



ESPE

UNIVERSIDAD DE LAS FUERZAS ARMADAS

INNOVACIÓN PARA LA EXCELENCIA

DEPARTAMENTO DE CIENCIAS DE LA VIDA Y DE LA AGRICULTURA

CARRERA DE INGENIERÍA EN BIOTECNOLOGÍA

TESIS PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERA EN
BIOTECNOLOGÍA

AUTOR: MARÍA FERNANDA PROAÑO CUENCA

TEMA: RELACIÓN FILOGENÉTICA DE CEPAS DE *Pythium irregulare*
sensu lato DE LONG ISLAND, NEW YORK EN BASE AL ANÁLISIS DE
SECUENCIAS DEL GEN β - TUBULINA.

DIRECTORA: ALMA ROSEL KOCH KAISER, MC.

CODIRECTOR: ING.-MAT PEDRO JOSÉ ROMERO SAKER.

SANGOLQUÍ, MARZO 2014

CERTIFICACIÓN

MC. Alma Koch Kaiser

Ing.- Mat. Pedro Romero Saker

Certifican:

Que el trabajo titulado “RELACIÓN FILOGENÉTICA DE CEPAS DE *Pythium irregulare* sensu lato DE LONG ISLAND-NEW YORK, EN BASE AL ANÁLISIS DE SECUENCIAS DEL GEN β -TUBULINA”, realizado por la Srta. María Fernanda Proaño Cuenca, ha sido guiado y revisado periódicamente y cumple las normas estatutarias establecidas por la Universidad de las Fuerzas Armadas-ESPE.

Debido a que este estudio es parte del convenio entre el laboratorio de Microbiología de las Universidad de las Fuerzas Armadas y el Departamento de Entomología y Patología de Plantas de la Universidad Estatal de Oklahoma (OSU), se deja en libertad de la autora y de OSU para su publicación.

El mencionado trabajo consta de un documento empastado y un disco compacto con los archivos en formato portátil de Acrobat (pdf). Autorizan a María Fernanda Proaño Cuenca que lo entregue a la MSc. María Augusta Chávez Larrea, en su calidad de Coordinadora de la Carrera.

Sangolquí, Marzo de 2014.

Alma Koch Kaiser, MC.
DIRECTORA

Ing.-Mat. Pedro Romero Saker
CODIRECTOR

DECLARACIÓN DE RESPONSABILIDAD

Yo, MARÍA FERNANDA PROAÑO CUENCA

Declaro que:

El proyecto de grado denominado “RELACIÓN FILOGENÉTICA DE CEPAS DE *Pythium irregulare* sensu lato DE LONG ISLAND-NEW YORK, EN BASE AL ANÁLISIS DE SECUENCIAS DEL GEN B-TUBULINA”, ha sido desarrollado con base a una investigación exhaustiva, respetando derechos intelectuales de terceros, conforme las citas que constan al pie de las páginas correspondientes, cuyas fuentes se incorporan en la bibliografía. Consecuentemente este trabajo es de mi autoría.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance científico del proyecto de grado en mención.

Sangolquí, 18 de Marzo de 2014

María Fernanda Proaño Cuenca

AUTORIZACIÓN

Yo, MARÍA FERNANDA PROAÑO CUENCA

Autorizo a la Universidad de las Fuerzas Armadas-ESPE la publicación, en la biblioteca virtual de la Institución del trabajo “RELACIÓN FILOGENÉTICA DE CEPAS DE *Pythium irregulare* sensu lato DE LONG ISLAND-NEW YORK, EN BASE AL ANÁLISIS DE SECUENCIAS DEL GEN β -TUBULINA”, cuyo contenido, ideas y criterios son de mi exclusiva responsabilidad y autoría

Sangolquí, 18 Marzo de 2014.

María Fernanda Proaño Cuenca

DEDICATORIA

A mi familia: Patricio, Doris, Gaby, Paty y María Gabriela, ustedes son mi todo.

A Dios que es en esencia la justicia, el amor, la seguridad y la libertad que tanto mencionaba sin entender.

María Fernanda Proaño Cuenca

AGRADECIMIENTOS

En primer lugar, deseo expresar mi agradecimiento al Departamento de Entomología y Patología de Plantas de la Universidad Estatal de Oklahoma (OSU), en especial a la Ph.D. Carla Garzón por haberme dado la oportunidad de realizar pasantías en sus laboratorios y de convertir la colaboración en una de sus investigaciones en mi tesis de grado; además por su guía y colaboración en este proyecto, pero sobre todo por ser un ejemplo y por mostrarme el potencial que existe en mí. A la M.S. Margery Daughtrey, del Centro de Horticultura de Long Island, Departamento de Fitopatología y Biología de Plantas y Microbios, Universidad de Cornell, por proveer las muestras analizadas en este estudio.

A la Ing. Patricia Garrido, por compartir sus conocimientos, guía y ayuda durante este proyecto, en especial por confiar en mis habilidades, enseñarme con su ejemplo a trabajar con amor y por convertirse en una amiga. Y a todas las demás personas que trabajan en el Departamento de Entomología y Patología de Plantas de OSU, quienes me brindaron su amistad y con quienes compartí gratos momentos.

A la Universidad de las Fuerzas Armadas- ESPE, Departamento de Ciencias de la Vida y de la Agricultura, Facultad de Ingeniería en Biotecnología, a sus autoridades y docentes quienes contribuyeron en mi formación académica, en especial a mi Directora MC. Alma Koch y mi Coodirector el Ing-Mat. Pedro Romero Saker, por apoyarme en la realización de esta tesis.

De manera muy especial a la MC. Alma Koch por su apoyo incondicional en la consecución del objetivo de graduarme, pero principalmente por su interés en mi desarrollo profesional y personal.

A la Lda. Jessica Maisincho por su ayuda constante durante el tiempo que colaboré en los laboratorios de Microbiología y por brindarme una linda amistad.

A mis padres: Doris y Patricio por el amor, comprensión, alegría, apoyo y valores que me han dado, a quienes les debo todo lo que soy, por no dejar de creer y estar junto a mí siempre que lo he necesitado.

A mis hermanas Gaby y Paty por compartir conmigo la más bonita relación de amistad, por las enseñanzas y los momentos hermosos que he vivido y seguiré viviendo junto a ellas.

A mi sobrina María Gabriela, mi niña dulce, linda y hermosa, por haber llegado a llenar mi vida y la de mi familia de muchas alegrías.

A mis compañeros y amigos de laboratorio de Microbiología con quienes compartí, en especial a Gaby, Alexita, Liz, Henry, Vane y Danny quienes me dieron su apoyo, amistad y además me brindaron momentos agradables que guardaré siempre.

Y a mis queridos amigos de Colegio y Universidad, especialmente a Faby, Edison, Santi, Pame y Erika, que pese a la distancia y circunstancias han sabido brindarme siempre una amistad sincera y verdadera. Gracias por ser como son y quererme tal cual.

MUCHAS GRACIAS

ÍNDICE DE CONTENIDOS

CERTIFICACIÓN	i
DECLARACIÓN DE RESPONSABILIDAD	ii
AUTORIZACIÓN	iii
DEDICATORIA	iv
AGRADECIMIENTOS	v
ÍNDICE DE CONTENIDOS	vii
LISTADO DE TABLAS	x
LISTADO DE FIGURAS	xi
RESUMEN	xvi
ABSTRACT	xvii
CAPÍTULO 1. INTRODUCCIÓN	1
1.1. Formulación del problema.....	1
1.2. Justificación del problema	2
1.3. Objetivos de la investigación.....	4
1.3.1. Objetivo general.....	4
1.3.2. Objetivos específicos	5
1.4. Marco teórico.....	6
1.4.1. <i>Pythium irregulare</i>	6
1.4.1.1. Introducción	6
1.4.1.2. Distribución e impacto en la agricultura.....	7
1.4.1.3. Aislamiento	8
1.4.1.4. Identificación morfológica	9
1.4.1.5. Reproducción	10
1.4.1.6. Caracterización molecular	12
1.4.1.7. Identificación de <i>Pythium irregulare</i> por métodos tradicional y moleculares	14
1.4.2. Análisis y ensamblaje de secuencias.....	19
1.4.2.1. Limpieza de las secuencias de ADN	19
1.4.2.2. Alineamiento y ensamblaje de secuencias	20
1.4.3. Filogenia molecular	20
1.4.3.1. Análisis filogenéticos	21
1.4.4. Estudios taxonómicos y filogenéticos del género <i>Pythium</i>	33

1.5. Hipótesis.....	35
CAPÍTULO 2. MATERIALES Y MÉTODOS	36
2.1. Participantes.....	36
2.1.1. Instituciones	36
2.1.2. Colaboradores científicos	36
2.2. Zona de estudio.....	36
2.3. Periodo de investigación.....	37
2.4. Análisis de datos	37
2.5. Procedimiento.....	38
2.5.1. Material de trabajo	38
2.5.1.1. Limpieza de muestras.....	39
2.5.1.2. Liofilización	39
2.5.2. Caracterización molecular	40
2.5.2.1. Extracción de ADN mediante DNeasy Plant Mini kit de QIAGEN.....	40
2.5.2.2. Determinación de la concentración y pureza de las muestras de ADN mediante espectrofotometría (NANODROP 1000).....	40
2.5.2.3. Dilución del ADN.....	41
2.5.2.4. Reacción en cadena de la polimerasa (PCR convencional)	41
2.5.3. Secuenciación.....	44
2.5.3.1. Preparación para la secuenciación.....	44
2.5.3.2. Envío	44
2.5.4. Análisis de secuencias.....	45
2.5.4.1. Edición y registro de secuencias	45
2.5.4.2. Análisis de similitud y homología	45
2.5.5. Análisis filogenético	46
2.5.5.1. Alineamiento múltiple.....	47
2.5.5.2. Selección del modelo de sustitución.....	48
2.5.5.3. Construcción del árbol filogenético	48
2.5.5.4. Evaluación estadística del árbol filogenético	49
CAPÍTULO 3. RESULTADOS	50
3.1. Caracterización molecular	50
3.1.1. Extracción de ADN y cuantificación	50
3.1.2. Amplificación y visualización del gen β _tubulina	50

3.1.3. Secuenciación del gen β -tubulina	56
3.1.4. Edición y registro de las secuencias obtenidas.....	56
3.1.5. Análisis de similitud y homología de secuencias.	59
3.1.6. Análisis filogenético	61
3.1.6.1. Análisis estadístico de las secuencias.....	61
3.1.6.2. Alineamiento múltiple.....	63
3.1.6.3. Selección del modelo de sustitución.....	65
3.1.6.4. Construcción y evaluación estadística del árbol filogenético	67
CAPÍTULO 4. DISCUSIÓN.....	69
CAPÍTULO 5. CONCLUSIONES.....	81
CAPITULO 6. RECOMENDACIONES	84
CAPÍTULO 7. BIBLIOGRAFÍA.....	85

LISTADO DE TABLAS

Tabla 2.1: Nombre, secuencia, coordenadas, temperatura de fusión (T _m) y referencia de los primers utilizados en las reacciones de PCR y de secuenciación.....	41
Tabla 2.2: Condiciones de reacción PCR para la amplificación del gen β -tubulina en un volumen final de 25 uL.....	42
Tabla 2.3: Ciclos de ampficación utilizados para este estudio	42
Tabla 2.4: Lista de las especies de <i>Pythium</i> incluidas en el análisis filogenético.....	47
Tabla 3.1: Resumen de los datos de composición de bases nucleotídicas y de longitud de las secuencias en estudio.....	61
Tabla 3.2: Datos de frecuencias dinucleótidos idénticos, transversiones, transiciones promedio de las secuencias en estudio.....	62
Tabla 3.3: Pares de nucleótidos y frecuencias promedio de las secuencias en estudio.....	62

LISTADO DE FIGURAS

- Figura 1.1:** Pudrición de raíces de geranio causada por *Pythium irregulare* (Katawczik, 2008)..... 7
- Figura 1.2:** Algunas características morfológicas de *Pythium irregulare*. (A) Hifas limoniformes hinchadas (B) Oospora aplerótica (C) Oogonio intercalar (D) Esporangio terminal (E) Oogonio con una proyección de pared irregular (Fuente: Katawczik, 2008). 10
- Figura 1.3:** Ciclo de enfermedad del ahogamiento y podredumbre de la raíz/semillas producidos por *Pythium* spp. (Agrios 2005). 12
- Figura 1.4:** Diagrama que resume el flujo de etapas que utiliza el algoritmo Muscle para el alineamiento múltiple de secuencias (Edgar, 2004)..... 26
- Figura 2.1:** Esquema del proceso de limpieza de cepas de *Pythium irregulare* en estudio. 39
- Figura 3.1:** Fragmentos de ADN tipo escalera que contiene el marcador de peso molecular 100 bp DNA ladder (VWR cat. no. 95039-808) utilizado para medir el largo en pares de bases (bp) del ADN amplificado por PCR..... 51
- Figura 3.2:** Visualización del producto de PCR del gen β -tubulina amplificado de 13 muestras de ADN (1:C8; 2:C9; 3:C14; 4:C20; 5:C24; 6: C32; 7: C37; 8:C38; 9:C39; 10:C41; 11:C42; 12:C43; 13:C44), en gel de agarosa al 1% p/v. M: Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de la reacción PCR. Las muestras se identifican por su código de estudio. 51
- Figura 3.3:** Visualización del producto de PCR del gen β -tubulina amplificado de 13 muestras de ADN (1:C54; 2:C60; 3:C61; 4:C62; 6:C65; 7: C66; 9:C68), en gel de agarosa al 1% p/v. M:

Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de la reacción PCR. Las muestras se identifican por su código de estudio. 52

Figura 3.4: Visualización del producto de PCR del gen β -tubulina amplificado de 16 muestras de ADN. (1:C69; 2:C71; 3:C74; 4:C79; 5:C80; 6: C81; 7: C82; 8:C83; 9:C84; 10:C87; 11:C90; 12:C92; 13:C93; 14:C94; 15:C95; 16:C96), en gel de agarosa al 1% p/v. M: Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de la reacción PCR. Las muestras se identifican por su código de estudio. 52

Figura 3.5: Visualización del producto de PCR del gen β -tubulina amplificado de 10 muestras de ADN. (1:C100; 2:C101; 3:C109; 4:C121; 5:C123; 6: C129; 7: C131; 8:C132; 9:C136; 10:C140), en gel de agarosa al 1% p/v. M: Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de la reacción PCR. Las muestras se identifican por su código de estudio. 53

Figura 3.6: Visualización del producto de PCR del gen β -tubulina amplificado de 17 muestras de ADN. (1:C145; 2:C147; 3:C149; 4:C152; 5:C153; 6: C156; 7: C158; 8:C160; 9:C161; 10:C162; 11:C164; 12:C165; 13:C167; 14:C169; 15:C171; 16:C178; 17:C179), en gel de agarosa al 1% p/v. M: Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de la reacción PCR. Las muestras se identifican por su código de estudio. 53

Figura 3.7: Visualización del producto de PCR del gen β -tubulina amplificado de 18 muestras de ADN. (1:C180; 2:C185; 3:C186; 4:C187; 5:C189; 6: C191; 7: C192; 8:C196; 9:C198; 10:C199; 11:C200; 12:C201; 13:C202; 14:C203; 15:C205; 16:C206; 17:C207; 18:C208), en gel de agarosa al 1% p/v. M: Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de

- la reacción PCR. Las muestras se identifican por su código de estudio.. 54
- Figura 3.8:** Visualización del producto de PCR del gen β -tubulina amplificado de 16 muestras de ADN (1:C209; 2:C210; 3:C211; 4:C212; 5:C213; 6: C214; 7: C216; 8:C220; 9:C223; 10:C228; 11:C230; 12:C232; 13:C234; 14:C235; 15:C236; 16:C188), en gel de agarosa al 1% p/v. M: Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de la reacción PCR. Las muestras se identifican por su código de estudio..... 54
- Figura 3.9:** Visualización del producto de PCR del gen β -tubulina amplificado de 18 muestras de ADN. (1:C239; 2:C240; 3:C241; 4:C242; 5:C243; 6: C244; 7: C246; 8:C247; 9:C248; 10:C254; 11:C255; 12:C256; 13:C257; 14:C258; 15:C259; 16:C260; 17:C261; 18:C264), en gel de agarosa al 1% p/v. M: Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de la reacción PCR. Las muestras se identifican por su código de estudio.. 55
- Figura 3.10:** Visualización del producto de PCR del gen β -tubulina amplificado de cinco muestras de ADN (1:C117; 2:C120; 3:C135; 4:C174; 5:C197), en gel de agarosa al 1% p/v. M: Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de la reacción PCR. Las muestras se identifican por su código de estudio. 55
- Figura 3.11:** Paneles que muestran los problemas encontrados con la cepa C109. **(A)** Calidad inaceptable de las secuencia sentido y antisentido de la cepa **(B)** Ruido de fondo en el cromatograma perteneciente a la secuencia de la cepa C109. Los cuatro tipos de nucleótidos están diferenciados por colores (C: azul, G: negro, T: rojo, A: verde) **(C)** Panel que muestra el alineamiento de la secuencia de la cepa C109. 56

- Figura 3.12:** Paneles que muestran los problemas encontrados con la cepa C117. **(A)** Calidad inaceptable de las secuencia sentido y antisentido de la cepa **(B)** Ruido de fondo en el cromatograma perteneciente a la secuencia de la cepa C117. Los cuatro tipos de nucleótidos están diferenciados por colores (C: azul, G: negro, T: rojo, A: verde) **(C)** Alineamiento de las secuencias de la cepa C117. 57
- Figura 3.13:** Paneles que muestran los problemas encontrados con la cepa C135. **(A)** Calidad inaceptable de las secuencia sentido y antisentido de la cepa **(B)** Ruido de fondo en el cromatograma perteneciente a la secuencia de la cepa C135. Los cuatro tipos de nucleótidos están diferenciados por colores (C: azul, G: negro, T: rojo, A: verde) **(C)** Alineamiento de las secuencias de la cepa C135..... 58
- Figura 3.14:** Paneles que muestran los problemas encontrados con la cepa C256. **(A)** Calidad inaceptable de las secuencia sentido y antisentido de la cepa **(B)** Secuencias superpuestas en el cromatograma perteneciente a la cepa C256. Los cuatro tipos de nucleótidos están diferenciados por colores (C: azul, G: negro, T: rojo, A: verde) **(C)** Alineamiento de las secuencias de la cepa C256..... 59
- Figura 3.15:** (A) Resultados del alineamiento local de una de las secuencias problema (C109) con secuencias existentes en la base de datos de Blast (B) Información asociada a la secuencia *Pythium irregulare* (Número de accesoión GenBank: AB512839) (C) Información asociada a la secuencia *Pythium irregulare* (Número de accesoión GenBank: AB512837 61
- Figura 3.16:** Ventana que muestra el alineamiento múltiple con el algoritmo Muscle en base a codones de las secuencias en estudio. 64

- Figura 3.17:** Mejores modelos de sustitución y evolución para las secuencias en estudio según el método estadístico de máxima verosimilitud..... 66
- Figura 3.18:** Árbol consenso en base a secuencias del gen β -tubulina utilizando el método de máxima verosimilitud basado en el modelo de Tamura de tres parámetros (T92+I). El árbol fue enraizado con *Pythium spinosum* y *Pythium sylvaticum* como grupos externos. Los números en las ramas representan los valores “bootstrap” calculados a partir de 1000 replicaciones (Sólo los valores superiores al 50% se muestran). 69

RESUMEN

Pythium irregulare es una importante especie patógena de plantas encontrada alrededor del mundo. Estudios de tipo morfológicos y moleculares han confirmado que *Pythium irregulare* sensu lato muestra altos niveles de diversidad morfológica y genética y que forma parte de un complejo de especies. Una colección internacional de 120 aislados de invernaderos de Long Island - New York colectados de diversos huéspedes fueron seleccionados para investigar su relación filogenética. El análisis filogenético fue realizado en base a la caracterización del gen β -tubulina. Fragmentos de aproximadamente 1200 pb del gen β -tubulina fueron amplificados con el primer sentido β tub_F1A y el primer antisentido β tub_R1A. En el análisis bajo el método de máxima verosimilitud se incluyeron seis secuencias, cuatro correspondientes a especies del complejo *Pythium irregulare* (*P. irregulare* sensu stricto, *P. cryptoirregulare*, *P. cylindrosporum*) y las secuencias de *Pythium spinosum* y *Pythium sylvaticum* usadas como grupos externos. El árbol filogenético consenso obtenido reveló la formación de tres grupos principales con un soporte "bootstrap" moderado-alto (Valor $p < 0.05$). Los resultados no mostraron una clara separación entre las especies del complejo *P. irregulare*, lo cual es consistente con otros estudios preliminares. El análisis del gen β -tubulina como loci individual fue útil pero no garantizó una filogenia de especies completamente segura por lo que un análisis concatenado de genes sería un mejor método para la reconstrucción filogenética. La discriminación entre las especies de este complejo es fundamental para el estudio de su epidemiología y para el manejo de las enfermedades que causa.

Palabras clave:

Pythium irregulare, Sensu lato, Filogenética, Máxima verosimilitud, Bootstrap

ABSTRACT

Pythium irregulare is an important plant pathogenic species found around the world. Morphological and molecular studies have confirmed that *Pythium irregulare* sensu lato shows high levels of morphological and genetic diversity and is part of a species complex. An international collection of 120 isolates from greenhouses of Long Island - New York collected from diverse host were selected to investigate their phylogenetic relationship. Phylogenetic analyses were performed based on the characterization of the β -tubulin gene. Fragments of approximately 1200 bp of β -tubulin gene were amplified with the forward primer β tub_F1A and reverse primer β tub_R1A. The analysis under the maximum likelihood method included six sequences, four corresponding to *Pythium irregulare* species complex (*P. irregulare* sensu stricto, *P. cryptoirregulare*, *P. cylindrosporum*) and sequences of *Pythium spinosum* and *Pythium sylvaticum* as outgroups. The consensus phylogenetic tree obtained revealed the formation of three major groups with a moderate-high "bootstrap" support (p-value <0.05). The results showed no clear separation between species within the *P. irregulare* complex, which is consistent with other preliminary studies. Analysis of the β -tubulin gene as single loci was useful but not guaranteed a completely and secure phylogeny of species so a concatenated gene analysis would be a better method for phylogenetic reconstruction. Discrimination among species of this complex is essential for the study of their epidemiology and management of diseases that cause.

