al alinear las secuencias de pectato liasas del metagenoma del bosque altoandino del clúster C2 (Figura 6) y verificar las secuencias en la página del NCBI Conserved Domains Database (<https://www.ncbi.nlm.nih.gov/cdd>) éstas mostraron corresponder al dominio **pfam00544** correspondiente a pectato liasas de la superfamilia PL1, además mostraron coincidencias con el dominio de un grupo de alergénicos ambientales **Amb a 1.05 (smart00656)** también pertenecientes a la superfamilia PL1 (Figura 6).

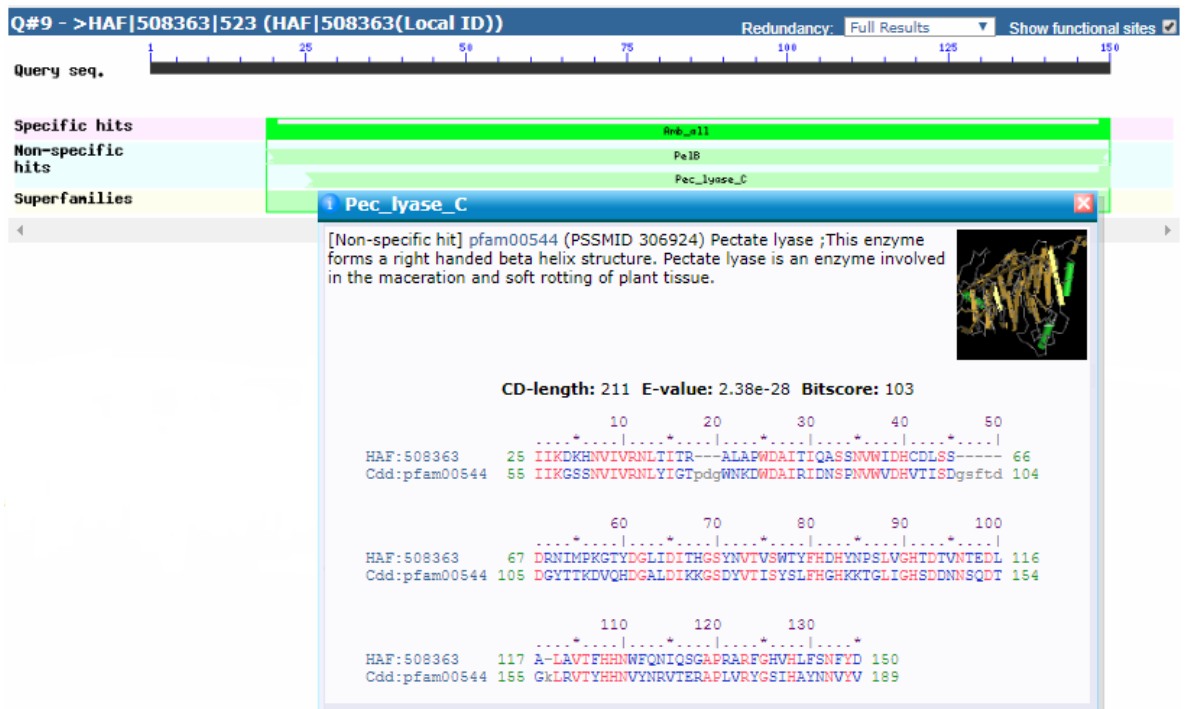
Para las secuencias del metagenoma HAF|99989|102568, HAF|359136|367984, HAF|405029|415893 y HAF|508363|523245 se obtuvo un e-value de coincidencia con el dominio pfam00544 correspondiente a PL1 de  $4.97e^{-22}$ ,  $1.03e^{-16}$ ,  $2.11e^{-26}$  y  $2.38e^{-28}$  respectivamente y con el dominio de smart00656 correspondiente a

alergénicos ambientales Amb a 1.05 se obtuvo un e-value de  $8.26e^{-41}$ ,  $3.98e^{-29}$ ,  $1.20e^{-25}$  y  $6.12e^{-44}$  respectivamente (Tabla 1).



**Figura 6** Logo del alineamiento de las secuencias del clúster 2.

En los recuadros de la figura 6 se encuentran los motivos característicos de PL1 (**vWIDH**, **DIKxxxxxVTxS** y **RxxPxxRxGxxHxxxN**), con estrellas se puntualizan los aminoácidos conservados en dichos motivos (W = triptófano; D = Aspartato; H = Histidina; I = Isoleucina; Y = Tirosina; S = Serina; R = Arginina; P = Prolina; N = Asparagina).



**Figura 7** Búsqueda en la herramienta CD-Search de una de las secuencias del grupo C1 de pectato liasas del metagenoma de boque altoandino que muestra la coincidencia (en rojo) con el dominio pfam00544 (pec\_lyase\_C en la imagen) y el dominio de alergénicos ambientales Amb a 1.05 (Amb\_all en la imagen, antes denominados Amb a 2 pero en la nomenclatura actual Amb a 1.05 (Radauer et al., 2014)) pertenecientes a la superfamilia PL1

**Tabla 1.** Puntajes de e-value de los dominios de cada una de las secuencias de pectato liasas del metagenoma del bosque altoandino, entre menor e-value, mayor probabilidad de similitud. Entre mayor bitscore, mayor similitud entre las secuencias.

Clúster	Secuencia	e-value	Bitscore	Domain ID	Nombre	Superfamilia
1	HAF 350840 359299	3.02e <sup>-58</sup>	182.183	pfam09492	Pectato liasa	PL10
	HAF 428026 439807	2.12e <sup>-70</sup>	214.154	pfam09492	Pectato liasa	PL10
	HAF 543601 559276	5.09e <sup>-46</sup>	150.596	pfam09492	Pectato liasa	PL10
2	HAF 99989 102568	8.26e <sup>-41</sup>	135.098	smart00656	Amb a 1.05	PL1
		4.97e <sup>-22</sup>	87.2628	pfam00544	pectato liasa C	PL1
	HAF 359136 367984	3.98e <sup>-29</sup>	104.282	smart00656	Amb a 1.05	PL1
		1.03e <sup>-16</sup>	72.6252	pfam00544	pectato liasa C	PL1
	HAF 405029 415893	2.11e <sup>-26</sup>	98.0484	pfam00544	pectato liasa C	PL1
		1.20e <sup>-25</sup>	95.4221	smart00656	Amb a 1.05	PL1
	HAF 508363 523245	6.12e <sup>-44</sup>	142.802	smart00656	Amb a 1.05	PL1
		2.38e <sup>-28</sup>	103.441	pfam00544	pectato liasa C	PL1

De los 60 péptidos de pectato liasas, 7 presentan los dominios conservados de pectato liasas.

- **Frecuencia PL con dominios conservados** =  $\frac{7}{60} = 0,12$
- **Frecuencia PL10** =  $\frac{3}{60} = 0,05$
- **Frecuencia PL1** =  $\frac{4}{60} = 0,06$

Por los valores reportados se observa que las pectato liasas presentan frecuencias muy bajas en el metagenoma.

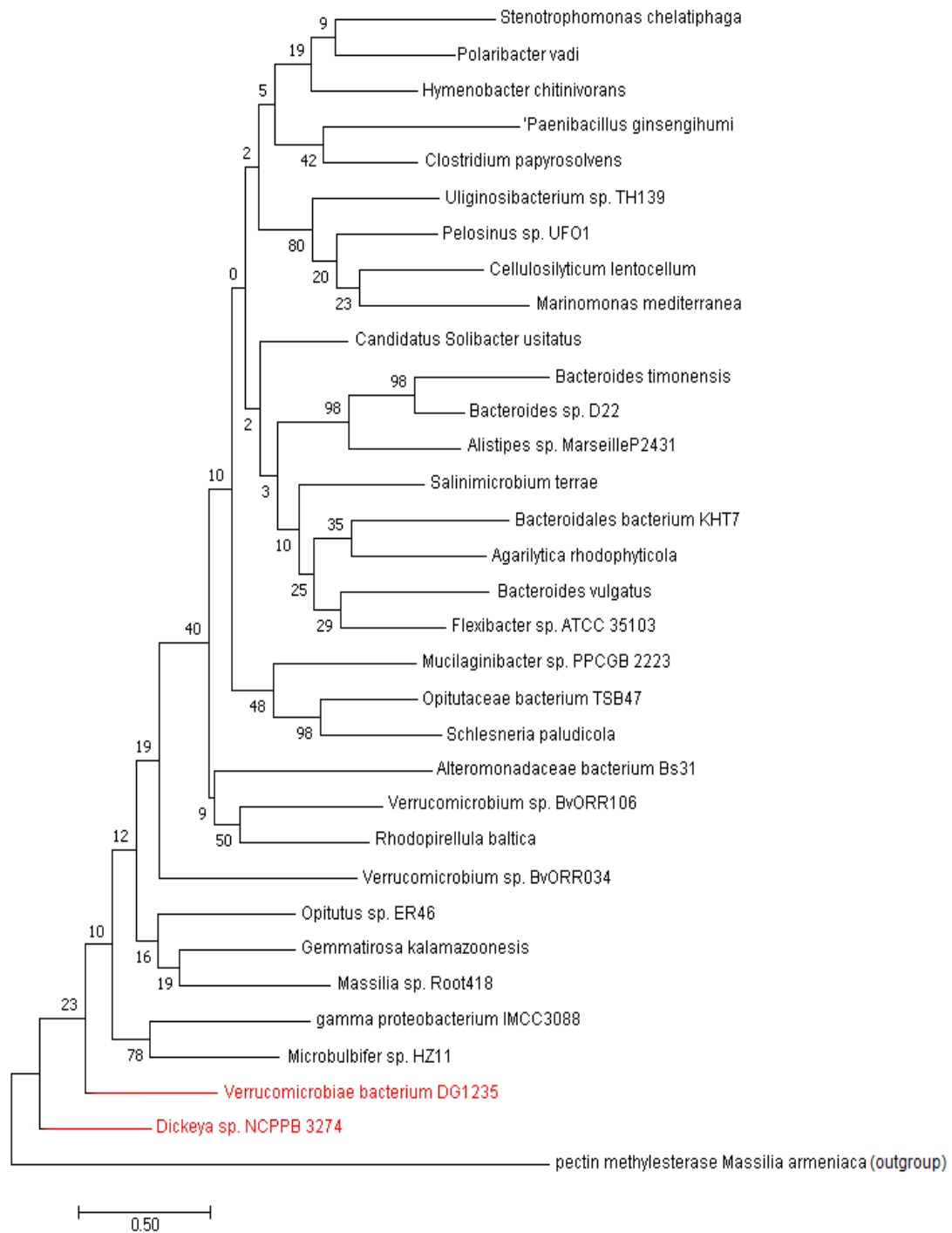
#### 7.4. RELACIONES FILOGENÉTICAS

**Árboles filogenéticos para búsqueda de secuencias ancestrales:** para cada uno de los clústeres C1 y C2 se realizaron arboles filogenéticos que mostraron las secuencias ancestrales de dichos grupos, la figura 8a corresponde al clúster C1, los números de accesión de las secuencias escogidas fueron: **WP\_084478753.1** (*Dickeya sp.*), **WP\_008101266.1** (*Verrucomicrobiae bacterium*) (en rojo en la

figura 8<sup>a</sup>), aunque no muestra unos grupos definidos, el árbol parece mostrar un agrupamiento jerárquico.

