

Centro de Investigación Científica de Yucatán, A.C.

Posgrado en Ciencias Biológicas

IDENTIFICACIÓN DE PROTEÍNAS EFECTORAS DE

Pseudocercospora fijiensis

Tesis que presenta

KARLA GISEL CARREÓN ANGUIANO

En opción al título de

DOCTOR EN CIENCIAS

(Ciencias Biológicas: OPCIÓN BIOTECNOLOGÍA)

Mérida, Yucatán, México

2021

*CENTRO DE INVESTIGACIÓN CIENTÍFICA DE YUCATÁN, A. C.
POSGRADO EN CIENCIAS BIOLÓGICAS*



RECONOCIMIENTO

Por medio de la presente, hago constar que el trabajo de tesis de **Karla Gisel Carreón Anguiano** titulado “**Identificación de proteínas efectoras de *Pseudocercospora fijiensis***” fue realizado en la **Unidad de Biotecnología**, en la línea de investigación **Agrobiotecnología**, en el laboratorio de **Biotecnología microbiana** del Centro de Investigación Científica de Yucatán, A.C. bajo la dirección de la **Dra. Blondy Canto Canché** y **Dr. Luis A. Sáenz Carbonell**, dentro de la opción de **Biotecnología**, perteneciente al Programa de Posgrado en Ciencias Biológicas de este Centro.

Este trabajo de investigación forma parte del proyecto de ciencia básica “**Enfoque agrogenómico en el estudio de la Sigatoka negra**” con clave **CB2013-01 220957**, financiado por el CONACYT.

Atentamente.

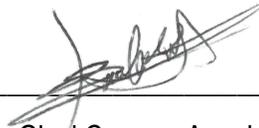


Dra. Cecilia Hernández Zepeda
Directora de Docencia

Mérida, Yucatán, México, a 28 de Mayo de 2021.

DECLARACIÓN DE PROPIEDAD

Declaro que la información contenida en la sección de Materiales y Métodos, los Resultados y Discusión de este documento proviene de las actividades de investigación realizadas durante el período que se me asignó para desarrollar mi trabajo de tesis, en las Unidades y Laboratorios del Centro de Investigación Científica de Yucatán, A.C., y que a razón de lo anterior y en contraprestación de los servicios educativos o de apoyo que me fueron brindados, dicha información, en términos de la Ley Federal del Derecho de Autor y la Ley de la Propiedad Industrial, le pertenece patrimonialmente a dicho Centro de Investigación. Por otra parte, en virtud de lo ya manifestado, reconozco que de igual manera los productos intelectuales o desarrollos tecnológicos que deriven o pudieran derivar de lo correspondiente a dicha información, le pertenecen patrimonialmente al Centro de Investigación Científica de Yucatán, A.C., y en el mismo tenor, reconozco que si derivaren de este trabajo productos intelectuales o desarrollos tecnológicos, en lo especial, estos se registrarán en todo caso por lo dispuesto por la Ley Federal del Derecho de Autor y la Ley de la Propiedad Industrial, en el tenor de lo expuesto en la presente Declaración.



Karla Gisel Carreon Anguiano

Este trabajo se llevó a cabo bajo la dirección de **la Dra. Blondy Canto Canché** y el **Dr. Luis A. Sáenz Carbonell**, en la **Unidad de Biotecnología** del Centro de Investigación Científica de Yucatán, el cual forma parte del proyecto titulado “Enfoque agrogenómico en el estudio de la Sigatoka Negra” con clave CB2013-01 220957, financiado por el CONACYT y cuya responsable técnica es la **Dra. Blondy Canto Canché**.

AGRADECIMIENTOS

Agradezco al Centro de Investigación Científica de Yucatán (CICY, A.C.), por la oportunidad de estudiar el Doctorado haciendo uso de sus instalaciones, equipos y facilidades académicas.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT), por la beca (No. 589301) durante el periodo 2015-2020 y también por el apoyo económico que proviene del proyecto FOSEC-SEP “Enfoque agrogenómico en el estudio de la Sigatoka negra” CB2013-01 (clave 220957).

A mis directores de tesis, la Dra. Blondy Beatriz Canto Canché por la asesoría durante el Doctorado, por su apoyo académico y moral, por todas las enseñanzas, las horas de estudio, por toda su dedicación, su paciencia, creer en mí y por su cariño. Al Dr. Luis Alfonso Sáenz Carbonell por su apoyo académico y moral.

A los miembros del comité tutorial, Dra. Blondy Beatriz Canto Canche, Dr. Luis Alfonso Sáenz Carbonell, Dr. Ignacio Islas Flores y al Dr. Julio Vega Arreguín por sus sugerencias, correcciones para mejorar la calidad del trabajo experimental y escrito.

A los miembros del comité predoctoral, Dr. Ignacio Islas Flores, Dr. Julio Vega Arreguín, Dr. Oscar Moreno Valenzuela, Dra. Marcela Gamboa Angulo y al Dr. Jairo Cristóbal Alejo, quienes contribuyeron a mi crecimiento profesional.

Infinito agradecimiento a los miembros del comité de revisión de tesis, Dra. Blondy Canto Canché, Dr. Ignacio Islas Flores, Dr. Luis A. Sáenz Carbonell, Dr. Julio Vega Arreguín, Dra. Ileana Echeverría Machado, Dr. Felipe Sánchez Teyer y al Dr. Roberto C. Vázquez Euán, por su paciencia y su apoyo en este proceso.

A los técnicos del laboratorio de Biotecnología microbiana y Laboratorio No. 06, al Dr. Miguel Tzec-Simá por el apoyo técnico, por su apoyo invaluable en las técnicas de cultivo de hongos, así como a su esposa Rosalba por el cariño y ser como mi familia. Al M.C. Bartolomé Chí Manzanero, por el apoyo técnico, su paciencia, sus enseñanzas, por estar en las buenas y en las malas, por estar el pendiente y creer en mí, por su cariño durante todos estos años, y a quien considero como mi familia.

A los técnicos, Q.F.B. Rosa Grijalva Arango, Dr. César de los Santos Briones, M.C. Mildred Carrillo, M.C. Bartolomé Chí Manzanero, por hacer de esta estancia amigable, por tantos recuerdos y tanto cariño.

A mis amigos del laboratorio de Biotecnología Microbiana, Dr. Miguel Ángel Canseco Pérez, M.C. Inés Arana Guevara, M.C. Sergio Ramos Jiménez, M.C. Jaime Castro Avilés, Biol. Jaime Muñoz López, Ing. Ricardo Barahona Cortes, Ing. Jair Díaz Ramírez, a las estudiantes de licenciatura Celina Asteria Ayil Naal, María Mercedes Rodríguez Argüello, la M.C. Denise Esther Gutiérrez, al M.C. Osvaldo Couoh y a la Ing. Mariel Rivero por su amistad y cariño, por el apoyo invaluable por todo este tiempo.

A mi sistema de apoyo, a la M.C. Gemaly Ek Hernández y a la M.C. Jewel Nicole Anna Todd por su amistad, su cariño, por ser como mis hermanas y por su apoyo incondicional, al M.C. Rufino Gómez Tah por ser ese hermano que nunca tuve, por el cariño y por ser un amigo incondicional, su apoyo ha sido invaluable e inolvidable.

A todos mis amigos y compañeros de los diferentes laboratorios y departamentos del CICY, por su amistad y compañerismo.

A mis amigos de otros institutos, el Dr. Orelvis Portal Villafaña, Dr. Max Apolinar Hernández, M.C. Rocío Sánchez por su amistad invaluable. Al Dr. Ramón Valenzuela Soto por ser el primero en darme la oportunidad de estar en el mundo de la ciencia, por sus ánimos y por su apoyo incondicional al iniciar éste camino.

A mis amigos Carla Alejandra López Córdoba, Alejandro Rodríguez Gonzáles, Brenda López, Enrique Torres López, Aldredo Ugarte, Ricardo Flores Gonzáles por el apoyo moral y su amistad incondicional e invaluable. Un agradecimiento especial para el Dr. Omar Alba López, un ser maravilloso, quien en todo momento fue un apoyo incondicional en mi vida, la motivación de cada día y quien siempre quiso lo mejor para mi porvenir.

A UNU-BIOLAC y BioAli-CYTED por la beca otorgada por la Universidad de Antioquia y la Universidad CES en Colombia, para asistir al curso “Producción Biotecnológica de Bioinsumos para el Control de Enfermedades Emergentes en Plantas”.

DEDICATORIAS

A Dios por brindarme salud y por guiarme en cada paso.

A mi familia, mis padres Rosa Icela Anguiano Valero y Carlos David Carreón Dávila por su apoyo incondicional, por creer en mí, por su sacrificio y todo su amor.

A mis hermanas que siempre estuvieron ahí apoyándome en la distancia, Nayeli Janeth Carreón Anguiano, Ma. Fernanda Carreón Anguiano, a mis sobrinas Danna y Karla, infinitas gracias.

A toda mi familia, tíos, primos y abuelitos que siempre estuvieron alentándome a seguir adelante, por sus palabras, sus bendiciones y cariño. Por ser cómplices y formar parte de todo esto.

A todos aquellos seres queridos que se adelantaron en el camino...

Dedicado especialmente a

Emma Valero Paniagua.

PRODUCTOS GENERADOS

Artículo 1

Karla Gisel Carreón-Anguiano, Ignacio Islas-Flores, Julio Vega-Arreguín, Luis Sáenz-Carbonell and Blondy Canto-Canché. EffHunter: A Tool for Prediction of Effector Protein Candidates in Fungal Proteomic Databases. *Biomolecules*. 2020 May 4; 10(5):712. <https://doi.org/10.3390/biom10050712>. PMID: 32375409; PMCID: PMC7277995.

Registro propiedad intelectual

Registro ante el INDAUTOR del pipeline EffHunter con el número de folio **03-2019-101809310300-01** a nombre de la Dra. Blondy Canto Canché y Karla Gisel Carreón Anguiano (23 Octubre del 2019).

ÍNDICE

ÍNDICE	I
ABREVIATURAS	V
LISTADO DE FIGURAS	VII
LISTADO DE CUADROS	IX
INTRODUCCIÓN.....	1
CAPÍTULO I	3
ANTECEDENTES	3
1.1 GENERALIDADES DEL BANANO	3
1.2 Sigatoka negra	4
1.2.1 <i>Pseudocercospora fijiensis</i>	4
1.2.2 Infección por <i>Pseudocercospora fijiensis</i>	5
1.3 Interacciones planta-patógeno	8
1.3.1 Inmunidad activada por PAMPs, MAMPs o DAMPs.....	9
1.3.2 Resistencia mediada por genes R	10
1.3.3 Respuesta hipersensible (HR)	10
1.3.4 Resistencia sistémica adquirida e inducida.....	11
1.4 Efectores involucrados en la infección del hospedante	14
1.4.1 Efectoroma	17
1.4.2 Búsqueda de efectores.....	18
JUSTIFICACIÓN	21
HIPÓTESIS	21
OBJETIVO GENERAL	21
OBJETIVOS ESPECÍFICOS	21
ESTRATEGIA EXPERIMENTAL	22
CAPÍTULO II	23
Este capítulo fue aceptado y publicado en la revista Biomoléculas del MDPI.....	23
CAPÍTULO III	53
IDENTIFICACIÓN <i>IN SILICO</i> DE PROTEÍNAS EFECTORAS DE <i>P. FIJIENSIS</i> . *	53
3.1 INTRODUCCIÓN	53
3.2 MATERIALES Y MÉTODOS.	55
3.2.1 Identificación y caracterización <i>in silico</i> de candidatos a efectores de <i>Pseudocercospora fijiensis</i>	55
3.2.2. Análisis de la longitud de aminoácidos y residuos de cisteínas de los efectores predichos.....	56
3.2.3. Anotación funcional de proteínas efectoras y su localización subcelular.	57
3.2.4 Localización genómica (scaffolds).	57
3.2.5 Búsqueda de homólogos de los efectores de <i>P. fijiensis</i>	57
3.3 RESULTADOS	58
3.3.1 Predicción de efectores en <i>P. fijiensis</i>	58
3.3.2 Caracterización de los efectores predichos de <i>P. fijiensis</i>	60
3.3.3 Anotación funcional y localización subcelular del efectoroma de <i>P. fijiensis</i> .	61
3.3.4 Localización genómica de los efectores de <i>P. fijiensis</i>	65
3.3.5 Homólogos de los efectores de <i>P. fijiensis</i>	66
3.4 DISCUSIÓN.....	67
CAPÍTULO IV	73