Key words:

Pythium irregulare, sensu lato, Phylogenetic, Maximum likelihood, Bootstrap.

CAPÍTULO 1. INTRODUCCIÓN

1.1. Formulación del problema

Pythium es un género de oomicetos patógenos de plantas encontrados alrededor del mundo. En invernaderos, viveros, campos agrícolas y bosques es conocido por causar la pudrición de semillas y de raíces de plantas, pudrición de frutos carnosos u otros órganos vegetales en contacto con el suelo y ahogamiento “damping-off” pre y post emergencia de plantas (Agrios, 2005).

Tradicionalmente la diferenciación e identificación de especies de *Pythium* se hace basada en criterios morfológicos. Los sistemas de clasificación basados en estos criterios han sido útiles, pero en las últimas décadas se ha hecho notorio que muchas de las especies morfológicas son parte de grupos polifiléticos, es decir, grupos que no descienden de un ancestro en común (Villa, Kageyama, Asano, & Sura, 2006).

La identificación de especies que pertenecen al género *Pythium* es difícil debido a similitudes y variaciones intraespecíficas en las características morfológicas que ocurren entre aislados individuales, por lo que, además de las características morfológicas, métodos moleculares son utilizados en estudios taxonómicos y han permitido incrementar la habilidad para identificar y caracterizar organismos que tienen pocos caracteres morfológicos variables (Garzón, Geiser, & Moorman, 2005), como en el caso de *Pythium irregulare*,

un oomicete patógeno de plantas que causa daños a una amplia variedad de cultivos de tipo vegetal.

Estudios de tipo morfológicos y moleculares han confirmado que *Pythium irregulare* sensu lato¹(s.l) muestra altos niveles de diversidad morfológica y genética lo que lleva a cuestionar si se trata de una sola especie o si es en realidad un complejo de especies morfológicamente similares: especies crípticas (Garzón *et al.*, 2005).

Si la variabilidad que *Pythium irregulare* exhibe se puede caracterizar, se dispondría de información útil para la implementación de estrategias en el control y gestión de enfermedades causadas por el patógeno (Garzón *et al.*, 2005).

1.2. Justificación del problema

Los oomicetos son un grupo importante de patógenos de plantas eucarióticas, responsables de grandes pérdidas económicas cada año debido a que actúan en el suelo y en la rizosfera causando una notable reducción en las cosechas y en la calidad de los cultivos (Meng, Torto Alalibo, Chibucos, Tyler, & Dean, 2009).

Pythium irregulare, un oomiceto, es el agente causal de enfermedades como el ahogamiento y pudrición tanto de la raíz como de las semillas y de plántulas. Es un organismo único al ser capaz de penetrar las superficies de

¹ Sensu lato: término usado para incluir a especies morfológicamente similares, significa en el sentido amplio.

plantas hospederas, estableciendo rápidamente infecciones que pueden tener consecuencias severas para la producción agrícola. La identificación de este organismo en cultivos es necesaria debido a que está involucrado en una amplia gama de actividades y porque su detección tiene aplicaciones en estudios de epidemiología (Agrios, 2005).

Antes de la aplicación de la biología molecular moderna, la taxonomía de *Pythium* se basó principalmente en aspectos morfológicos y fisiológicos. Ahora los análisis filogenéticos de secuencias en base a regiones o genes la han transformado y se han convertido en herramientas importantes pues no están sujetos a interacciones ambientales lo que los hace estables y reproducibles (Belbahri *et al.*, 2008).

El trabajo con cepas de *Pythium irregulare* sensu lato colabora con una parte de una investigación llevada a cabo en la Universidad Estatal de Oklahoma (OSU) y cuyo objetivo es reconstruir la relación evolutiva de las especies pertenecientes al complejo *Pythium irregulare* sensu lato. La discriminación entre las especies de este complejo es fundamental para el estudio de su epidemiología y para el manejo de las enfermedades que causa.

P.Garrido (comunicación personal, Mayo 6, 2013), responsable de la investigación, busca reconstruir esta relación mediante un análisis multigénico. Se requiere el análisis de genes independientes y de la comparación de filogenias derivadas de los mismos para intentar construir la historia evolutiva de un organismo, ya que la evolución de un solo gen no

representa la evolución de todo un genoma (Villa *et al.*, 2006). La contribución de este estudio es la caracterización de un gen, el gen β -tubulina.

β -tubulina es un gen altamente conservado y común a todos los eucariotas usado para establecer comparaciones entre miembros de un mismo género, permitiendo la diferenciación de especies. Este gen ha sido utilizado en la caracterización de patógenos y en la reconstrucción de relaciones filogenéticas (McCartney, Foster, Fraaije, & Ward, 2003).

De acuerdo con Garzón (2007), la identificación de patógenos suele resultar difícil debido a que los organismos pueden tener una morfología confusa o poco estudiada. Es entonces que el desarrollo de técnicas moleculares para la identificación de patógenos brinda la posibilidad de diferenciar especies que pueden parecer ser una sola pero que genéticamente son dos diferentes.

1.3. Objetivos de la investigación

1.3.1. Objetivo general

Establecer la relación filogenética de cepas de *Pythium irregulare* sensu lato de Long Island - New York, en base a secuencias del gen β -tubulina.

1.3.2. Objetivos específicos

- Limpiar y mantener aisladas 120 muestras de *Pythium irregulare* sensu lato de Long Island, New York -Estados Unidos, mediante microbiología tradicional.
- Amplificar y secuenciar el gen β -tubulina en las cepas de *Pythium irregulare* sensu lato.
- Crear secuencias consenso a partir del análisis, edición y ensamblaje de los datos generados después de la secuenciación, con el programa BioNumerics versión 7.01.
- Construir un alineamiento múltiple de las secuencias en estudio con el algoritmo Muscle del programa MEGA versión 5.2.
- Elegir un modelo de sustitución y de evolución para las secuencias en estudio con el programa MEGA versión 5.2.
- Construir un árbol filogenético basado en el correspondiente alineamiento múltiple y el modelo de sustitución-evolución elegido, con el método estadístico de máxima verosimilitud.
- Evaluar estadísticamente el árbol filogenético construido mediante un análisis “bootstrap” con 1000 repeticiones.

1.4. Marco teórico

1.4.1. *Pythium irregulare*

1.4.1.1. Introducción

El género *Pythium* pertenece a la familia Pythiaceae, orden Pythiales, clase Oomycetes, Phylum Oomycota y reino Chromista (Nzungize, y otros, 2011). Es un género de oomicetos que incluye más de 120 especies (Dick, 1990).

Los oomicetos son un grupo importante de patógenos de plantas, que tienen la habilidad de infectar a un amplio rango de especies causando pérdidas cualitativas y cuantitativas en cultivos sembrados en diferentes ecosistemas (Almodóvar, 1996). Son predominantemente diploides durante sus ciclos de vida y producen una fase haploide transitoria muy breve antes de la fertilización (Blair, Coffey, Park, Geiser, & Kang, 2008).

Pythium irregulare es un oomicete patógeno del suelo encontrado alrededor del mundo en muchas plantas hospederas y es el agente de causal de enfermedades que causan podredumbre de plántulas, semillas y raíces (figura 1.1), lesiones del tallo, podredumbre de frutos y otros órganos vegetales que están en contacto con el suelo (Garzón, Yáñez, & Moorman, 2007; Agrios, 2005).



Figura 1.1: Pudrición de raíces de geranio causada por *Pythium irregulare* (Katawczik, 2008).

1.4.1.2. Distribución e impacto en la agricultura

Las especies de *Pythium* están ampliamente distribuidas en el mundo. Muchas son saprofitas en el suelo y el agua así como en restos de plantas muertas y animales. Otras son agentes de control biológico que protegen contra hongos patógenos, mientras que otras son patógenos de plantas y animales (Long *et al.*, 2011). Muchas especies saprofitas de *Pythium* tienen la capacidad de volverse patogénicas bajo condiciones de estrés para el huésped.

Pythium irregulare tiene una amplia gama de huéspedes y ocupa una gran variedad de hábitats ecológicos terrestres y acuáticos (Nzungize *et al.*, 2011). Se ha reportado en todos los continentes excepto en la Antártida y se ha encontrado en césped, cereales, pimiento, piñas, lentejas, cítricos, apio, fresas, maíz, soya, zanahoria, cebolla y en cultivos de floricultura y silvicultura.

En invernadero se encuentra como patógeno debido a fuentes de irrigación de agua contaminadas y posiblemente también por la introducción de material vegetal infectado (Katawczik, 2008).

Comúnmente en la mayoría de suelos cultivados causa serios problemas en cultivos agrónomicamente importantes. Afecta principalmente a semilleros de diferentes cultivos causando pudrición de semillas y de la raíz de plantas adultas. Los órganos de plantas que han estado en contacto con el suelo, son a veces infectados por este oomicete durante períodos de humedad extendidas (Palmucci & Grijalba, 2007).

El desarrollo de *Pythium* se ve favorecido por el drenaje deficiente del suelo, alta proporción de arcilla, mala circulación de aire entre las plantas, condiciones de alta humedad y agua libre que facilitan la propagación de zoosporas. Además la introducción de inóculos en cultivos puede deberse a tierra, sustratos, herramientas de riego, agua o semillas infectadas (Palmucci & Grijalba, 2007).

1.4.1.3. Aislamiento

P. irregulare puede ser aislado del tejido de una planta mediante esterilización de la superficie del tejido, lavado en agua estéril y sembrado en placas de agar agua o medio PARP (Jeffers & Martin, 1986), un medio semi-selectivo con una base de agar de harina de maíz y modificada con agentes antifúngicos y antibacterianos (pimaricina, ampicilina, rifampicina, y pentacloronitrobenzeno). El aislamiento a partir del suelo puede hacerse

usando tejido vivo sobre el suelo y su posterior incubación y transferencia del tejido a un medio selectivo. Otra forma de aislamiento puede ser por una dilución del suelo en placas, lo que implica pesar una cantidad de suelo secado al aire, suspenderlo en una cantidad conocida de agua y sembrar diferentes volúmenes diluidos en medios selectivos (Katawczik, 2008).

1.4.1.4. Identificación morfológica

P. irregulare se diferencia de otras especies por su morfología estructural. Tiene hifas típicamente de 5 μm de diámetro, las cuales pueden tener hinchazones limoniformes. Los esporangios, las estructuras asexuales, son esféricos, pueden ser terminales o intercalares y suelen producir zoosporas de cerca de 7-10 μm cuando está enquistado (Katawczik, 2008). El oogonio, la estructura sexual femenina, es esférica, lisa, ligeramente irregular en su forma y puede formar una o más proyecciones cilíndricas en la pared ogonial. El diámetro de los oogonios varía en un rango de 15 a 25 μm ; la posición ogonial puede ser intercalar o terminal; con la formación de uno o tres anterios por oogonio los cuales pueden ser monoicos, hipóginos o dióicos (Garzón, Yáñez, & Moorman, 2007).

El anteridio, la estructura sexual masculina, puede ser sésil o poseer alguna estructura de soporte y su tallo puede ser ramificado o no ramificado. Las células de anteridios pueden tener entre 10-15 x 4-5 μm , además hacen contacto apical con el oogonio. Las oosporas, el resultado de la reproducción sexual, son por lo general apleróticas y ocasionalmente pleróticas y varían entre 14-20 μm en diámetro con paredes de 1.0-1.5 μm de espesor (Garzón

et al., 2007). Algunas características morfológicas se observan en la figura 1.2.

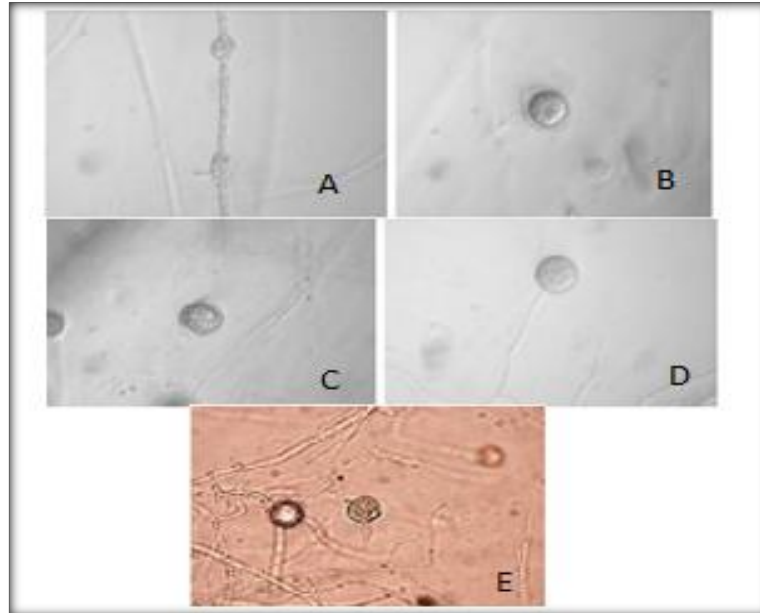


Figura 1.2: Algunas características morfológicas de *Pythium irregulare*. **(A)** Hifas limoniformes hinchadas **(B)** Oospora aplerótica **(C)** Oogonio intercalar **(D)** Esporangio terminal **(E)** Oogonio con una proyección de pared irregular (Katawczik, 2008).

1.4.1.5. Reproducción

Pythium irregulare tiene tanto una etapa sexual como asexual. La reproducción sexual se realiza a través de oogonios y anteridios. Los órganos femeninos, oogonios, son esféricos o limoniformes y pueden ser intercalares o terminales. Los anteridios, los órganos masculinos, consisten de una célula anteridial, que puede ser sésil o una hifa intercalada o terminal. Durante la reproducción sexual, la célula anteridial toca el oogonio y forma un tubo de fertilización, que penetra en el oogonio. Los anteridios se denominan monoicos si se originan en el tallo oogonial y dioicos si lo hacen de una hifa

diferente no conectada con el oogonio. Después de la fecundación, el contenido oogonial forma un cigoto que se convierte en una oospora. Para la maduración de las oosporas, suele ser necesaria una fase latente antes de que la germinación tenga lugar. En la germinación, las oosporas producen un tubo germinal (figura 1.3). Las oosporas son resistentes tanto a las temperaturas como a la humedad y sirven como etapa de supervivencia y de reposo del oomicete. Además la liberación de semillas o exudados de la raíz causa que la oospora germine (Rusagara, Lyumugabe, Busogoro, & Baudoin, 2012; Katawczik, 2008).

La reproducción asexual es a través de la formación de un esporangio. El esporangio libera zoosporas móviles en agua, que enquistan e infectan por la formación de un tubo germinal. Las zoosporas no se forman en el propio esporangio, sino en una vesícula fuera de ella. El esporangio se separa del resto del micelio por una pared transversal (figura 1.3). Las zoosporas sólo son liberadas en condiciones de humedad. La producción de esporangios puede ser estimulada por iones Mg, K, y Ca. Además los exudados de raíces y semillas en germinación tienen un efecto estimulante sobre la germinación de los esporangios y en el crecimiento del micelio (Katawczik, 2008).

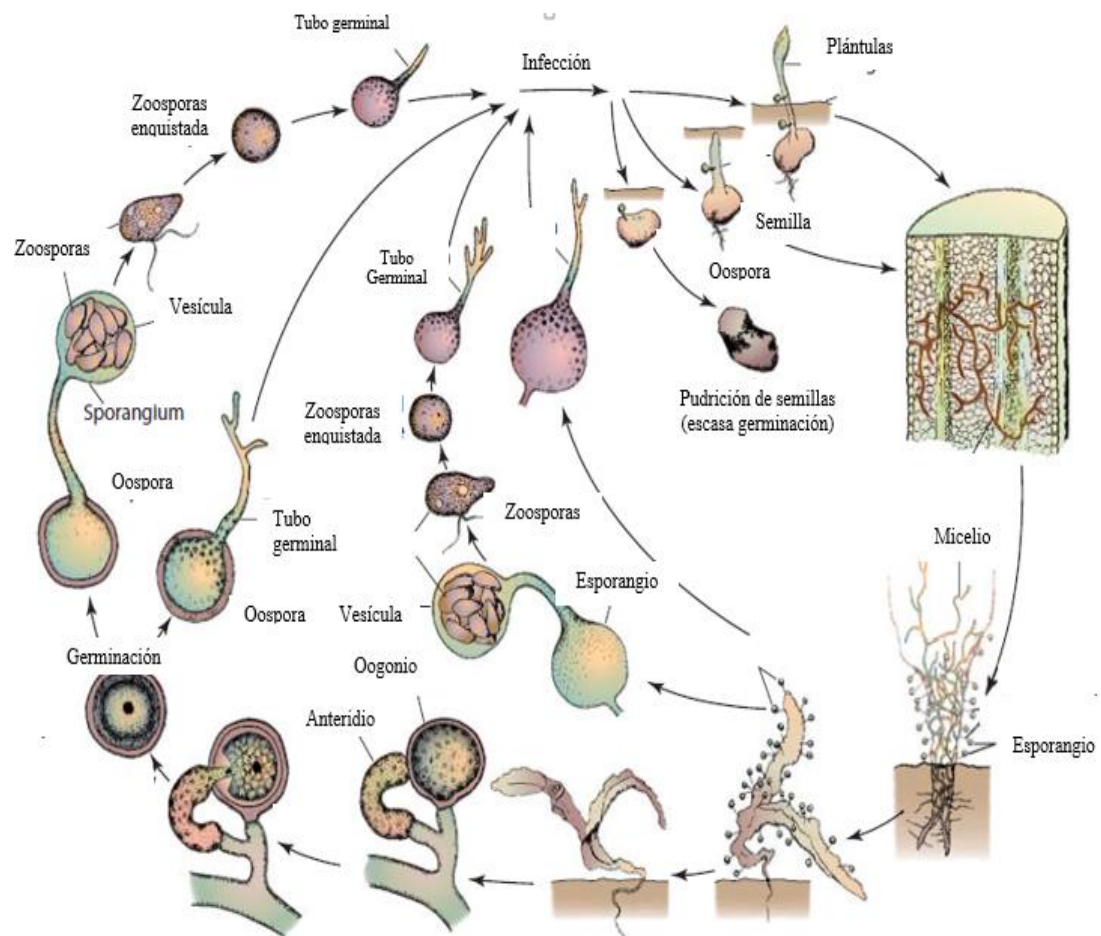


Figura 1.3: Ciclo de enfermedad del ahogamiento y podredumbre de la raíz/semillas producidos por *Pythium* spp. (Agris 2005).

1.4.1.6. Caracterización molecular

La identificación a nivel de especie del género *Pythium* por muchas décadas se basó en criterios morfológicos. Los criterios de identificación morfológicos están basados en características que varían dependiendo de las condiciones de los medios de cultivos utilizados y del aislado en sí (Kageyama, Nakashima, Kajihara, Suga, & Nelson, 2005).

Las características morfológicas son a menudo variables en *Pythium* y algunas especies pueden o no desarrollar estructuras reproductivas sobre un

medio de cultivo; lo cual ha sido motivo de identificaciones erróneas entre especies con morfología similar (Al-Sa'di, Drenth, Deadman, de Cock, & Aitken, 2007).

En el caso de *P. irregulare*, existe una gran variación en las estructuras, cuyo apareamiento puede verse influenciado por el medio ambiente y la planta huésped. Esta variación morfológica ocurre entre aislados individuales por lo que además de las características morfológicas, métodos moleculares son utilizados en estudios taxonómicos. Estudios morfológicos y moleculares han revelado que *Pythium irregulare* exhibe una gran variabilidad morfológica y genética, además parece ser parte de un complejo de especies (Garzón *et al.*, 2007).

Las regiones ITS (espaciador transcrito interno) que incluye ITS1, 5.8 S e ITS2 del ADN ribosomal (ADNr) han sido usadas como herramientas taxonómicas para la identificación molecular de oomicetes. Los datos de secuencia de estas regiones se han utilizado para identificar y clasificar especies de *Pythium* e investigar relaciones filogenéticas. Otros estudios han incorporado el análisis de genes que codifican para proteínas metabólicas y estructurales como citocromo oxidasa mitocondrial II (coxII, incluyendo el espaciador entre coxI y coxII) y β -tubulina (Villa *et al.*, 2006). Ambos son genes conservados que permiten establecer comparaciones entre miembros de un mismo género pero suficientemente variables para contribuir a la diferenciación de especies (Moralejo *et al.*, 2008).

Distintos grupos en *P. irregulare* se han reportado gracias a la aplicación de análisis moleculares como ITS-RFLP, RAPD, AFLP fingerprinting y secuenciación de ADN, lo que ha llevado a sugerir que la especie *Pythium irregulare* ha sufrido un proceso de especiación (Garzón *et al.*, 2007).

1.4.1.7. Identificación de *Pythium irregulare* por métodos tradicionales y moleculares

Introducción

La biología molecular y las técnicas de secuenciación han revolucionado la taxonomía de los oomicetes y han cambiado los conocimientos sobre las relaciones existentes entre ellos. En la actualidad, se tiende a la elaboración de esquemas de clasificación basados en criterios filogenéticos, mientras se descubren nuevas especies de oomicetos patógenos. El conocimiento y aplicación de estos avances permite la mejora en el tratamiento de muchas enfermedades, las cuales dependen de la correcta identificación de las especies implicadas. Sin la habilidad para identificar los organismos responsables de enfermedades en cultivos específicos, resulta imposible entender la epidemiología de una enfermedad o manejarla. Todos los aspectos de epidemiología de enfermedad de una planta, de propagación de la enfermedad, de estimación de las relaciones campo-pérdida/enfermedad, requieren la capacidad de identificar el patógeno con certeza (McCartney *et al.*, 2003).

Detección mediante técnicas tradicionales

El desarrollo de metodologías que permitan la diferenciación de microorganismos resulta importante para su identificación. Estas metodologías se basan tradicionalmente en la observación de los síntomas que presentan los hospederos, en la observación macroscópica o microscópica de los microorganismos, en la observación de las estructuras reproductivas, en el desarrollo del patógeno en un medio de cultivo específico o de la tinción sobre el tejido infectado (Rodríguez *et al.*, 2009).