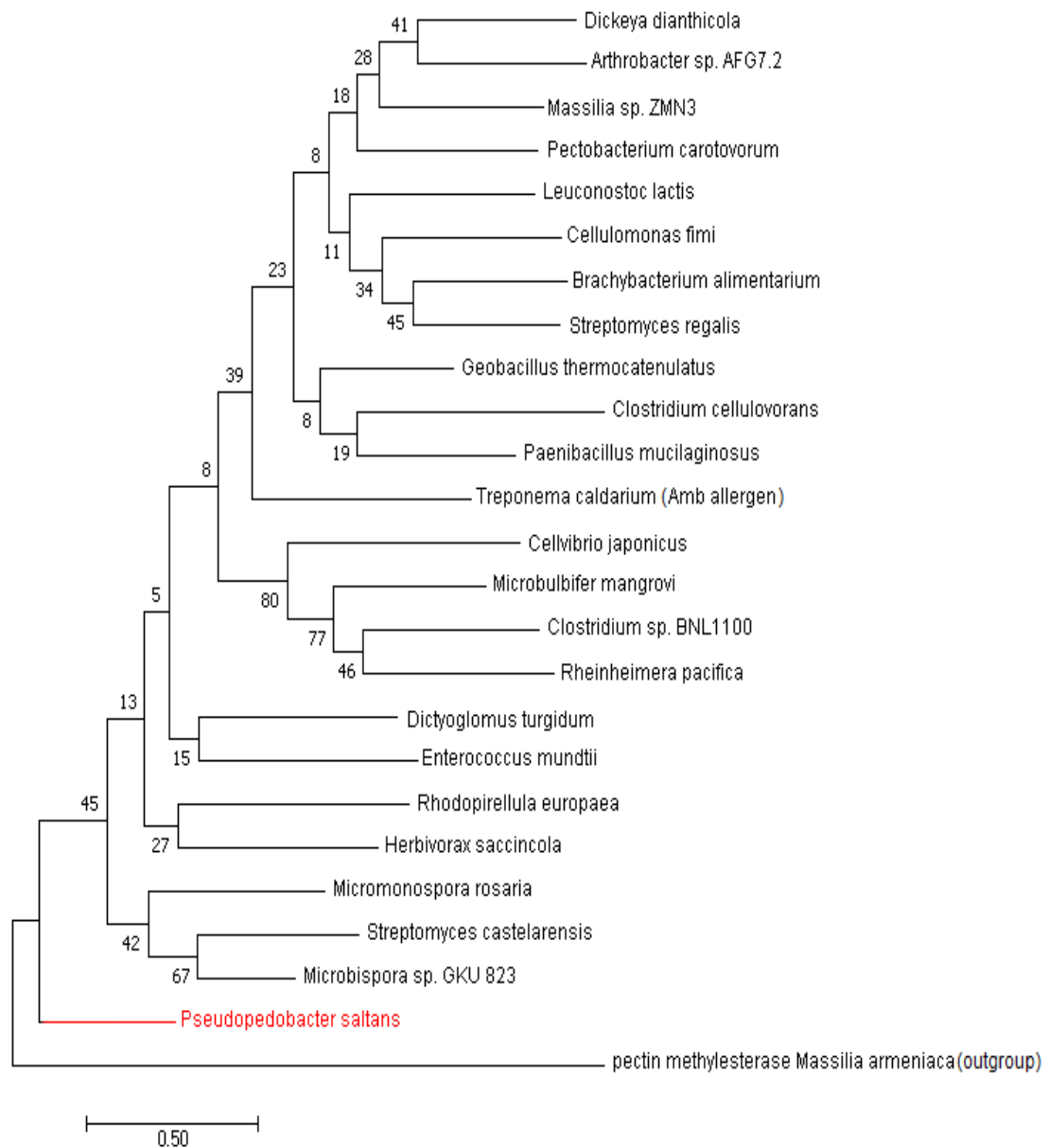
La figura 8b corresponde al clúster C2, los números de accesión de la secuencia escogida fue: **WP\_013631903.1 (*Pseudobacter saltans*)** (en rojo en la figura 8b) En este caso el árbol muestra también un agrupamiento jerárquico pero en algunos puntos se muestran grupos definidos.

En ambas figuras (8a y 8b) se muestra el árbol con el logaritmo de mayor probabilidad, el porcentaje de los árboles en los cuales las secuencias se agrupan se muestra al lado de las ramas, aquellos que tienen un número más alto tienen mayor probabilidad de reagrupamiento. El árbol se encuentra en escala y el tamaño de las ramas indica el número de sustituciones por sitio. Se usó una secuencia de pectin metilesterasa como outgroup.



**Figura 8a** Árbol filogenético del grupo C1 realizado con el método de Maximum Likelihood (ML) basado en un modelo de matriz JTT con 500 bootstrap en el programa MEGA 7.2. En rojo se muestran las secuencias ancestrales del grupo C1.





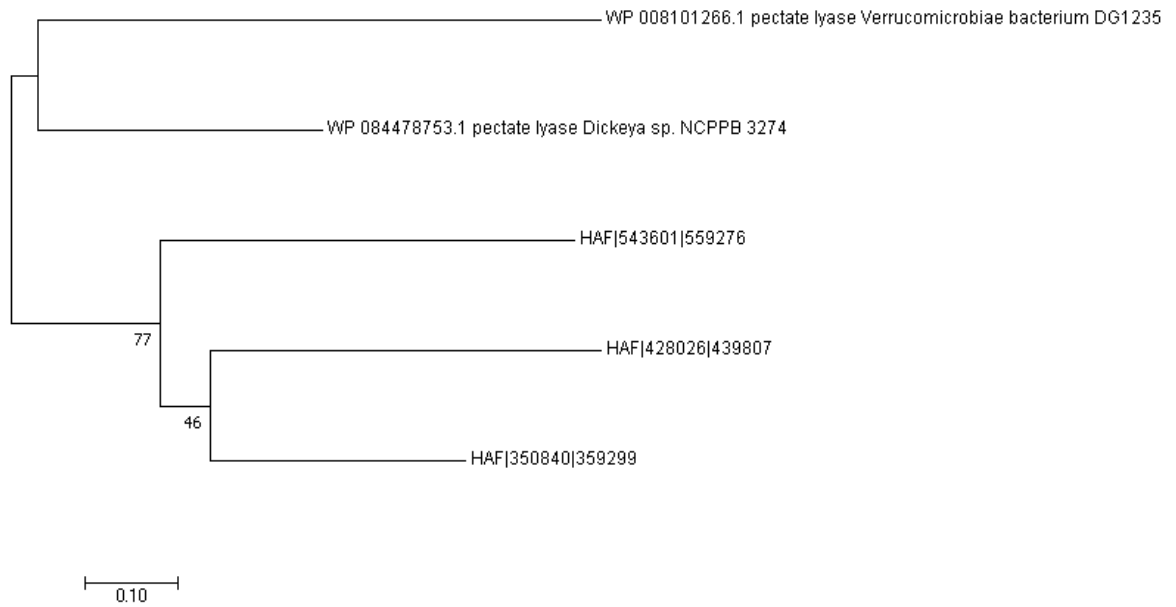
**Figura 8b.** Árbol filogenético del grupo C2 realizado con el método de Maximum Likelihood (ML) basado en un modelo de matriz JTT con 500 bootstrap en el programa MEGA 7.2. En rojo se muestra las secuencias ancestrales.

**7.4.1. Verificación de filogenia por taxonomía:** al buscar en las bases de datos del NCBI taxonomy e ITIS, se pudo observar que los arboles filogenéticos generados no proporcionaban una relación a nivel de taxones por lo que el análisis que se realiza se basa solo en homologías de secuencias mas no en filogenia,

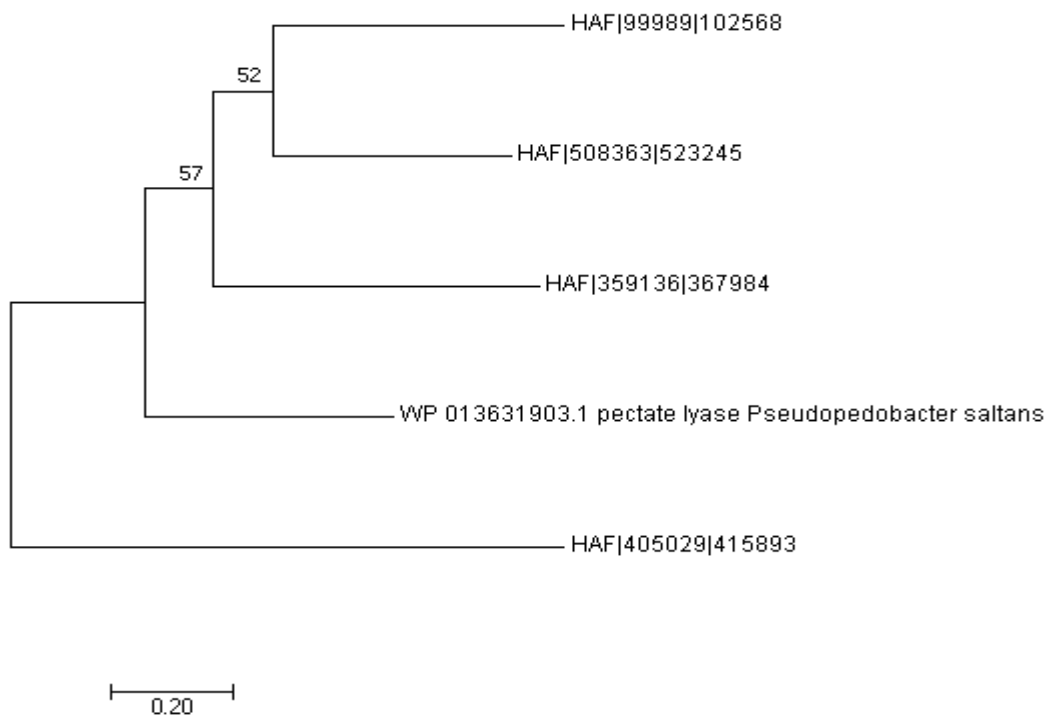
esto se muestra en los anexos S 4 y S 5, en las cuales se lista la taxonomía de dichas bacterias en el orden en el que están en los árboles.