SELECCIÓN Y CLONACIÓN DE LOS CANDIDATOS DE EFECTORES DE <i>P. FIJIENSIS</i>	73
4.1 INTRODUCCIÓN	73
4.2 MATERIALES Y MÉTODOS	76
4.2.1 Selección de genes candidatos de efectores de <i>P. fijiensis</i>	76
4.2.2 Condiciones de crecimiento de microorganismos	78
4.2.3 Obtención de material de interacción <i>Musa spp-P. fijiensis</i>	78
4.2.4 Extracción de ARN y síntesis de ADNc	79
4.2.5 Amplificación y clonación de los candidatos a efectores de <i>P. fijiensis</i>	79
4.2.6 Transformación de <i>Escherichia coli</i> TOP10 y confirmación de los genes de interés en el vector pGEM-Teasy	80
4.2.7 Purificación y secuenciación de los genes clonados en pGEM-Teasy	80
4.3 RESULTADOS	81
4.3.1 Selección de efectores candidatos en <i>P. fijiensis</i>	81
4.3.2 Amplificación y clonación de los candidatos a efectores de <i>P. fijiensis</i>	84
4.3.3 Confirmación <i>in silico</i> de los candidatos de efectores de <i>P. fijiensis</i>	87
4.3.4 Modelamiento <i>in silico</i> del efector candidato PfID77734 de <i>P. fijiensis</i>	88
4.4 DISCUSIÓN	90
CAPÍTULO V	95
EXPRESIÓN HETERÓLOGA DE LAS PROTEÍNAS EFECTORAS PfAVR4 Y PfECP2 DE <i>P. FIJIENSIS</i>	95
5.1 INTRODUCCIÓN	95
5.2 MATERIALES Y MÉTODOS	100
5.2.1 Secuencias de los efectores reportados de <i>P. fijiensis</i> : PfAvr4 y PfEcp2	100
5.2.2 Amplificación y clonación de los fragmentos del efector PfEcp2	100
5.2.3 Secuenciación de los plásmidos pLATE-PfEcp2 (+SP) y pLATE-PfEcp2 (-SP)	102
5.2.4 Transformación de <i>E. coli</i> cepa Rosetta gami B con los constructos pLATE-PfEcp2 (+SP) y pLATE-PfEcp2 (-SP)	102
5.2.5 Expresión heteróloga de los efectores PfEcp2 (+SP) y PfEcp2 (-SP)	103
5.2.6 Purificación de PfAvr4 y PfEcp2 de <i>P. fijiensis</i>	104
5.2.7 Solubilización de cuerpos de inclusión	104
5.2.8 Cuantificación de proteínas por método Bradford	105
5.3 RESULTADOS	105
5.3.1 Amplificación de los CDSs del efector PfEcp2 (+SP) y PfEcp2 (-SP) con los adaptadores LIC y generación de los constructos pLATE-efector	105
5.3.2 Secuenciación de los vectores de expresión pLATE-PfEcp2 (+SP) y pLATE-PfEcp2 (-SP)	106
5.3.3 Transformación de <i>E. coli</i> Rosetta gami B con los vectores pLATE-PfEcp2 (+SP) y pLATE-PfEcp2 (-SP)	107
5.3.4 Expresión de PfEcp2 (+SP) y PfEcp2 (-SP) en <i>E. coli</i> Rosetta gami B	107
5.4 DISCUSIÓN	113
CAPÍTULO VI	117
DISCUSIÓN GENERAL	117
CAPÍTULO VII	123
CONCLUSIONES Y PERSPECTIVAS GENERALES	123
7.1 CONCLUSIONES	123
7.2 PERSPECTIVAS	125
ANEXOS	127

BIBLIOGRAFÍA..... 129

ABREVIATURAS

ADN	Ácido desoxirribonucleico
ADNc	Ácido desoxirribonucleico complementario
ARN	Ácido ribonucleico
Avr	Factor de avirulencia
BLASTP	Herramienta bioinformática para alineamiento local “Basic Local Alignment Search Tool”.
BSA	Albúmina de suero bovino
CFEM	Commonly found in extracellular membranes
DUF	Domain of unknown function
ECP	Extracellular protein
ETI	Effector triggered immunity
e.g.	Por ejemplo (latín)
<i>et al</i>	Y otros (latín)
FDR	False Discovery Rate
g	Fuerza de gravedad
HR	Respuesta hipersensible “Hypersensitive response”
IDT	Integrated DNA Technologies
JGI	Join Genome Institute
kDa	Kilodalton
MAMP	Microbe-associated molecular patterns
mM	Milimolar
NCBI	National Centre for Biotechnology Information
ng	Nanogramos
PAMP	Pathogen-associated molecular patterns
PCR	Polymerase Chain Reaction
PDA	Potato dextrose agar
Pf	<i>Pseudocercospora fijiensis</i>
PFAM	Protein families database
PHI	Pathogen-Host Interaction database
PMSF	Fluoruro de fenilmetilsulfonilo
PRR	Pattern recognition receptor
PTI	Pathogen-triggered immunity

rpm	Revoluciones por minuto
SDS	Dodecil sulfato de sodio
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
TMD	Transmembranal Domain
µg	Microgramos
µL	Microlitros

LISTADO DE FIGURAS

Figura 1.1 Ciclo de vida de <i>Pseudocercospora fijiensis</i> Morelet, agente causal de la Sigatoka negra del banano. (Churchil, 2011)	7
Figura 1.2 Estados del avance de la infección causada por Sigatoka negra (Fouré,1987;Díaz y Almodóvar, 2007; CIAT, 2013).	8
Figura 1.3 Proceso de infección de <i>Cladosporium fulvum</i>	15
Figura 3.1 Esquematización de la caracterización in silico del efectoroma de <i>P. fijiensis</i> .58	
Figura 3.2 Distribución genómica del efectoroma predicho de <i>P. fijiensis</i>	66
Figura 4.1 Filtros aplicados para priorizar los análisis funcionales de efectores de <i>P. fijiensis</i>	77
Figura 4.2 Diagrama de los tiempos de infección de <i>P. fijiensis</i> y <i>Musa spp.</i>	79
Figura 4.3 Plantas de banano M. acuminata cv. Calcutta4 infectados con <i>P. fijiensis</i> C1233.	84
Figura 4.4 Fotografía gel de agarosa al 0.8% de control Actina de <i>P. fijiensis</i>	85
Figura 4.5 Amplificación por PCR de los efectores candidatos de <i>P. fijiensis</i>	86
Figura 4.6 Marcos de lectura abierto de los efectores candidatos que fueron amplificados y clonados.	88
Figura 4.7 Modelamiento tridimensional del traslape del dominio PAN/Apple de las estructuras de PfID77734 de <i>P. fijiensis</i> y Microneme protein 4 de <i>Toxoplasma gondii</i> . .	89
Figura 5.1 Diagrama experimental aplicado para la ligación y clonación de los fragmentos de PfEcp2 en vector pLATE.	102
Figura 5.2 Proceso de inducción con IPTG 1mM en <i>E. coli</i> Rosetta gami B.	104
Figura 5.3 Amplificación de la región codificante de PfEcp2 (+SP) y PfEcp2 (-SP) de transformantes de <i>E. coli</i> TOP10.	106
Figura 5.4 Secuencias nucleotídicas de los constructos pLATE-PfEcp2 (+SP) y pLATE-PfEcp2 (-SP).	106
Figura 5.5 Confirmación por PCR en colonia de las cepas transformadas de <i>E. coli</i> Rosetta gami B con los constructos de expresión pLATE-PfEcp2 (+SP) y pLATE-PfEcp2 (-SP). 107	

Figura 5.6 Curva estándar de cuantificación de proteínas por el método de Bradford...	108
Figura 5.7 Análisis SDS-PAGE del proceso de purificación de PfEcp2 (+SP) y de PfEcp2 (-SP).....	109
Figura 5.8 Análisis electroforético de las muestras colectadas en la purificación de la proteína recombinante PfAvr4 (-SP) en gel de 15% SDS-PAGE.....	110
Figura 5.9 Perfil de la fracción soluble de <i>E. coli</i> Rosetta gami B sin transformar inducida con IPTG de 12% SDS-PAGE.....	112

LISTADO DE CUADROS

Cuadro 1.1 Proteínas relacionadas con la patogénesis.....	11
Cuadro 1.2 Proteínas efectoras de hongos y oomicetos patogénicos en plantas (Ali y Bakkeren, 20112)	16
Cuadro 3.1 Predicción y clasificación del efectoroma de <i>P. fijiensis</i>	59
Cuadro 3 2 Clasificación del efectoroma de <i>P. fijiensis</i> de acuerdo al contenido de cisteínas.	60
Cuadro 3 3 Dominios funcionales identificados en la base de datos de efectores canónicos de <i>P. fijiensis</i>	61
Cuadro 3 4 Dominios funcionales identificados en la base de datos de efectores no canónicos de <i>P. fijiensis</i>	63
Cuadro 3 5 Dominios funcionales compartidos entre las bases de efectores canónica y no canónica de <i>P. fijiensis</i>	63
Cuadro 3 6 Predicción de localización subcelular de los efectores de <i>P. fijiensis</i>	64
Cuadro 3 7 Distribución de efectores candidatos con homología con otras especies.....	67
Cuadro 4 1 Candidatos a efectores de <i>P. fijiensis</i> seleccionados para los análisis funcionales.....	83
Cuadro 4 2 Cebadores utilizados para la amplificación de los candidatos a efectores de <i>P. fijiensis</i>	85
Cuadro 5 1 Distribución del efector Ecp2 en patógenos de clase de Dotideomicetos....	96
Cuadro 5 2 Cebadores utilizados para el desarrollo de este trabajo.....	101
Cuadro 5.3 Cuantificación de proteína en las alícuotas recuperadas durante la purificación de PfEcp2 (+SP) y PfEcp2 (-SP).....	107

RESUMEN

El banano forma parte de los principales cultivos en la producción y el comercio agrícola a nivel mundial, siendo parte de los primeros 5 de cultivos con más importancia a nivel mundial, producidos por países de regiones tropicales y subtropicales. Este cultivo es importante para muchos países como producto de alimentación básica y para exportación.

Sin embargo, el cultivo de bananos y plátanos son propensos a muchas plagas y enfermedades que pueden reducir drásticamente la producción y, por lo tanto, tales retos representan una amenaza para la seguridad alimentaria mundial. Uno de los problemas fitosanitarios más comunes es la enfermedad de la Sigatoka negra, la cual es causada por el hongo ascomiceto hemibiotrófico *Pseudocercospora fijiensis*.

Durante la fase de infección, *P. fijiensis* amplía la comunicación extracelular con la células huésped, lo que sugiere un papel importante para las proteínas fúngicas secretadas, muchas de estas, como celulasas, hidrolasas, oxidorreductasas, juegan un importante papel en la nutrición y crecimiento del hongo. Por otra parte, se conoce que el genoma de *P. fijiensis* contiene genes que codifican para proteínas relacionadas con patogénesis, entre ellas, las proteínas pequeñas ricas en cisteínas con actividad enzimática desconocida, mejor conocidas como efectores.