La identificación de un microorganismo (patógeno) basada solo en la sintomatología suele dar origen a errores durante el diagnóstico de una enfermedad dado que dos patógenos diferentes pueden causar los mismos síntomas en el mismo hospedero, el mismo patógeno puede provocar diferentes síntomas en diferentes hospederos y los síntomas pueden alterarse dependiendo del medio ambiente donde se desarrolla el proceso patológico (Rodríguez *et al.*, 2009).

La detección de un microorganismo basada en la observación macroscópica o microscópica es más eficiente que la anterior. Es un procedimiento simple que permite diferenciar organismos morfológicamente distintos. Suele ser un procedimiento lento, y tedioso, debido a que requiere aislar el microorganismo y ponerlo a crecer en un medio de cultivo hasta que se alcance una etapa de crecimiento adecuada que permita su identificación. En ocasiones la estructura de un microorganismo puede cambiar debido a la influencia del ambiente y contenido nutricional del medio de cultivo y se

requiere de experiencia para diferenciar especies (Rodríguez *et al.*, 2009). Los métodos convencionales de identificación de patógenos de plantas se basan en la interpretación visual de síntomas, en el aislamiento, cultivo e identificación de patógenos en laboratorio, cuya precisión es dependiente de la experiencia de la persona que hace el diagnóstico (McCartney *et al.*, 2003).

Estas metodologías, a pesar de contribuir a la identificación, pueden ser deficientes por lo que conviene, además, desarrollar técnicas moleculares de detección e identificación, las cuales se basan en el análisis de los ácidos nucleicos extraídos de los microorganismos en forma directa o bien de una muestra que contenga el microorganismo (Rodríguez *et al.*, 2009).

Detección mediante técnicas moleculares

La detección de microorganismos mediante técnicas moleculares es trascendental pues para el ser humano algunos microorganismos suelen ser benéficos para su salud y economía o pueden resultar ser patógenos no solo para el ser humano sino también para animales y plantas económicamente importantes (Rodríguez *et al.*, 2009).

Existen muchas estrategias y metodologías para complementar los enfoques tradicionales usados, las cuales incluyen técnicas específicas de ADN. Algunas ventajas de las técnicas moleculares son: su especificidad, sensibilidad, rapidez, reproducibilidad y que pueden ser automatizadas. El uso de técnicas moleculares ha permitido identificar nuevos microorganismos, en los cuales no había sido posible su cultivo e identificación por técnicas

tradicionales. Sin embargo, los enfoques taxonómicos basados únicamente en datos moleculares, pueden presentar algunos problemas debido a que los bancos de datos pueden estar incompletos y a que las secuencias de ADN se encuentren mal identificadas (Pfennig & Magalhães, 2008; Rodríguez *et al.*, 2009). Aun así, las técnicas basadas en ácidos nucleicos, que dependen de las variaciones en las secuencias nucleotídicas de ADN del patógeno, se han convertido en las preferidas debido a su rapidez, sensibilidad, especificidad y reproducibilidad (Narayanasamy, 2013).

Técnicas basadas en PCR

La reacción en cadena de la polimerasa (PCR) o la amplificación de un segmento de ADN específico de un microorganismo, es la técnica más utilizada para la detección de microorganismos (Rodríguez *et al.*, 2009) y es una técnica disponible para la detección de patógenos de plantas. Una PCR permite la amplificación de millones de copias de un ADN específico gracias a ciclos repetidos de denaturación, polimerización y elongación a diferentes temperaturas mediante la utilización de oligonucleótidos específicos (primers), trifosfatos desoxirribonucleótidos (dNTPs), y una ADN *Taq* polimerasa en un tampón adecuado. Los fragmentos de ADN amplificados son visualizados por electroforesis en geles de agarosa teñidos con una molécula capaz de intercalarse en la doble hélice de ADN. La banda específica de ADN con el tamaño esperado indica la presencia del patógeno en la muestra (Capote , Pastrana , Aguado , & Sánchez-Torres, 2012).

La precisión y reproducibilidad de los ensayos PCR dependen de la experiencia técnica del analista. Si los iniciadores (primers) no son específicos, si las condiciones de la PCR no son óptimas o si hay contaminación de la muestra con inhibidores de la enzima polimerasa, se pueden amplificar productos no específicos que guían a la detección de falsos negativos. La flexibilidad, automatización, rapidez y confiabilidad de las técnicas moleculares basadas en PCR son las técnicas moleculares con más aceptación en la actualidad para la detección de microorganismos en general (Rodríguez *et al.*, 2009; Ramsden, 2009).

Secuenciación

Las características morfológicas no siempre son suficientes para identificar un patógeno. Uno de los métodos más directos de identificación consiste en la amplificación por PCR de un gen diana con primers universales, seguido por la secuenciación y la comparación con bases de datos de acceso público. El aumento de la capacidad de secuenciación y la disminución de los costos han permitido la acumulación de un gran número de secuencias de oomicetes en bases de datos de acceso público. Las secuencias de genes seleccionados han sido ampliamente utilizadas para la identificación de patógenos específicos y para el desarrollo de métodos de diagnóstico basados en la secuencia, pues el análisis comparativo de secuencias es básico en la identificación de muchos patógenos. Es una estrategia rápida, adecuada, pero es dependiente de la calidad de las secuencias existentes en bases de datos (Capote *et al.*, 2012).

El uso de bases de datos para identificar organismos basados en la similitud de ADN puede tener algunos riesgos que incluyen secuencias erróneas e incompletas, secuencias relacionadas con organismos mal identificados, la imposibilidad de cambiar o actualizar los datos fácilmente y problemas asociados con la definición de límites de las especies. Todo ello conduce a una interpretación errónea de los resultados (Tsui *et al.*, 2013; Ramsden, 2009).

1.4.2. Análisis y ensamblaje de secuencias

El proceso de secuenciación determina el orden de los nucleótidos (A, C, G y T) y produce fragmentos a partir de los cuales se debe reconstruir la secuencia del fragmento analizado, ya sea un genoma, un cromosoma artificial bacteriano (BAC) o un producto de PCR (Cañizares & Blanca, 2013).

El ensamblaje de una secuencia se refiere a la reconstrucción de la secuencia del fragmento analizado a partir de la información de un grupo de secuencias. El proceso de ensamblaje consta de tres fases: limpieza de secuencias, alineamiento de secuencias y ensamblaje (Cañizares & Blanca, 2013).

1.4.2.1 Limpieza de las secuencias de ADN

La limpieza de secuencias consiste en eliminar o bloquear diferentes regiones que contengan bases con mala calidad de lectura, bases que pertenezcan a vectores o adaptadores, bases que estén altamente repetidas

(secuencias simples o elementos transponibles). Estas regiones son errores accidentales durante la secuenciación que interfieren con el alineamiento de secuencias y es necesario eliminarlas (Cañizares & Blanca, 2013).

1.4.2.2 Alineamiento y ensamblaje de secuencias

Consiste en alinear todas las secuencias para identificar las secuencias que se solapan entre sí. El alinear se refiere a una forma de comparar dos o más secuencias o cadenas de ácidos nucleicos, o estructuras primarias proteicas para resaltar sus zonas de similitud. Con el alineamiento se logran identificar pares o grupos de secuencias con secuencias similares y con el ensamblaje se obtienen secuencias consenso que representan cada conjunto de secuencias o contig (Cañizares & Blanca, 2013).

1.4.3 Filogenia molecular

La filogenética es una rama de la biología que se centra en la inferencia de relaciones evolutivas entre grupos de especies. Tradicionalmente, los estudios se han basado en caracteres morfológicos y fisiológicos, pero con la obtención de las primeras secuencias de genes y proteínas la filogenia molecular tomó auge. Esta disciplina se basa en la comparación de secuencias biológicas para llevar a cabo una clasificación jerárquica entre organismos (Mina & Romero-Campero, 2012).

1.4.3.1. Análisis filogenéticos (Mina & Romero-Campero, 2012)

La determinación de la relación jerárquica entre especies y la estimación del tiempo de divergencia entre ellas en base a secuencias biológicas como proteínas o ADN se llevan a cabo para estudios filogenéticos. Los estudios filogenéticos normalmente se dividen en cinco fases:

Fase 1: Obtención de las secuencias biológicas a ser analizadas.

Fase 2: Construcción de un alineamiento múltiple de las secuencias biológicas seleccionadas.

Fase 3: Selección de los modelos de sustitución y de los modelos estadísticos de la evolución molecular más apropiados para las secuencias correspondientes.

Fase 4: Construcción de los árboles filogenéticos basados en los correspondientes alineamientos múltiples y modelos de sustitución.

Fase 5: Evaluación estadística de los árboles filogenéticos.

Selección de las secuencias biológicas

Dos enfoques generales se utilizan para seleccionar secuencias de ADN diana para su utilización en el diagnóstico y análisis filogenéticos. Uno es el desarrollo de un método que utiliza genes conocidos conservados, que son comunes a todos los organismos, que tienen una secuencia de variación útil que pueden ser explotada, el otro enfoque consiste en buscar partes

aleatorias del genoma de un organismo para encontrar regiones que muestran la especificidad requerida (Capote *et al.*, 2012).

Genes con alta variabilidad se están utilizando ampliamente para desarrollar métodos de diagnóstico para oomicetos. Los genes β -tubulina son quizás los segundos objetivos más comunes, después de las secuencias ITS ribosómicas, para el desarrollo de diagnóstico en hongos y oomicetos. Aunque la base de datos de secuencias de β -tubulina no es tan grande pueden ser útiles cuando la variación de las secuencias ITS no es adecuada para la elaboración de un diagnóstico de cada taxón (McCartney *et al.*, 2003).

Alineamiento múltiple

La alineación implica la determinación de las posiciones a lo largo de las secuencias de ADN o de secuencias de proteínas que derivan de una posición ancestral común (Doyle & Gaut, 2000). Un alineamiento múltiple permite revelar relaciones evolutivas entre distintas secuencias y constituye una base para la inferencia filogenética. Permite extrapolar la información de secuencias conocidas hacia las secuencias desconocidas, identificar y comparar “motivos” conservados, y ayudan a diseñar iniciadores o primers para PCR (Poot, 2013). Es una metodología poderosa que subyace en muchos análisis bioinformáticos. Se puede utilizar en conjunción con una amplia gama de análisis evolutivos y funcionales para identificar desde relaciones evolutivas, motivos funcionales y dominios de interacción. Un alineamiento múltiple de secuencias fiable identifica residuos homólogos entre

las secuencias relacionadas, datos que son absolutamente esenciales para los análisis filogenéticos (Guttman, Provar, Austin, Hug, & Price , 2014).

Para la construcción de un alineamiento múltiple se necesita un sistema de puntuación; típicamente se usa la suma de pares (sum of pairs scoring) como puntaje estándar. Los métodos usados para la construcción de un alineamiento múltiple son generalmente métodos progresivos basados en perfiles, conjunto de secuencias alienadas, o métodos iterativos. De ambos métodos los iterativos corrigen muchos problemas realineando repetidamente subgrupos y luego realineando estos subgrupos en una alineación global (Guttman *et al.*, 2014).

Algoritmo Muscle para el alineamiento múltiple (Edgar, 2004)

Muscle (multiple sequence alignment by log-expectation) es un método basado en iteraciones. Se utiliza como algoritmo en la creación de alineamientos múltiples de secuencias. Utiliza dos medidas de distancia para un par de secuencias, una es la distancia *kmer* (para un par no alineado) y la distancia *Kimura* (para un par alineado). Un *kmer* es una subsecuencia contigua de longitud *k*, también conocido como *k-tupla*. Secuencias relacionadas tienden a tener más *kmers* en común de lo esperado al azar. Esta medida no requiere una alineación, proporcionando una ventaja de velocidad significativa. A partir de un par alineado de secuencias, se calcula la identidad par y se convierte en una estimación de la distancia, aplicando la corrección *Kimura* para múltiples sustituciones en un solo sitio. Se generan matrices de distancia que se agrupan mediante el algoritmo para análisis

filogenético: UPGMA (Unweighted Pair Group Method with Arithmetic Mean; Método de pares de grupos no ponderado con media aritmética) con la expectativa de estimar un árbol evolutivo más fiable (Edgar, 2004).

Hay tres etapas principales usadas para el alineamiento con el algoritmo Muscle las cuales se detallan a continuación y se pueden observar en la figura 1.4.

Etapas 1. Inicio progresivo: El objetivo de la primera etapa es producir un alineamiento múltiple, destacando la velocidad por encima de la exactitud.

1.1. La distancia *kmer* se calcula para cada par de secuencias de entrada, generando la matriz de distancia D1.

1.2. La matriz D1 se agrupan por UPGMA, produciendo un árbol binario (Árbol1).

1.3. Una aproximación progresiva se construye siguiendo el orden de ramificación del árbol 1. Se produce un alineamiento múltiple de todas las secuencias de entrada: MSA1 (Alineamiento múltiples de secuencias 1), en la raíz.

Etapas 2. Mejora progresiva: La principal fuente de error en la etapa uno es la medida de distancia de aproximación *kmer*, lo que se traduce en un árbol subóptimo. Muscle estima nuevamente el árbol usando la distancia *Kimura*, que es más precisa y que requiere una alineación.

- 2.1. La distancia *Kimura* para cada par de secuencias de entrada se calcula a partir del MSA1, generando la matriz de distancia D2.
- 2.2. La matriz D2 se agrupa por UPGMA, produciendo un árbol binario (Árbol 2)
- 2.3. Un alineamiento progresivo se conduce produciendo un segundo alineamiento múltiple (MSA2).

Etapas 3. Refinamiento

- 3.1. Del árbol se selecciona un punto con el fin de disminuir la distancia desde la raíz.
 - 3.2. El árbol 2 se divide en dos subárboles mediante la supresión del punto elegido. El perfil de la alineación múltiple en cada subárbol se calcula.
 - 3.3. Un nuevo alineamiento múltiple se produce por re-alineación de los dos perfiles.
 - 3.4. Si la puntuación SP (suma de pares) mejora, la nueva alineación se mantiene, de lo contrario, se descarta.
- Los pasos 3.1 a 3.4 se repiten hasta la convergencia o hasta que se alcanza un límite definido por el usuario.

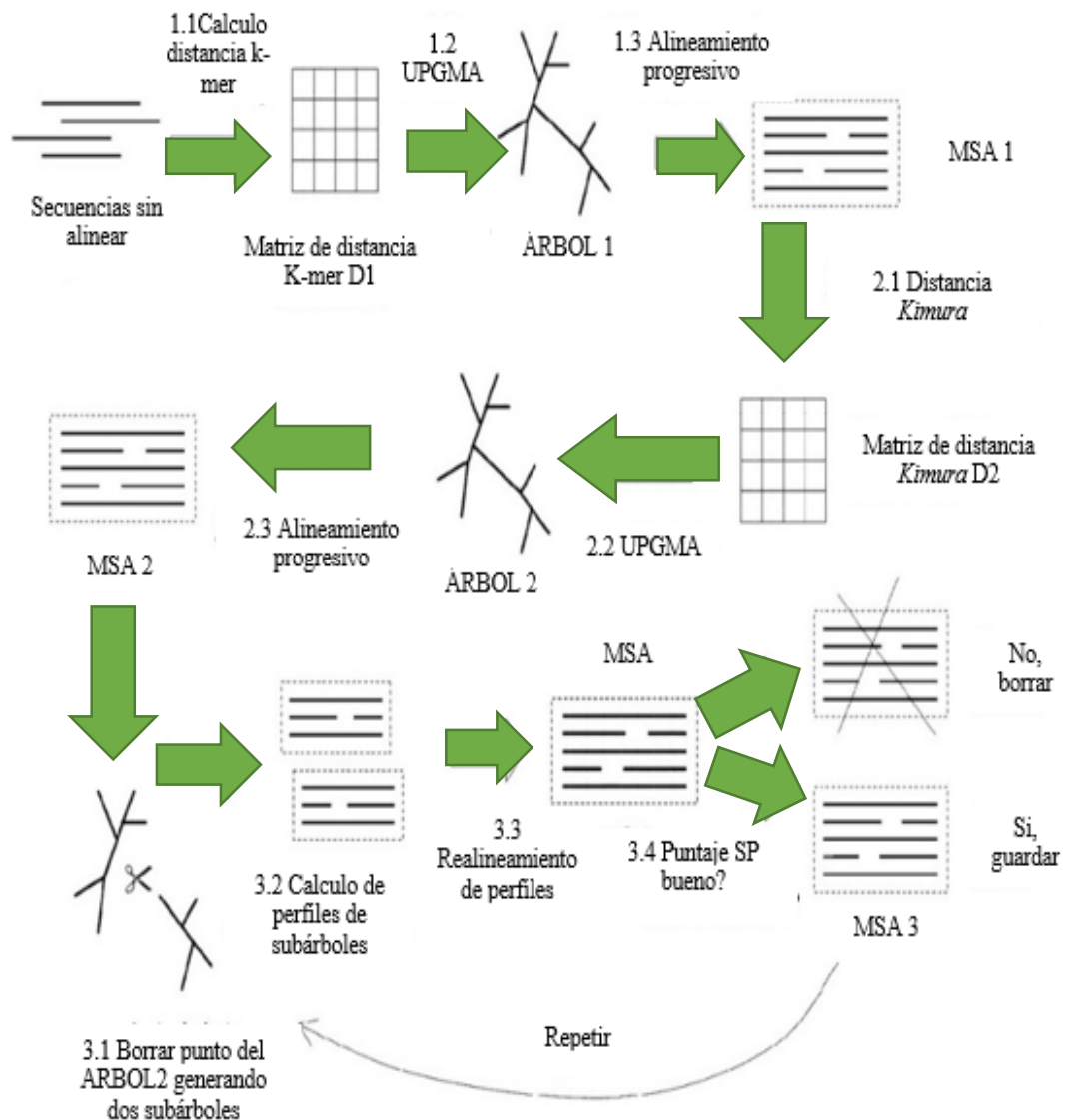


Figura 1.4: Diagrama que resume el flujo de etapas que utiliza el algoritmo Muscle para el alineamiento múltiple de secuencias (Edgar, 2004).

Modelos de sustitución

La distancia genética entre dos secuencias homólogas se define como el número de sustituciones acumuladas entre ellos, que divergieron de un ancestro común (Mina & Romero, 2012).

Un análisis filogenético se basa en la elección correcta del modelo de sustitución adecuado. Existen varios modelos de sustitución, los más comunes son:

- Modelo JC69 (Jukes y Cantor, 1969): Es el modelo más simple que propone una corrección del número de sustituciones observadas. Se asume que la probabilidad de mutación de un nucleótido por otro es independiente de la posición de dicho nucleótido y las secuencias de nucleótidos en sí.
- Modelo K80 (Kimura, 1980): Kimura propone un refinamiento del modelo Jukes-Cantor. Toma en cuenta una mayor probabilidad de observar transiciones que transversiones.
- Modelo T92 (Tamura, 1992): Es un método matemático simple desarrollado para estimar el número de sustituciones de nucleótidos por sitio entre dos secuencias de ADN, mediante la extensión del método de dos parámetros de Kimura (1980) para el caso que exista un sesgo en el contenido G + C.
- Modelo TN93 (Tamura y Nei, 1993): Distingue entre los dos tipos diferentes de transiciones, por ejemplo, (A \leftrightarrow G) se le permite tener una tasa diferente a (C \leftrightarrow T). Se asume que las transversiones ocurren en la misma proporción, pero a esa tasa, se le permite ser diferente de las dos tasas de transiciones.

Construcción de árboles filogenéticos

El principal objetivo de un estudio de filogenia molecular es el establecimiento de relaciones jerárquicas entre las especies, u otros taxones, en función de su relación evolutiva y del tiempo estimado de divergencia entre las especies, lo cual es a través de árboles filogenéticos (Mina & Romero, 2012).

Un árbol filogenético es una relación jerárquica en la cual se infiere una hipótesis sobre el pasado común de un grupo de organismos a través de cambios detectados en un carácter estudiado. A un grupo que contiene a todos los descendientes de un mismo antepasado se lo denomina monofilético (Caujape-Castells, Pedrola-Monfort, & Membrives, 2003).

Los métodos para la construcción de un árbol se pueden clasificar en dos categorías: métodos fenéticos basados en distancias y métodos cladísticos basados en caracteres (Mina & Romero, 2012).

Métodos basados en distancias

A partir de una matriz de distancia se pueden utilizar distintos métodos estadísticos para reconstruir un árbol. Los más utilizados son el UPGMA y el Neighbor-Joining. Son métodos comúnmente utilizados cuando se pretende obtener una clasificación jerárquica. Son rápidos, permiten utilizar pocos recursos computacionales, lo cual permite analizar secuencias muy largas y se pueden completar en tiempos relativamente cortos (Cañizares & Blanca, 2013; Lemey, Salemi & Vandame, 2009).

Métodos basados en caracteres: Máxima verosimilitud

Los métodos basados en caracteres tienen como principal ventaja que hacen uso de toda la información evolutiva conocida, es decir, de las sustituciones individuales entre las secuencias, para determinar las relaciones ancestrales más probables (Mina & Romero, 2012). Uno de los métodos basados en caracteres es el de máxima verosimilitud el cual se detalla a continuación:

La estimación de la máxima verosimilitud es un método con enfoque estadístico usado en ciencias biológicas. Es abreviado comúnmente como EMV (MLE, por las siglas en inglés de Maximum Likelihood Estimation) y fue propuesto por Fisher (1922; Aldrich, 1997). Su aplicación en la filogenética implica estimar la probabilidad de observar que un conjunto particular de secuencias alineadas, tengan un modelo de sustitución de nucleótidos y den una topología de árbol. Dada una topología de árbol, las longitudes de las ramas se estiman de acuerdo con el modelo y los parámetros elegidos, y la probabilidad de obtener estados especiales de caracteres se estima para cada posición de nucleótido. El método de máxima verosimilitud emplea modelos probabilísticos para seleccionar el mejor árbol; aquel que tenga la más alta probabilidad (verosimilitud) de reflejar el proceso evolutivo real (Doyle & Gaunt, 2000; Lemey *et al.*, 2009).