**7.4.2. Cladograma de secuencias de pectato liasas y péptidos del metagenoma de bosque altoandino:** se realizaron los cladogramas con las secuencias ancestrales de la base de datos local y las secuencias del metagenoma con las que tenían similitud, en éste caso se realizaron arboles sin raíz con el método de Neighbor-Joining para así evidenciar los agrupamientos entre las secuencias de pectato liasas. Para C1 se observan los péptidos del metagenoma como una rama aparte de las dos secuencias de la base de datos local. HAF|428026|439807 y HAF|350840|359299 muestran mayor similitud entre sí con un 46% de probabilidad de agrupamiento (Figura 9a), al ser estas secuencias de péptidos, se podría inferir que las tres secuencias mostradas pueden formar parte de la misma proteína o forman parte proteínas homologas.

Para C2 los péptidos de pectato liasas del metagenoma forman un grupo entre HAF|99989|102568, HAF|508363|523245 y HAF|359136|367984 con probabilidad de agrupamiento de 52% para los dos primeros y 57% para el último y los dos anteriores, el péptido HAF|405029|415893 se ubica más alejado de este grupo y más basal incluso que la proteína “ancestral” (Figura 9b).

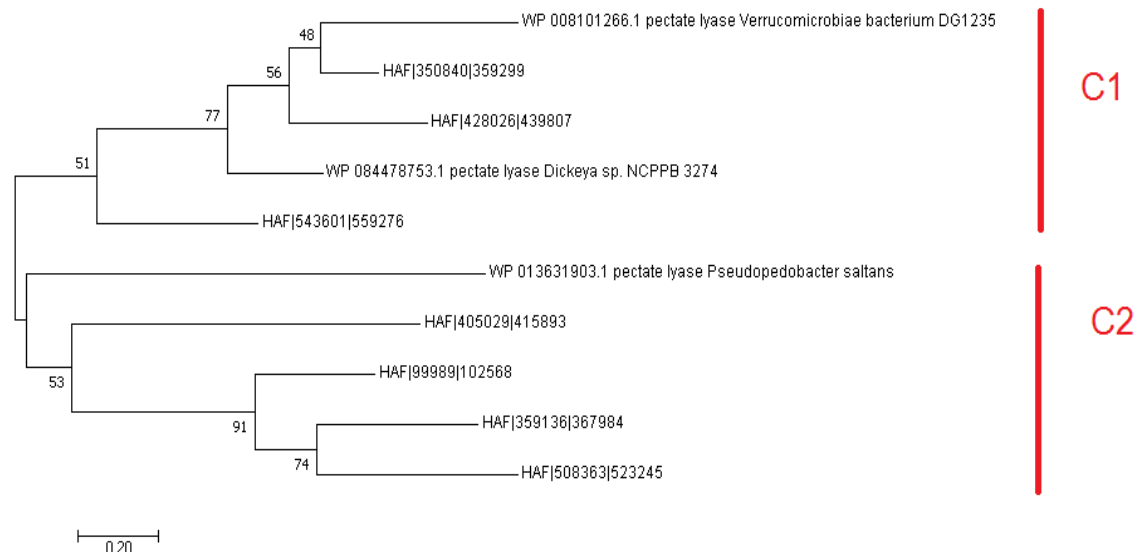


**Figura 9a** Cladograma de ancestros del grupo C1 y las secuencias del metagenoma de Bosque Alto Andino realizado con el método de Neighbor-Joining (NJ) en el programa MEGA 7. 2.



**Figura 9b.** Cladograma de ancestros del grupo C2 y las secuencias del metagenoma de Bosque Alto Andino realizado con el método de Neighbor-Joining (NJ) en el programa MEGA 7. 2.

Al hacer el árbol filogenético con ambos clústeres (C1 y C2) juntos se observa una separación que corresponde precisamente a dichos grupos (Figura 10). En este caso las secuencias muestran mayores probabilidades de agrupamiento en cada caso.



**Figura 10** Cladograma de C1 y C2 realizado con el método de Neighbor-Joining (NJ) en el programa MEGA 7. 2.

## 8. DISCUSIÓN

### 8.1. BASE DE DATOS DE ENZIMAS PUTATIVAS DE PECTATO LIASAS Y HEATMAP.

Debido a la alta capa de materia orgánica que puede contener los suelos de bosque altoandino (Bautista Zúñiga *et al.*, 1998) se espera que haya una gran riqueza de pectato liasas, pues estas enzimas son principales en la degradación de material vegetal. Sin embargo, la base de datos de pectato liasas putativas contiene 60 péptidos de las mismas que representa solo el 0,0097% de los péptidos en el metagenoma, aunque es poco esta base de datos de pectato liasas reportada es un pilar importante para el profundo estudio de estas enzimas en Colombia y su aplicación en diferentes campos tanto biotecnológicos como industriales.

Debido a que los péptido de pectato liasas encontrados en la base de datos corresponden a fragmentos de las proteínas se analizaron solo aquellos que contenían más información de acuerdo al agrupamiento con el heatmap ya que el análisis no se podía realizar con los 60 péptidos juntos. El heatmap es una herramienta permite una mejor visualización de los resultados del BLASTp al agrupar estos resultados y por lo tanto facilitar su análisis (Babicki et al., 2016). El análisis se realizó finalmente con los 7 péptidos que mostraban mayor identidad con la base de datos de pectato liasas descargada. Sería interesante hacer el trabajo con las secuencias de ADN y analizar el codón preferencia.

## 8.2. DOMINIOS DE PECTATO LIASAS PUTATIVAS

El primer grupo analizado, el **clúster C1** (Figura 1), presentó dominio de la familia PL10, ésta familia es particular porque tiene el mismo mecanismo catalítico de  $\beta$ -eliminación pero su estructura tridimensional no es la típica  $\beta$  hélice sino que presenta una forma de  $\alpha$ -hélice en forma de barril ( $\alpha/\alpha$ )<sub>3</sub> (Soriano Lasheras, 2004), ésta familia esta recientemente definida con algunos representante entre ellos Pel-15E del fitopatógenos *Bacillus sp.* KSM-P15 (Sawada et al., 2000), Pel10A de la bacteria del suelo *Cellvibrio japonicus* (Charnock et al., 2002), PelA del fitopatógeno *Bacillus alcalophilus* (Zhai, Cao, & Wang, 2003), PelA de *Treponema pectinovorum* ATCC 33768 la cual es una bacteria que habita en la placa dental sublingual humana (Walker & Ryan, 2003), PelA de *Azospirillum irakense* KBC1 (Bekri et al., 1999) la cual es una bacteria del suelo que fija nitrógeno y la cual no se ha reportado como patógena en plantas (De Armas et al., 2004) y PL D del fitopatógeno *Xanthomonas campestris* (Yuan et al., 2012).

Debido a que la familia PL10 de pectato liasas es relativamente nueva, los motivos no han sido reportados en la literatura, sin embargo Soriano Lasheras en 2004 presenta alineamientos de las pectato liasas PelA de *Azospirillum irakense*, PelA de *Bacillus alcalophilus*, Pel-15E de *Bacillus sp.* KSM-P15 y Pel10A de *Cellvibrio japonicus*, mostrando motivos que también se pueden observar en la base de datos Pfam en el alineamiento que corresponde a la familia PL10 (PF09492) y los

cuales también corresponden al alineamiento del Web Logo realizado en este trabajo, mostrando así la similitud entre las enzimas putativas con las reportadas en la literatura de la familia PL10.

El **clúster C2** (Figura 1) presentó los motivos representativos de la familia PL1 **vWiDH, DIKxxxxxVTxS** y **RxPxxRxGxxHxxxN**, el primer residuo de arginina del último motivo se identifica como el encargado de extraer protones en esta familia (Scavetta, 1999), ésta familia es una de las más ampliamente distribuidas ya que se puede encontrar en bacterias, hongos y plantas. La familia PL1 también se caracteriza por presentar algunas pectato liasas que tienen efectos alérgicos. El clúster C2 presentó dominios Amb a 1.05 nombrado así por la planta *Ambrosia artemisiifolia*, que pertenecen a dichas enzimas, éstos forman parte de los alérgicos más fuertes de ésta planta (Bordas-Le Floch et al., 2015), presentes en el polen de muchas plantas, tienen similitudes a algunas pectato liasas en bacterias fitopatógenas pertenecientes a PL1 debido a que son necesarias para “invadir” el tejido de la planta para fertilizarla de manera análoga a la interacción huésped-patógeno (Asam, Hofer, Wolf, Aglas, & Wallner, 2015; Wing, Yamaguchi, Larabell, Ursin, & McCormick, 1990). (Anexo S 2). Estas evidencias pueden indicar que las enzimas putativas de pectato liasas del grupo C2 pueden o bien ser de bacterias fitopatógenas o tener efectos alérgicos.

Charnock y colaboradores, (2002) demostraron que hay evolución convergente entre PL1 y PL10 con las pectato liasas Pel10Acm de *Cellvibrio japonicus* y Pel1C del mutante R218K, ambas enzimas presentan anti  $\beta$ -eliminación abstrayendo el hidrógeno 5 del poligalacturonato y la eliminación del oxígeno 4 generando un producto de oligogalacturonato 4,5 –insaturado. Ésta puede ser una de las razones por las se encontraron estas dos familias en el metagenoma, ya que pueden tener una función parecida en éste ecosistema.

### **8.3. ANÁLISIS CON ÁRBOLES FILOGENÉTICOS**

Uno de los objetivos de este trabajo constaba en establecer relaciones filogenéticas entre las enzimas de pectato liasas encontradas en el metagenoma

de bosque altoandino, al hacer los árboles filogenéticos se observó que las secuencias que se agrupaban no mostraban ninguna relación de taxones, es decir que las enzimas no mostraban ningún tipo de relación filogenética que permitiera hacer diferenciación entre taxones de bacterias (en este caso en específico).

Para el caso de los grupos de ancestros con las pectato liasas del metagenoma se decidió hacer cladogramas con el algoritmo de Neighbor-Joining que reflejen el grado de similitud de los terminales (Peña, 2011) debido a que este método no puede sesgarse por la poca información de la secuencia (Moret, Roshan, & Warnow, 2002), el tamaño de las secuencias del metagenoma era cortas en comparación con las de los “ancestros” (de 130 a 160 aminoácidos en las secuencias del metagenoma y 406 a 506 para los “ancestros”). Se debe tener en cuenta que entre las secuencias del metagenoma pueden haber duplicados o péptidos que se complementan con el otro y hace parte de la misma proteína por lo que este método agrupa las secuencias por similitud y permite visualizar los dos clústeres que corresponden a los dos dominios PL10 y PL1.

Los suelos del ecosistema de bosque altoandino se caracterizan por tener un pH ácido (~5.2), sin embargo para una actividad óptima de las pectato liasas se requiere un ion de  $\text{Ca}^{2+}$  y un pH alcalino de entre 8 y 10 (Bekri et al., 1999; Charnock et al., 2002; Soriano Lasheras, 2004) la actividad de dichas enzimas decrece cuando se disminuye el pH, Ali et al. en 2015 demostraron en una pectato liasa de *Bacillus subtilis* que cuando el pH disminuye a niveles ácidos (pruebas en pH 7 y 4) se pierde actividad catalítica de la enzima debido a que ocurren fallas en la unión de los iones de calcio entre el sustrato y la enzima, esto indica una peculiaridad en las pectato liasas encontradas ya que pueden ser funcionales a pH bajo como lo es el de bosque altoandino, ésto demuestra un alto potencial de éstas enzimas para procesos de bioprospección.