Los efectores son proteínas que cumplen funciones importantes de la patogenicidad y en *P. fijiensis* en la actualidad sólo se han caracterizado tres de ellas, por lo que es importante ampliar el número de proteínas efectoras caracterizadas. Este trabajo tiene el interés de profundizar en el conocimiento de la interacción de *P. fijiensis* con su hospedero empleando el uso de herramientas bioinformáticas, con el fin de encontrar alternativas de manejo de la enfermedad. Por medio del desarrollo de un pipeline bioinformático basado en lenguaje Perl, realizamos la identificación del efectograma de *P. fijiensis*, reportamos la predicción de 264 proteínas efectoras putativas para eventualmente analizar funcionalmente algunas de ellas.

ABSTRACT

Bananas are part of the main crops in agricultural production and trade worldwide, being part of the first five of the most important crops worldwide, produced by countries in tropical and subtropical regions. This crop is important to many countries as a staple food product and for export.

However, the cultivation of bananas and plantains are prone to many pests and diseases that can drastically reduce production and, therefore, such pose a threat to global food security. One of the most common phytosanitary problems is Black Sigatoka disease, which is caused by the hemibiotrophic ascomycete fungus *Pseudocercospora fijiensis*.

During the infection phase, *P. fijiensis* expands extracellular communication with host cells, suggesting an important role for secreted fungal proteins, many of these, such as cellulases, hydrolases, oxidoreductases, play an important role in nutrition and growth of the fungus. On the other hand, it is known that the *P. fijiensis* genome contains genes that code for proteins related to pathogenesis, among them, small proteins rich in cysteines with unknown enzymatic activity, better known as effectors.

Effectors are proteins that fulfill important functions of pathogenicity and in *P. fijiensis*, but at present only three of them have been characterized, so it is important to expand the number of characterized effector proteins. This work has the interest of deepening the knowledge of the interaction of *P. fijiensis* with its host using the use of bioinformatics tools, in order to find alternatives for managing the disease. Through the development of a bioinformatic pipeline based on Perl language, we performed the analysis of the putative effectorome, we reported the prediction of 264 putative effector proteins to eventually functionally analyze some of them.

INTRODUCCIÓN

La enfermedad Sigatoka negra, producida por *Pseudocercospora fijiensis*, hongo ascomiceto hemibiotrófico, es considerada una de las enfermedades más costosas y dañinas para el cultivo de banano en todo el mundo (Torrado *et al.*, 2008). La enfermedad afecta rápidamente el tejido foliar reduciendo la tasa fotosintética y en ausencia de control, puede causar la pérdida de más de 50% del peso del racimo, donde se deteriora la calidad del fruto por problemas de maduración de los mismos en campo y durante el proceso de transporte a los mercados destino (Martínez *et al.*, 2011). La infección por este hongo puede causar la reducción de los campos de banano en más de 50% y se controla principalmente en las variedades susceptibles con el uso extensivo de fungicidas. Estas aspersiones incrementan los costos de producción hasta en un 40% con un costo global cercano a los \$520 millones de dólares por año (Friesen, 2016).

El banano, así como otros hospedantes, son organismos sésiles que carecen de un sistema circulatorio, de células móviles para la defensa y de un sistema inmune adaptativo. Además, están sujetas al cambio de condiciones ambientales, incluyendo el constante ataque de agentes patógenos. Sin embargo, poseen mecanismos de defensa que van desde barreras físicas hasta potentes mecanismos moleculares de resistencia en cada célula y señales sistemáticas provenientes del sitio de la infección y que tienen marcadas similitudes con la inmunidad innata de los animales (Ramos y Portal, 2010).

Por ello, en la actualidad nos encontramos con reportes donde se plantean diversas estrategias para el reconocimiento e identificación de las moléculas que participan en el proceso de patogenicidad, moléculas conocidas como efectores.

Los efectores se describen como moléculas que incluyen proteínas, microARN, carbohidratos y metabolitos secundarios, potencialmente involucrados en el proceso de infección (Jones *et al.*, 2018; Sonah *et al.*, 2016; Sperschneider *et al.*, 2015; Win *et al.*, 2012). Las moléculas efectoras, generalmente proteínas, son importantes factores de virulencia que usan los hongos fitopatógenos, estos son secretados hacia la célula de la planta para modular la defensa y la fisiología del hospedador (Wang *et al.*, 2020).

Trabajos referentes a la identificación de efectores proteicos en *P. fijiensis*, se han llevado a cabo en modelos *in silico*. Stergiopoulos *et al.*, 2010 reporta los primeros ortólogos funcionales a efectores en *P. fijiensis*, PfAvr4 y PfEcp2, realizando la identificación con BlastP, obteniendo un

porcentaje de identidad del 42% y 57% respectivamente, usando como secuencias problema los efectores ya caracterizados de *Cladosporium fulvum*, Avr4 y Ecp2. Mientras que Chang et al., 2016, reportan el primer efectograma predicho de *P. fijiensis*, comprendido por 105 candidatos con un tamaño menor a 250 aminoácidos (Chang et al., 2016; Stergiopoulos et al., 2010).

En este capítulo, se presenta el desarrollo de un pipeline bioinformático que permite realizar la identificación de efectores de naturaleza proteica usando el proteoma total reportado en el JGI MycoCosm de *P. fijiensis*. En la predicción se emplearon criterios establecidos de acuerdo a la tendencia y descubrimiento de nuevos efectores. De la base de datos generada, se seleccionaron 12 candidatos para clonación y para análisis funcionales posteriores.

Se espera que con un mejor entendimiento de los efectores, se pueda profundizar el conocimiento de cómo se da el desarrollo de la enfermedad y con ello identificar o proponer nuevos blancos moleculares para la detección y/o control de la enfermedad.

CAPÍTULO I

ANTECEDENTES

1.1 GENERALIDADES DEL BANANO

Los bananos y plátanos (*Musa* spp) son plantas herbáceas, poliploides y perennes ampliamente adaptadas a regiones tropicales y subtropicales. En cuanto a banano, se distinguen tres partes importantes: el corno con hijuelos y el sistema radicular, el pseudotallo con el sistema foliar y el racimo o inflorescencia (Simmonds y Shepherd, 1995).

Los ancestros de todas las variedades triploides y tetraploides conocidas son dos especies diploides con 22 cromosomas cada una, *M. acuminata* y *M. balbisiana* Colla. El tamaño del genoma de *Musa* fue determinado como 550Mpb, en *M. balbisiana* y 600Mpb en *M. acuminata*. Este es más grande que los genomas de algunas especies como el arroz (*Oryza sativa*) o *Arabidopsis thaliana*, pero más pequeño que algunos cereales como la cebada o el centeno. Los 11 cromosomas que conforman el conjunto haploide son relativamente pequeños y todos tienen un tamaño similar con 50 Mpb (Osuji *et al*, 1997).

Éste fruto se produce en 16 entidades del país, Campeche, Colima, Chiapas, Guerrero, Hidalgo, Jalisco, Estado de México, Michoacán, Morelos, Nayarit, Oaxaca, Puebla, Quintana Roo, Tabasco, Veracruz y Yucatán. El cultivo ocupa una superficie de alrededor de 78,600 hectáreas, con una producción promedio en los últimos cuatro años de 2, 243,000 toneladas (SAGARPA, 2017).

Los cinco principales estados productores son Chiapas, Tabasco, Veracruz, Jalisco y Colima, con una producción estimada en 2, 045,000 toneladas en 2016, lo que significa el 84.5% del volumen nacional registrado (SAGARPA, 2017).

En lo que comprende a comercio exterior, el banano es exportado a 34 países, entre ellos Estados Unidos, Guatemala, Países Bajos, Reino Unido, España, Japón, Rusia, Corea del Sur, Albania, Túnez, Singapur, Nueva Zelandia, Portugal, Hong Kong, Bélgica, Italia, Cánada, Alemania, Irlanda, Lituania y Turquía, entre otros (SAGARPA, 2017).

1.2 Sigatoka negra

La Sigatoka negra es una enfermedad foliar del banano causada por el hongo ascomiceto *Pseudocercospora fijiensis* Morelet y constituye el principal problema fitopatológico del cultivo. La Sigatoka negra deteriora el área foliar de la planta, retrasa la floración y la cosecha, reduciendo tanto el llenado del racimo como la vida verde de la fruta (Crous, 2013; Osuji *et al.*, 1997).

Ésta enfermedad es la forma más severa del complejo Sigatoka, la cual comprende tres patógenos estrechamente relacionados: *Pseudocercospora musae*, causante de la Sigatoka amarilla, identificado por primera vez en la Isla de Java en 1902, *P. fijiensis*, agente causal de la enfermedad de la Sigatoka Negra, descubierto en Fiji en 1963 y *Pseudocercospora eumusae* que causa la enfermedad de la mancha foliar eumusae del banano (Churchill, 2011).

La Sigatoka es una enfermedad policíclica, en la que tanto conidios, producto del ciclo de reproducción asexual, como las ascosporas, producto del ciclo sexual, participan en su dispersión (Rodríguez y Cayón, 2008).

La disminución en la producción de frutos de plátano o bananos es debida principalmente a la pérdida de la superficie funcional de la hoja, que a su vez ocasiona la producción de fruta pequeña y de baja calidad y la maduración prematura del fruto con destino a la exportación (Rodríguez y Cayón, 2008).

La enfermedad se ha dispersado por todo el continente americano y el Caribe. En todos los países productores de bananos han ocurrido severas epidemias, que han obligado a intensificar las medidas de combate y han puesto de manifiesto la necesidad de estrategias de manejo integrado de la enfermedad, para mitigar su impacto negativo (Rodríguez y Cayón, 2008).

1.2.1 *Pseudocercospora fijiensis*

Pseudocercospora fijiensis es un hongo ascomiceto haploide, hemibiotrófico que causa la enfermedad de la raya negra de la hoja de banano BLSD (de las siglas de “*The Black Leaf Streak Disease*”), conocida también como Sigatoka negra, y es considerada como una de las enfermedades más costosas y dañinas del banano y plátano en todo el mundo (Torrado y Castaño, 2008).

El nombre de *M. fijiensis* le fue dado a la forma sexual (teleomorfa) del patógeno en el año de 1963 en muestras de Fiji (primer reporte de la enfermedad) (Leiva y Acosta, 2010; Rodríguez y Cayón, 2008). Sin embargo, la enfermedad estaba ampliamente difundida en el pacífico antes de su descubrimiento en la isla.

La cronología de la enfermedad, destaca diferentes registros alrededor del mundo, la primera aparición de la Sigatoka negra en América fue en Honduras en 1972, de ahí se fue extendiendo a toda América Latina, hasta llegar a México en el año de 1981 (Torrado y Castaño, 2008).

Recientemente, se acordó que el agente causal de la Sigatoka negra debe ser llamado *Pseudocercospora fijiensis* porque en la mayor parte del tiempo se encuentra en su estadio haploide y el género *Mycosphaerella* debe restringirse únicamente a especies que tienen teleomorfos (Crous *et al.*, 2013). Sin embargo, el cambio está siendo lento y la mayoría de los reportes referentes a la Sigatoka negra continúan llamando *M. fijiensis* al agente causal de la Sigatoka negra (Crous *et al.*, 2016).

En el “DOE Joint Genome Institute (JGI)”, se ha secuenciado el genoma de *P. fijiensis* (cepa CIRAD86). La versión 2.0 del ensamble de *P. fijiensis* obtuvo una mejora considerable en comparación con la versión 1.0. El genoma principal está en 56 estructuras de secuenciación ó “scaffolds” y aproximadamente la mitad del genoma está contenida en cinco de estas estructuras de mínimo 5.9Mb de longitud cada uno. El tamaño estimado del genoma es de 74.1 Mpb con un total de 13,107 genes estructural y funcionalmente anotados (Churchil, 2011; Kema, 2009).

1.2.2 Infección por *Pseudocercospora fijiensis*

Según Washington *et al.* (1998) el sitio de infección primaria de *P. fijiensis* es el envés de la hoja; esto concuerda con observaciones de otros autores, que reportan que la densidad de estomas en hojas de musáceas es mayor en el envés y en el ápice, sitios donde el tubo de germinación del hongo tiene una mayor probabilidad de penetrar (Leyva *et al.*, 2010).