Definición de máxima verosimilitud (Opperdoes, 1997)

La probabilidad de un conjunto de datos, D , es la probabilidad de los datos, dada una hipótesis. La hipótesis por lo general viene en forma de diferentes parámetros. La probabilidad, L , de un conjunto de datos, D , se denota de la siguiente manera:

$$L = P(D/\theta)$$

Según Posada y Buckley (2014), se define la verosimilitud como la probabilidad de los datos (D) dado un modelo de evolución (M), un vector del modelo de k parámetros $\theta = (\theta_1, \theta_2, \dots, \theta_k)$, una topología de árbol (τ), y un vector para la longitud de ramas S , $u = (u_1, u_2, \dots, u_k)$:

$$L = P(D/M\theta\tau v)$$

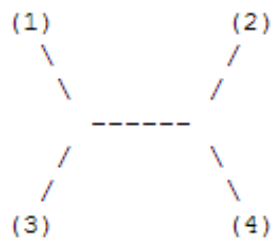
Cálculo de un árbol filogenético con máxima verosimilitud (Opperdoes, 1997)

Cuando se infiere un árbol filogenético utilizando máxima verosimilitud, se busca la probabilidad de que, a partir de un árbol y un modelo de evolución dados (hipótesis), se observe un alineamiento determinado (datos).

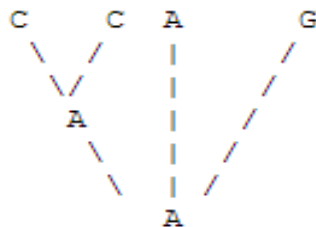
Si se supone que se tiene secuencias de nucleótidos alineados para cuatro taxones como se ilustra en la imagen:

	1	jN							
(1)	A	G	G	C	U	C	C	A	AA
(2)	A	G	G	U	U	C	G	A	AA
(3)	A	G	C	C	C	A	G	A	A A
(4)	A	U	U	U	C	G	G	A	A C

Y que se quiere evaluar la probabilidad de obtener un árbol sin raíz representado por los nucleótidos del sitio j en la secuencia mostrada a continuación:



Dado que la mayoría de modelos son reversibles en el tiempo, la probabilidad de un árbol es generalmente independiente de la posición de la raíz. Por ello, es conveniente crear un nodo interno arbitrario como se hace en la imagen de a continuación:



Bajo el supuesto de que los sitios de los nucleótidos evolucionan independientemente (modelo de evolución de Markov), se puede calcular la probabilidad para cada sitio por separado y combinar la probabilidad en un

valor total hacia el final. Para calcular la probabilidad del sitio j , se consideran todos los escenarios posibles por los cuales los nucleótidos presentes podrían haber evolucionado. Por lo tanto, la probabilidad para un sitio en particular es la suma de las probabilidades de cada posible reconstrucción de los estados ancestrales, dado un cierto modelo de sustitución de bases.

La probabilidad para el árbol completo, entonces es producto de la probabilidad en cada sitio:

$$\ln L = \ln L(1) \times \ln L(2) \dots \times \ln L(N) = \sum_{j=1}^N \ln L(j)$$

Dado que las probabilidades individuales son números extremadamente pequeños es conveniente reportar la probabilidad de todo el árbol como el logaritmo de verosimilitud:

$$\ln L = \ln L(1) + \ln L(2) \dots + \ln L(N) = \sum_{j=1}^N \ln L(j)$$

Evaluación estadística de un árbol filogenético

La verificación de la fiabilidad de un árbol construido se puede realizar con un análisis llamado “bootstrap”, que es la forma más simple de probar si el conjunto de datos apoyan estadísticamente al árbol obtenido. Es un método estadístico que estima las distribuciones por creación repetida y es, en sí, un alineamiento artificial que conserva las características del original. A cada alineamiento aleatorio se le aplica el algoritmo a evaluar, obteniendo un árbol. Después se construye un árbol consenso con todos los árboles obtenidos, en

donde, el porcentaje de veces que una ramificación aparece es el valor de “bootstrap”. Valores de “bootstrap” > 70% suelen tomarse como suficientemente robustos estadísticamente y equivalen a un nivel de significatividad $p < 0.05$ (Santamaria, 2006).

1.4.4 Estudios taxonómicos y filogenéticos del género *Pythium*

El mejoramiento de la especificidad de un ensayo de diagnóstico requiere de la combinación de múltiples regiones de diagnóstico. Muchos autores han seguido la estrategia multi-locus, porque permite la detección de la variación genética (Capote *et al.*, 2012).

La principal región que se ha usado en análisis filogenéticos del género *Pythium* corresponde al ADN ribosomal nuclear (ADNr), incluyendo los separadores intergénicos transcritos 1 y 2 (internal transcribed spacers, ITS1 and ITS2). Estudios han secuenciado y analizado la subunidad pequeña ribosomal nuclear y los dominios D1 y D2 de la subunidad grande ribosomal nuclear. Estos estudios han demostrado que *Pythium* fue filogenéticamente aislado y se diferencia de otros géneros de oomicetes. La región ITS es importante para estudios taxonómicos, principalmente porque los primers desarrollados por White *et al.* (1990) amplifican una región altamente variable en todos los taxones, incluidos los oomicetes (Lévesque & de Cock, 2004).

Distintos grupos en *P. irregulare* se han reportado por análisis ITS-RFLP (Polimorfismos de longitud de fragmentos de restricción), RAPD (Amplificación aleatoria de ADN polimórfico), evaluaciones de diversidad

genética basados en isoenzimas, AFLP fingerprinting (Polimorfismos en la longitud de fragmentos amplificados) y secuenciación de ADN de la región ITS (Garzón *et al.*, 2007). Sin embargo, como más datos de secuencias se han recogido a partir de una gama más amplia de aislados, la utilidad de loci alterativos para la identificación exacta de especies es cada vez mayor. En la actualidad, genes con alta variabilidad son cada vez más ampliamente utilizados como dianas para el desarrollo de métodos de diagnóstico y para la caracterización genética de patógenos. Algunos de estos genes son genes nucleares como β -tubulina, factor traductor de elongación 1 alfa (TEF 1 α), calmodulina, genes de avirulencia y genes mitocondriales como COX I multicopia y COX II y su región intergénica (Capote *et al.*, 2012; Martin & Tooley, 2003).

El gen β -tubulina es de importancia en estudios filogenéticos porque es altamente conservado, pero con variabilidad suficiente para distinguir las especies de algunos organismos. Codifica para una de las dos familias conservadas de tubulinas, subunidades de los microtúbulos, que forman el citoesqueleto (provee el soporte interno de las células), husos mitóticos y flagelos de las células eucariotas. Es un gen útil en la reconstrucción de relaciones filogenéticas entre los hongos y organismos similares en todos los niveles (Villa *et al.*, 2006). Aunque la base de datos de secuencias β -tubulina no es tan grande como es el caso de las secuencias ITS ribosómicas, pueden ser particularmente útiles cuando la variación de la secuencia ITS no es adecuada en la elaboración de un diagnóstico de cada taxón (McCartney *et al.*, 2003).

1.5. Hipótesis.

Las cepas de *Pythium irregulare* sensu lato de Long Island, NY, son monofiléticas de acuerdo a las secuencias de β -tubulina.

CAPÍTULO 2. MATERIALES Y MÉTODOS

2.1. Participantes

2.1.1 Instituciones

Esta investigación es parte del convenio entre el laboratorio de Microbiología, perteneciente al Departamento de Ciencias de la Vida y de la Agricultura de la Universidad de las Fuerzas Armadas-ESPE y el Departamento de Entomología y Patología de Plantas de la Universidad Estatal de Oklahoma (OSU), USA.

2.1.2. Colaboradores científicos

MC. Alma Koch Kaiser Directora de la tesis.

Ing.-Mat. Pedro Romero Codirector de la tesis.

Ph.D. Carla Garzón Responsable del laboratorio de patógenos de
suelo en OSU.

2.2. Zona de estudio

La caracterización molecular se realizó en el laboratorio de patógenos del suelo de la Universidad Estatal de Oklahoma (OSU) del Departamento de Entomología y Fitopatología de OSU.

2.3. Periodo de investigación

La investigación se desarrolló en un periodo de ocho meses, desde marzo del 2013 a noviembre del 2013.

2.4. Análisis de datos

Después de la caracterización y secuenciación de las cepas de *Pythium irregulare* sensu lato en base a las secuencias del gen β -tubulina, el análisis de datos involucró la limpieza, alineamiento y ensamblaje de las secuencias nucleotídicas con el programa BioNumerics versión 7.1 (Applied Maths). BioNumerics es un software que permite conducir análisis de secuencias y que cuenta con una herramienta de ensamblaje que permite la creación de secuencias consenso.

Las secuencias consenso creadas con BioNumerics junto con seis secuencias nucleotídicas de otras especies de *Pythium* para el gen β -tubulina, se compilaron en un archivo formato FASTA.

Las secuencias de estas seis especies de *Pythium* fueron obtenidas de la base de datos Nucleotide del Centro Nacional de Información Biotecnológica de los Estados Unidos (NCBI, sitio web:<http://www.ncbi.nlm.nih.gov/Genbank/index.html>).

Se utilizó la herramienta de alineamiento local Blastn (Basic local alignment search tool) de la NCBI, para comparar y homologar las secuencias ensambladas con secuencias que se encuentren en la base de datos. La

comparación permitió elegir algunas de las secuencias homólogas que se incluyeron en el análisis filogenético.

El análisis filogenético se hizo con el software bioinformático MEGA (Molecular Evolutionary Genetics Analysis) versión 5.2 (Tamura, Peterson, Peterson, Stecher, Nei & Kumar, 2013). El análisis incluyó las fases de alineamiento múltiple con el algoritmo Muscle, la elección del mejor modelo de sustitución de nucleótidos bajo el análisis estadístico de máxima verosimilitud, la construcción del árbol filogenético con el método de máxima verosimilitud y la verificación de la fiabilidad del árbol construido mediante un análisis “bootstrap” con 1000 repeticiones.

2.5. Procedimiento

2.5.1. Material de trabajo

Las 120 cepas de *Pythium irregulare* sensu lato utilizadas en este estudio (anexo A), fueron recolectadas de diferentes huéspedes y de diferentes invernaderos ubicados en Long Island, New York por la M.S. Margery Daughtrey (Centro de Horticultura de Long Island, Departamento de Fitopatología y Biología de Plantas y Microbios, Universidad de Cornell), y sometidas al laboratorio de patógenos de suelo de OSU para su caracterización genética. Estas cepas fueron previamente identificadas morfológicamente por los científicos del Centro de Horticultura de Long Island (estado de New York), y fueron recibidas en OSU en forma de micelio liofilizado.

2.5.1.1 Limpieza de muestras

Las muestras se recibieron en el laboratorio de patógenos del suelo de OSU, en tubos de 2mL que contenían el micelio congelado de cada una. Cuando hubo la necesidad de cultivar cepas de *P. irregulare* para obtener ADN, se sembraron en medio PARP (anexo B), luego se traspasaron al medio sólido agar agua (WA) y, por último, se colocó un pedazo de WA en matraces que contenían el medio de cultivo líquido PDB (Caldo papa dextrosa); otros pequeños pedazos de agar con micelio crecido se transfirieron a agua desionizada autoclavada para preservar al microorganismo (Garrido, 2013).

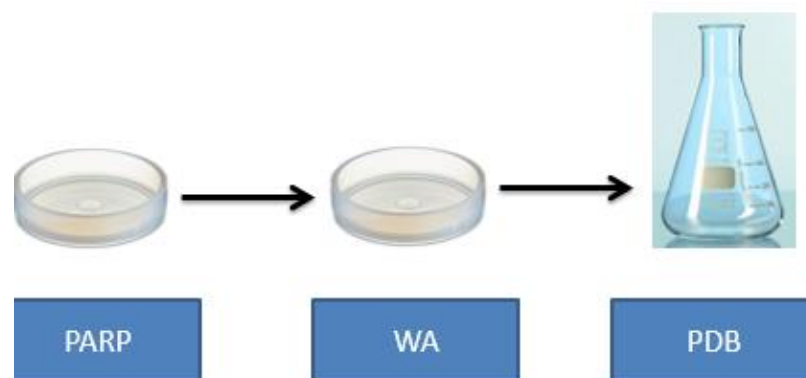


Figura 2.1: Esquema del proceso de limpieza de cepas de *Pythium irregulare* en estudio.

2.5.1.2 Liofilización

Sobre los tubos que contenían los micelios cubiertos con papel parafilm, se hicieron pequeños huecos con ayuda de una aguja. Todos los tubos se colocaron dentro de un frasco de liofilización y se insertaron en el equipo. Las muestras se liofilizaron durante 24 horas (Garrido, 2013).

2.5.2. Caracterización molecular

2.5.2.1 Extracción de ADN mediante DNeasy Plant Mini kit de QIAGEN

A partir de las cepas liofilizadas se procedió a la extracción de ADN. Se empleó el protocolo de DNeasy Plant Mini kit, descrito por el fabricante QIAGEN en el 2000 (anexo C).

2.5.2.2 Determinación de la concentración y pureza de las muestras de ADN mediante espectrofotometría (NANODROP 1000)

Se utilizó el procedimiento del manual del equipo NANODROP 1000 para la determinación de la concentración y la pureza de las muestras de ADN extraídas. Se inició la programación del equipo en el modo ADN y con unidades de cuantificación en ng/ μ L. Se limpió el lector con papel absorbente y con una micropipeta se colocó 1.5 μ L de buffer de elución AE, como blanco. Después se colocó 1.5 μ L de cada una de las muestras y se realizó la medición registrando los valores de concentración de ADN (ng/ μ L) y la relación de absorbancias a 260 y 280 nm (260/280).

La relación 260/280 se utilizó para evaluar la pureza del ADN. El ADN posee máximos de absorción alrededor de 260 nm y las proteínas a 280 nm, una relación aproximada entre 1.8 y 2.1 es generalmente aceptada como "pura" para el ADN. Si la relación es apreciablemente inferior a este valor, esto puede ser indicativo de contaminación con proteínas, fenoles u otros compuestos; mientras que valores por encima de ese rango sugiere contaminación con RNA (Thermo Fisher Scientific Inc, 2008).

2.5.2.3 Dilución del ADN

El ADN extraído y cuantificado se diluyó en agua estéril a una concentración de 10 ng/μL en un volumen final de 30 μL, con la siguiente formula:

$$V1 = \frac{V2 \times C2}{C1}$$

Donde V2 es el volumen final de dilución (30 μL), C2 es la concentración final (10 ng/μL), C1 es la concentración de ADN obtenida por espectrofotometría y V1 es el volumen de ADN en ng/μL que se tomó para llegar a la concentración esperada.

2.5.2.4 Reacción en cadena de la polimerasa (PCR convencional)

Una porción del gen β-tubulina fue amplificado con el par de primers βtub_F1A y βtub_R1A (tabla 2.1). Los primers se utilizaron a una concentración de 5ng/μL.

Tabla 2.1: Nombre, secuencia, coordenadas, temperatura de fusión (Tm) y referencia de los primers utilizados en las reacciones de PCR y de secuenciación.

Nombre del primer	Secuencia (5-3´)	Coordenadas del primer	Tm (°C)	Referencia
βtub_F1A	GCCAAGTTCTGGGARGTSAT	52-71	56.1	Blair <i>et al</i> (2008)
βtub_R1A	CCTGGTACTGCTGGTAYTCMGA	1278-1258	58.1	Blair <i>et al</i> . (2008)

Las amplificaciones se realizaron en un volumen final de reacción de 25 μ L. Para cada reacción se preparó un master mix con GoTaq® Green Master Mix (Mx), 2X (Promega); que contiene en solución: ADN Taq polimerasa, dNTPs, MgCl₂ y tampones de reacción a condiciones óptimas para la amplificación de templates de ADN por PCR (tabla 2.2).

Tabla 2 2: Condiciones de reacción PCR para la amplificación del gen β -tubulina en un volumen final de 25 μ L.

Reactivos	Cantidad por reacción (μL)
Go Taq Green Mx	12.5
Primer β tub_F1A	1.25
Primer β tub_R1A	1.25
Agua	8.0
ADN	2.0
Volumen total.	25.0

En las reacciones se incluyeron controles negativos para controlar que no exista contaminación. El control negativo contenía todos los componentes de la PCR excepto el ADN que se reemplazó con agua desionizada.

La amplificación se realizó en un termociclador bajo las condiciones detalladas en la tabla 2.3.

Tabla 2.3: Ciclos de amplificación utilizados para este estudio

Fase	Temperatura	Tiempo	Ciclos
Denaturación inicial	95 °C	2 min.	1
Denaturación	94 °C	30 s.	35
Hibridación	59 °C	30 s.	35

Continúa...

Fase	Temperatura	Tiempo	Ciclos
Extensión	72 °C	1 min.	35
Extensión final	72 °C	5 min.	1
Espera final	4 °C	Hasta el análisis de los productos de PCR.	1

Preparación del gel de agarosa al 1%

El gel de agarosa se preparó con solución tampón TAE (Tris base, EDTA y ácido acético) 1X en un matraz. La solución se calentó por un minuto y medio aproximadamente en un microondas. Se dejó enfriar a temperatura ambiente y, cuando la temperatura fue tolerable al tacto en la superficie del matraz, se añadió 3 μ L de bromuro de etidio (Garrido, 2013).

Electroforesis

La visualización de los productos amplificados se hizo a través de su migración en geles de agarosa al 1%. La migración se llevó a cabo por electroforesis horizontal en tampón de corrida TAE 1X. Se utilizó un marcador de peso molecular de 100 pb (VWR cat. no. 95039-808). Se cargó directamente en el gel de agarosa 5 μ L de los productos de PCR, incluido el control negativo, además de 5 μ L del marcador DNA ladder 100bp. Se conectó la cámara de electroforesis a una fuente de poder y se dejó correr durante una hora a 90 voltios. Terminada la corrida, el gel se reveló mediante luz ultravioleta (UV) dentro de un transiluminador para su documentación fotográfica (Garrido, 2013).

2.5.3. Secuenciación

2.5.3.1 Preparación para la secuenciación

La cuantificación se realizó por electroforesis en geles de agarosa al 1% con el marcador DNA ladder 100bp (VWR).

Los productos de PCR se diluyeron a una concentración de 5 ng/ μ L en un volumen de 20 μ L y se prepararon para la secuenciación con el sistema de purificación enzimática: ExoSAP-IT PCR Product Clean-Up (Affymetrix), con el fin de remover dNTPs y primers indeseados que pueden interferir con la reacción de secuenciación. Se colocó directamente sobre las muestras de trabajo 2 μ L de la solución ExoSAP-IT y se incubó las muestras a 37 °C durante 15 min, seguido de un calentamiento a 80 °C durante 15 min. Las muestras se mantuvieron a 4 °C hasta su envío (Garrido, 2013).

2.5.3.2 Envío

Las reacciones de secuenciación fueron realizadas por el laboratorio central de análisis (Core Facility) de la Universidad Estatal de Oklahoma en el equipo 3730 DNA Analyzer (Applied Biosystems). Se envió 20 μ L de cada producto de PCR y 10 μ L de los primers sentido (Btub_F1A) y antisentido (Btub_R1A). Se hizo una doble secuenciación por muestra (Garrido, 2013).

2.5.4. Análisis de secuencias

2.5.4.1 Edición y registro de secuencias

Las secuencias de ADN obtenidas se limpiaron, alinearon y ensamblaron en el programa BioNumerics versión 7.01 de prueba y disponible en: <http://www.applied-maths.com/download/software>. Después de importar los archivos de las secuencias a analizar, se evaluó y verificó la calidad y la asignación correcta de las bases inspeccionando los cromatogramas de cada una de las secuencias. Las secuencias sentido y antisentido se alinearon y se crearon secuencias consenso utilizando los parámetros de alineamiento por defecto. Las secuencias consenso se guardaron en archivos formato FASTA. FASTA es un formato que comienza con un '>' seguido de un texto que describe la secuencia, y que luego describe a la secuencia real a partir de la siguiente línea (Guttman *et al.*, 2014).

2.5.4.2 Análisis de similitud y homología

Las secuencias editadas se compararon con las secuencias disponibles en la base de datos Nucleotide del NCBI (National Centre for Biotechnology Information) utilizando la aplicación de alineamiento local Blast (Basic local alignment search tool) disponible en: http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST_PROGRAMS=megaBlast&PAGE_TYPE=BlastSearchn

Se utilizó la opción para secuencias nucleotídicas y el algoritmo Megablast, con el fin de comparar las secuencias en estudio con otras

secuencias con altos niveles de similitud y que han sido previamente caracterizadas. Se usó los parámetros por defecto, y se incluyó el filtro para las regiones de baja complejidad (low complexity regions) con el fin de filtrar secuencias que no son informativas; dejando disponibles regiones biológicamente más interesantes para la coincidencia específica contra las secuencias de la bases de datos (Guttman *et al.*, 2014).

Para evaluar cuan buenos fueron los hits del Blast se evaluó (Guttman *et al.*, 2014):

- Identidad: El grado en que dos secuencias no varían.
- Cobertura de las secuencias de interés: Representa el porcentaje de similitud entre las secuencias de interés y las secuencias de la base de datos.
- La puntuación (Bit score) o la puntuación del alineamiento (S): Una medida precisa que se normaliza sobre el sistema de puntuación particular empleado.
- El valor E - valor esperado. Un valor de probabilidad que se basa en el número de alineaciones diferentes con puntuaciones al menos tan buenas como la que se observa, y que se espera se produzcan simplemente por casualidad.

2.5.5. Análisis filogenético

La relación filogenética de las muestras se determinó mediante el análisis de las secuencias obtenidas con los primers Btub_F1A y Btub_R1A. Además se incluyeron seis secuencias correspondientes a las diferentes especies del

complejo *Pythium irregulare* (*P. irregulare sensu stricto*, *P. cryptoirregulare*, *P. cylindrosporium*) para el gen β -tubulina y las secuencias de las especies *Pythium spinosum* y *Pythium sylvaticum* como grupos externos o “outgroups” (tabla 2.4).

Tabla 2.4: Lista de las especies de *Pythium* incluidas en el análisis filogenético.