La búsqueda de pectato liasas en el metagenoma de bosque altoandino resultó exitosa mostrando que si hay presencia de las mismas, de acuerdo a los valores de baja temperatura y pH del ecosistema y teniendo en cuenta las características

generales de las pectato liasas (pH básico principalmente), el estudio específico y más a fondo de las enzimas putativas de este metagenoma puede dar luz a nuevas características que ayuden a mejorar procesos en industrias donde éstas son necesarias como la industria del algodón, clarificación de jugos y vinos, tratamiento de aguas pépticas residuales entre otros, y además de esto permite ampliar el conocimiento de los microorganismos de uno de los ecosistemas claves en nuestro país, el bosque altoandino.

## **9. CONCLUSIONES**

Se consolidó una base de datos de 60 péptidos de genes putativos de pectato liasas de suelo de bosque altoandino de los cuales 7 péptidos muestran motivos altamente funcionales lo cual representa un avance en el estudio de estas enzimas en Colombia.

Se reportaron dos dominios de pectato liasas presentes en el metagenoma de bosque altoandino, PL1 y PL10, ésta información resulta un aporte para el diseño de cebadores que puedan ser usados en la biblioteca genómica almacenada en el Laboratorio de GeBiX y así por medio de técnicas moleculares aislar los genes putativos de pectato liasas

Se determinó la homología de las secuencias de pectato liasas del metagenoma con las de las bases de datos públicas evidenciando en los cladogramas de distancia las relaciones entre las familias de pectato liasas PL1 y PL10 y se evidencio que no se puede usar esta enzima para estudios filogenéticos que distinguan taxones.



## BIBLIOGRAFÍA

- Abascal, F., Aguirre, J., Andres-Leon, E., Bajic, D., Baú, D., Bueren-Calabuig, J. A., ... Zardoya, R. (2014). *Bioinformática con N* (Á. Sebastián & A. Pascual-García, Eds.). CreateSpace.
- Ali, S., Søndergaard, C. R., Teixeira, S., & Pickersgill, R. W. (2015). Structural insights into the loss of catalytic competence in pectate lyase activity at low pH. *FEBS Letters*, *589*(21), 3242–3246. <https://doi.org/10.1016/j.febslet.2015.09.014>
- Alkorta, I., Garbisu, C., Llama, M. J., & Serra, J. L. (1998). Industrial applications of pectic enzymes: a review. *Process Biochemistry*, *33*(1), 21–28. [https://doi.org/10.1016/S0032-9592\(97\)00046-0](https://doi.org/10.1016/S0032-9592(97)00046-0)
- Araújo, R., Casal, M., & Cavaco-Paulo, A. (2008). Application of enzymes for textile fibres processing. *Biocatalysis and Biotransformation*, *26*(5), 332–349. <https://doi.org/10.1080/10242420802390457>
- Asam, C., Hofer, H., Wolf, M., Aglas, L., & Wallner, M. (2015). Tree pollen allergens - An update from a molecular perspective. *Allergy: European Journal of Allergy and Clinical Immunology*, *70*(10), 1201–1211. <https://doi.org/10.1111/all.12696>
- Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J. R., Maciejewski, A., & Wishart, D. S. (2016). Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Research*, *44*(W1), W147–W153. <https://doi.org/10.1093/nar/gkw419>
- Bautista Zúñiga, F., Rivas Solórzano, H., Durán de Bazúa, C., & Palacio, G. (1998). Caracterización y clasificación de suelos productivos en Córdoba, Veracruz, México. *Boletín de Investigaciones Geográficas*, *36*.
- Bekri, M. A., Desair, J., Keijers, V., Proost, P., Leeuwen, S., Vanderleyden, J., ... Leeuwen, M. S. (1999). *Azospirillum irakense Produces a Novel Type of Pectate Lyase* *Azospirillum irakense Produces a Novel Type of Pectate Lyase*. *181*(8), 2440–2447.

- Benen, J. A. E., Kester, H. C. M., Pařenicov, L., & Visser, J. (2000). Characterization of *Aspergillus niger* Pectate Lyase A † , ‡. *Biochemistry*, 39(50), 15563–15569. <https://doi.org/10.1021/bi000693w>
- Bolaños, I. A. (2016). *IDENTIFICACIÓN DE GENES PUTATIVOS DE LACASAS DEL METAGENOMA DE SUELO DE BOSQUE ALTO ANDINO - PARQUE NACIONAL NATURAL “LOS NEVADOS”- COLOMBIA*.
- Borda-Molina, D., Montaña, J. S., Zambrano, M. M., & Baena, S. (2017). Mining lipolytic enzymes in community DNA from high Andean soils using a targeted approach. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, 110(8). <https://doi.org/10.1007/s10482-017-0877-8>
- Bordas-Le Floch, V., Groeme, R., Chabre, H., Baron-Bodo, V., Nony, E., Mascarell, L., & Moingeon, P. (2015). New Insights into Ragweed Pollen Allergens. *Current Allergy and Asthma Reports*, 15(11). <https://doi.org/10.1007/s11882-015-0565-6>
- Castillo Rosales, E., & Rodríguez Alegra, M. E. (2014). Enzimas aplicadas en procesos industriales. *Revista Digital Universitaria*, 15(12), 1607–6079. Retrieved from <http://www.revista.unam.mx/vol.15/num12/art96>
- Charnock, S. J., Brown, I. E., Turkenburg, J. P., Black, G. W., & Davies, G. J. (2002). Convergent evolution sheds light on the anti-beta -elimination mechanism common to family 1 and 10 polysaccharide lyases. *Proceedings of the National Academy of Sciences of the United States of America*, 99(19), 12067–12072. <https://doi.org/10.1073/pnas.182431199>
- Daniel, R. (2005). The metagenomics of soil. *Nature Reviews. Microbiology*, 3(6), 470–478. <https://doi.org/10.1038/nrmicro1160>
- De Armas, H. N., Verboven, C., De Ranter, C., Desair, J., Vande Broek, A., Vanderleyden, J., & Rabijns, A. (2004). *Azospirillum irakense* pectate lyase displays a toroidal fold. *Acta Crystallographica Section D: Biological Crystallography*, 60(6), 999–1007.

<https://doi.org/10.1107/S090744490400602X>

- Dea, I. C. M. (1989). Industrial polysaccharides. *Pure and Applied Chemistry*, 61(7), 1315–1322. <https://doi.org/https://doi.org/10.1351/pac198961071315>
- Demirjian, D. C., Morís-Varas, F., & Cassidy, C. S. (2001). Enzymes from extremophiles. *Current Opinion in Chemical Biology*, 5(2), 144–151. [https://doi.org/10.1016/S1367-5931\(00\)00183-6](https://doi.org/10.1016/S1367-5931(00)00183-6)
- Dowell, K. J. (2008). *Molecular Phylogenetics An introduction to computational methods and tools for analyzing evolutionary relationships*. Retrieved from <https://www.semanticscholar.org/paper/Molecular-Phylogenetics-An-introduction-to-methods-Dowell/4811bea505c8fa681c8c43baee7b78aa402b2f63>
- Dubey, A. K., Yadav, S., Kumar, M., Singh, V. K., Sarangi, B. K., & Yadav, D. (2010). In Silico Characterization of Pectate Lyase Protein Sequences from Different Source Organisms. *Enzyme Research*, 2010, 1–11. <https://doi.org/10.4061/2010/950230>
- Dubey, A., Yadav, S., Kumar, M., Anand, G., & Yadav, D. (2016). Molecular Biology of Microbial Pectate Lyase: A Review. *British Biotechnology Journal*, 13(1), 1–26. <https://doi.org/10.9734/BBJ/2016/24893>
- E. BROWN, I., H. MALLEEN, M., J. CHARNOCK, S., J. DAVIES, G., & W. BLACK, G. (2001). Pectate lyase 10A from *Pseudomonas cellulosa* is a modular enzyme containing a family 2a carbohydrate-binding module. *Biochemical Journal*, 355(1), 155. <https://doi.org/10.1042/0264-6021:3550155>
- Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), 1792–1797. <https://doi.org/10.1093/nar/gkh340>
- Eriksson, M., Ka, J., & Mohn, W. W. (2001). Effects of Low Temperature and Freeze-Thaw Cycles on Hydrocarbon Biodegradation in Arctic Tundra Soil Effects of Low Temperature and Freeze-Thaw Cycles on Hydrocarbon

- Biodegradation in Arctic Tundra Soil. *Appl. Environ. Microbiol.*, 67(11), 5107–5112. <https://doi.org/10.1128/AEM.67.11.5107>
- Ferreira, V., Da Silva, R., Silva, D., & Gomes, E. (2010). Production of pectate lyase by *penicillium viridicatum* RFC3 in solid-state and submerged fermentation. *International Journal of Microbiology*, 2010. <https://doi.org/10.1155/2010/276590>
- Ferrer, M., Martínez-Abarca, F., & Golyshin, P. N. (2005). Mining genomes and “metagenomes” for novel catalysts. *Current Opinion in Biotechnology*, 16(6), 588–593. <https://doi.org/10.1016/j.copbio.2005.09.001>
- GEBIX. (2009). *Second Progress Report - Colciencias. Colombian Center for Genomics and Bioinformatics of Extreme Environments*. 75.
- Gummadi, S. N., & Kumar, D. S. (2008). Batch and fed batch production of pectin lyase and pectate lyase by novel strain *Debaryomyces nepalensis* in bioreactor. *Bioresource Technology*, 99(4), 874–881. <https://doi.org/10.1016/j.biortech.2007.01.022>
- Hassan, S., Shevchik, V. E., Robert, X., & Hugouvieux-Cotte-Pattat, N. (2013). PelN is a new pectate lyase of *dickeya dadantii* with unusual characteristics. *Journal of Bacteriology*, 195(10), 2197–2206. <https://doi.org/10.1128/JB.02118-12>
- Hernandez Camacho, J., & Sánchez Páez, H. (1992). Biotas terrestres de Colombia. In G. Halffter (Ed.), *La diversidad biológica de Colombia* (Primera, pp. 153–175). Mexico.
- Hugouvieux-Cotte-Pattat, N., Condemine, G., & Shevchik, V. E. (2014). Bacterial pectate lyases, structural and functional diversity. *Environmental Microbiology Reports*, 6(5), 427–440. <https://doi.org/10.1111/1758-2229.12166>
- Jiménez, D. J., Montaña, J. S., Álvarez, D., & Baena, S. (2012). A novel cold active esterase derived from Colombian high Andean forest soil metagenome. *World Journal of Microbiology and Biotechnology*, 28(1), 361–370.