Después de la penetración de las hojas por vía estomática, *P. fijiensis* inicialmente crece como un microorganismo biotrófico, exclusivamente colonizando el espacio intercelular entre las células mesófilas, obteniendo los nutrientes del hospedero en el apoplasto sin la formación de haustorios. Sin embargo, más tarde se convierte en necrotrófico, ocasionando el colapso del tejido del hospedero (Leiva *et al.*, 2010; Patiño *et al.*, 2006).

El ciclo de vida de la infección se compone de cuatro etapas distintas que están directamente relacionadas con el ciclo de vida del hongo, siendo éstas, supervivencia, multiplicación, dispersión e infección (Agrios, 2005). El hongo *M. fijiensis*, es el estado perfecto y pertenece a la clase Ascomycetae; mientras que *Pseudocercospora fijiensis*, es el estado imperfecto y pertenece a la clase de Hyphomycetae (Crous, 2009).

La infección se inicia con la deposición de las esporas, ya sean ascosporas o conidios, sobre las hojas. Los procesos de germinación y penetración del inóculo solo ocurren cuando hay condiciones favorables con humedad relativa de 90 a 100%, temperaturas de 26 a 28°C y, sobre todo, presencia de agua libre sobre las hojas. Bajo tales condiciones, las esporas germinan en un periodo de 2 a 6 horas.

De este modo se inicia el proceso de penetración que tarda de 2 a 3 días si las condiciones se mantienen favorables para el patógeno (Bornacelly, 2009).

Robert et al. (2012), reportaron que estudiaron 735 individuos de 37 países y se identificaron 25 poblaciones distribuidas en el mundo, divididas en cuatro áreas geográficas (Sudeste Asiático, Oceanía, América y África), donde se muestra que las poblaciones del patógeno mantienen un alto nivel de diversidad genética en el Sudeste Asiático, siendo éste la zona de origen, y que la recombinación desempeña un papel importante en la evolución de este patógeno. Conocer la magnitud y distribución de la variación genética de *P. fijiensis* es necesaria para los programas de mejoramiento y el manejo de la resistencia de la Sigatoka negra (Santana et al., 2012).

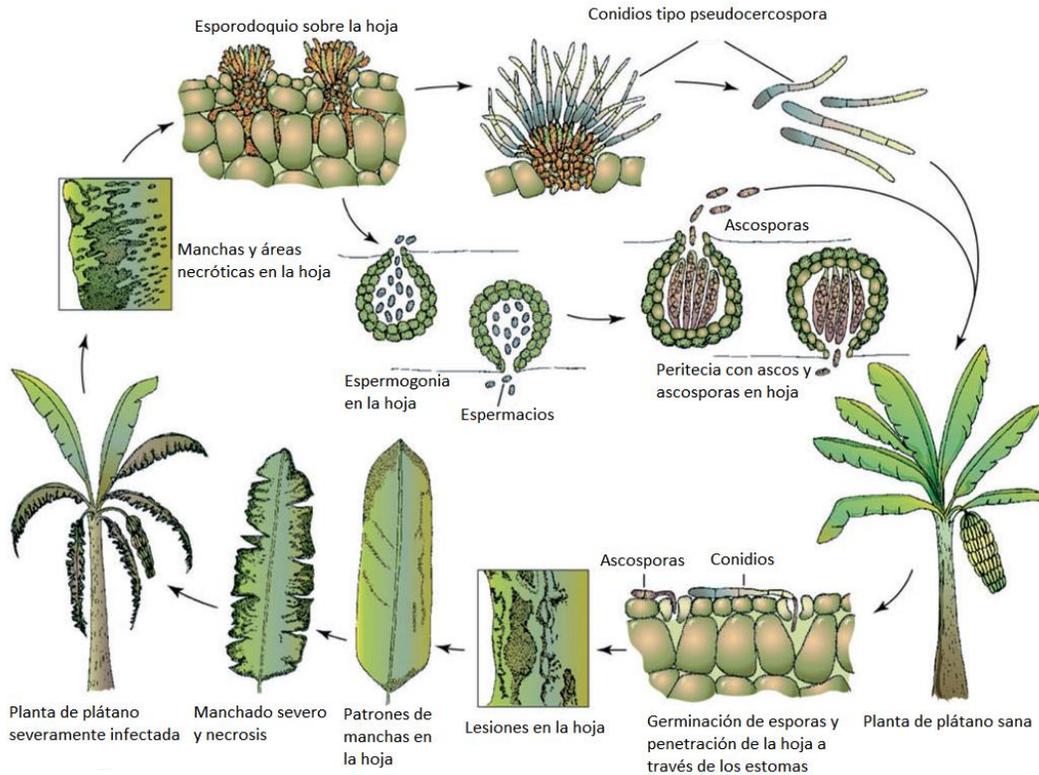


Figura 1.1 Ciclo de vida de *Pseudocercospora fijiensis* Morelet, agente causal de la Sigatoka negra del banano, tomado de Churchill, 2011.

La Sigatoka es una enfermedad policiclica, en la que tanto los conidios como las ascosporas, participan de la dispersión. Los conidios se forman en lesiones jóvenes como estrias y aparecen en el primer estadio de mancha. Se producen más abundantemente en periodos de alta humedad y especialmente si se cuenta con una película de agua libre sobre las hojas; se asocian principalmente con infecciones a corta distancia, entre hojas de una misma planta, de la planta madre a los hijos y entre plantas cercanas. Las ascosporas en cambio, son los medios primarios de dispersión sobre distancias más largas entre las plantaciones durante los periodos de lluvia (Agrios, 2004; Ploetz, 1994)

Los síntomas de la enfermedad se han clasificado en seis etapas que reflejan la severidad de la enfermedad: etapa 1, manchas cafés pequeñas de 0.25mm de diámetro que aparecen sobre el envés de la hoja; etapa 2, las manchas se elongan a rayas de 1mm de ancho por 2mm de largo; etapa 3, las rayas se vuelven más largas, oscuras en el envés y visibles en la superficie superior

como rayas amarillas; etapa 4, sobre la superficie superior, las rayas cafés oscuras aparecen rodeadas por una zona amarillenta pálida; etapa 5, los centros negros de las manchas comienzan a destruirse mientras que lo rodea forma un anillo clorótico; etapa 6, los centros de las manchas se secan formando una mancha gris rodeada por un anillo negro bien definido que es a su vez rodeado por un anillo clorótico (CIAT, 2013; Fouré,1987).

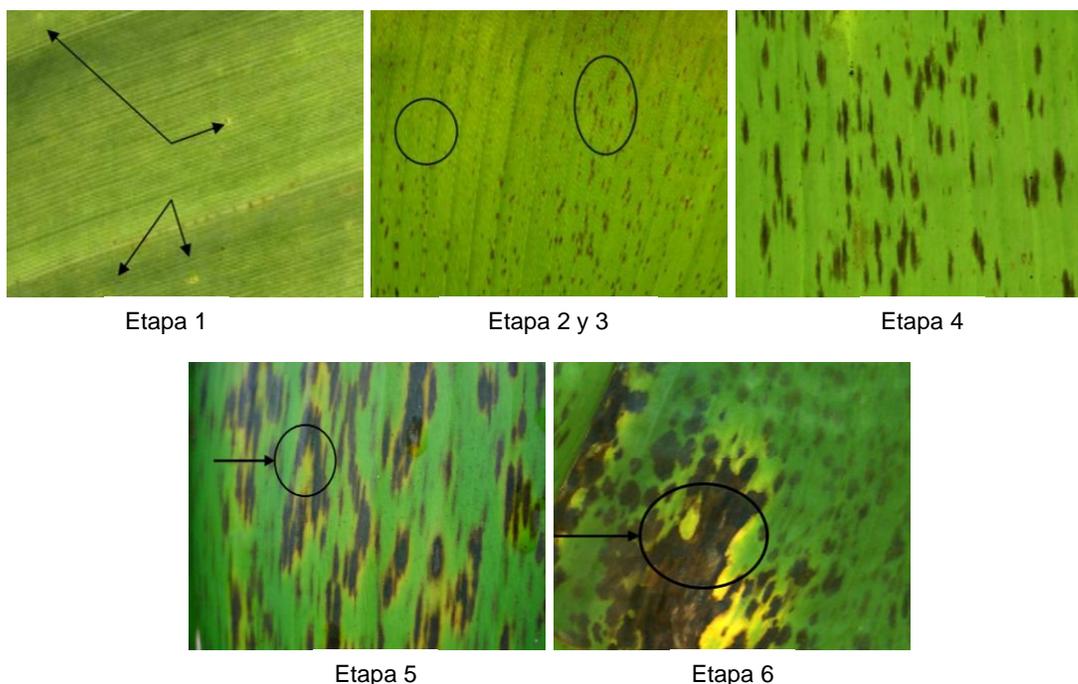


Figura 1.2 Estados visuales del avance de la infección causada por la Sigatoka negra (CIAT, 2013; Díaz y Almodóvar, 2007; Fouré,1987).

1.3 Interacciones planta-patógeno

Las plantas a diferencia de los animales, son organismos sésiles que carecen de un sistema circulatorio, de células móviles para la defensa y de un sistema inmune adaptativo. Además, están sujetas al cambio de condiciones ambientales, incluyendo el constante ataque de agentes patógenos. Sin embargo, poseen mecanismos de defensa que van desde barreras físicas (películas de cera en la superficie de sus órganos, paredes celulares rígidas, etc.), hasta potentes mecanismos moleculares de resistencia en cada célula y señales sistémicas provenientes del sitio de la infección que tienen marcadas similitudes con la inmunidad innata de los animales (Ramos y Portal 2010).

Como parásitos, los fitopatógenos están obligados a obtener los nutrientes de fuentes existentes y para ello necesitan invadir y adaptarse al tejido del hospedante para eventualmente

reproducirse (Sperschneider *et al.*, 2015). Por ello, la habilidad que posee un patógeno para causar enfermedad a una planta hospedera es medida según la respuesta de la planta. Si la infección toma lugar con un subsecuente desarrollo de la enfermedad, la planta es considerada como susceptible al patógeno. La susceptibilidad puede ser causada por una falta de capacidad de la planta para reconocer al patógeno y/o producir una respuesta de defensa efectiva y rápida (Agrios, 2004). En este caso, la interacción entre la planta y el patógeno es llamada interacción compatible. En caso contrario, las plantas son capaces de restringir el desarrollo del patógeno en el sitio inicial de infección, son llamadas resistentes y la interacción es incompatible (Agrios, 2004).

1.3.1 Inmunidad activada por PAMPs, MAMPs o DAMPs

La resistencia vegetal se basa en mecanismos complejos de reconocimiento a nivel bioquímico del potencial del patógeno y después se desencadena la respuesta molecular dando la consecuente transducción de señales al interior de la célula vegetal. Las plantas han desarrollado varias estrategias para detectar a los patógenos (Ramos y Portal, 2010; Jones y Dangl, 2006). Sobre la superficie de la membrana de la célula de la planta existen receptores que reconocen componentes típicos de los patógenos, que son importantes para la penetración y colonización del patógeno. Estos componentes se denominan inductores y más recientemente se han nombrado como patrones moleculares asociados al patógeno o al microorganismo (PAMPs del inglés *Pathogen-Associated Molecular Patterns*, o MAMPs, del inglés *Microbe Associated Molecular Patterns*). Los PAMPs o MAMPs, son componentes esenciales de microorganismos, tales como la flagelina bacterial o la quitina en los hongos. Las plantas también responden a moléculas endógenas relacionadas con la invasión del patógeno, tales como fragmentos de la pared celular o cuticular llamados patrones moleculares asociados al daño (DAMPs del inglés, *Damage Associated Molecular Patterns*). Los PAMPs, MAMPs y DAMPs, son reconocidos por proteínas receptoras llamadas receptores de reconocimiento de patrones (PRRs) (Boller y Felix 2009). La estimulación de los PRRs induce resistencia del tipo inmunidad basal activada por PAMP, MAMP o DAMP (PTI, del inglés, *PAMP- Triggered Immunity*). La segunda percepción es llevada a cabo por receptores intracelulares e involucra un reconocimiento de moléculas de avirulencia del patógeno llamadas proteínas efectoras, este reconocimiento induce la inmunidad activada por efectores (ETI, del inglés, *Effector Triggered Immunity*). ETI usualmente induce la respuesta hipersensible, la cual es una forma de muerte celular programada, similar a la apoptosis presente en mamíferos. Este modelo de reconocimiento lleva una dinámica co-evolutiva entre la planta y el patógeno, éste último busca evadir el reconocimiento de la planta y

la primera busca reconocer al patógeno en los estados más tempranos de la infección (Lethi *et al.*, 2009; Jones y Dangl, 2006).