Muestra	Número de acceso GenBank (β -tubulina)	Referencia
<i>Pythium sylvaticum</i> ^a	GU071880.1	Spies <i>et al.</i> (2013)
<i>Pythium spinosum</i> ^a	GU071881.1	Spies <i>et al.</i> (2013)
<i>Pythium cylindrosporium</i>	GU071877.1	Spies <i>et al.</i> (2013)
<i>Pythium cryptoirregulare</i>	GU071888.1	Spies <i>et al.</i> (2013)
<i>Pythium irregulare</i>	AB512837.1	Senda & Kageyama (2009)
<i>Pythium irregulare</i>	AB512837.1	Senda & Kageyama (2009)

^a Especies usadas como “outgroups”.

2.5.5.1 Alineamiento múltiple

El alineamiento múltiple se realizó con el software bioinformático MEGA (Molecular Evolutionary Genetics Analysis) versión 5.2 (Tamura *et al.*, 2013).

Se ingresaron en el programa un total de 123 secuencias de ADN, sin alinear y en formato FASTA. Se eligió Muscle como el algoritmo para la construcción del alineamiento múltiple bajo la configuración predeterminada. El alineamiento se hizo en base a codones por ser secuencias codificantes. Se hizo además un análisis de la composición nucleotídica de cada secuencia y de la frecuencia de pares de nucleótidos.

2.5.5.2 Selección del modelo de sustitución

La elección del modelo de sustitución de nucleótidos se hizo con la opción: *Find Best DNA/Protein Models (ML)* del menú *Models* del programa MEGA y bajo las opciones siguientes: método estadístico: máxima verosimilitud; tipo de sustitución: nucleótidos y tratamiento de datos faltantes/gaps: delección parcial, y en base al alineamiento múltiple

2.5.5.3 Construcción del árbol filogenético

El árbol filogenético se construyó con el método estadístico de máxima verosimilitud con el programa MEGA versión 5.2. La construcción se hizo con la opción: *Construct/Test Maximum Likelihood Tree* del menú *Phylogeny* del programa MEGA.

Se modificaron tres parámetros del cuadro de diálogo que se desplegó:

- Para el parámetro modelo/método se escogió el método que describió mejor el patrón de sustitución según la lista de modelos de la fase de elección del modelo de sustitución.
- Para el cambio entre sitios se escogió la mejor opción que se mostró en la lista de modelos de la fase de elección del modelo de sustitución.
- Para los Gap/tratamiento de datos desaparecidos, se escogió la opción: delección parcial, en la cual los sitios con datos faltantes son

removidos solo cuando es necesario. Esta opción retiene mayor información.

2.5.5.4 Evaluación estadística del árbol filogenético

La estimación de la fiabilidad del árbol construido se hizo mediante un análisis “bootstrap” con 1000 repeticiones con la opción incorporada en el programa MEGA versión 5.2, a partir de los cuales se elaboró un árbol consenso.

CAPÍTULO 3. RESULTADOS

3.1. Caracterización molecular

3.1.1. Extracción de ADN y cuantificación

Con el protocolo DNeasy Plant Mini kit de QIAGEN se extrajo el ADN genómico de las 120 cepas de *Pythium irregulare* sensu lato recibidas. El ADN de cada una de las muestras se cuantificó por espectrofotometría con el equipo Nanodrop 1000, dando como resultado un rango de concentración de 3-800 $\mu\text{L}/\text{ng}$ y una relación de absorbancias promedio (260/280) de 1,8. En el Anexo D se adjunta la tabla que detalla las concentraciones y relaciones de absorbancias 260/280 para cada una de las muestras.

3.1.2. Amplificación y visualización del gen β _tubulina

La cuantificación de los fragmentos de ADN amplificados se hizo con el marcador de peso molecular 100 bp DNA ladder (VWR cat. no. 95039-808). Este marcador tipo escalera contiene 12 fragmentos de ADN que van desde 100 hasta 3000 pb y que permite una fácil determinación del tamaño de las bandas (figura 3.1).

La reacción de PCR fue eficiente con la metodología que se empleó. Los primers $\beta\text{tub_F1A}$ y $\beta\text{tub_R1A}$ permitieron amplificar una parte del gen β _tubulina de las 120 muestras de ADN. Se observaron fragmentos de

aproximadamente 1200 pb. Los fragmentos se visualizaron en geles de agarosa al 1% p/v (figuras 3.2 a la 3.10)

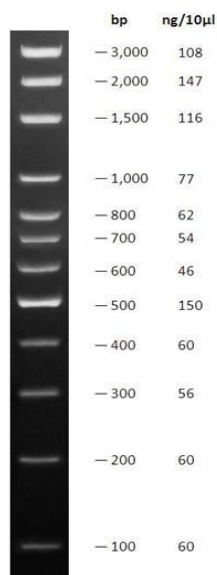


Figura 3.1: Fragmentos de ADN tipo escalera que contiene el marcador de peso molecular 100 bp DNA ladder (VWR cat. no. 95039-808) utilizado para medir el largo en pares de bases (bp) del ADN amplificado por PCR.

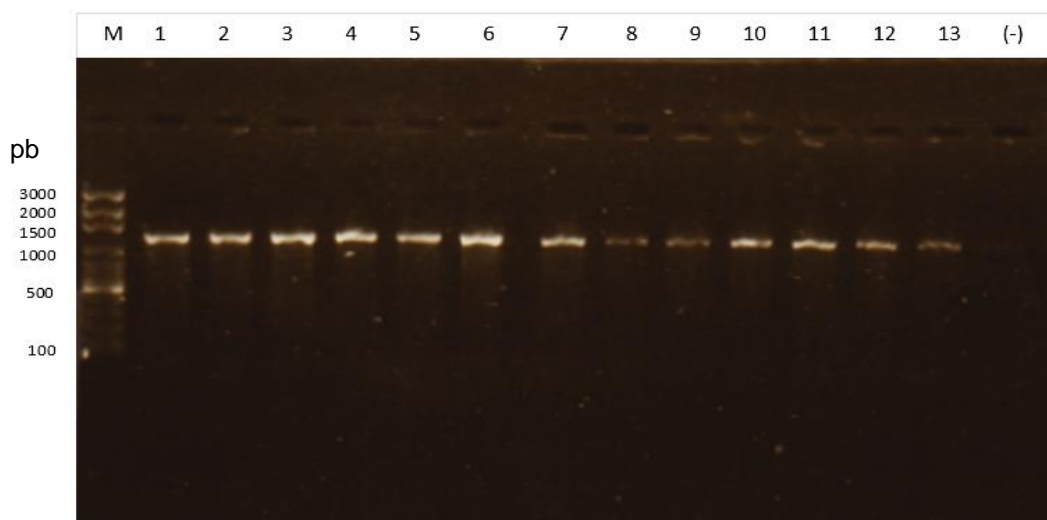


Figura 3.2: Visualización del producto de PCR del gen β -tubulina amplificado de 13 muestras de ADN (1:C8; 2:C9; 3:C14; 4:C20; 5:C24; 6: C32; 7: C37; 8:C38; 9:C39; 10:C41; 11:C42; 12:C43; 13:C44), en gel de agarosa al 1% p/v. M: Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de la reacción PCR. Las muestras se identifican por su código de estudio.

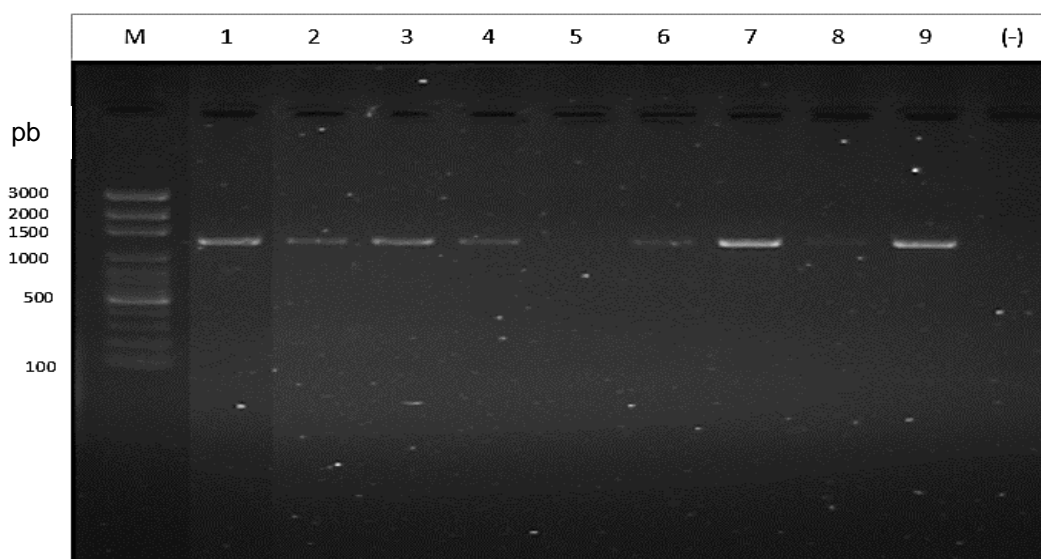


Figura 3.1: Visualización del producto de PCR del gen β -tubulina amplificado de siete muestras de ADN (1:C54; 2:C60; 3:C61; 4:C62; 6:C65; 7: C66; 9:C68), en gel de agarosa al 1% p/v. M: Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de la reacción PCR. Las muestras se identifican por su código de estudio.

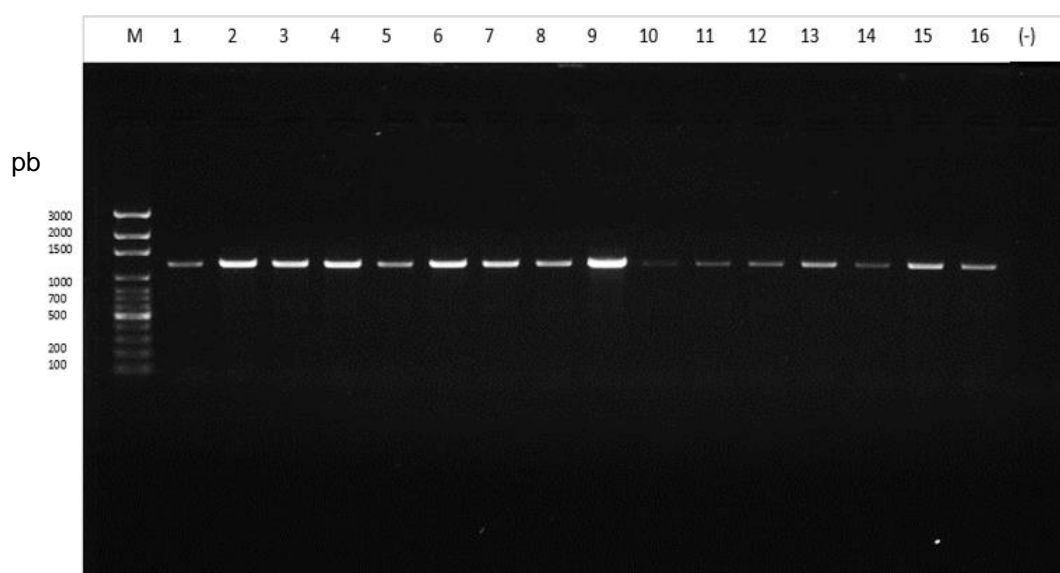


Figura 3.2: Visualización del producto de PCR del gen β -tubulina amplificado de 16 muestras de ADN. (1:C69; 2:C71; 3:C74; 4:C79; 5:C80; 6: C81; 7: C82; 8:C83; 9:C84; 10:C87; 11:C90; 12:C92; 13:C93; 14:C94; 15:C95; 16:C96), en gel de agarosa al 1% p/v. M: Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de la reacción PCR. Las muestras se identifican por su código de estudio.

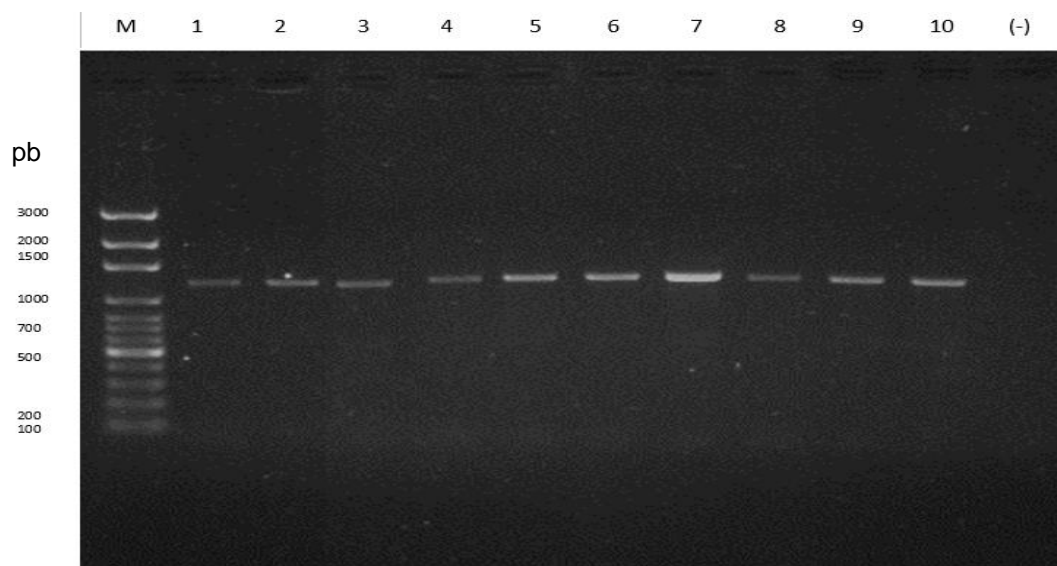


Figura 3.5: Visualización del producto de PCR del gen β -tubulina amplificado de 10 muestras de ADN. (1:C100; 2:C101; 3:C109; 4:C121; 5:C123; 6: C129; 7: C131; 8:C132; 9:C136; 10:C140), en gel de agarosa al 1% p/v. M: Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de la reacción PCR. Las muestras se identifican por su código de estudio.

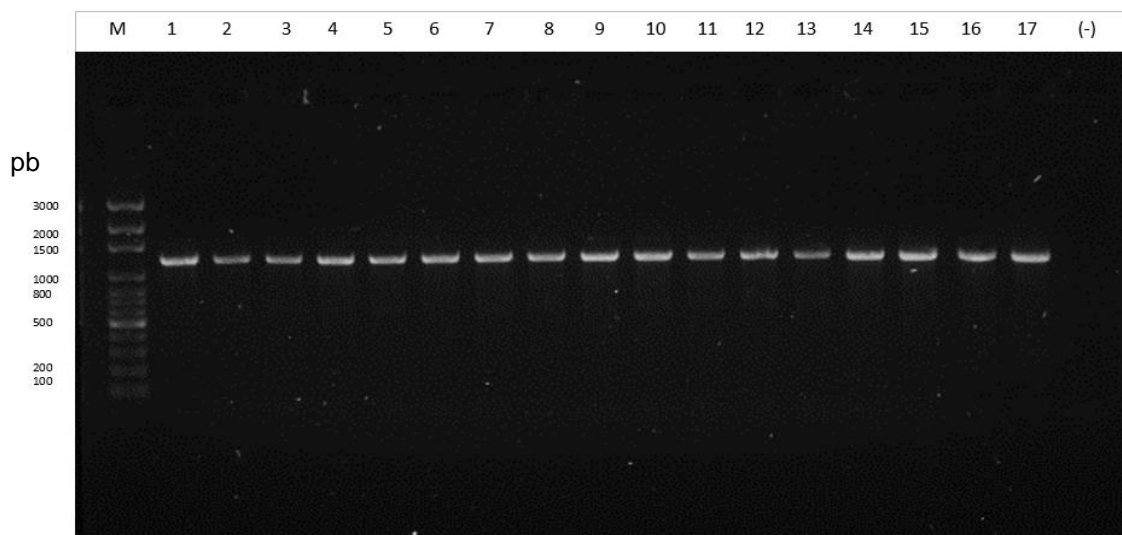


Figura 3.6: Visualización del producto de PCR del gen β -tubulina amplificado de 17 muestras de ADN. (1:C145; 2:C147; 3:C149; 4:C152; 5:C153; 6: C156; 7: C158; 8:C160; 9:C161; 10:C162; 11:C164; 12:C165; 13:C167; 14:C169; 15:C171; 16:C178; 17:C179), en gel de agarosa al 1% p/v. M: Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de la reacción PCR. Las muestras se identifican por su código de estudio.

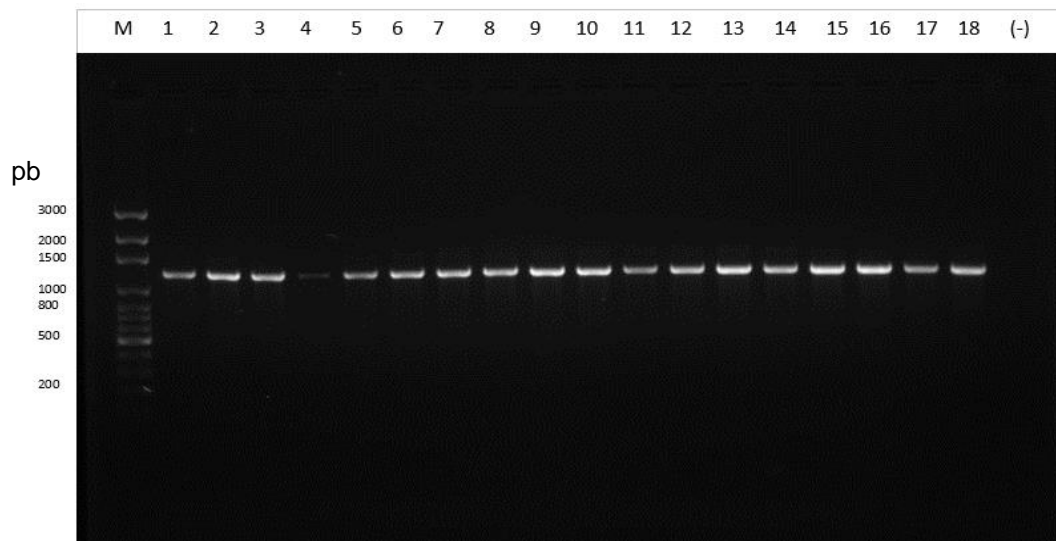


Figura 3.7: Visualización del producto de PCR del gen β -tubulina amplificado de 18 muestras de ADN. (1:C180; 2:C185; 3:C186; 4:C187; 5:C189; 6: C191; 7: C192; 8:C196; 9:C198; 10:C199; 11:C200; 12:C201; 13:C202; 14:C203; 15:C205; 16:C206; 17:C207; 18:C208), en gel de agarosa al 1% p/v. M: Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de la reacción PCR. Las muestras se identifican por su código de estudio.

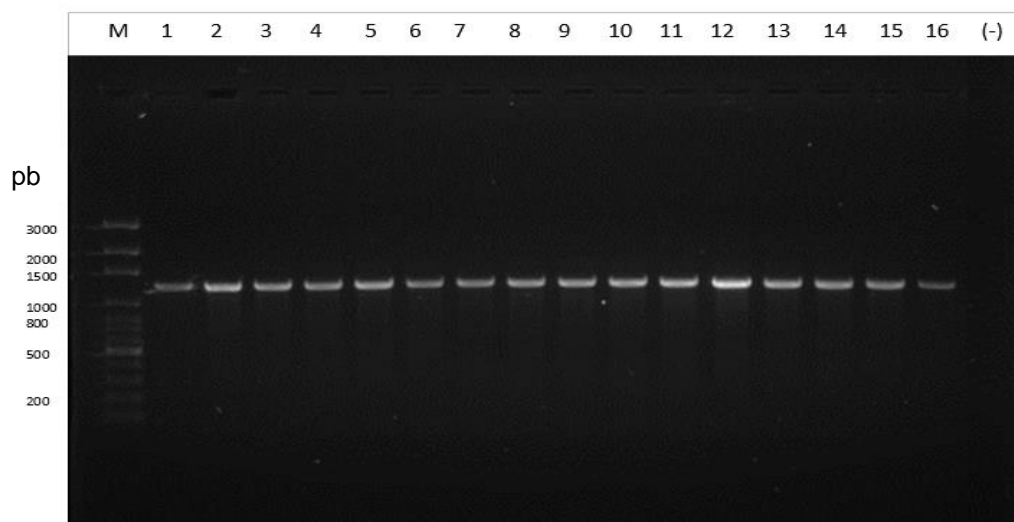


Figura 3.8: Visualización del producto de PCR del gen β -tubulina amplificado de 16 muestras de ADN (1:C209; 2:C210; 3:C211; 4:C212; 5:C213; 6: C214; 7: C216; 8:C220; 9:C223; 10:C228; 11:C230; 12:C232; 13:C234; 14:C235; 15:C236; 16:C188), en gel de agarosa al 1% p/v. M: Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de la reacción PCR. Las muestras se identifican por su código de estudio.

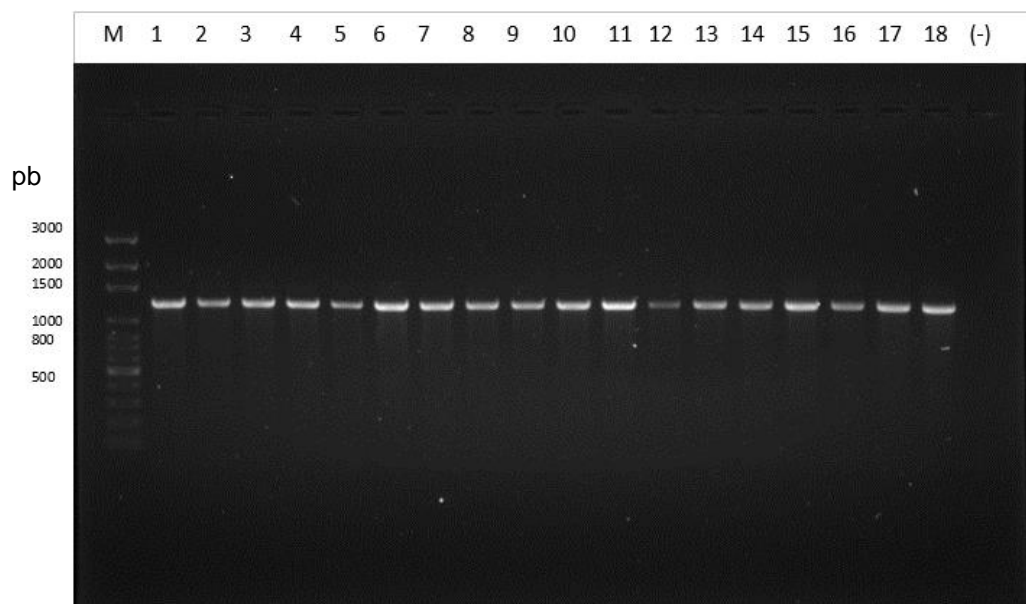


Figura 3.9: Visualización del producto de PCR del gen β -tubulina amplificado de 18 muestras de ADN. (1:C239; 2:C240; 3:C241; 4:C242; 5:C243; 6: C244; 7: C246; 8:C247; 9:C248; 10:C254; 11:C255; 12:C256; 13:C257; 14:C258; 15:C259; 16:C260; 17:C261; 18:C264), en gel de agarosa al 1% p/v. M: Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de la reacción PCR. Las muestras se identifican por su código de estudio.