<https://doi.org/10.1007/s11274-011-0828-x>

- Kashyap, D. R., Vohra, P. K., Chopra, S., & Tewari, R. (2001). Applications of pectinases in the commercial sector: a review. *Bioresource Technology*, *Amsterdam*, *77*, 215–227.
- Kennedy, J., Marchesi, J. R., & Dobson, A. D. (2008). Marine metagenomics: strategies for the discovery of novel enzymes with biotechnological applications from marine environments. *Microbial Cell Factories*, *7*, 27. <https://doi.org/10.1186/1475-2859-7-27>
- Kikuchi, T., Shibuya, H., Aikawa, T., & Jones, J. T. (2006). Cloning and characterization of pectate lyases expressed in the esophageal gland of the pine wood nematode *Bursaphelenchus xylophilus*. *Molecular Plant-Microbe Interactions : MPMI*, *19*(3), 280–287. <https://doi.org/10.1094/MPMI-19-0280>
- Lee, M. H., & Lee, S.-W. (2013). Bioprospecting Potential of the Soil Metagenome: Novel Enzymes and Bioactivities. *Genomics & Informatics*, *11*(3), 114. <https://doi.org/10.5808/GI.2013.11.3.114>
- Lim, H. K., Chung, E. J., Kim, J.-C., Choi, G. J., Jang, K. S., Chung, Y. R., ... Lee, S.-W. (2005). Characterization of a Forest Soil Metagenome Clone That Confers Indirubin and Indigo Production on *Escherichia coli*. *Applied and Environmental Microbiology*, *71*(12), 7768–7777. <https://doi.org/10.1128/AEM.71.12.7768-7777.2005>
- Lombard, V., Bernard, T., Rancurel, C., Brumer, H., Coutinho, P. M., & Henrissat, B. (2010). A hierarchical classification of polysaccharide lyases for glycogenomics. *Biochemical Journal*, *432*(3), 437–444. <https://doi.org/10.1042/BJ20101185>
- López-López, O., Cerdán, M. E., & González Siso, M. I. (2014). New extremophilic lipases and esterases from metagenomics. *Current Protein & Peptide Science*, *15*(5). <https://doi.org/10.2174/1389203715666140228153801>
- López, G., Chow, J., Bongen, P., Lauinger, B., Pietruszka, J., Streit, W. R., &

- Baena, S. (2014). A novel thermoalkalostable esterase from *Acidicaldus* sp. strain USBA-GBX-499 with enantioselectivity isolated from an acidic hot springs of Colombian Andes. *Applied Microbiology and Biotechnology*, 98(20), 8603–8616. <https://doi.org/10.1007/s00253-014-5775-7>
- Lorenz, P., & Schleper, C. (2002). Metagenome—a challenging source of enzyme discovery. *Journal of Molecular Catalysis B: Enzymatic*, 19–20, 13–19. [https://doi.org/10.1016/S1381-1177\(02\)00147-9](https://doi.org/10.1016/S1381-1177(02)00147-9)
- Madhavan, A., Sindhu, R., Parameswaran, B., Sukumaran, R. K., & Pandey, A. (2017). Metagenome Analysis: a Powerful Tool for Enzyme Bioprospecting. *Applied Biochemistry and Biotechnology*. <https://doi.org/10.1007/s12010-017-2568-3>
- Mayans, O., Scott, M., Connerton, I., Gravesen, T., Benen, J., Visser, J., ... Jenkins, J. (1997). Two crystal structures of pectin lyase A from *Aspergillus* reveal a pH driven conformational change and striking divergence in the substrate-binding clefts of pectin and pectate lyases. *Structure (London, England: 1993)*, 5(5), 677–689. [https://doi.org/10.1016/S0969-2126\(97\)00222-0](https://doi.org/10.1016/S0969-2126(97)00222-0)
- Mendes, R., Kruijt, M., De Bruijn, I., Dekkers, E., Van Der Voort, M., Schneider, J. H. M., ... Raaijmakers, J. M. (2011). Deciphering the rhizosphere microbiome for disease-suppressive bacteria. *Science*, 332(6033), 1097–1100. <https://doi.org/10.1126/science.1203980>
- Minambiente, D. T. A. O. (2017). *Plan de Manejo - Parque Nacional Natural Los Nevados* (pp. 1–238). pp. 1–238. Retrieved from <http://www.parquesnacionales.gov.co/portal/es/ecoturismo/region-andina/parque-nacional-natural-los-nevados/>
- Montaña, J. (2015). *Aproximación metagenómica para la identificación de enzimas lipolíticas en suelo de bosque alto andino del parque nacional natural los nevados*. 215.

- Montaña, J. S., Jiménez, D. J., Hernández, M., Ángel, T., & Baena, S. (2012). Taxonomic and functional assignment of cloned sequences from high Andean forest soil metagenome. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, 101(2), 205–215. <https://doi.org/10.1007/s10482-011-9624-8>
- Moret, B., Roshan, U., & Warnow, T. (2002). Sequence-length requirements for phylogenetic methods. *Workshop on Algorithms in Bioinformatics*, 343–356. [https://doi.org/10.1007/3-540-45784-4\\_26](https://doi.org/10.1007/3-540-45784-4_26)
- Moya A., L. A., & Torres R., E. (2012). Hydrolysis of cellulose and oil palm empty fruit bunches by using consortia of fungi isolated from the soil of Colombian high andean forest. *Soils, Fertilization and Management of Water*, 30(3), 411–418.
- National Research Council (US) Committee on Metagenomics: Challenges and Functional Applications., & Council, N. R. (2007). *The New Science of Metagenomics*. <https://doi.org/10.17226/11902>
- Pandey, A., Negi, S., & Soccol, C. R. (2017). Current Developments in Biotechnology and Bioengineering: Production , Isolation and Purification of Industrial Products. In *Current Developments in Biotechnology and Bioengineering*. <https://doi.org/10.1016/B978-0-444-63662-1.00001-4>
- Parques Nacionales Naturales. (2006). *Plan De Manejo Parque Nacional Natural*. 202. Retrieved from <http://www.parquesnacionales.gov.co/portal/wp-content/uploads/2013/12/LosNevados.pdf>
- Payasi, A., Sanwal, R., & Sanwal, G. G. (2009). Microbial pectate lyases: Characterization and enzymological properties. *World Journal of Microbiology and Biotechnology*, 25(1), 1–14. <https://doi.org/10.1007/s11274-008-9870-8>
- Peña, C. (2011). Métodos de inferencia filogenética. *Revista Peruana de Biología*, 18(2), 265–267. <https://doi.org/10.15381/rpb.v18i2.243>
- Peng, H., Cui, J., Long, H., Huang, W., Kong, L., Liu, S., ... Peng, D. (2016). Novel

- pectate lyase genes of heterodera glycines play key roles in the early stage of parasitism. *PLoS ONE*, 11(3). <https://doi.org/10.1371/journal.pone.0149959>
- Pevsner, J. (2009). Bioinformatics and Functional Genomics. In Intergovernmental Panel on Climate Change (Ed.), *Climate Change 2013 - The Physical Science Basis* (Vol. 53). <https://doi.org/10.1002/9780470451496>
- Radauer, C., Nandy, A., Ferreira, F., Goodman, R. E., Larsen, J. N., Lidholm, J., ... Breiteneder, H. (2014). Update of the WHO/IUIS Allergen Nomenclature Database based on analysis of allergen sequences. *Allergy: European Journal of Allergy and Clinical Immunology*, 69(4), 413–419. <https://doi.org/10.1111/all.12348>
- Reginatto, V., Amante, E. R., Gerhardy, K., Kunst, S., & Duran, N. (2009). Biodegradation and ecotoxicological assessment of pectin production wastewater. *Journal of Environmental Sciences*, 21(11), 1613–1619. [https://doi.org/10.1016/S1001-0742\(08\)62463-8](https://doi.org/10.1016/S1001-0742(08)62463-8)
- Roy, C., Kester, H., Visser, J., Shevchik, V., Hugouvieux-Cotte-Pattat, N., Robert-Baudouy, J., & Benen, J. (1999). Modes of action of five different endopectate lyases from *Erwinia chrysanthemi* 3937. *Journal of Bacteriology*, 181(12), 3705–3709.
- Sasaki, M., Koide, O., Kobayashi, T., Usami, R., & Horikoshi, K. (2015). A pectate lyase from a deep subseafloor *Georgenia muralis* with unusual molecular characteristics. *Extremophiles*, 19(1), 119–125. <https://doi.org/10.1007/s00792-014-0691-4>
- Sawada, K., Ogawa, A., Ozawa, T., Sumitomo, N., Hatada, Y., Kobayashi, T., & Ito, S. (2000). Nucleotide and amino-acid sequences of a new-type pectate lyase from an alkaliphilic strain of *Bacillus*. *European Journal of Biochemistry*, 267(5), 1510–1515. <https://doi.org/10.1046/j.1432-1327.2000.01146.x>
- Scavetta, R. D. (1999). Structure of a Plant Cell Wall Fragment Complexed to Pectate Lyase C. *The Plant Cell Online*, 11(6), 1081–1092.