1.3.2 Resistencia mediada por genes R

El modelo denominado gen por gen, propuesto originalmente por Harold Flor (Flor, 1942), postula que la resistencia de una planta a un patógeno ocurre cuando la planta posee un gen dominante de resistencia llamado R y el patógeno expresa un gen complementario dominante de avirulencia llamado Avr. La infección sólo ocurre cuando la planta carece del alelo dominante del gen R. En este caso se habla de interacción compatible, mientras que, cuando la infección no progresa, la interacción se denomina incompatible. A partir de la década de 1990, se comenzaron a identificar numerosos genes de resistencia y en la actualidad se conoce un gran número de ellos. La mayor parte de ellos codifican proteínas R de localización citoplasmática que contienen un dominio llamado NBS (por sus siglas en inglés *Nucleotide Binding Site*) que une nucleótidos y otro dominio rico en leucinas llamado LRR (por sus siglas en inglés *Leucine-Rich Repeats*). Otros genes de resistencia que codifican proteínas R capaces de participar en la detección de proteínas de avirulencia del patógeno, codifican receptores del tipo RLK (por sus siglas en inglés *Receptor-Like Kinase*) y RLP (por sus siglas en inglés *Receptor-Like Protein*) (Gudesblat, 2007).

1.3.3 Respuesta hipersensible (HR)

La reacción de hipersensibilidad se considera como la máxima expresión de resistencia de las plantas al ataque por patógenos. Durante la RH se disparan reacciones que llevan a la muerte a las células que rodean el sitio donde penetró el patógeno y eso detiene el avance del patógeno y la infección. Así la RH forma parte de los mecanismos de defensa de la planta.

La muerte de células durante la RH se considera una muerte celular programada (MCP). Normalmente muchas de las células de los organismos eucariotes mueren y son removidas de manera programada a través de una serie de cambios morfológicos y bioquímicos. La MCP es fundamental en procesos relacionados tanto con el crecimiento y desarrollo normal del organismo, así como con la respuesta a estrés por factores bióticos y abióticos (Coll y Dangl, 2011).

La RH se define como una muerte rápida de células asociada con la restricción del crecimiento de patógenos y generalmente se reconoce por la presencia de una o varias células muertas con coloración café en el sitio de infección. La lesión café puede visualizarse a nivel macroscópico

cuando involucra a un número suficiente de células; no obstante, en algunos casos la necrosis solamente es visible al microscopio cuando son pocas las células involucradas en la RH.

Trabajos previos han mostrado que las plantas en las que se dispara la RH, muestran cierto grado de resistencia a patógenos en tejidos distantes al sitio donde ocurrió la reacción; este tipo de protección se le conoce como resistencia sistémica adquirida (RSA) y el ácido salicílico (AS) es, al parecer esencial para su inducción (Sanzón y Zavaleta, 2011).

Estudios con mutantes de maíz (*Zea mays*), tomate (*Solanum lycopersicum*) y *Arabidopsis thaliana* (L) Heynh, han revelado que la RH también depende de genes adicionales que parecen estar presentes tanto en individuos resistentes como susceptibles, y que le confieren a la planta la habilidad de responder hipersensitivamente aún en situaciones donde no existe la relación R-Avr (Coll y Dangl, 2011; Sanzón y Zavaleta, 2011).

1.3.4 Resistencia sistémica adquirida e inducida

El contacto con microorganismos patogénicos y no patogénicos puede activar un amplio rango de mecanismos de defensa en partes de la planta distantes al sitio de reconocimiento. Los mecanismos de este tipo más estudiados son la resistencia sistémica adquirida (SAR, del inglés *Systemic Acquired resistance*) y la resistencia sistémica inducida (ISR, del inglés *Induced systemic Resistance*). La SAR puede ser activada por una infección local o por moléculas químicas, generando una resistencia, esta respuesta es correlacionada con la activación de genes PR (del inglés, *Pathogenesis Related*), y requiere la participación del ácido salicílico (SA). La ISR es producida como resultado de la colonización de raíces por algunas bacterias no patogénicas de la rizósfera. ISR requiere componentes de la ruta de señalización del ácido jasmónico (JA) y del etileno (ET) (Durrant y Dong, 2004). El Cuadro 1.1, hecho por Sels y colaboradores en 2008, presenta la lista de proteínas que han sido descritas como relacionadas a la patogénesis (PR protein families).

Cuadro 1.1 Proteínas relacionadas con la patogénesis.

Familia	Miembro	kDa	Función	Objetivo
PR-1	Tabaco PR-1a	15	Antifúngico	Desconocido

PR-2	Tabaco PR-2	30	B-1,3-Glucanasa	B-1,3-Glucano
PR-3	Tabaco P,Q	25-30	Quitinasa (Clase I, II ,IV ,V ,VI)	Quitina
PR-4	Tabaco 'R'	15-20	Quitinasa clase I,II	Quitina
PR-5	Tabaco S	25	Thaumatín-like	Membrana
PR-6	Inhibidor I tomate	8	Inhibidor de proteínasa	n_
PR-7	P69 tomate	75	Endoproteínasa	
PR-8	Quitinasa de pepino	28	Quitinasa clase III	Quitina
PR-9	Tabaco 'peroxidasa formadora de lignina'	35	Peroxidasa	n_
PR-10	'PR1' de Perejil	17	'ribonucleasa-like'	n_
PR-11	Tabaco 'clase V' quitinasa	40	Quitinasa clase I	Quitina
PR-12	Rs-AFP3 de Rabano	5	Defensina	Membrana
PR-13	THI2.1 <i>Arabidopsis</i>	5	Tionina	Membrana
PR-14	LTP4 cebada	9	Proteína transferencia de lípidos	Membrana
PR-15	OxOa Cebada(germen)	20	Oxalato oxidasa	n_
PR-16	OxOLP cebada	20	Similar a Oxalato oxidasa	n_
PR-17	Tobaco PRp27	27	Desconocido	n_

n_ No hay reporte de ensayo *in vitro* de la actividad antimicrobiana

En el reporte de Sels y colaboradores en 2008 se puede ver que las proteínas PR contienen diferentes dominios funcionales, como oxidasas, inhibidores de proteínasa hasta defensinas.

Las plantas están expuestas a innumerables patógenos por lo que se han generado varios mecanismos de inmunidad, lo que lleva a que sólo un grupo de plantas se infecten con un determinado patógeno.

Kirankumar y Choong-Min (2004), propusieron que éste tipo de resistencia exhibida contra bacterias, hongos y oomicetos, puede presentarse de dos maneras. El Tipo I es aquella resistencia que no produce síntomas visibles, al contrario del tipo II, donde resulta en una rápida RH seguida por la muerte celular. La resistencia del hospedero es codificada por los genes R. Por otro lado, la resistencia “por el no hospedero” (NHR) es la forma de resistencia que presentan todas las especies contra los patógenos no-adaptados (Nürnberg y Lipka, 2005). La resistencia no hospedera es la forma más común de resistencia ante enfermedades exhibidas contra la mayoría de los microorganismos con potencial patogénico (Sumit, 2013).

En la respuesta de defensa inducida, se comprometen las barreras físicas (constitutivas), químicas y reacciones inducibles. Las barreras constitutivas comprenden las capas de ceras, cutículas, pared celular y metabolitos secundarios con actividad antimicrobiana como las fitoanticipinas. Adicionalmente, el citoesqueleto junto con la conectividad de la membrana plasmática con la pared celular son un complejo preformado esencial sensible a la resistencia del no hospedero (*Nonhost resistance*, NHR) (Mysore y Ryu, 2004). La acumulación de especies reactivas de oxígeno, fitoalexinas y productos de la traducción de genes relacionados a patogénesis (PR) y la respuesta hipersensible (muerte celular programada), son algunas de las respuestas químicas y reacciones inducibles presentes en la interacción del tipo “no hospedera” (Johansson, 2015). Es decir, la inoculación de un patógeno en una planta no hospedera genera la RH asociada con “la resistencia no-hospedero”. Sin embargo en algunos casos ésta resistencia no está asociada con la inducción de RH.

1.4 Efectores involucrados en la infección del hospedante

Se ha descrito que la colonización exitosa del tejido del hospedante por el patógeno, depende de si éste tiene la capacidad de ejercer por lo menos alguna de las siguientes estrategias: a) capacidad de evadir el sistema de detección de la planta de las moléculas inductoras “que lo pueden delatar” , o producirlas bajo “camuflaje”, de forma que no sean detectadas por el hospedante; b) producir las moléculas inductoras para poder interferir con las respuestas de defensa, por ejemplo PsIsc1 y VdIsc1 (efectores de *Phytophthora sojae* y *Verticillium dhaliae* respectivamente), que interfieren y bloquean la producción del ácido salicílico (Liu *et al.*, 2014); c) detoxificación de compuestos antimicrobianos; d) poseer una tasa alta de crecimiento (o de movilización) de modo que pueda alejarse rápidamente del sitio en el que se están dando con mayor intensidad la respuesta de defensa; e) tener un hábito alimenticio que se acerque más al extremo de los necrotróficos, de manera que al matar un área extensiva de células hospedantes por la producción de grandes cantidades de toxinas o enzimas interfiera con las respuestas de defensa activa (producción de fitoalexinas y otros metabolitos antimicrobianos, engrosamiento de paredes celulares, acumulación de calosa, entre otras) que llevan a cabo las células vivas vecinas (Stergiopoulos *et al.*, 2010).

Un patosistema ampliamente estudiado como modelo del proceso de defensa, es el de *Cladosporium fulvum* y tomate, el cual cumple con el sistema de reconocimiento gen-por-gen. Durante la infección del tomate, *C. fulvum* secreta proteínas efectoras en el apoplasto, las cuales actúan como factores de virulencia (Figura 1.3). Sin embargo, estos efectores son reconocidos por los cognados (Cf) de *C. fulvum*, proteínas de resistencia (R) presentes en líneas resistentes de tomate, y ese reconocimiento conduce a una rápida respuesta de defensa de la planta que culmina en la respuesta hipersensible (Rovenich, 2014; Stergiopoulos *et al.*, 2010).

Se han clonado cuatro genes de avirulencia (Avr) de *C. fulvum*, los cuales codifican para proteínas pequeñas ricas en cisteínas que son secretadas durante la infección. Las proteínas efectoras son Avr2, Avr4, Avr4E y Avr9, dándose el reconocimiento de estas en tomate mediante los cognados proteínas Cf respectivas, es decir, Cf-2, Cf-4, Cf-4E y Cf-9. Se han descrito también seis proteínas extracelulares, llamadas Ecp1, Ecp2, Ecp4, Ecp5, Ecp6 y Ecp7, que han sido caracterizadas en *C. fulvum* y provocan una RH en líneas de tomate, que poseen los respectivos cognados de resistencia Cf-Ecp (Rovenich, 2014; Stergiopoulos *et al.*, 2010).

Stergiopoulos *et al.* (2010), reportaron los homólogos funcionales de *P. fijiensis*, Avr4, Ecp2 y Ecp6, hasta el momento han sido estudiados los efectores Avr4 y Ecp6 (Kohler *et al.*, 2016; Rovenich, 2014; Stergiopoulos *et al.*, 2010; Bolton *et al.*, 2008). El Cuadro 1.3 presenta las principales características de los efectores elucidados en varios microorganismos (Basado en Ali y Bakkeren en 2012).