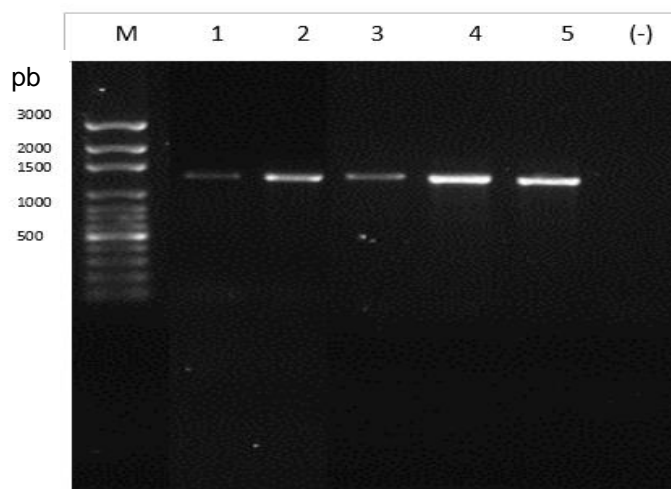


Figura 3.10: Visualización del producto de PCR del gen β -tubulina amplificado de cinco muestras de ADN (1:C117; 2:C120; 3:C135; 4:C174; 5:C197), en gel de agarosa al 1% p/v. M: Marcador de peso molecular 100bp DNA ladder. (-): Control negativo de la reacción PCR. Las muestras se identifican por su código de estudio.

3.1.3. Secuenciación del gen β -tubulina

Los productos amplificados del gen β -tubulina de las 120 cepas de *Pythium irregulare* sensu lato fueron cuantificados en geles de agarosa al 1% p/v, obteniendo una concentración aproximada de 75 ng/uL en la mayoría de muestras analizadas aunque no en todas (figuras 3.2- 3.10). Los productos fueron enviados a secuenciar en la Universidad Estatal de Oklahoma (OSU Core Facility) y los resultados fueron remitidos en formato AB1 y FASTA.

3.1.4. Edición y registro de las secuencias obtenidas

Las secuencias y los cromatogramas enviados por el OSU Core Facility fueron editados en el programa BioNumerics versión 7.01. La calidad de las secuencias se observaron en una escala de colores que iban del verde (calidad aceptable) al amarillo, naranja y rojo (calidad inaceptable). Los cromatogramas de las secuencias se inspeccionaron, así como las correctas asignaciones de bases y el alineamiento de las secuencias (figura 3.11).

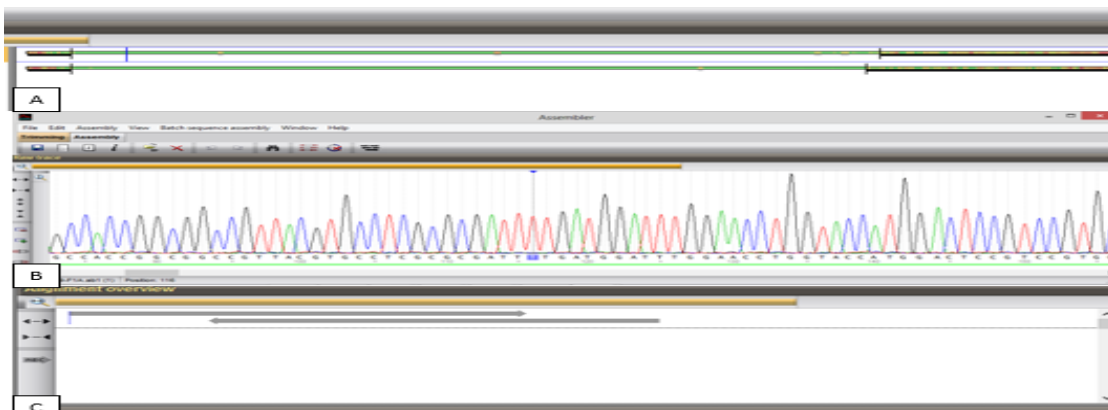


Figura 3.11: (A) Panel que muestra la calidad aceptable de una secuencia, cepa C109, (color verde) (B) Cromatograma obtenido para la secuencia del gen tubulina de la cepa C109 con el primer F1A. Los cuatro tipos de nucleótidos están diferenciados por colores (C: azul, G: negro, T: rojo, A: verde) (C) Panel que muestra el alineamiento de la secuencia de la cepa C109.

El resultado fue el alineamiento de las secuencias del primer sentido con las secuencias del primer antisentido de cada muestra. Se obtuvo un correcto alineamiento para 117 secuencias del total de 120 secuencias en estudio, con las que se crearon secuencias consenso en formato FASTA.

Tres secuencias, del total de 120, generaron inconvenientes durante su edición y alineamiento. Las secuencias, según su código de estudio, correspondieron a las cepas C117, C135 y C256. Los inconvenientes fueron:

La cepa C117 mostró una calidad de secuencias inaceptable por lo que el programa cortó partes de las secuencias (figura 3.12A). Por otra parte, aunque en el cromatograma se observaron picos separados y con una altura adecuada, hubo mucho ruido de fondo (figura 3.12B), además el alineamiento de las secuencias no fue posible (figura 3.12C).

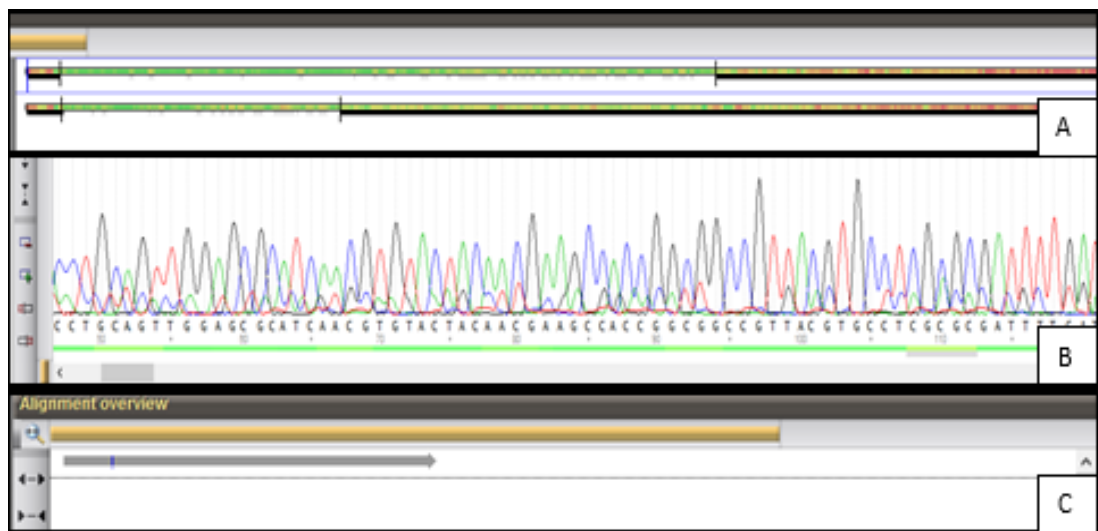


Figura 3.12: Paneles que muestran los problemas encontrados con la cepa C117. **(A)** Calidad inaceptable de las secuencia sentido y antisentido de la cepa **(B)** Ruido de fondo en el cromatograma perteneciente a la secuencia de la cepa C117. Los cuatro tipos de nucleótidos están diferenciados por colores (C: azul, G: negro, T: rojo, A: verde) **(C)** Alineamiento de las secuencias de la cepa C117.

El programa cortó partes de las secuencias pertenecientes a la cepa C135 porque mostró una calidad de secuencias inaceptable (figura 3.13A), en el cromatograma se observaron picos con mucho ruido de fondo (figura 3.13B) y por último el alineamiento de las secuencias no fue posible (figura 3.13C).

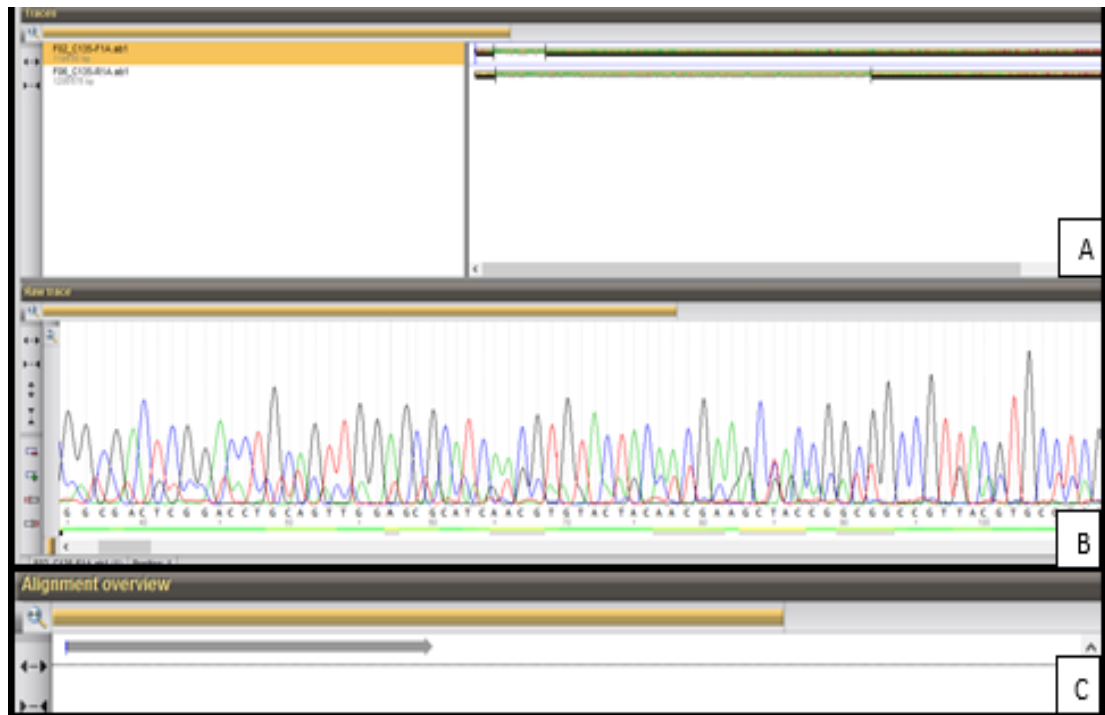


Figura 3.13: Paneles que muestran los problemas encontrados con la cepa C135. **(A)** Calidad inaceptable de las secuencia sentido y antisentido de la cepa **(B)** Ruido de fondo en el cromatograma perteneciente a la secuencia de la cepa C135. Los cuatro tipos de nucleótidos están diferenciados por colores (C: azul, G: negro, T: rojo, A: verde) **(C)** Alineamiento de las secuencias de la cepa C135.

La cepa C256 mostró una calidad de secuencias inaceptable, por lo que el programa cortó partes de las secuencias (figura 3.14A), en el cromatograma se observaron secuencias superpuestas (figura 3.14B) y por último el alineamiento de las secuencias no fue posible (figura 3.14C).

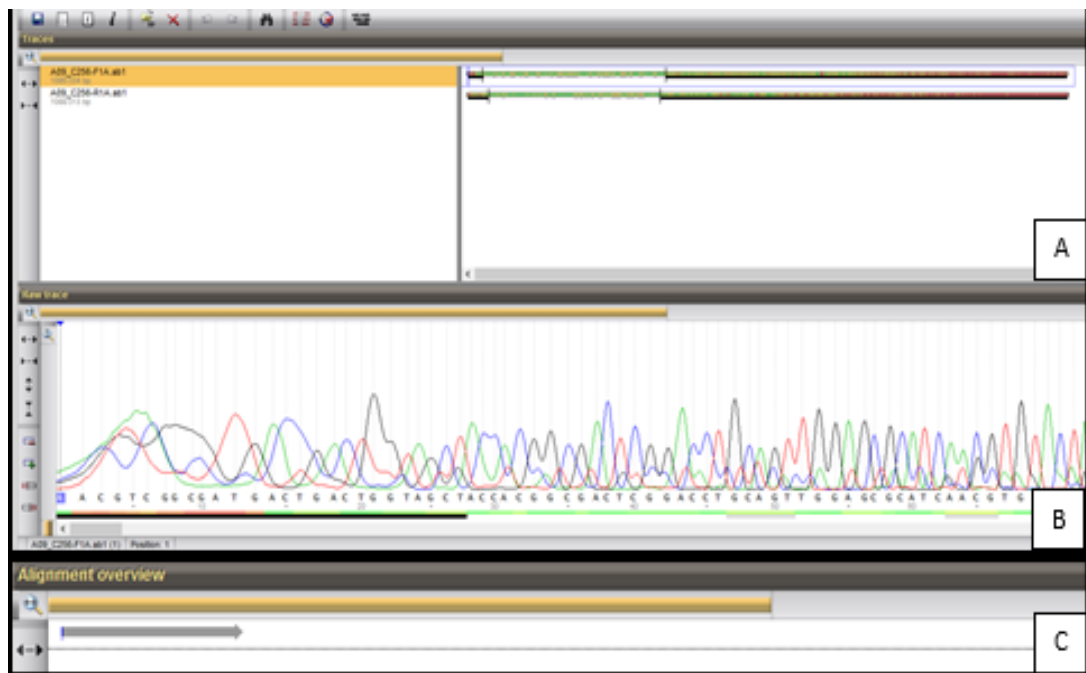


Figura 3.14: Paneles que muestran los problemas encontrados con la cepa C256. **(A)** Calidad inaceptable de las secuencia sentido y antisentido de la cepa **(B)** Secuencias superpuestas en el cromatograma perteneciente a la cepa C256. Los cuatro tipos de nucleótidos están diferenciados por colores (C: azul, G: negro, T: rojo, A: verde) **(C)** Alineamiento de las secuencias de la cepa C256.

Los problemas con las secuencias nucleotídicas de las tres cepas mencionadas hicieron que estas no formaran parte de los análisis posteriores reduciéndose el total de secuencias a 117.

3.1.5. Análisis de similitud y homología de secuencias

La comparación de las secuencias editadas con las secuencias existentes en la base de datos del programa Blast permitió encontrar secuencias similares significativas para cada secuencia; además permitió una identificación a nivel de género y de especie de las mismas. Blast es quizás

la herramienta más importante para la búsqueda de secuencias en bases de datos que son similares a las secuencias de interés.

De forma general, las secuencias asignadas por Blast más representativas a nivel de especie fueron: una cepa de *Pythium irregulare* (número de accesión GenBank: AB512837) y otra cepa de *Pythium irregulare* (número de accesión GenBank: AB512839). El porcentaje de identidad tuvo un valor del 99% para todas las cepas. El porcentaje de identidad representa la medida en que las secuencias de interés y la secuencia de la base de datos devuelta por Blast tienen los mismos residuos en las mismas posiciones, lo que permitió asegurar que el gen analizado pertenece taxonómicamente al género asignado *Pythium*.

El porcentaje de cobertura de las secuencias analizadas dio en todos los casos un valor del 100 %, lo que indicó que las secuencias de consulta tenían una similitud de un 100 % con las secuencias de la base de datos y que el alineamiento de los nucleótidos (aproximadamente 1100) coincidió totalmente con los del género asignado. El puntaje de alineamiento viene dado en el valor de Max score que fue diferente en todos los casos ya que este valor es dependiente de la longitud de las secuencias de interés. La significancia del alineamiento se midió con el valor E que fue 0.0, lo que hizo válida la comparación estadísticamente (figura 3.15).

A	Description	Max score	Total score	Query cover	E value	Ident	Accession
<input checked="" type="checkbox"/>	Pythium irregulare gene for beta-tubulin, partial cds, strain: PoinB	2017	2017	100%	0.0	99%	AB512839.1
<input checked="" type="checkbox"/>	Pythium irregulare gene for beta-tubulin, partial cds, strain: CBS 250.28	2006	2006	100%	0.0	99%	AB512837.1

Pythium irregulare gene for beta-tubulin, partial cds, strain: PoinB					
GenBank: AB512839.1					
FASTA Graphics					
Go to: <input type="checkbox"/>					
LOCUS	AB512839	1287 bp	DNA	linear	PLN 24-JUL-2012
		↑	↑	↑	↑
		Longitud	Fuente	Tipo	Grupo taxonómico del NCBI

Pythium irregulare gene for beta-tubulin, partial cds, strain: CBS 250.28					
GenBank: AB512837.1					
FASTA Graphics					
Go to: <input type="checkbox"/>					
LOCUS	AB512837	1287 bp	DNA	linear	PLN 24-JUL-2012
		↑	↑	↑	↑
		Longitud	Fuente	Tipo	Grupo taxonómico del NCBI

PLN: Secuencias de plantas, hongos y algas

Figura 3.15: (A) Resultados del alineamiento local de una de las secuencias problema (C109) con secuencias existentes en la base de datos de Blast (B) Información asociada a la secuencia *Pythium irregulare* (Número de accesoión GenBank: AB512839) (C) Información asociada a la secuencia *Pythium irregulare* (Número de accesoión GenBank: AB512837).

El anexo E muestra los resultados del análisis de Blastn de todas las secuencias de β -tubulina en estudio.

3.1.6. Análisis filogenético

3.1.6.1 Análisis estadístico de las secuencias

La tabla 3.1 muestra los datos promedio de longitud y composición de bases nucleotídicas de las secuencias en estudio.

Tabla 3.1: Resumen de los datos de composición de bases nucleotídicas y de longitud de las secuencias en estudio.

Cepa analizada	Composición de bases ^a (%)				Longitud (pb)
	T(U)	C	A	G	
Promedio de todas las cepas	22.9	29.0	21.2	26.9	1102.0

a: Las frecuencias están dadas en porcentajes.

Según la tabla 3.1 hay una longitud de secuencias promedio de 1102 pb y un mayor contenido de las bases Citocina (C) y Guanina (G), 29 % y 26.9 % respectivamente.

El anexo F muestra los datos de longitud y composición de bases nucleotídicas de todas las secuencias analizadas. De manera general no hay un comportamiento regular o longitud parecida en todas las secuencias.

Las tablas 3.2 y 3.3 muestran datos de pares de nucleótidos idénticos, transversiones, transiciones y las frecuencias de dinucleótidos en las secuencias de estudio. Los dinucleótidos más significativos fueron TT, CC, GG, AA y los menos significantes fueron TA y AT.

Tabla 3.2: Datos de frecuencias dinucleótidos idénticos, transversiones, transiciones promedio de las secuencias en estudio.

Datos	Número de frecuencias
Pares idénticos	573
Pares por transición (si)	148
Pares por transversión (sv)	297
R: si/sv.	0.50

Tabla 3.3: Pares de nucleótidos y frecuencias promedio de las secuencias en estudio.

Pares de nucleótidos	Número de frecuencias
TT	133
TC	37
TA	25
TG	38
CT	37
CC	174
CA	42

Continúa...

Pares de nucleótidos	Número de frecuencias
CG	42
AT	26
AC	42
AA	111
AG	37
GT	38
GC	42
GA	38
GG	155

3.1.6.2 Alineamiento múltiple

El resultado del alineamiento múltiple con el algoritmo Muscle, el método estadístico de máxima verosimilitud y bajo la configuración predeterminada en el programa MEGA fue el que se observa en la figura 3.16. La matriz de datos comprendió 1475 caracteres de los cuales 365 eran conservados, 1090 eran caracteres variables y 1040 tenían parsimonia informativa.

Por ser secuencias codificantes el alineamiento se hizo en base a codones, lo que asegura un enfoque mucho más realista porque evita la introducción de espacios en el alineamiento o gaps en posiciones que podrían dar lugar a cambios de marco en las secuencias reales (Hall, 2013).

Las columnas marcadas con un asterisco (*) indican los residuos con posiciones estructurales comunes y/o residuos ancestrales (Fassler & Cooper, 2013). Los cuatro tipos de nucleótidos están diferenciados por colores (T: rojo; G: Morado; C: celeste y A: verde), las columnas sin color indican la introducción de gaps, identificados por guiones (-), en el alineamiento con el fin de mantener la correspondencia entre sitios homólogos (figura 3.16).

3.1.6.3 Selección del modelo de sustitución.

El modelo que mejor describió la sustitución de nucleótidos fue el de Tamura de tres parámetros (T92) con un ajuste por la presencia de sitios invariables (+ I) (figura 3.17).

La bondad de ajuste de cada modelo se mide por el criterio de información Bayesiano (BIC) y el criterio de información de Akaike corregido (Tamura *et al.*, 2013). El modelo con el puntaje BIC (Criterio de Información Bayesiana) más bajo se considera que describe el mejor modelo de sustitución. Para cada modelo, el valor AICc (Criterio de Información de Akaike, corregido), el valor de máxima verosimilitud ($\ln L$), y el número de parámetros (incluyendo las longitudes de las ramas) se muestran (Nei & Kumar, 2000).

En la figura 3.17 cuando fue aplicable las estimaciones del parámetro gamma y / o la fracción estimada de sitios invariantes se observan para cada modelo. Los valores asumidos o estimados de transición / transversión se muestran para cada modelo. Les siguen las frecuencias de nucleótidos (f) y las tasas de sustituciones de bases (r) para cada par de nucleótidos.