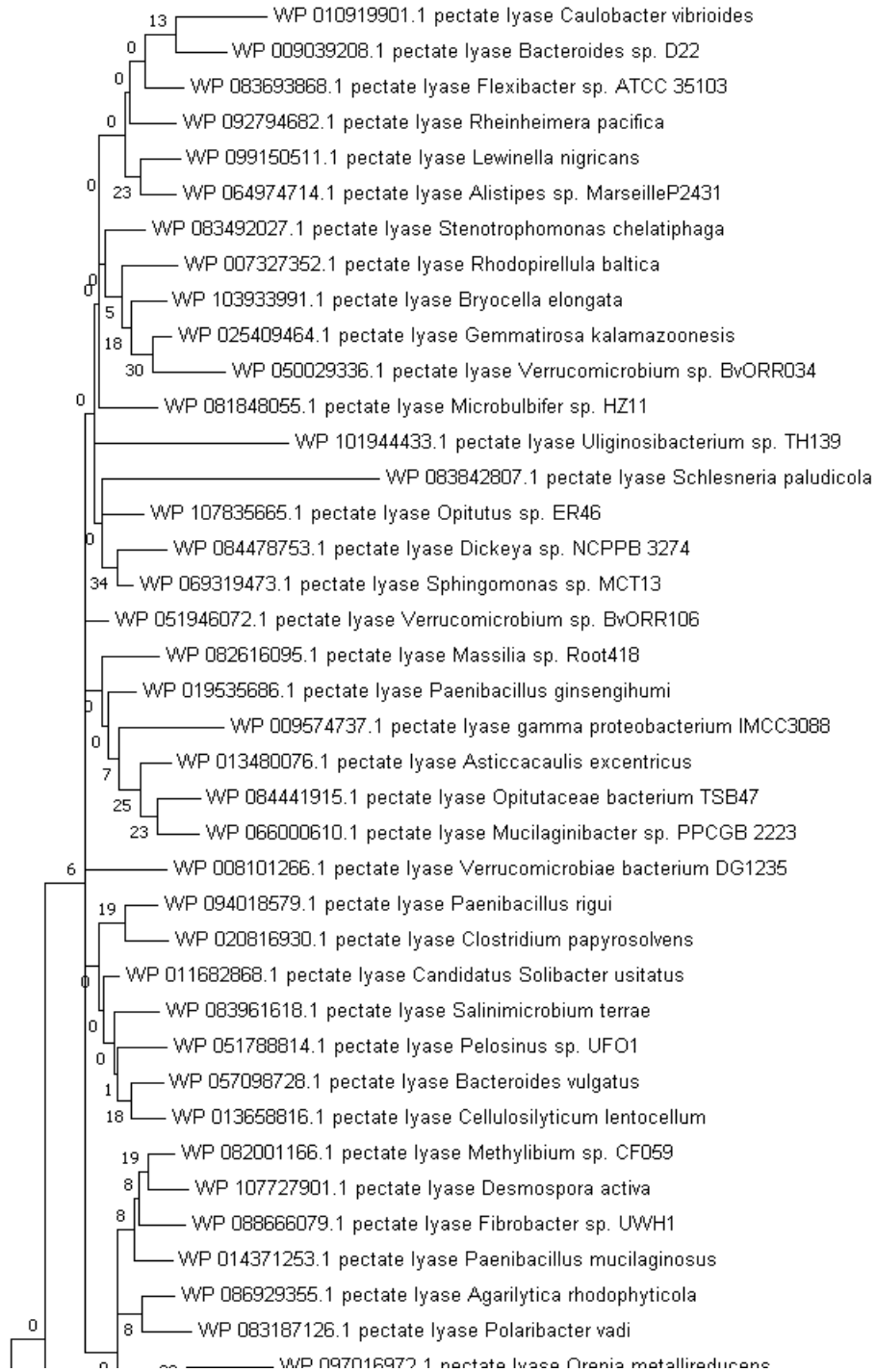
<https://doi.org/10.1105/tpc.11.6.1081>

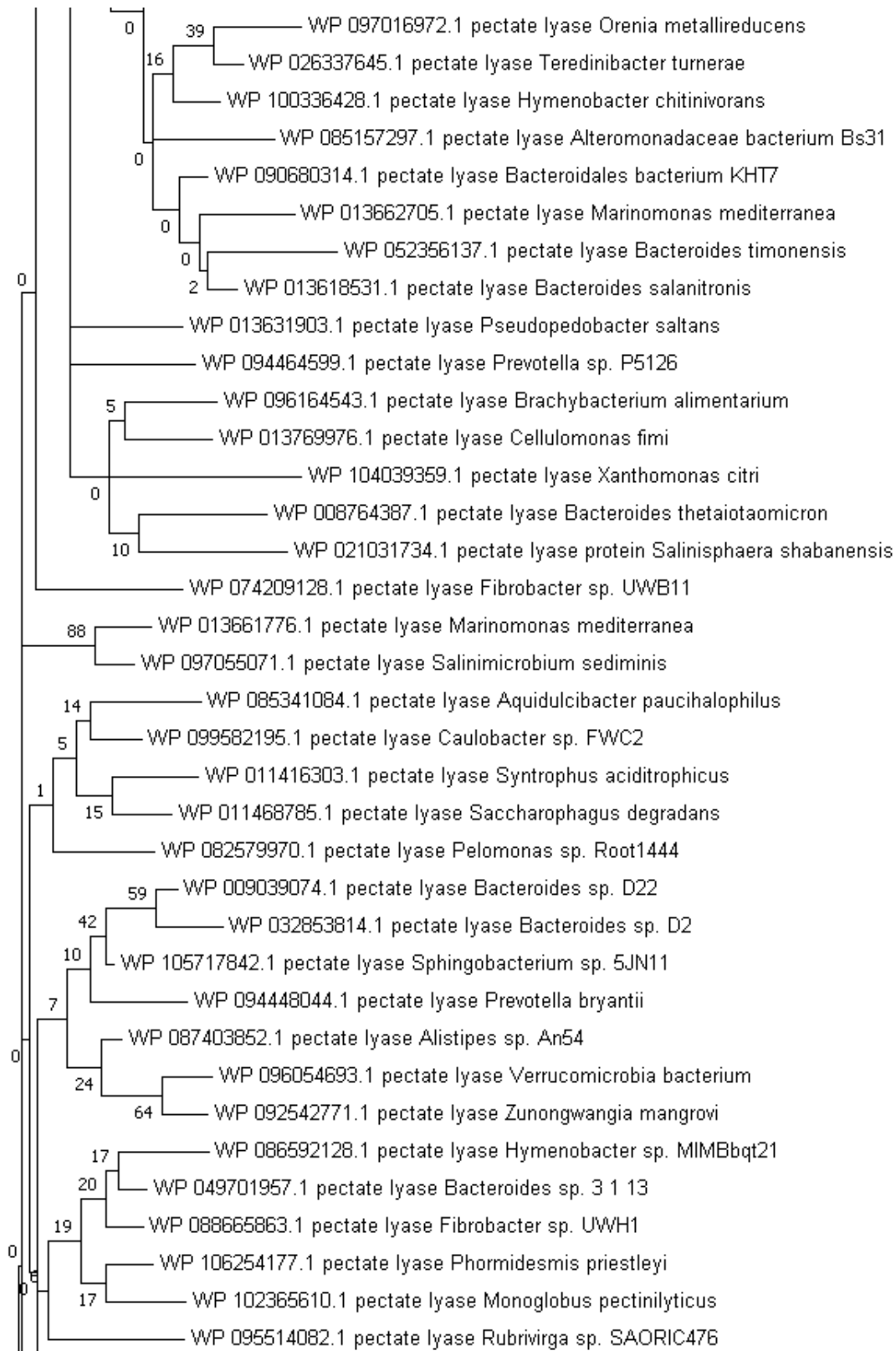
- Schloss, P. D., & Handelsman, J. (2003). Biotechnological prospects from metagenomics. *Current Opinion in Biotechnology*, 14(3), 303–310. [https://doi.org/10.1016/S0958-1669\(03\)00067-3](https://doi.org/10.1016/S0958-1669(03)00067-3)
- Shevchik, V. E., Condemine, G., Robert-Baudouy, J., & Hugouvieux-Cotte-Pattat, N. (1999). The Exopolygalacturonate Lyase PelW and the Oligogalacturonate Lyase Ogl, Two Cytoplasmic Enzymes of Pectin Catabolism in *Erwinia chrysanthemi* 3937. *Journal of Bacteriology*, 181(13), 3912–3913.
- Slater, K. (2003). *Environmental impact of textiles*. Cambridge, England: Woodhead Publishing Ltd.
- Solbak, A. I., Richardson, T. H., McCann, R. T., Kline, K. A., Bartnek, F., Tomlinson, G., ... Kerovuo, J. (2005). Discovery of pectin-degrading enzymes and directed evolution of a novel pectate lyase for processing cotton fabric. *Journal of Biological Chemistry*, 280(10), 9431–9438. <https://doi.org/10.1074/jbc.M411838200>
- Soriano Lasheras, M. (2004). *Análisis de sistemas pectinolíticos bacterianos. Aislamiento y caracterización de las pectinasas PelA de Paenibacillus sp. BP-23 e YvpA de Bacillus subtilis*. 255.
- Streit, W. R., Daniel, R., & Jaeger, K. E. (2004). Prospecting for biocatalysts and drugs in the genomes of non-cultured microorganisms. *Current Opinion in Biotechnology*, 15(4), 285–290. <https://doi.org/10.1016/j.copbio.2004.05.006>
- Tamaru, Y., & Doi, R. H. (2001). Pectate lyase A, an enzymatic subunit of the *Clostridium cellulovorans* cellulosome. *Proceedings of the National Academy of Sciences of the United States of America*, 98(7), 4125–4129. <https://doi.org/10.1073/pnas.071045598>
- Tamaru, Yutaka, Ui, S., Murashima, K., Kosugi, A., Chan, H., Doi, R. H., & Liu, B. (2002). Formation of protoplasts from cultured tobacco cells and *Arabidopsis thaliana* by the action of cellulosomes and pectate lyase from *Clostridium*