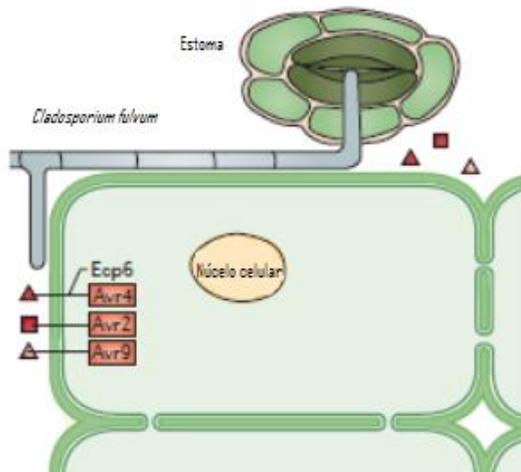


Figura 1.3 Proceso de infección de *Cladosporium fulvum*. Hongo ascomiceto biotrófico del tomate, entra en el tejido de la hoja a través de los estomas y crece como hifas extracelulares; Ecp6, secuestra oligosacáridos de quitina de la pared celular del hongo; Avr2, inhibe una proteasa de cisteínas secretada por las plantas de tomate; Avr4, contiene un dominio diferente de unión a quitina y funciona para proteger la pared celular fúngica de la degradación de la planta por quitinasas; Avr9, aún no se conoce su función.

Tomado de Giraldo y Valent, 2013.

Cuadro 1.2 Proteínas efectoras de hongos y oomicetos patogénicos en plantas (Ali y Bakkeren, 20112)

Proteína	Organismo	Longitud de residuos de aa (maduro)	Número de cisteínas	Longitud SP (aa)	Actividad Biológica/ Homología	Localización de la proteína	Rol de virulencia/ Patogenicidad	Gen R de resistencia	Referencia
AVR 2	<i>C. fulvum</i>	78 (58)	8	20	Induce RH; inhibidor de proteasa	Apoplasto	Inhibe Rcr3 y otras proteasas	<i>Cf-2</i>	Stergiopoulos <i>et al.</i> , 2010
AVR 4	<i>C. fulvum</i>	135 (86)	8	18	Induce RH; Unión a quitina, ortólogo en algunos otros hongos	Apoplasto; célula fúngica, pared de quitina	Protege contra quitinasas	<i>Cf-4</i>	Joosten <i>et al.</i> , 1994; Stergiopoulos <i>et al.</i> , 2010; Van den Burg <i>et al.</i> , 2006
AVR 9	<i>C. fulvum</i>	63 (28)	6	23	Induce RH; Inhibidor de carboxypeptidasa	Apoplasto	Desconocido	<i>Cf.9</i>	van den Ackerveken <i>et al.</i> , 1993
NIP 2	<i>R. sacalis</i>	109 (?)	7 (6)	16	Toxina no específica / induce necrosis en numerosas plantas spp.	Probablemente apoplasto	No necesaria para la total virulencia	<i>Desconocido</i>	Rohe <i>et al.</i> , 1995; Stergiopoulos y de Wit, 2009
NIP 3	<i>R. sacalis</i>	115 (?)	9 (8)	17	Toxina no específica / induce necrosis en numerosas especies de plantas	Probablemente apoplasto	No necesaria para la total virulencia	<i>Desconocido</i>	Rohe <i>et al.</i> , 1995; Stergiopoulos y de Wit, 2009
PWL 1	<i>M. oryzae</i>	147 (124)	2	23	Proteína hidrofílica rica en glicerina	Complejo interfacial biotrófico	Desconocido	<i>Desconocido</i>	Kang <i>et al.</i> , 1995; Khang <i>et al.</i> , 2010
PWL 2	<i>M. oryzae</i>	145 (124)	2	21	Proteína hidrofílica rica en glicerina	Citoplasma	Desconocido	<i>Desconocido</i>	Sweigard <i>et al.</i> , 1995; Khang <i>et al.</i> , 2010
PWL 3	<i>M. oryzae</i>	137 (116)	0	21	Proteína hidrofílica rica en glicerina	Probablemente apoplasto	No funcional	<i>Desconocido</i>	Kang <i>et al.</i> , 1995
AVR 3 (SIX 1)	<i>F. oxysporum f. sp. lycopersici</i>	283 (122)	6 u 8	21	Desconocido	Xilema	Requerido para la total virulencia	<i>I-3</i>	Rep <i>et al.</i> , 2004; Stergiopoulos y de Wit, 2009
AVR 4 (SIX2)	<i>F. oxysporum f. sp. lycopersici</i>	232 (172)	8	20	Desconocido	Xilema	Requerido para la total virulencia	<i>Desconocido</i>	Houterman <i>et al.</i> , 2007; Stergiopoulos y de Wit, 2009
AVR 2 (SIX 3)	<i>F. oxysporum f. sp. lycopersici</i>	163 (144)	3 (2)	19	Induce RH; desconocido; motivo funcional RXLR	Xilema	Requerido para la total virulencia	<i>I-2</i>	Houterman <i>et al.</i> , 2007; Stergiopoulos y de Wit, 2009
MfAVR4	<i>P. fijiensis</i>	121 (100)	10	21	Peritrofina-A con unión a quitina, inducción de RH o necrosis en tomate Cf4	Probablemente apoplasto	Protección del hongo contra quitinasas	<i>Cf-4 y Hcr9</i>	Stergiopoulos <i>et al.</i> , 2010
MfECP2	<i>P. fijiensis</i>	161 (142)	4	19	Inducción de RH o necrosis en tomate CfEcp2	Probablemente apoplasto	Promueve virulencia por la interacción con la célula hospedera causando necrosis	<i>CfEcp-2</i>	Stergiopoulos <i>et al.</i> , 2010

1.4.1 Efectoroma

Los hongos ocupan diversos nichos ambientales y muchos han evolucionado para vivir un estilo de vida patogénica, causando enfermedades devastadoras en las plantas y los animales. Como se han mencionado, la interacción entre el huésped y el patógeno es complejo y se encuentra en constante evolución. Se sabe que los patógenos secretan proteínas efectoras que “manipulan” al hospedero en su beneficio.

Dependiendo de la estrategia de infección, los microorganismos patogénicos pueden expresar efectores en el apoplasto y/o efectores citoplasmáticos que son introducidos a las células de la planta (Giraldo y Valent, 2013).

En el caso de la familia de los efectores Crinkler (CRN) presentes en *Phytophthora infestans*, se han predicho con el uso de herramientas bioinformáticas, alrededor de 196, en otras especies como *P. ramorum* 60, *P. sojae* 202 y 26 en *Pythium altimum* (Stam *et al.*, 2013). Las cepas de bacterias fitopatógenas pueden tener más de 30 efectores cada una; miles de nuevos efectores pueden existir en bacterias, hongos y nematodos (Alfano, 2009).

Los estudios moleculares han identificado más de 60 efectores fúngicos; Sin embargo, esto representa sólo la punta del iceberg. Por ejemplo, hasta ahora sólo seis efectores se han caracterizado de tres hongos de la roya, mientras que más de 30 Avr han sido identificadas en la roya del lino y alrededor de 50 en cada una de las royas del tallo, roya amarilla y la roya de la hoja. Similarmente, a partir de estudios transcriptómicos se han identificado alrededor de 40 Avr en las interacciones entre *Magnaporthe oryzae* y arroz (Sperschneider *et al.*, 2015).

A partir del análisis y comparación de los genomas y de transcriptomas obtenidos a partir de micelios cultivados *in vitro* de los hongos del complejo Sigatoka (*P. musae*, *P. eumusae* y *P. fijiensis*), se predijeron 612, 638 y 585 proteínas secretadas respectivamente, entre las cuales se identificaron 110, 112, 105 probables efectores, con un tamaño comparable a la mayoría de los efectores de otros hongos hemibiotróficos (Chang *et al.*, 2016).

Kettles y colaboradores en 2017, usando análisis bioinformáticos y funcionales, reportaron alrededor de 63 posibles efectores de *Zymoseptoria tritici*, de los cuales, 14 indujeron muerte celular o clorosis en *N. benthamiana*. Estos resultados indican que hay un reconocimiento de los efectores apoplásticos de *Z. tritici* por presumibles receptores en la superficie celular de *N. benthamiana*, una especie vegetal no huésped y taxonómicamente distante del trigo. Esto

sugiere que en el apoplasto existe un reconocimiento de múltiples efectores de “patógenos no adaptados” y que pueden contribuir a la RH de los “no hospederos”.

1.4.2 Búsqueda de efectores

Actualmente se han desarrollado herramientas que ayudan en la identificación de nuevos efectores, una de estas herramientas es la bioinformática. El desarrollo de la bioinformática y sus técnicas han contribuido al rápido desarrollo de investigaciones sobre las interacciones moleculares planta-microorganismo (Humira *et al.*, 2016).

Éste tipo de análisis se basa en las características reportadas de los efectores, dada a la falta de secuencias conservadas, los enfoques de identificación se han basado en criterios como el tener un péptido señal de secreción y ser proteínas pequeñas ricas en cisteínas. Las predicciones computacionales sirven como un excelente punto de partida, para posteriormente continuar con un análisis funcional para la validación, evolución y caracterización (Humira *et al.*, 2016; Sperschneider *et al.*, 2015).

Trabajos como el de Stergiopoulos y colaboradores en 2010, identificaron el homólogo funcional de Avr4 de *C. fulvum* en *P. fijiensis* y otros tres homólogos funcionales de Ecp2, donde uno de ellos inducía RH en una línea de tomate transformada con el gen de resistencia Cf-Ecp2.

Chang y colaboradores en 2016, identificaron *in silico* efectores en el complejo Sigatoka, usando un amplio arsenal de programas (*The SignalP, TMHMM, TargetP, Phobius, ProScan, The WoLF PSORT*), para identificar proteínas con péptido señal, dominios transmembranales, dominios funcionales, localización celular. Obtuvieron como candidatos putativos a 110 efectores en *P. musae*, 112 en *P. eumusae* y 105 en *P. fijiensis*.

Dados esos antecedentes, en el presente trabajo se emplearon herramientas bioinformáticas para la identificación y caracterización de probables proteínas efectoras presentes en *P. fijiensis*, se espera que los resultados e información obtenida de estos análisis contribuya en un mejor entendimiento de la patogénesis de este importante fitopatógeno hemiobiotrófico, sin embargo, cabe mencionar que la etapa de los análisis funcionales usando plantas de *N. benthamiana* quedaron incompletos puesto que con la pandemia originada por COVID-19 no pudieron ser completados (Figura 1.4).