El análisis involucró 123 secuencias nucleotídicas. Todas las posiciones con menos del 95% de cobertura fueron eliminadas. Hubo un total de 901 posiciones en el conjunto de datos finales.

Model	Parameters	BIC	AICc	lnL	(+I)	(+G)	R	f(A)	f(T)	f(C)	f(G)	r(AT)	r(AC)	r(AG)	r(TA)	r(TC)	r(TG)	r(CA)	r(CT)	r(CG)	r(GA)	r(GT)	r(GC)
T92+I	246	26686.367	24322.723	-11914.811	0.05	n/a	0.52	0.221	0.221	0.279	0.279	0.072	0.091	0.096	0.072	0.096	0.091	0.072	0.076	0.091	0.076	0.072	0.091
T92+G+I	247	26705.554	24332.307	-11918.598	0.05	200.00	0.52	0.221	0.221	0.279	0.279	0.072	0.091	0.096	0.072	0.096	0.091	0.072	0.076	0.091	0.076	0.072	0.091
JC+I	244	26707.188	24362.752	-11936.834	0.05	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
K2+I	245	26718.381	24364.341	-11936.624	0.05	n/a	0.52	0.250	0.250	0.250	0.250	0.082	0.082	0.085	0.082	0.085	0.082	0.082	0.085	0.082	0.085	0.082	0.082
JC+G+I	245	26721.164	24367.124	-11938.015	0.05	200.00	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
GTR+I	252	26724.340	24303.074	-11898.959	0.04	n/a	0.52	0.213	0.229	0.291	0.267	0.058	0.111	0.101	0.054	0.092	0.100	0.082	0.073	0.078	0.081	0.086	0.085
T92	245	26726.916	24372.875	-11940.891	n/a	n/a	0.52	0.221	0.221	0.279	0.279	0.073	0.091	0.096	0.073	0.096	0.091	0.073	0.076	0.091	0.076	0.073	0.091
K2+G+I	246	26732.353	24368.709	-11937.803	0.05	200.00	0.52	0.250	0.250	0.250	0.250	0.082	0.082	0.085	0.082	0.085	0.082	0.082	0.085	0.082	0.085	0.082	0.082
GTR+G+I	253	26738.175	24307.305	-11900.070	0.04	200.00	0.52	0.213	0.229	0.291	0.267	0.058	0.111	0.101	0.054	0.092	0.100	0.082	0.073	0.078	0.081	0.086	0.085
T92+G	246	26741.983	24378.339	-11942.619	n/a	200.00	0.52	0.221	0.221	0.279	0.279	0.073	0.091	0.096	0.073	0.096	0.091	0.073	0.076	0.091	0.076	0.073	0.091
JC	243	26743.545	24408.713	-11960.819	n/a	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
K2	244	26754.746	24410.310	-11960.613	n/a	n/a	0.52	0.250	0.250	0.250	0.250	0.082	0.082	0.085	0.082	0.085	0.082	0.082	0.085	0.082	0.085	0.082	0.082
GTR	251	26756.785	24345.123	-11920.988	n/a	n/a	0.52	0.213	0.229	0.291	0.267	0.058	0.113	0.103	0.054	0.090	0.098	0.083	0.071	0.078	0.082	0.084	0.085
GTR+G	252	26768.833	24347.566	-11921.205	n/a	200.00	0.52	0.213	0.229	0.291	0.267	0.058	0.113	0.103	0.054	0.090	0.098	0.083	0.071	0.078	0.082	0.084	0.085
K2+G	245	26776.261	24422.221	-11965.564	n/a	200.00	0.52	0.250	0.250	0.250	0.250	0.082	0.082	0.085	0.082	0.085	0.082	0.082	0.085	0.082	0.085	0.082	0.082
JC+G	244	26862.681	24518.245	-12014.580	n/a	200.00	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
HKY+I	248	30381.896	27999.044	-13750.962	0.06	n/a	0.58	0.213	0.229	0.291	0.267	0.072	0.092	0.098	0.067	0.107	0.084	0.067	0.084	0.084	0.078	0.072	0.092
TN93+I	249	30386.583	27994.128	-13747.499	0.06	n/a	0.56	0.213	0.229	0.291	0.267	0.073	0.093	0.099	0.068	0.103	0.085	0.068	0.081	0.085	0.079	0.073	0.093

Figura 3.17: Mejores modelos de sustitución y evolución para las secuencias en estudio según el método estadístico de máxima verosimilitud.

3.1.6.4. Construcción y evaluación estadística del árbol filogenético

El árbol filogenético se construyó en base a las secuencias del gen β -tubulina, con el método estadístico de máxima verosimilitud, en base al modelo de Tamura de tres parámetros (T92) y asumiendo que una cierta fracción de los sitios eran evolutivamente invariables (4.5221% de los sitios) (Tamura *et al.*, 2013).

El árbol consenso bootstrap inferido con 1000 repeticiones representa la historia evolutiva de los taxones analizados. El porcentaje de árboles replicados en el que los taxones se agruparon en el test bootstrap se muestran sobre las ramas (Felsenstein, 1985).

La topología del árbol condensado obtenido en base al gen β -tubulina, indicó la formación de tres grupos principales (I, II y III). En el grupo I se incluyen 73 taxones, algunos de ellos agrupados con un buen soporte y separados por pequeñas ramas internas, además sobresalen ramas que fueron condensadas resultando en politomías. En el grupo II se incluyen 45 taxones, en este grupo sobresalen también ramas condensadas que resultaron en politomías. El grupo III incluye tres taxones agrupados y con un buen soporte (figura 3.18). El enraizamiento del árbol se consiguió incluyendo como grupos externos a secuencias de *Pythium spinosum* (GU071881.1) y *Pythium sylvaticum* (GU071880.1). Las relaciones de los tres grupos recibieron un soporte bootstrap moderado-alto (63-100%) en el análisis de máxima verosimilitud.

También se pudo observar que de las cuatro secuencias referencia de *Pythium* para el gen β -tubulina, tres de ellas: *Pythium irregulare* (AB512837.1), *Pythium irregulare* (AB512839.1) y *Pythium cylindrosporium* (GU071877.1) se agruparon en el grupo II y la secuencia de *Pythium cryptoirregulare* (GU071888.1) se agrupó en el grupo I.

El árbol consenso que se obtuvo provee información para pensar que las cepas de *Pythium irregulare* sensu lato tienen un ancestro en común, a pesar de no mostrar un porcentaje de ocurrencia mayor al 50% en todas sus ramas. El enraizamiento permitió observar que las cepas de *Pythium irregulare* sensu lato son monofiléticas en base a las secuencias del gen β -tubulina.

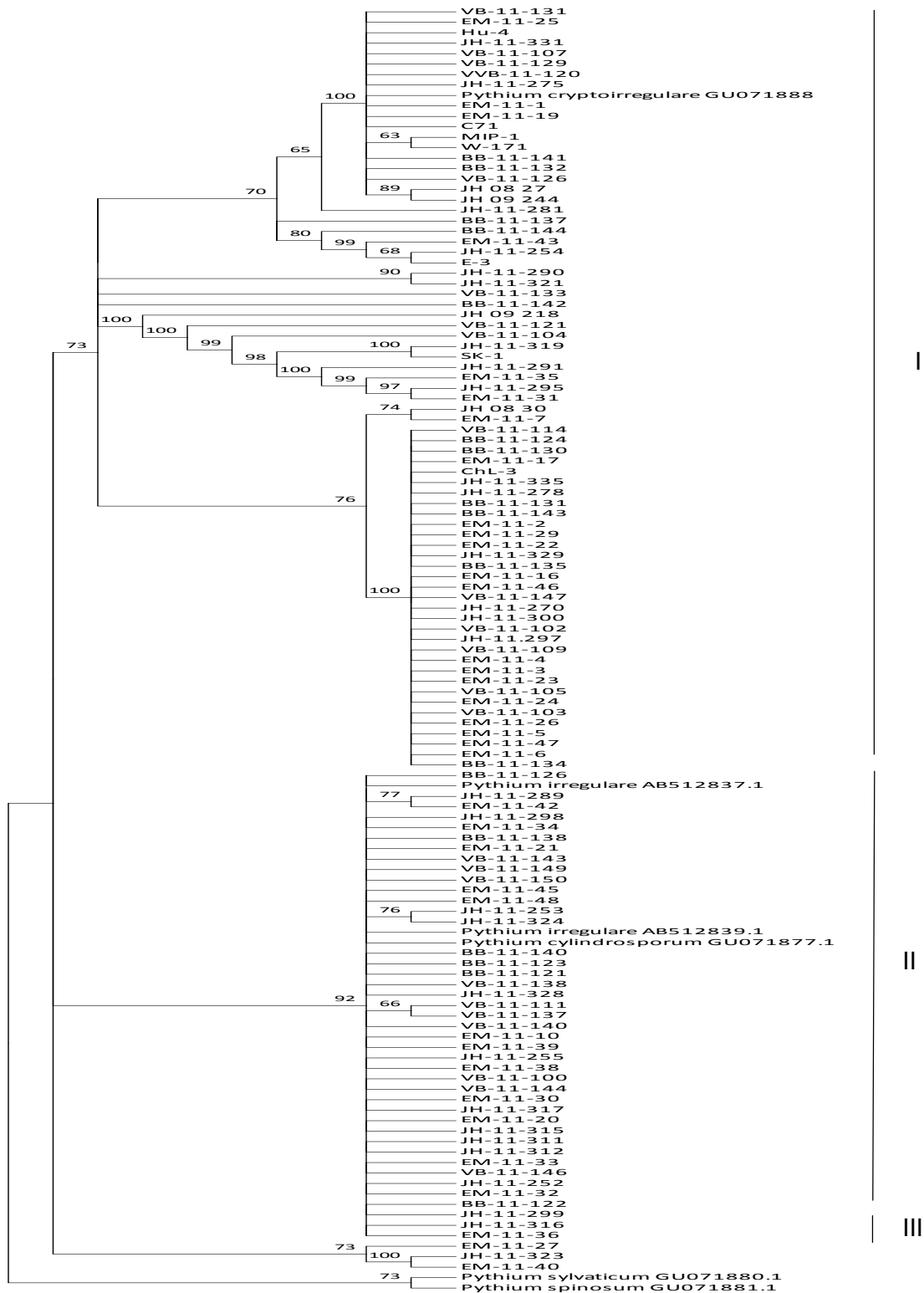


Figura 3.18: Árbol consenso en base a secuencias del gen β -tubulina utilizando el método de máxima verosimilitud basado en el modelo de Tamura de tres parámetros (T92+I). El árbol fue enraizado con *Pythium spinosum* y *Pythium sylvaticum* como grupos externos. Los números en las ramas representan los valores “bootstrap” calculados a partir de 1000 replicaciones (Sólo los valores superiores al 50% se muestran).

CAPÍTULO 4. DISCUSIÓN

Los estudios que investigan la asociación de las especies de *Pythium* con cultivos de plantas utilizan métodos basados en cultivos para su aislamiento y características morfológicas para identificar aislados a nivel de especie. La taxonomía de *Pythium* se ha basado principalmente en características morfológicas, algunas de las cuales son altamente variables o están ausentes en muchas especies. Esta variación no ha permitido llegar a un consenso para establecer las características morfológicas más importantes para la identificación de las especies (Long *et al.*, 2013; Senda, Kageyama, Suga & Lévesque, 2009).

El uso de técnicas moleculares tanto para la identificación y detección de especies *Pythium* tiene varias ventajas sobre las técnicas convencionales, incluyendo una mayor precisión, resolución y rapidez (Jakobus, Mazzola, & McLeod, 2013). Gracias al avance de los métodos moleculares se ha logrado estudiar relaciones filogenéticas, identificar especies y contribuir a la sistemática de *Pythium* (Long *et al.*, 2013).

Los métodos basados en ácidos nucleicos son cada vez más utilizados para desarrollar pruebas de diagnóstico de patógenos de plantas. Particularmente los métodos basados en la reacción en cadena de la polimerasa, son sensibles y tienen el potencial de ser específicos. Permite sintetizar rápidamente (amplificar) millones de copias del ADN específico de una secuencia y al ser una técnica sensible, sólo se requieren de pequeñas cantidades de ADN (Ward, Foster, Fraaije, & Alastair McCartney, 2003).

El aislamiento de ADN es un primer paso para el estudio de secuencias de ADN. La cantidad, calidad e integridad del ADN afecta los resultados de un ensayo (Surzycki, 2000). La utilización de protocolos de extracción de ADN rápidos y simples que sean capaces de obtener una alta cantidad y calidad de ADN es esencial para estudios filogenéticos de patógenos de plantas. kits comerciales se utilizan ampliamente, en estos estudios por ser rápidos y permiten obtener ADN de calidad, adecuada sin embargo, suelen ser caros, requieren la ejecución de muchos pasos en un corto tiempo y regularmente se obtienen bajas concentraciones de ADN (Zelaya-Molina , Ortega, & Dorrance, 2010).

En esta investigación con el protocolo del kit DNeasy Plant Mini de QIAGEN se obtuvieron concentraciones de ADN en rangos desde 3.6 a 880 ng/uL, y una calidad promedio de ADN evaluada con la relación A260/A280 de 1.79. Generalmente un ensayo de tipo PCR con ADN de oomicetos requiere una concentración mínima de ADN de 10 ng/uL (Surzycki, 2000). En la mayoría de muestras se logró extraer la concentración adecuada excepto por unas cuantas muestras con una concentración inferior que, a pesar de eso, amplificaron con la PCR. En el caso de la calidad del ADN el valor obtenido de la relación A260/A280 indicó que el ADN estaba de calidad aceptable para la PCR (Glasel, 1995; Maniatis, Fritsch, & Sambrook,, 1982).

Algunos análisis han explorado las relaciones entre especies de *Pythium* usando la región ITS del ADN ribosomal (rADN) nuclear, una región que amplifica fácilmente y que contiene un largo número de sitios variables (Blair *et al.*, 2008). En la actualidad, se incorporan en estudios de filogenia genes

con el fin de representar la evolución del genoma de los oomicetes mediante el análisis y comparación de la filogenia derivada de cada uno de ellos. Genes como β -tubulina y citocromo oxidasa II, que codifican para proteínas estructurales y metabólicas respectivamente, están recibiendo cada vez más atención (Villa *et al.*, 2006; Blair *et al.*, 2008).

El objetivo de este estudio fue utilizar el gen β -tubulina como un loci informativo a nivel filogenético. B-tubulina es un gen que no contiene intrones, que es conservado en los eucariotas y que permite a la identificación de especies en muchos grupos biológicos (Villa *et al.*, 2006; Blair *et al.*, 2008).

En este estudio se usó una porción del gen β -tubulina (1226 pb). Su amplificación se hizo con el par de primers Btub_F1A y Btub_R1A que fueron diseñados para amplificar secuencias parciales del gen β -tubulina en aislados de *Phytophthora* según el estudio realizado por Blair *et al.* (2008) para el análisis filogenético del género *Phytophthora*. Con estos primers se logró amplificar la porción del gen β -tubulina en las 120 cepas de *Pythium irregulare* s.l., provenientes de invernaderos de Long Island New York y de distintos huéspedes, las cuales estuvieron en el rango de peso molecular esperado (1226 pb).

Se empleó la secuenciación del gen para caracterizar las cepas en estudio. En la actualidad, la secuenciación se ha convertido en uno de los más importantes avances científicos, al permitir conocer el orden de los nucleótidos de un gen en particular. Gracias al despliegue de la secuenciación de ADN, a su bajo costo y alto rendimiento ha sido posible examinar de la biología y de

la evolución de la patogenicidad de algunos oomicetos, además de la resolución de su filogenia usando múltiples alineamientos concatenados (Fitzpatrick, Logue, Stajich, & Butler, 2006; James, 2006).

Pese a sus ventajas, durante la secuenciación de ADN hay muchas variables que puede afectar la calidad de la información, por eso resulta importante evaluar la información generada mediante herramientas y software que ayuden a determinar errores que pudieron ocurrir durante el proceso (The Perkin-Elmer Corporation., 2008). En esta investigación, las secuencias enviadas por OSU fueron analizadas evaluando sus cromatogramas, alineando la cadena forward y reverse y verificando la correcta asignación de bases entre ambas cadenas (sentido y antisentido) a través del programa BioNumerics versión 7.01.

Ciento diecisiete de las 120 secuencias fueron editadas, alineadas y ensambladas. Tres de las secuencias presentaron problemas debido a la mala calidad de las mismas, lo que no permitió su alineamiento; además en los cromatogramas de cada una se observó principalmente problemas como ruido de fondo y secuencias superpuestas que pudieron aparecer por una baja concentración del ADN, por la presencia de cantidades importantes de contaminantes o por la presencia de primers de PCR de pasos anteriores que no se eliminaron en el paso de purificación (Soria, 2000).

Para las 117 secuencias se pudo generar secuencias consenso que representan a cada conjunto de secuencias o contig (Cañizares y Blanca, 2013). Estas secuencias fueron sometidas a un análisis Blast. Blast es quizás

la herramienta más importante para la identificación de secuencias que comparten una similitud local con secuencias registradas en la base de datos (Guttman *et al.*, 2014). Con el fin de conocer el grado de similitud entre las secuencias en estudio (secuencias de interés) y las secuencias anotadas y registradas en la base de datos Blast, se prestó atención a los parámetros de porcentaje de cobertura, porcentaje de identidad máxima y valor E. El valor de porcentaje de cobertura de las secuencias de interés fue de un 100 %, indicando así que las secuencias analizadas tenían una similitud de un 100 % con el mejor resultado devuelto por la base de datos y que el alineamiento de los nucleótidos (aproximadamente 1100) coincidió totalmente con los del género y especie asignado. Por otro lado, el valor de porcentaje de identidad del 99% permitió asegurar que el gen analizado pertenece taxonómicamente al género asignado (Hall, 2013) y que las secuencias problema son similares en un 99 % a secuencias de *Pythium irregulare* para el gen β -tubulina. El valor E tuvo un valor de 0.0, siendo este el mejor valor posible, ya que cuanto más bajo sea el valor E, más significativa es la puntuación. Según el análisis Blast, las secuencias son homologas en el sentido de que los emparejamientos resultantes tienen un alto nivel de significancia y no presentan ningún tipo de degeneración en los extremo 5' y 3' del gen según la gráfica de distribución de hits del Blast. La identificación de secuencias similares es importante pues la similitud es el principal predictor de homología, la homología es el principal predictor de función y un alineamiento de secuencias permite identificar secuencias similares (Guttman *et al.*, 2014; Rahman, 2007).

El alineamiento múltiple de las secuencias en estudio fue analizado con el método de máxima verosimilitud. Los datos filogenéticos usualmente consisten en alineamientos múltiples de secuencias que es el paso más crítico en un análisis filogenético debido a que establece las correspondencias posicionales en la evolución; permite realizar análisis de evolución, identificar homología, construir filogenias y probar modelos evolutivos (Guttman *et al.*, 2014). Un buen alineamiento produce inferencias filogenéticas correctas (Rodríguez, 2013).

A menos que las relaciones filogenéticas reales se conozcan de antemano, no hay manera de establecer el mejor procedimiento de alineación para un análisis filogenético dado. En general, es aconsejable simplemente someter a la alineación generada por computadora a un procedimiento de construcción de árbol. Esta opción se aplica en especial a los programas de construcción de árboles que incluyen paquetes de alineación (Por ejemplo: Muscle y ClustalW en MEGA) (Brinkman & Leipe, 2000).

El alineamiento múltiple en esta investigación se hizo con el algoritmo Muscle (comparación de secuencias múltiples por log-expectativa). Este algoritmo tiene como paso fundamental la alineación de perfiles por parejas que después se utilizan para una alineación progresiva, refinamiento y la construcción de un árbol guía (Edgar, 2004).

El alineamiento múltiple de las secuencias con el algoritmo Muscle del programa MEGA versión 5.01 y bajo el método estadístico de máxima verosimilitud, evidenció diferencias entre las secuencias de cada cepa. Solo

cierto número de posiciones alinearon y conservaron bases entre todas las secuencias de las cepas en estudio. Las columnas, en un alineamiento múltiple de secuencias, son indicadoras de homología. Rodríguez (2013) menciona que el número de sustituciones observadas pueden no reflejar los verdaderos eventos evolutivos que ocurrieron, lo que dificulta la estimación de las verdaderas distancias evolutivas entre las secuencias estudiadas. Un nucleótido idéntico observado en el alineamiento puede deberse a mutaciones en paralelo en ambas secuencias, efecto conocido como homoplastia, que para su corrección requiere modelos estadísticos (modelos de evolución) que permitan inferir las verdaderas distancias evolutivas entre secuencias.

El proceso de encontrar un árbol filogenético utilizando máxima verosimilitud implica la búsqueda de una topología y de las longitudes de las ramas de un árbol que den la mayor probabilidad de observar las secuencias de ADN en nuestros datos (Cho, 2012). En los últimos años, se ha dado un creciente interés en producir árboles filogenéticos precisos utilizando este enfoque estadístico, principalmente por su versatilidad, por su aplicación en una gran cantidad de situaciones y por ser un procedimiento que permite estimar los parámetros de un modelo, de tal manera, que sean los más probables a partir de los datos obtenidos (Keane, Naughton, Travers, McInerney, & McCormack, 2005).

El árbol que dio la mayor probabilidad es el que se examinó, el cual tuvo un valor de log likelihood superior de -11917.62. El valor log likelihood representa la función de probabilidad o verosimilitud. El ajuste del modelo a los datos se mide por la función de probabilidad (Edwards A. , 1972). En el

cálculo de la función de verosimilitud interviene el producto de las probabilidades individuales, por lo que interesa tomar logaritmos, que transforman los productos en sumas y los cocientes en restas. Al tratarse de productos de probabilidades, la función de verosimilitud será siempre menor que 1 y por tanto su logaritmo será negativo (Duduk, 2009).

El modelo de sustitución y evolución que resultó ajustarse de mejor manera a los datos fue el de Tamura de tres parámetros (T92). La elección de este modelo indica que hay fuertes eventos de transición-transversion y sesgos en el contenido de G+C (Mina y Romero-Campero, 2012).