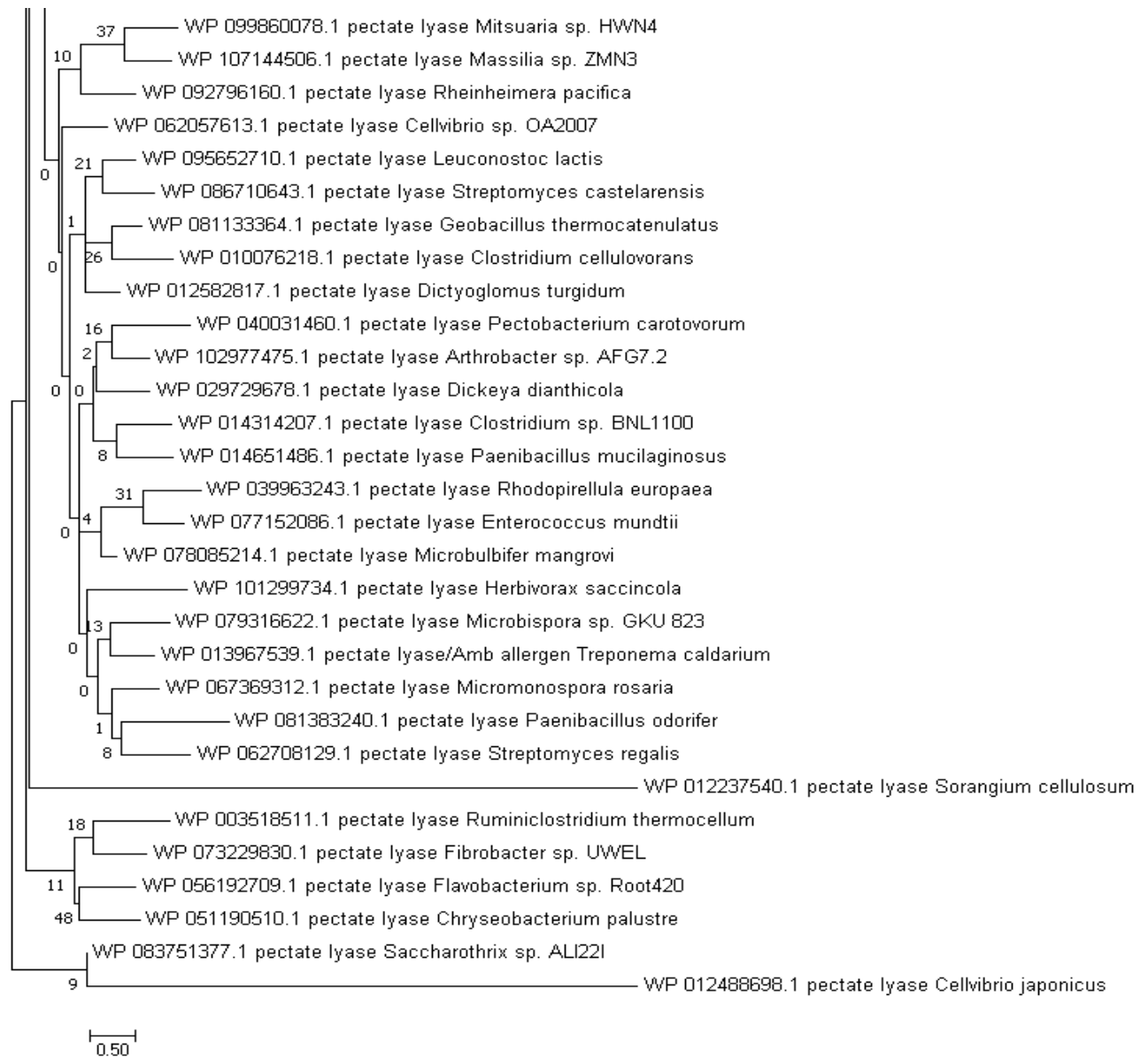
- cellulovorans. *Applied and Environmental Microbiology*, 68(5), 2614–2618. <https://doi.org/10.1128/AEM.68.5.2614-2618.2002>
- Tanabe, H., Kobayashi, Y., & Akamatsu, I. (1986). Pretreatment of Pectic wastewater from orange canning by soft-rot *Erwinia corotovor*a. *Journal of Fermentation Technology*, 64(3), 265–268. [https://doi.org/10.1016/0385-6380\(86\)90110-X](https://doi.org/10.1016/0385-6380(86)90110-X)
- 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. <https://doi.org/10.1093/nar/22.22.4673>
- Timmermans, K. (2001). *Trips, CBD and Traditional Medicines: Concepts and Questions. Report of an ASEAN Workshop on the TRIPS Agreement and Traditional Medicine*. Retrieved from [https://books.google.com.co/books?id=Hp\\_-HAAACAAJ](https://books.google.com.co/books?id=Hp_-HAAACAAJ)
- Torres, O. D., & Velho, L. (2009). La bioprospección como un mecanismo de cooperación internacional para fortalecimiento de capacidades en ciencia y tecnología en Colombia. *Ciencia Da Informacao*, 38(3), 96–110. <https://doi.org/10.1590/S0100-19652009000300007>
- Torsvik, V. (2002). Microbial diversity and function in soil: from genes to ecosystems. *Current Opinion in Microbiology*, 5(3), 240–245. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12057676> <http://www.sciencedirect.com/science/article/pii/S1369527402003247>
- Uchiyama, T., & Miyazaki, K. (2009). Functional metagenomics for enzyme discovery: challenges to efficient screening. *Current Opinion in Biotechnology*, 20(6), 616–622. <https://doi.org/10.1016/j.copbio.2009.09.010>
- Van den Burg, B. (2003). Extremophiles as a source for novel enzymes. *Current Opinion in Microbiology*, 6(3), 213–218. [https://doi.org/10.1016/S1369-5274\(03\)00060-2](https://doi.org/10.1016/S1369-5274(03)00060-2)

- Voget, S., Leggewie, C., Uesbeck, A., Raasch, C., Jaeger, K. E., & Streit, W. R. (2003). Prospecting for Novel Biocatalysts in a Soil Metagenome. *Applied and Environmental Microbiology*, 69(10), 6235–6242. <https://doi.org/10.1128/AEM.69.10.6235-6242.2003>
- Walker, S. G., & Ryan, M. E. (2003). Cloning and expression of a pectate lyase from the oral spirochete *Treponema pectinovorum* ATCC 33768. *FEMS Microbiology Letters*, 226(2), 385–390. [https://doi.org/10.1016/S0378-1097\(03\)00639-6](https://doi.org/10.1016/S0378-1097(03)00639-6)
- Wang, H., Li, X., Ma, Y., & Song, J. (2014). Characterization and high-level expression of a metagenome-derived alkaline pectate lyase in recombinant *Escherichia coli*. *Process Biochemistry*, 49(1), 69–76. <https://doi.org/https://doi.org/10.1016/j.procbio.2013.10.001>
- Wing, R. A., Yamaguchi, J., Larabell, S. K., Ursin, V. M., & McCormick, S. (1990). Molecular and genetic characterization of two pollen-expressed genes that have sequence similarity to pectate lyases of the plant pathogen *Erwinia*. *Plant Molecular Biology*, 14(1), 17–28. <https://doi.org/10.1007/BF00015651>
- Yuan, P., Meng, K., Luo, H., Shi, P., Huang, H., Bai, Y., ... Yao, B. (2011). A novel low-temperature active alkaline pectate lyase from *Klebsiella* sp. Y1 with potential in textile industry. *Process Biochemistry*, 46(10), 1921–1926. <https://doi.org/10.1016/j.procbio.2011.06.023>
- Yuan, P., Meng, K., Wang, Y., Luo, H., Shi, P., Huang, H., ... Yao, B. (2012). A Low-Temperature-Active Alkaline Pectate Lyase from *Xanthomonas campestris* ACCC 10048 with High Activity over a Wide pH Range. *Applied Biochemistry and Biotechnology*, 168(6), 1489–1500. <https://doi.org/10.1007/s12010-012-9872-8>
- Zhai, C., Cao, J., & Wang, Y. (2003). Cloning and expression of a pectate lyase gene from *Bacillus alcalophilus* NTT33. *Enzyme and Microbial Technology*, 33(2–3), 173–178. [https://doi.org/10.1016/S0141-0229\(03\)00091-7](https://doi.org/10.1016/S0141-0229(03)00091-7)

## ANEXOS







**Anexo S 1** Árbol filogenético de la base de datos local completa para identificar el grupo más divergente. El árbol filogenético se realizó con método de Maximum Likelihood (ML) basado en un modelo de matriz JTT con 500 bootstrap en el programa MEGA 7.2. Se muestra el árbol con el logaritmo de mayor probabilidad, el porcentaje de los árboles en los cuales las secuencias se agrupan se muestra al lado de las ramas. El árbol se encuentra en escala y el tamaño de las ramas indica el número de sustituciones por sitio.

**Anexo S 2** Tabla de algunos organismos representativos miembros de las diferentes familias de pectato liasas. Tomado de (Hugouvieux-Cotte-Pattat, Condemine, & Shevchik, 2014). Modificado con la base de datos CAZY (<http://www.cazy.org>, Lombard et al., 2010).

	<b>Organismo</b>	<b>Estilo de Vida</b>	<b>PL1</b>	<b>PL2</b>	<b>PL3</b>	<b>PL9</b>	<b>PL10</b>
	Haloferax volcanii DS2	Halófilo, agua	1	0	0	0	0
<b>Archae</b>	Halorhabdus utahensis DSM 12940	Halófilo, agua	2	0	0	0	0
	Haloterrigena turkmenica DSM 5511	Halófilo, suelo	4	1	0	0	0
	Dickeya dadantii 3937	Patógeno de plantas	8	1	2	3	0
	Pectobacterium atrosepticum SCRI1043	Patógeno de plantas	6	2	2	2	0
	Xanthomonas campestris pv. campestris ATCC33913	Patógeno de plantas	3	0	1	0	1
<b>Bacteria</b>	Rhizobium etli CIAT 652	Simbionte de plantas	1	0	1	0	0
	Paenibacillus mucilaginosus 3016	Rizosferica	12	0	1	0	1
	Sorangium cellulosum So0157-2	Suelo	6	0	3	3	2
	Bacillus subtilis subsp. subtilis str.168	Suelo	2	0	1	1	0

	<i>Clostridium cellulovorans</i> 743B	Ambiental	4	0	0	7	2
	<i>Saccharophagus degradans</i> 2-40	Marina	9	0	4	1	2
	<i>Yersinia enterocolitica</i> subsp. <i>enterocolitica</i> 8081	Patógeno humano	0	2	0	0	0
	<i>Bacteroides thetaiotamicron</i> VPI-5482	Intestino humano	5	0	0	2	1
<b>Fungi</b>	<i>Aspergillus nidulans</i> FGSC A4	Patógeno de plantas	8	0	5	1	0
	<i>Aspergillus niger</i> CBS 513.88	Patógeno de plantas	6	0	0	0	0
<b>Plantas</b>	<i>Arabidopsis thaliana</i>	Planta	26	0	0	0	0
	<i>Ambrosia artemisiifolia</i>	Planta (Alergénicos)	5	0	0	0	0
<b>Animals</b>	<i>Meloidogyne incognita</i>	(Nemátodo) Patógeno de plantas	0	0	4	0	0
	<i>Globodera rostochiensis</i> Trenta	(Nemátodo) Patógeno de plantas	0	0	4	0	0
	<i>Aphelenchus avenae</i>	(Nemátodo) Suelo	0	0	2	0	0
	<i>Bursaphelenchus mucronatus</i>	(Nemátodo) Patógeno de plantas	0	0	2	0	0



	Heterodera glycines	(Nemátodo) Patógeno de plantas	0	0	7	0	0
<b>Oomycetes</b>	Phytophthora capsici	Patógeno de plantas	18	0	7	0	0

**Anexo S 3** Antecedentes.

Año	Autores	Aporte	
1986	Tanabe, Kobayashi, & Akamatsu	Tratamiento de aguas residuales en 24 horas con pectato liasa de <i>Erwinia carotovora</i>	Internacional
1999	Ron et al.	Acción de pectato liasas en sustratos poliméricos y oligoméricos	
2000	Benen, Kester, Pařenicová, & Visser	Actividad específica de pectato liasas de <i>Aspergillus niger</i>	
2001	Tamaru & Doi	Gen de pectato liasa de <i>Clostridium cellulovorans</i> codifica para el complejo enzimático llamado celulosoma	
2001	Brown, Mallen, Charnock, Davies, &	Identificación de pectato liasa Pel10A de <i>Pseudomonas cellulosa</i> con resistencia a pH 10 y temperaturas de 62 °C	

	Black	
2002	Yutaka Tamaru et al.	Extracción de protoplastos por medio de cultivos
2003	Voget et al.	Identificación de genes de 2 pectato liasas (PelA y Uvs119) que hidrolizan la goma gellan.
2005	Solbak et al.	40 pectato liasas nuevas que puedan ser utilizadas en el biodegradado del algodón.
2006	Kikuchi, Shibuya, Aikawa, & Jones	Analizaron pectato liasas en nematodos y su interacción nematodo-planta
2007	Gummadi & Kumar	Analizaron los efectos del cambio de pH, agitación y maceración para la máxima producción de pectin y pectato liasas de la levadura <i>Debaryomyces nepalensis</i> en un biorreactor.
2007	Ferreira, Da Silva, Silva, & Gomes	reportaron las propiedades fisicoquímicas y la producción de una pectato liasa obtenida de <i>Penicillium viridicatum</i>
2010	Dubey et al.	Análisis in silico 121 secuencias de pectato liasas descargadas de GenBank a Homologías, alineamiento múltiple, construcción de árbol filogenético y

		análisis de motivos conservados.
2011	Yuan et al.	Pectato liasas de <i>Klebsiella</i> sp. eficaces para industria textil.
2013	Hassan et al.	Caracterizacion de pectato liasas con rol en la infeccion de plantas
2015	Sasaki et al.	Purificaron una nueva enzima de la bacteria de suelo submarino con actividad a pH de 10 y una temperatura de 50 °C.
2015	Ali et al.	Demostraron perdida de actividad catalitica de pecato liasa a bajo pH.
2016	Peng et al.	Reducción del numero de nematodos en su planta hospedera tras represion de pectato liasa del mismo
2012	Montaña, Jiménez, Hernández, Ángel, & Baena	Análisis composición funcional y taxonómica de suelos de bosque altoandino.
2012	Jiménes, Montaña, Álvarez & Baena	Nueva lipasa termoestable.
2012	Moya A. & Torres R.	Eficiencia de actividad hidrolítica de hongos de bosque altoandino del

Colombia

2014	López et al.	PNN. Nueva esterasa termoalcaloestable aislada a partir de suelo de bosque altoandino.
2017	Borda-Molina, Montaña, Zambrano, & Baena	Nueva lipasa termoestable y alcaloestable.