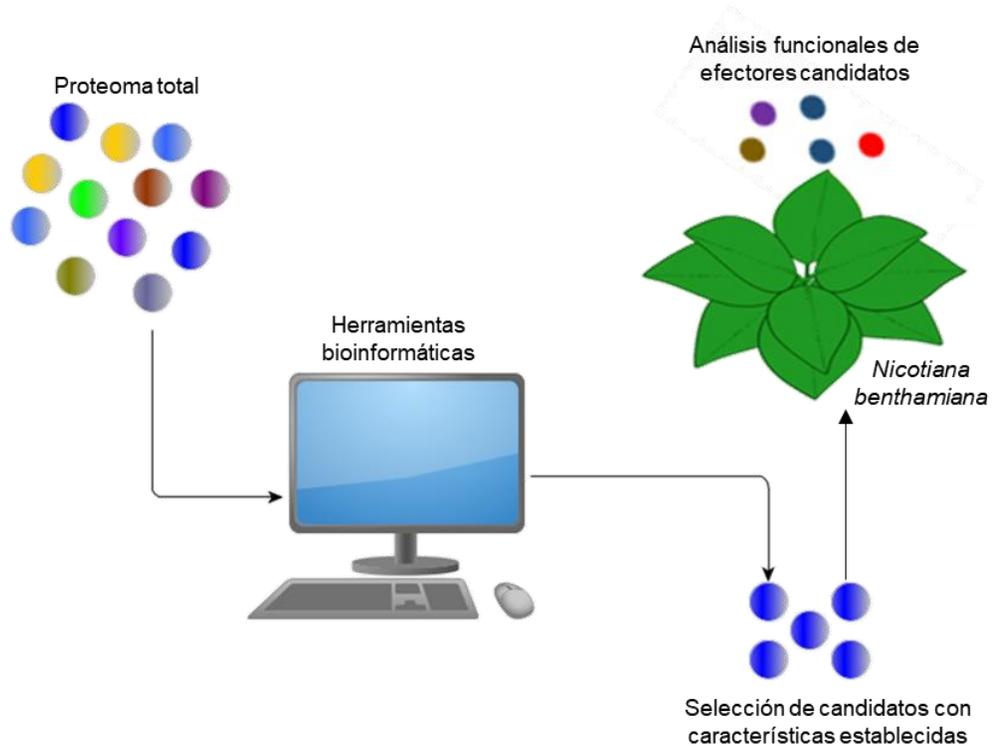


Figura 1.4 Esquematización del análisis in silico del proteoma total de *P. fijiensis*. Esquema representativo del flujo de trabajo empleado para la identificación del repertorio de posibles efectores del genoma de *P. fijiensis*.

JUSTIFICACIÓN

Pseudocercospora fijiensis es uno de los hongos patógenos más importantes para el cultivo de banano, pero hasta el momento solo se han caracterizado funcionalmente tres proteínas efectoras, por lo que es importante identificar nuevos efectores, para profundizar en el conocimiento del patosistema que permita identificar nuevos blancos moleculares para el control de la enfermedad de la Sigatoka negra.

La presente investigación se enfoca a la identificación mediante la aplicación de herramientas bioinformáticas y funcionales de genes que codifican presuntas proteínas con función efectora. Para ello se toman en cuenta las características antes reportadas en trabajos desarrollados en microorganismos modelo como *Cladosporium fulvum*, *Fusarium oxysporum*, *Magnaporthe oryzae*, *Colletotrichum higginsianum*, *Phytophthora infestans*, entre otros, en los cuales se han identificado numerosas proteínas involucradas en el proceso de patogenicidad en las interacciones de los patógenos con su hospedero.

HIPÓTESIS

Pseudocercospora fijiensis tiene múltiples proteínas efectoras.

OBJETIVO GENERAL

Identificar proteínas efectoras de *Pseudocercospora fijiensis* mediante análisis *in silico* y funcional.

OBJETIVOS ESPECÍFICOS

1. Identificar *in silico* proteínas de *Pseudocercospora fijiensis* como efectores.
2. Amplificar y clonar las secuencias codificantes de los candidatos a efectores.
3. Expresar *in vitro* las proteínas efectoras ya conocidas.
4. Realizar análisis funcional de los presuntos efectores mediante agroinfiltración y coinfiltración en hoja de banano cv Calcuta-4 y plantas de *Nicotiana benthamiana*.

ESTRATEGIA EXPERIMENTAL

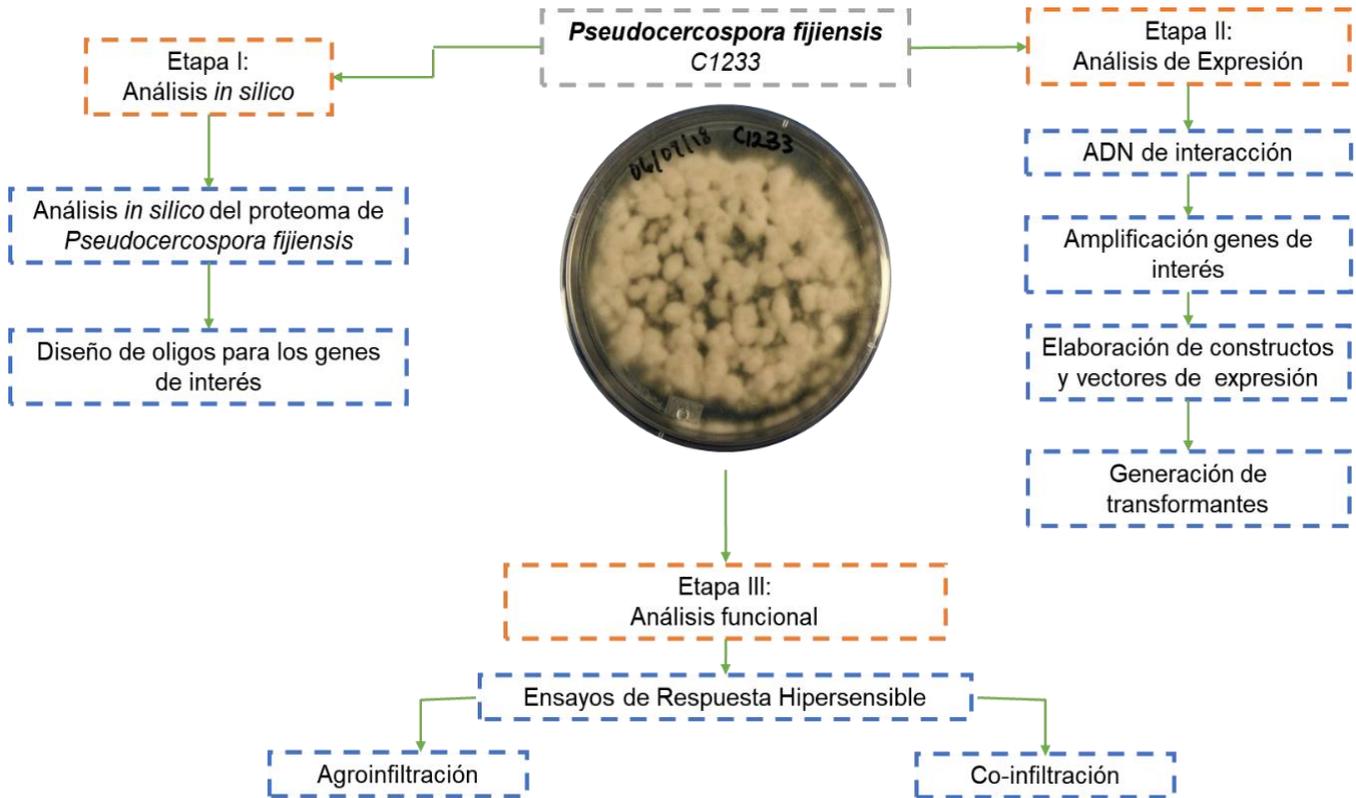


Figura 1.5 Esquema general del diseño experimental. El trabajo de tesis se llevará a cabo en tres etapas, I Análisis *in silico*; II Expresión de las proteínas y III Análisis funcional.

CAPÍTULO II**Este capítulo fue aceptado y publicado en la revista Biomoléculas del MDPI**

Karla Gisel Carreón-Anguiano, Ignacio Islas-Flores, Julio Vega-Arreguín, Luis Sáenz-Carbonell and Blondy Canto-Canché. EffHunter: A Tool for Prediction of Effector Protein Candidates in Fungal Proteomic Databases. *Biomolecules*. 2020 May 4; 10(5):712. <https://doi.org/10.3390/biom10050712>. PMID: 32375409; PMCID: PMC7277995.

*Correspondence: cantocanche@cicy.mx; Tel.: +52-999-942-8330

Received: 9 January 2020; Accepted: 21 March 2020; Published: 4 May 2020

Abstract: Pathogens are able to deliver small-secreted, cysteine-rich proteins into plant cells to enable infection. The computational prediction of effector proteins remains one of the most challenging areas in the study of plant fungi interactions. At present, there are several bioinformatic programs that can help in the identification of these proteins; however, in most cases, these programs are managed independently. Here, we present EffHunter, an easy and fast bioinformatics tool for the identification of effectors. This predictor was used to identify putative effectors in 88 proteomes using characteristics such as size, cysteine residue content, secretion signal and transmembrane domains.

Keywords: computational prediction; host-pathogen interaction; effector proteins; fungal secretome

1. Introduction

Fungal phytopathogens are a major threat to food security since they can cause devastating losses to important crops in agriculture.

These pathogenic fungi secrete diverse small proteins in the infection process which are pathogenicity or virulence determinants that manipulate the interaction, and thus are commonly referred to as “effectors” [1–4]. Small, secreted, cysteine-rich proteins constitute a common source of fungal effectors [3,5–7].

Fungal effector proteins are poorly conserved, and in contrast to oomycetes effectors, where the presence of conserved amino acid motifs (e.g., RxLR, dEER) has created profile Hidden Markov Models (HMM), the prediction of fungal effectors has been more challenging. In general, most

fungus effectors do not share significant sequence similarity with each other, which can be attributed to rapid divergence and host specialization. However, they share structural properties such as a signal peptide for secretion, absence of transmembrane domains, presence of some motifs, small-medium molecular weight sizes and cysteine-rich content [8–10]. Additional fungal effector features have been reported for specific subclasses of effectors, for example, particular genomic locations such as gene clusters, gene-sparse regions or localization in dispensable chromosomes [11].

Major efforts have been devoted to *in silico* identification of secreted effectors in large-scale genome studies. Three principal approaches have been commonly used by different research groups: a) analyzing proteomes in different bioinformatic programs that help to filter the secretome, for instance, SignalP 4.1 [12], WoLFPSORT [13] and TMHMM 2.0 [14]. The use of these programs is usually done in a separate manner and not as a package; b) machine learning approaches, which can predict new effector proteins based on the extracted features of reported and confirmed effectors [15,16]; and c) comparative genomics to search those effectors that belong to families, or to find rare ortholog effector candidates that might be transferred horizontally [17].

EffectorP 1.0 and 2.0, the first reported machine learning classifiers of fungal effectors [15,16], have been the most preferred fungal effector prediction tools used to date. Although these approaches are often successful at identifying effectors, criteria such as small size and enrichment in cysteine require thresholds manually set by researchers. Therefore, finding an easy and quick tool that led to adjusting the search criteria used for the prediction of effectors becomes a niche of opportunity.

The present work introduces EffHunter, a pipeline developed to integrate SignalP 4.1 [12], Phobius [18], TMHMM 2.0 [14] and WoLFPSORT [13] with Perl/BioPerl scripts for filtering protein size and cysteine content. Running the analysis in a single step, the sensitivity of the prediction was 70%, defined as the proportion of positives that are correctly identified, and a specificity of 100%, which is the proportion of negatives that are correctly identified. These values obtained were the same as those obtained with the step-by-step method of prediction and similar to EffectorP 2.0 values with a sensitivity of 68% and specificity of 98%.

2. Materials and Methods

2.1. Architecture of EffHunter Pipeline

EffHunter is based on the available software SignalP 4.1 [12], Phobius [18], TMHMM 2.0 [14] and WoLFPSORT [13]. A set of Perl/BioPerl (v5.18.2) scripts was designed to perform these complementary tasks: the protein length analysis (<400 amino acids), the cysteine count (>4 cysteine residues) and connecting the individual steps into a single pipeline for proteome scale analysis. EffHunter works as follows: first, all the sequences are put in FASTA format, where each sequence is filtered with respect to the size indicated by the user, and the result is stored in a new FASTA file. Second, the first list obtained from the filter by protein size is submitted to the filter by number of cysteines; in this step, the user indicates the value to filter those proteins that have greater or equal to the number indicated. Once obtained, the proteins that meet the criteria are stored in a new FASTA file. Third, the retrieved sequences are searched for signal peptide signatures with SignalP 4.1 and Phobius programs. Fourth, the resulting FASTA file of protein sequences are searched for transmembrane domains with TMHMM 2.0, and then WoLFPSORT predicts the subcellular localization. Each analysis uses a FASTA file as input and generates an output in FASTA format too.