El árbol filogenético obtenido fue de tipo cladograma, un diagrama de ramificación donde las ramas tienen la misma longitud y en donde se muestra un ancestro en común, pero no se indica el tiempo evolutivo que separa a cada taxón (Duduk, 2009). Un primer árbol construido mostraba que no todas las ramas eran significativas, por lo que fue necesario construir un árbol condensado para dar más significancia a las ramas.

El árbol consenso que se obtuvo provee información para pensar que las cepas de *Pythium irregulare* sensu lato tienen un ancestro en común, a pesar de lo cual, el árbol no muestra un porcentaje de ocurrencia mayor al 50% en todas sus ramas. A cada rama se le asigna un porcentaje de ocurrencia y, de esta manera, se asume que una rama es significativa si aparece más de 50% o 70% de las veces, las ramas restantes pueden ser condensadas resultando en politomias (Mina y Romero-Campero, 2012).

La topología del árbol consenso indicó la formación de tres grupos principales (I, II y III). Las relaciones de los tres grupos recibieron un soporte bootstrap moderado-alto (63-100%) en el análisis de máxima verosimilitud. En los grupos I y II sobresalieron ramas que fueron condensadas resultando en politomías. También se pudo observar que de las cuatro secuencias referencia de *Pythium* para el gen β -tubulina, tres de ellas: *Pythium irregulare* (AB512837.1), *Pythium irregulare* (AB512839.1) y *Pythium cylindrosporium* (GU071877.1) se agruparon en el grupo II y la secuencia de *Pythium cryptoirregulare* (GU071888.1) se agrupó en el grupo I.

La condensación de las ramas que resultaron en politomías, nodos con más de dos linajes descendientes, puede ser indicio de dos cosas: falta de conocimiento o eventos de especiación rápida. La falta de conocimiento debido a que no se cuenta con los datos suficientes para entender cómo se relacionan esos linajes lo que impide sacar conclusiones. La especiación rápida representa la posibilidad de múltiples episodios de especiación al mismo tiempo. En este caso, todos los linajes descendientes estarían igualmente relacionados entre sí. Las politomías pueden resolverse cuando se reúnen más datos (Blanca & Cañizares, 2010).

La filogenia se basa explícitamente en el modelo de un solo linaje ancestral común que se divide en el tiempo en múltiples linajes. Para las especies, estas divisiones son eventos de especiación. Para moléculas (genes o secuencias de proteínas), estas divisiones también pueden ser eventos de duplicación (Edwards, 2012). La definición de la posición del

ancestro hipotético en el árbol, conocido como enraizamiento de un árbol es una de las cosas que causa confusión en filogenética.

Una de las estrategias para el enraizamiento de un árbol filogenético es utilizar grupos externos (outgroups). Un outgroup es una o más especies que se saben están alejadas de todos los demás taxones en el árbol (Edwards, 2012), pero están emparentadas a ellos, y su inclusión facilita el determinar si un grupo, en este caso las cepas en estudio, forman grupos monofiléticos. La mayoría de métodos de construcción de árboles generan árboles sin raíz y gracias a un conocimiento previo se coloca la raíz en el lugar correcto. En este estudio el enraizamiento se consiguió incluyendo como outgroups a *Pythium spinosum* y *Pythium sylvaticum*, dos especies alejadas de *Pythium irregulare*. El enraizamiento permitió observar que las cepas de *Pythium irregulare* sensu lato estudiadas son monofiléticas en base a las secuencias del gen β -tubulina.

Esta investigación permitió establecer la relación filogenética de cepas de *Pythium irregulare* sensu lato, un oomicete y agente causal de algunas de las enfermedades de raíz más serias en plantas, mediante la caracterización del gen β -tubulina, con el propósito de diferenciar y estimar su relación filogenética. Se determinó que las cepas al compartir un ancestro, en común, se agrupan cercanamente entre sí. Esta investigación también proporcionó información para la identificación de *Pythium irregulare* sensu lato como patógeno utilizando métodos moleculares basados en la reacción en cadena de la polimerasa (PCR) y la secuenciación.

La reconstrucción filogenética es un campo de rápido crecimiento que se enriquece con diferentes enfoques estadísticos y que tiene aplicaciones en una amplia gama de áreas biológicas. Estudios de este tipo permiten determinar la relación jerárquica entre especies en función de su relación evolutiva y contribuyen con información útil para la implementación de estrategias en el control y gestión de enfermedades causadas por patógenos de plantas (Garzón, Geiser, & Moorman, 2005).

CAPÍTULO 5. CONCLUSIONES

1. El protocolo del kit QIAGEN para la obtención de ADN a partir de micelio liofilizado permitió la obtención de ADN de alta calidad y concentración, facilitando la amplificación del gen β -tubulina en todas las cepas de estudio.
2. La amplificación y secuenciación el gen β -tubulina de las cepas de *Pythium irregulare* sensu lato permitió observar fragmentos de aproximadamente 1200 pb que están en el rango de peso molecular esperado (1226 pb), según el estudio de Blair et al (2008).
3. Ciento diecisiete secuencias consenso, de un total de 120, se registraron en archivos formato FASTA a partir del análisis, edición y ensamblaje de los datos generados después de la secuenciación, con el programa BioNumerics versión 7.01.
4. El alineamiento múltiple de las secuencias con el algoritmo Muscle del programa MEGA versión 5.01 y bajo el método estadístico de máxima verosimilitud, evidenció diferencias entre las secuencias de cada cepa. Solo cierto número de posiciones alinearon y conservaron bases entre todas las secuencias de las cepas en estudio.
5. El modelo de sustitución y evolución que mejor se ajustó a los datos fue el de Tamura de tres parámetros (T92). Con este modelo se construyó un árbol filogenético en el cual destacan los eventos fuertes de transición-transversión y el sesgo en el contenido de G +C. Se observó desigualdad

en las frecuencias de nucleótidos en base al número de sustituciones encontradas durante el alineamiento.

6. El árbol filogenético basado en el correspondiente alineamiento múltiple, modelo de evolución de Tamura de tres parámetros (T92), con base al método de máxima verosimilitud y evaluado estadísticamente mediante una prueba “bootstrap” con 1000 repeticiones, reveló la presencia de tres grupos principales (I, II y III) y de ramas restantes condensadas que resultaron en politomías.
7. Las relaciones de los tres grupos principales formados en el árbol filogenético recibieron un soporte “bootstrap” moderado-alto (63-100%) en el análisis de máxima verosimilitud. No hubo una clara separación entre las especies dentro del complejo *P. irregulare*, lo cual es consistente con otros estudios preliminares.
8. El árbol consenso obtenido para las cepas de *Pythium irregulare* s.l. de Long Island New York, proporcionó información para asegurar que comparten un ancestro en común ya que se agrupan cercanamente entre sí. Las cepas de *Pythium irregulare* s.l. son monofiléticas en base al análisis de secuencias del gen β -tubulina, utilizando las especies *P. sylvaticum* y *P. spinosum* como outgroups.
9. El análisis del gen β -tubulina como loci individual es útil para resolver relaciones entre especies pero no garantiza una filogenia de especies completamente segura, particularmente si las especies son muy cercanas. El análisis filogenético de genes individuales solo provee una estimación

de las relaciones evolutivas entre las secuencias de ese gen obtenidas de un grupo de muestras. Por lo tanto, un análisis concatenado de genes es un mejor método para la reconstrucción filogenética de una especie, pues hay una mayor probabilidad de descubrir el árbol correcto.

10. Esta investigación proporcionó información para la identificación de *Pythium irregulare* como patógeno utilizando métodos moleculares basados en la reacción en cadena de la polimerasa y la secuenciación del gen de beta tubulina.

CAPITULO 6. RECOMENDACIONES

1. Obtener el resultado del alineamiento múltiple de secuencias de varias fuentes y compararlas cuidadosamente para identificar el mejor.
2. Construir alineamientos y otros árboles filogenéticos con diferentes métodos estadísticos como máxima parsimonia (MP) y análisis bayesianos (BA), con el fin de comparar si todos los árboles producen una misma topología.
3. Construir árboles multilocus para obtener mayor evidencia estadística que muestren un porcentaje de ocurrencia mayor al 70% en todas sus ramas para asegurar que los organismos comparten un mismo ancestro en común.
4. Realizar un análisis concatenado de genes para la reconstrucción filogenética de las cepas de *Pythium irregulare* s.l. de Long Island New York, porque este estudio evidenció que con el análisis de un solo gen, β -tubulina, se obtienen árboles con baja resolución entre especies muy cercanas.
5. Validar los métodos empleados en este estudio usando muestras del Ecuador a partir del aislamiento de cepas de *Pythium* spp. en medios selectivos, identificación en base a morfología e identificación molecular con la técnica PCR, con el fin de identificar las especies de *Pythium* encontradas en campos del país.

CAPÍTULO 7. BIBLIOGRAFÍA

Agrios, G. N. (2005). *Plant Pathology*. USA: Elsevier.

Al-Sa'di, A. M., Drenth, A., Deadman, M. L., de Cock, A., & Aitken, E. (2007). Molecular characterization and pathogenicity of *Pythium* species associated with damping-off in greenhouse cucumber (*Cucumis sativus*) in Oman. *Plant Pathology*, *56*, 140 – 149. doi:10.1111/j.1365-3059.2006.01501.x

Belbahri, L., McLeod, A., Paul, B., Calmin, G., Moralejo, E., Spies, C., . . . Lefort, F. (2008). Intraspecific and within-isolate sequence variation in the ITS rDNA region of *Pythium mercuriale* sp. nov., FEMS. *Microbiology Letters*, *284*(1), 17-27.

Blair, J., Coffey, M., Park, S.-Y., Geiser, D., & Kang, S. (2008). A multi-locus phylogeny for *Phytophthora* utilizing markers derived from complete genome sequences. *Fungal Genetics and Biology*, *45*, 266–277. doi:10.1016/j.fgb.2007.10.010

Blanca, J., & Cañizares, J. (2010). *Bioinformatics at COMAV*. Recuperado el 7 de Noviembre de 2013, de http://bioinf.comav.upv.es/courses/intro_bioinf/filogenias.html

Brinkman, F., & Leipe, D. (2000). Bioinformatics: A Practical Guide to the Analysis of Genes and ProteinS. En *PHYLOGENETIC ANALYSIS* (Second Edition ed.).

Cañizares, J., & Blanca, J. (2013). *Ensamblaje y mapeo de secuencias tipo Sanger*. Recuperado el 17 de Octubre de 2013, de <http://personales.upv.es/jcanizar/bioinformatica/ensamblaje.html>

- Capote , N., Pastrana , A., Aguado , A., & Sánchez-Torres, P. (2012). Molecular Tools for Detection of Plant Pathogenic Fungi and Fungicide Resistance. *Plant Pathology*, 151-202.
- Caujape-Castells, J., Pedrola-Monfort, J., & Membrives, N. (2003). ASPECTOS FILOSÓFICOS DEL ANÁLISIS DE DATOS EN SISTEMÁTICA MOLECULAR. *Bot. Macaronésica*, 24, 127-141.
- Cho, A. (2012). Constructing Phylogenetic Trees Using Maximum Likelihood. . *Scripps Senior Theses*(46). Recuperado el 8 de Enero de 2014
- Doyle, J., & Gaut, B. (2000). Evolution of genes and taxa: a primer . *Plant Molecular Biology*(42), 1-23.
- Duduk, B. (2009). Phylogenetic tree construction. *Institute for Pesticides and Environmental Protection*.
- Edgar, R. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), 1792-1797. doi:10.1093/nar/gkh340
- Edwards, A. (1972). Likelihood. Cambridge , Cambridge, UK. Recuperado el 2013
- Edwards, R. (7 de Junio de 2012). *How to root a phylogenetic tree*. Recuperado el 21 de Octubre de 2013, de The cabbages of doom: <http://cabbagesofdoom.blogspot.com/2012/06/how-to-root-phylogenetic-tree.html>

- Felsenstein, J. (1985). Phylogenies and the Comparative Method. *The American Naturalist*, 125(1), 1-15.
- Fitzpatrick, D., Logue, M., Stajich, J., & Butler, G. (2006). A fungal phylogeny based on 42 complete genomes derived from supertree and combined gene analysis. *BMC Evol. Biol*, 22-99.
- Garzón, C. D., Geiser, D. M., & Moorman, G. W. (2005). Amplified Fragment Length Polymorphism Analysis and Internal Transcribed Spacer and *coxII* Sequences Reveal a Species Boundary Within *Pythium irregulare*. *The American Phytopathological Society*, 95(12), 1489-1498. doi:10.1094/PHYTO-95-1489
- Garzón, C. D., Geiser, D. M., & Moorman, G. W. (2005). Diagnosis and Population Analysis of *Pythium* Species Using AFLP Fingerprinting. *The American Phytopathological Society*, 89, 81-89. doi:10.1094/PD-89-0081
- Garzón, C. D., Yáñez, J. M., & Moorman, G. W. (2007). *Pythium cryptoirregulare*, a new species within the *P. irregulare* complex. *Mycologia*, 99(2), 291–301.
- Glasel, J. (1995). Validity of nucleic acid purities monitored by A260/A280 absorbance ratios. *Biotechniques*.
- Guttman, D., Provart, N., Austin, R., Hug, L., & Price, M. (2014). Bioinformatic Methods I. LAB 1a — EXPLORING NCBI. 1-22.

- Jakobus, C., Mazzola, M., & McLeod, A. (2013). Characterization and detection of *Pythium* and *Phytophthora* species associated with grapevines in South Africa. *Eur J Plant Pathol*(131), 103-119. doi:10.1007/s10658-013-9791-5
- James, T. (2006). Reconstructing the early evolution of fungi using a six-gene phylogeny. *Nature*, 443, 818–822.
- Kageyama, K., Nakashima, A., Kajihara, Y., Suga, H., & Nelson, E. N. (2005). Phylogenetic and morphological analyses of *Pythium graminicola* and related species. *J Gen Plant Pathol*, 71, 174–182. doi:10.1007/s10327-005-0184-5
- Katawczik, M. (2008). *Pythium irregulare*. *Soilborne Plant Pathogens*. Recuperado el 15 de Junio de 2013, de http://www.cals.ncsu.edu/course/pp728/Pythium_irregulare/Pythium_irregulare.html
- Keane, T., Naughton, T., Travers, S., McInerney, J., & McCormack, G. (2005). DPRml: distributed phylogeny reconstruction by maximum likelihood. *Phylogenetics*, 21(7), 969–974. doi:10.1093/bioinformatics/bti100
- Lemey, P., Salemi, M., & Vandame, A.-M. (2009). *The Phylogenetic Handbook. A Practical Approach to Phylogenetic Analysis and Hypothesis Testing* (Second Edition ed.). New York: Cambridge University Press.
- Lévesque, A., & de Cock, A. (2004). Molecular phylogeny and taxonomy of the genus *Pythium*. *Mycol. Res*, 108(12), 1363-1383. doi:10.1017/S0953756204001431

- Long, Y.-Y., Wei, J.-G., Sun, X., He, Y.-Q., Luo, J.-T., & Guo, L.-D. (2011). Two new *Pythium* species from China based on the morphology and DNA sequence data. *Mycol Progress*. doi:10.1007/s11557-011-0778-6
- Maniatis, T., Fritsch, R., & Sambrook, J. (1982). Molecular cloning a laboratory manual. *Cold Spring Harbor Laboratory*.
- Martin, F., & Tooley, P. (2003). Phylogenetic relationships among *Phytophthora* species inferred from sequence analysis of mitochondrially encoded cytochrome oxidase I and II genes. *Mycologia*, 95(2), 269–284.
- McCartney, H., Foster, S., Fraaije, B., & Ward, E. (2003). Molecular diagnostics for fungal plant pathogens†. *Pest Management Science*(59), 129–142. doi:10.1002/ps.575
- Meng, S., Torto Alalibo, T., Chibucos, M. C., Tyler, B. M., & Dean, R. A. (2009). Common processes in pathogenesis by fungal and oomycete plant pathogens, described with Gene Ontology terms. *BMC Microbiology*, 9. doi:10.1186/1471-2180-9-S1-S7
- Mina, J., & Romero-Campero, F. (2012). *Molecular Evolution and Phylogenetics*. Recuperado el 18 de Octubre de 2013, de <http://www.cs.us.es/~fran/students/julian/phylogenetics/phylogenetics.html>
- Moralejo, E., Clemente, A., Descals, E., Belbahri, L., Calmin, C., Lefort, F., . . . McLeod, A. (2008). *Pythium recalcitans* sp. nov. revealed by multigene phylogenetic analysis. *Mycologia*, 100(2), 310–319.

- Narayanasamy, P. (2013). *Microbial Plant Pathogens- Detection and Disease Diagnosis: Fungal Pathogens*. Springer, 1.
- Nzungize, J., Gepts, P., Buruchara, R., Buah, S., Ragama, P., Busogoro, J., & Baudoin, J. (2011). Pathogenic and molecular characterization of *Pythium* species inducing root rot symptoms of common bean in Rwanda. *African Journal of Microbiology Research*, 5(10), 1169-1181.
- Opperdoes, F. (1997). *Maximum Likelihood*. Recuperado el 22 de Agosto de 2013, de http://www.icp.ucl.ac.be/~opperd/private/max_likeli.html
- Palmucci, H., & Grijalba, P. (2007). Root and stem rot caused by *Pythium aphanidermatum* on poinsettia in a soil-less culture system in Buenos Aires Province, Argentina. *Australasian Plant Disease Notes*, 2, 139-140. doi:10.1071/DN07055
- Pfennig, L., & Magalhães, L. (2008). *Saprophytic and Plant Pathogenic Soil Fungi*. "A Handbook of Tropical Soil Biology: Sampling and Characterization of Below-ground Biodiversity." (M. Moreira, J. Huising , & D. Bignell. , Edits.)
- Poot, A. (2013). *Alineamientos múltiples*. Recuperado el 12 de Septiembre de 2013, de <http://ftp.fisio.cinvestav.mx/Default/Fisiologia/Cursos/.../AlinMult.pdf>
- Posada, D., & Bickley , T. (2004). Model Selection and Model Averaging in Phylogenetics: Advantages of Akaike Information Criterion and Bayesian Approaches Over Likelihood Ratio Tests. *Syst. Biol*, 53(5), 593-808. doi:10.1080/10635150490522304

- Posada, D., & Crandall, K. (2001). Selecting the Best-Fit Model of Nucleotide Substitution. *Syst. Biol.*, 4(50), 580-601.
- Rahman, N. N. (2007). *Bioinformatics: Analyzing DNA Sequence using BLAST*. DHAKA: BRAC University.
- Ramsden, J. (2009). *Bioinformatics: An Introduction* (Second Edition ed.). London, UK: Springer.
- Rodríguez, E. (2013). *Construcción de árboles filogenéticos*. Recuperado el 13 de Noviembre de 2013, de <http://www.tamps.cinvestav.mx/~ertello/bioinfo/sesion13.pdf>
- Rodríguez, R., Aguilar, C., Ayala, L., Rocha, J., Padilla, V., & Espinosa, T. (2009).
- Rusagara, J., Lyumugabe, F., Busogoro, J., & Baudoin, J. (2012). Pythium root rot of common bean: biology and control methods. A review. *Biotechnol. Agron. Soc. Environ.*, 16(3), 405-413.
- Santamaria, R. (2006). *Predicción filogenética*. Recuperado el 15 de Agosto de 2013, de http://vis.usal.es/rodrigo/documentos/bioinformatica/temas/6_Predicción%20Filogenética.pdf
- Senda, M., Kageyama, K., Suga, H., & Lévesque, A. (2009). Two new species of Pythium, *P. senticosum* and *P. takayamanum*, isolated from cool-temperate forest soil in Japan. *Mycologia*, 101(4), 439-448. doi:10.3852/08-104.

Shonkwiler, R. W., & Herod, J. (2009). *Mathematical Biology An Introduction with Maple and Matlab*. New York, USA: Springer.

Soria, M. (2000). *Curso Introducción a la Bioinformática. Aplicaciones en Agricultura. Clase 6. Análisis de secuencias de ADN. Programas para diseño de primers para PCR*. Recuperado el 10 de Noviembre de 2013, de <https://sites.google.com/a/agro.uba.ar/bioinformatica/Home/clase-6-primers-para-pcr>

Surzycki, S. (2000). *Basic Techniques in Molecular Biology*. (126), 18-21.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2013). MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol. Biol. Evol*, 28(10), 2731-2739. doi:10.1093/molbev/msr121

The Perkin-Elmer Corporation. (2008). *Automated DNA sequencing. Chemistry guide. PE Applied Bio systems*. Recuperado el 2013

Thermo Fisher Scientific Inc. (2008). *NanoDrop 1000 Spectrophotometer V3.7 User's Manual*.

Tsui, C., Woodhall, J., Chen, W., Lévesque, A., Lau, A., Schoen, C., . . . Sybren de Hoog, G. (2013). Molecular techniques for pathogen identification and fungus detection in the environment. *IMA Fungus*, 2(2), 177-189. doi:10.5598/imafungus.2013.02.02.09

- Villa, N. O., Kageyama, K., Asano, T., & Sura, H. (2006). Phylogenetic relationships of *Pythium* and *Phytophthora* species based on ITS rDNA, cytochrome oxidase II and β -tubulin gene sequences. *The Mycological Society of America*, 410–422.
- Ward, E., Foster, S. J., Fraaije, B. A., & Alastair McCartney, H. (2003). Plant pathogen diagnostics: immunological and nucleic acid-based approaches. *Ann. appl. Biol.*, 145, 1-16.
- Zelaya-Molina , L., Ortega, M., & Dorrance, A. (2010). Easy and efficient protocol for oomycete DNA extraction suitable for population genetic analysis. *Biotechnol Lett*, 33, 715-720. doi:10.1007/s10529-010-0478-3

HOJA DE LEGALIZACIÓN DE FIRMAS

ELABORADO POR:

María Fernanda Proaño Cuenca

DIRECTORA DE LA CARRERA DE INGENIERÍA EN BIOTECNOLOGÍA

Dra. María Augusta Chávez Larrea, MSc

Lugar y fecha: SANGOLQUÍ, Marzo de 2014