**Anexo S 4** Taxonomía de las bacterias correspondientes al grupo C1. (V) Validadas por el código. (NV) No validadas.

Nombre	Filo	Clase	Orden	Familia	Genero
<b>NV</b> Dickeya sp. NCPPB 3274, Samson et al. 2005	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Dickeya
<b>NV</b> Verrucomicrobiae, Hedlund et al., 1998 bacterium DG1235	Verrucomicrobia	Verrucomicrobiae	-	-	-
<b>NV</b> gamma proteobacterium	Proteobacteria	Gammaproteobacteria	Celvibrionales	Haliaceae	-

	IMCC3088, Jang Y et al. 2011					
<b>V</b>	Microbulbifer HZ11, al. 1997	sp.Proteobacteria	Gammaproteobacteria	Alteromonadales	Alteromonadaceae	Microbulbifer
<b>NV</b>	Gemmatirosa kalamazoonesis, DeBruyn et al. 2013	Gemmatimonadetes	Gemmatimonadetes	Gemmatimonadales	Gemmatimonadaceae	Gemmatirosa
<b>V</b>	Massilia La Scola et al. 2000	sp. Root418, Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Massilia
<b>V</b>	Opitutus Chin et al. 2001	sp. ER46, Verrucomicrobia	Opitutae	Opitales	Opitutaceae	Opitutus
<b>V</b>	Verrucomicrobium BvORR034, Schlesner, 1988	sp.Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	Verrucomicrobium
<b>V</b>	Alteromonadaceae bacterium Bs31, Ivanova and Mikhailov,	Proteobacteria	Gammaproteobacteria	Alteromonadales	Alteromonadaceae	-

2001						
<b>V</b>	Rhodopirellula báltica, Schlesner et al., 2004	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Rhodopirellula
<b>V</b>	Verrucomicrobium sp. BvORR106, Schlesner, 1988	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	Verrucomicrobium
<b>V</b>	Mucilaginibacter sp. PPCGB 2223, Pankratov et al. 2007	Bacteroidetes	Sphingobacteriia	Sphingobacteriales	Sphingobacteriaceae	Mucilaginibacter
<b>NV</b>	Opitutaceae bacterium TSB47 (Verrucomicrobia sp. Hedlund, 2012)	Verrucomicrobia	Opitutae	Opitutales	Opitutaceae	-
<b>V</b>	Schlesneria paludícola, Kulichevskaya et al. 2007	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Schlesneria
<b>NV</b>	Flexibacter sp. ATCC	Bacteroidetes	Cytophagia	Cytophagales	Cytophagaceae	Flexibacter

35103, Soriano 1945						
<b>V</b>	Bacteroides vulgatus, Eggerth and Gagnon, 1933	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides
<b>V</b>	Agarilytica rhodophyticola, Ling et al. 2017	Proteobacteria	Gammaproteobacteria	Celvibrionales	Cellvibionaceae	Agarilytica
<b>NV</b>	Bacteroidales bacterium, Krieg et al., 2012	Bacteroidetes	Bacteroidia	Bacteroidales	-	-
<b>V</b>	Salinimicrobium terrae, Chen et al., 2008	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Salinimicrobium
<b>V</b>	Alistipes MarseilleP2431, Rautio et al. 2003	sp. Bacteroidetes	Bacteroidia	Bacteroidales	Rikenellaceae	Alistipites
<b>V</b>	Bacteroides Castellani and Chalmers 1919	sp., Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides

<b>NV</b>	Bacteroides timonensis, Ramasamy D et al. 2014	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides
<b>NV</b>	Candidatus Solibacter usitatus, Murray et al. 1995	Acidobacteria	Solibacteres	Solibacterales	Solibacteraceae	Candidatus Solibacter
<b>V</b>	Cellulosilyticum lentocellum, (Murray et al. 1987) Cai and Dong, 2010	Firmicutes	Clostridia	Clostridiales	Lechnospiraceae	Cellulosilyticum
<b>V</b>	Marinomonas mediterránea, Solano and Sanchez-Amat, 1999	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Oceanospirallaceae	Marinomonas
<b>NV</b>	Pelosinus sp. UFO1; Pelosinus lilaceae	Firmicutes	Negativicutes	Selenomodales	Veillonellaceae	Pelocinus
<b>V</b>	Uliginosibacterium TH139, Weon et	Proteobacteria	Betaproteobacteria	Rhodocyclales	Rhodocyclaceae	Uliginosibacterium



	al. 2008					
<b>V</b>	Clostridium papyrosolvens, Madden et al., 1982	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	Clostridium
<b>V</b>	Paenibacillus ginsengihumi, Kim et al., 2008	Firmicutes	Bacilli	Bacillales	Paenibacillaceae	Paenibacillus
<b>V</b>	Hymenobacter chitinivorans, Buczolits et al., 2006	Bacteroidetes	Cytophagia	Cytophagales	Cytophagaceae	Hymenobacter
<b>V</b>	Polaribacter vadi, Kim et al. 2017	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Polaribacter
<b>V</b>	Stenotrophomonas chelatiphaga, Kaparullina et al., 2010	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Stenotrophomonas

**Anexo S 5** Taxonomía de las bacterias correspondientes al grupo C2. (V) Validadas por el código. (NV) No Validadas.

<b>Name</b>	<b>Filo</b>	<b>Clase</b>	<b>Orden</b>	<b>Familia</b>	<b>Genero</b>
<b>NV</b> Pseudopedobacter saltans (sym. Pedobacter saltans), Cao et al. 2014	Bacteroidetes	Sphingobacteriia	Sphingobacteriales	Sphingobacteriaceae	Pseudopedobacter
<b>V</b> Streptomyces castelarensis, (Cercos, 1954) Kumar and Goodfellow, 2008	Actinobacteria	-	Actinomycetales	Streptomycetaceae	Streptomyces
<b>V</b> Microbispora sp. 823, Nonomura and Ohara 1957	GKU Actinobacteria	-	Actinomycetales	Streptoporangiaceae	Microbispora
<b>V</b> Micromonospora rosaria, Horan and Brodsky, 1986	Actinobacteria	-	Actinomycetales	Micromonosporaceae	Micromonospora
<b>NV</b> Rhodopirellula	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Rhodopirellula

europaea						
<b>NV</b>	Herbivorax (Hungateiclostridium saccincola), Koeck <i>et</i> <i>al.</i> 2016	saccincola Firmicutes	Clostridia	Clostridiales	Hungateiclostridiaceae	Herbivorax (Hungateiclostridium)
<b>V</b>	Dictyoglomus Svetlichny Svetlichnayà, 1995	turgidum, Dictyoglomi and	Dictyoglomia	Dictyoglomales	Dictyoglomaceae	Dictyoglomus
<b>V</b>	Enterococcus Collins <i>et al.</i> 1986	mundtii, Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	Enterococcus
<b>V</b>	Clostridium BNL1100, 1880	sp. Firmicutes Prazmowski	Clostridia	Clostridiales	Clostridiaceae	Clostridium
<b>V</b>	Rheinheimera Romanenko <i>et al.</i> , 2003	pacifica, Proteobacteria	Gammaproteobacteria	Chromatiales	Chromatiaceae	Rheinheimera
<b>NV</b>	Microbulbifer Vashist <i>et al.</i> 2013	mangrovi, Proteobacteria	Gammaproteobacteria	Alteromonadales	Alteromonadaceae	Microbulbifer
<b>V</b>	Cellvibrio	japonicus, Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Cellvibrio

Humphry et al. 2003		s				
<b>V</b>	Treponema caldarium Spirochaetae (Spirochaeta caldaria), corrig. (Pohlschröder et al. 1994) Abt, Göker and Klenk 2013	Spirochaetes	Spirochaetales	Spirochaetaceae	Treponema	
<b>V</b>	Geobacillus thermocatenulatus (Syn Geobacillus gargensis), (Golovacheva, 1991) Nazina et al., 2001 emend. Dinsdale et al., 2011)	Firmicutes Bacilli	Bacillales	Bacillaceae	Geobacillus	
<b>V</b>	Clostridium cellulovorans, Sleat et al. 1985	Firmicutes Clostridia	Clostridiales	Clostridiaceae	Clostridium	
<b>V</b>	Paenibacillus mucilaginosus, (Avakyan et al., 1998) Hu et al., 2010	Firmicutes Bacilli	Bacillales	Paenibacillaceae	Paenibacillus	

<b>V</b>	Brachybacterium alimentarium, Schubert et al., 1996	Actinobacteria	-	Actinomycetales	Dermabacteraceae	Bachybacterium
<b>V</b>	Leuconostoc (Syn. Leuconostoc argentinum), Garvie, 1960 (Dicks et al., 1993)	Firmicutes	Bacilli	Lactobacillales	Leuconostocaceae	Leuconostoc
<b>NV</b>	Streptomyces regalis	Actinobacteria	-	Actinomycetales	Streptomycetaceae	Streptomyces
<b>V</b>	Cellulomonas (McBeth and Scales 1913) Bergey et al. 1923)	Firmicutes	Bacilli	Actinomycetales	Cellulomonadaceae	Cellulomonas
<b>V</b>	Dickeya dianthicola, Samson et al., 2005	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Dickeya
<b>V</b>	Pectobacterium carotovorum, (Jones, 1901) Waldee, 1945 emend. Hauben et al.,	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Pectobacterium

---

	1998 emend. Gardan et al., 2003				
<b>V</b>	Massilia sp. ZMN3, La Scola et al., 2000 emend. Kämpfer et al., 2011	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Massilia
<b>V</b>	Arthrobacter AFG7.2, vConn and Dimmick, 1947 emend. Koch et al., 1995	sp. Actinobacteria -	Actinomycetales	Micrococcaceae	Arthrobacter

---