EffHunter is distributed as a compressed file in ZIP format. The source code is available for download at <https://www.cicy.mx/unidad-de-biotecnologia/investigador/blondy-beatriz-cantocanche> and https://github.com/GisCarreon/EffHunter_v.1.0 GitHub repository. Docker image is available at <https://hub.docker.com/r/giscarreon/ubuntu-effhunter>. Once the EffHunter_1.0 directory has been uncompressed, it shows the bin directory, in which the empty subdirectories SignalP, Phobius, TMHMM and WoLFPSORT are located. The user must download each program from the platforms indicated in Table 1 and uncompress and install them in each empty subdirectory mentioned above for the correct execution of EffHunter. To use EffHunter pipeline, the code indicates the necessary path to install each program and to compile and take the modules. The command to execute EffHunter once it is installed in a linux/unix terminal is `sh ./EffHunter.sh`.

2.2. Validation of EffHunter Pipeline in ab Initio Approach

EffHunter was challenged using different protein databases designated as positive and negative control sets. The positive data set contained a total of 150 effector proteins (Supplementary data set S1), of which 94 effectors were collected from the literature by Sperschneider *et al.* [15], and the other 56 were effector candidates retrieved from the Pathogen-Host Interaction database (PHI-base) (Supplementary Table S1). These databases comprise, to date, more than 4000 proteins involved in pathogenicity, from more than 260 plant and animal pathogens; 70% of them

being phytopathogens [19,20]. The search for protein effectors in the PHI base for the positive data set was done using the following criteria: length <400 amino acids [5,21–25], >4 cysteine residues [26,27], presence of signal peptide and lack of transmembrane domains [1,27,28].

The capability of EffHunter pipeline to exclude non-effectors was challenged with a large list of negative control proteins (4530 proteins). The negative control set comprises well-known families of proteins, ABC transporter proteins (2329) (Supplementary data set S2) [29,30], Cyt P450 proteins (476) (Supplementary data set S3) [31], and 1725 proteins classified as major facilitator transporters (MFTS) (Supplementary data set S4) [32–34]. The set of negative controls comprises varied proteins: from 73 to 5000 amino acids; from 0 to 74 cysteines; from 0 to 23 TMDs; with or without signal peptide. The great variability makes these proteins a robust set of negative controls. In addition, none are extracellular, and more importantly, no member of these protein families has been described as a fungal effector. The fasta sequences of all of these candidates were downloaded from the GenBank at NCBI (<https://www.ncbi.nlm.nih.gov/>); positive and negative controls were pooled in a single database containing in total 4680 proteins (Supplementary data set S5).

The search for effectors in this database was performed in the traditional way using each program separately by sequential analyses and in a single step, using the EffHunter pipeline with the command described above.

To assess the predictive ability of the EffHunter prediction, the variables sensitivity, specificity, precision and accuracy were calculated to measure the performance of the EffHunter pipeline:

$$\text{Sensitivity/Recall} = \frac{TP}{(TP + FN)} \quad (1)$$

$$\text{Specificity} = \frac{TN}{(TN + FP)} \quad (2)$$

$$\text{PPV/Precision} = \frac{TP}{(TP + FP)} \quad (3)$$

$$\text{ACC} = \frac{TP + TN}{(TP + FP + FN + TN)} \quad (4)$$

$$\text{FPR} = \frac{FP}{(FP + TN)} \quad (5)$$

$$\text{F1} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (6)$$

Sensitivity is defined as the proportion of positives that are correctly identified. Specificity is the proportion of negatives that were correctly identified. Precision or positive predictive value, PPV, is a measure which captures the proportion of positive predictions that are true. Accuracy analysis can be used to evaluate the overall performance of a method. In the equations, TP, true positives; TN, true negatives; FP, false positives; and FN, false negatives. Recall is defined as the proportion of the positives that are successfully retrieved. F1 Score is widely used to measure the success of binary classifier and compare performance of different software/pipelines. F1 Score reaches its best value at 1 and the worst score at 0 [16,25,28,35–37].

2.3. Validation of EffHunter Pipeline in Comparative Approach

Criteria to identify effectors are largely discrepant in the literature; therefore, in order to continue the evaluation of EffHunter, its predictions were compared with three published fungal effector datasets. Each dataset was obtained by using different strategies to identify the effectoromes.

Two approaches were followed. In the first one, we compared EffHunter prediction with previous reports of effector prediction, by sequential/separate analyses with different bioinformatics programs. Reports were for *Blumeria graminis* f. sp. *hordei* [38], *Pseudocercospora fijiensis* [39] and *Mycosphaerella graminicola* [40]. It is important to emphasize that in those reports, different combinations of bioinformatic tools were used. Details about such tools are provided in Table 4.

For *Blumeria graminis* f. sp. *hordei* and *Pseudocercospora fijiensis*, EffHunter analyses ran on the same bioinformatic databases used by the authors, but for *Mycosphaerella graminicola*, EffHunter ran on the non-redundant dataset of proteins for this pathogen at JGI (<https://mycocosm.jgi.doe.gov/Mycfi2/Mycfi2.home.html>) because the database reported by authors is not publicly available. In the search for effectors, the length cutoff was set equal to that used in each one of these reports, respectively.

The second approach was a comparison between the resulting list of protein effectors produced by EffHunter with those of Sonah *et al.* [10], since the authors used the bioinformatics tool, SECRETOOL [41], a program that integrates the use of SignalP 4.1, TMHMM 2.0 and WoLFPSORT to analyze and identify secreted proteins. Using SECRETOOL, the authors predicted effectoromes in 12 proteomes: *Alternaria brassicicola*, *Blumeria graminis*, *Cladosporium fulvum*, *Colletotrichum graminicola*, *Fusarium oxysporum*, *Leptosphaeria maculans*, *Magnaporthe oryzae*, *Mycosphaerella graminicola*, *Ustilago maydis*, *Puccinia graminis* f. sp. *tritici*, *Pyrenophora tritici-repentis* and *Phytophthora infestans*. Comparison among *in silico*

identification of effectors by EffHunter, SECRETOOL and EffectorP 2.0 was carried out for these pathogens. For this and further analyses, cutoff size was fixed at <400 amino acids and the cysteine residues >4 in EffHunter analyses.

As a final test for validation, the predictive performance and limitations of EffHunter were compared with EffectorP 2.0 on non-canonical effectors (which are effectors hard to predict), such as SAD1, Mg3LysM, BEC1054, BEC1011, BEC1019, CSEP0055, Bcg1, CSEP0105, SIS1, Xyla, PIIN 08944, AvrPm3 and AvrSr35.

2.4. Prediction of Effector Proteins in Fungal Genomes

Deduced proteomes from 87 fungi and 4 oomycetes were downloaded from the databases of Broad Institute and Joint Genome Institute (<https://jgi.doe.gov/>) [42] and from the resource for genome-scale data, Ensembl Genomes (<https://fungi.ensembl.org/index.html>), developed by the EBI and the Wellcome Trust Sanger Institute [43] (Supplementary Table S2).

3. Results

3.1. EffHunter: A Pipeline to Predict Fungal Effectors Proteins

The EffHunter pipeline was constructed for the *in silico* identification effectors in fungal proteomes. The architecture of EffHunter consists of four modules: 1) analysis of the protein length and cysteine count, 2) detection of signal peptide, 3) transmembrane domains and 4) subcellular localization.

Length estimation and the counting of cysteines in each protein are performed by a set of Perl/BioPerl scripts on the subject proteins; the programs listed in Table 1 accomplished the other bioinformatics analysis.

Table 1. Bioinformatics tools integrated into EffHunter.

Program	Features	Website	Reference
Perl/BioPerl	International association of users and developers of open-source Perl tools for bioinformatics, genomics and life science.	https://bioperl.org/	[44]
SignalP 4.1	Predicts the presence of signal peptides and the location of their cleavage sites in proteins from gram-positive bacteria, gram-negative bacteria and eukarya.	http://www.cbs.dtu.dk/services/SignalP/index.php	[12]

			[18]
Phobius	This server is for the prediction of transmembrane topology and signal peptides from the amino acid sequence of a protein.	http://phobius.sbc.su.se/	
WoLFPSORT	Converts protein amino acid sequences into numerical localization features, based on sorting signals, amino acid composition and functional motifs, to predict protein subcellular location.	https://wolfpsort.hgc.jp/	[13]
TMHMM 2.0	Predicts trans-membrane (TM) domain helices in proteins.	http://www.cbs.dtu.dk/services/TMHMM/	[14]

The EffHunter pipeline compiles the SignalP 4.1, Phobius, TMHMM 2.0 and WoLFPSORT programs, together with a set of scripts in Perl, to execute the analyses of length and content of cysteine residues in a single step for each polypeptide sequence of fungal proteomes (Figure 1).

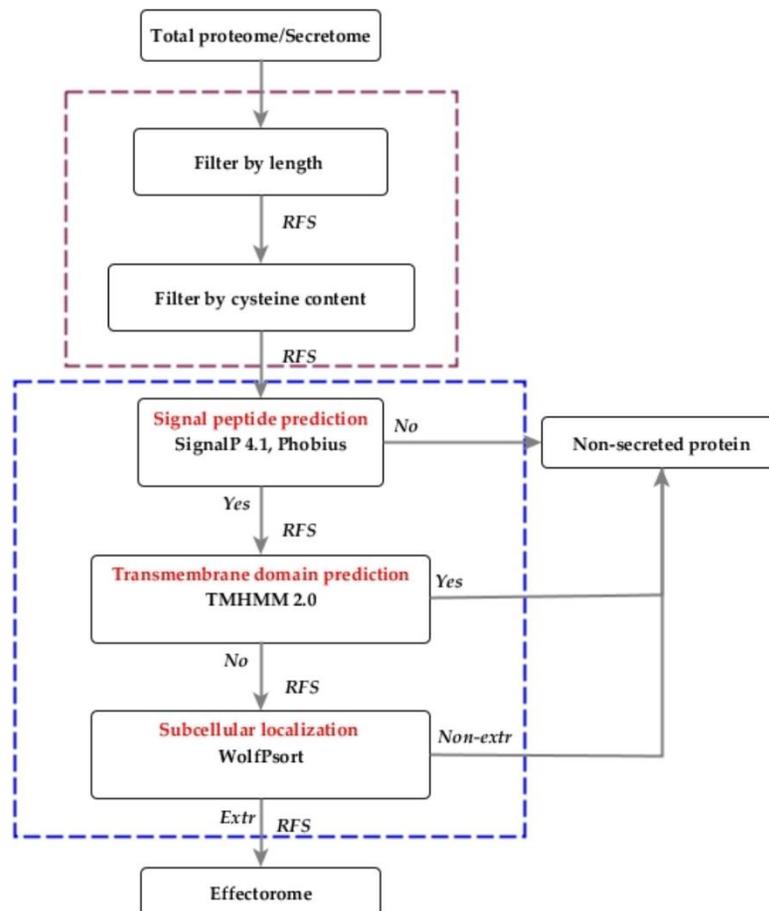


Figure 1. Complete EffHunter workflow for the prediction of effectors in fungal proteomes. Red square, custom tools (user parameters); blue square, pre-installed prediction tools. RFS, retrieved FASTA sequences. The pipeline works correctly either on total proteomes or secretomes.

3.2. Validation *ab Initio*

The evaluation of EffHunter's performance was done by employing the positive data set (Table 2), comprising 56 effectors available in the PHI-database and the 94 effectors used for the initial positive training of EffectorP 2.0 [15]. The final list comprises 150 protein effectors, which include 140 effectors from fungi and 10 protein effectors from oomycetes (Supplementary data set S1).

Table 2. Positive data set of effector proteins used in this work.

Species	Effector proteins
<i>Alternaria alternata</i>	Aapg1
<i>Alternaria citri</i>	Acpg1
<i>Aspergillus flavus</i>	PECA
<i>Aspergillus fumigatus</i>	CfmB, CfmC
<i>Beauveria bassiana</i>	BbCHIT1
<i>Bipolaris maydis</i>	Ppt1
<i>Bipolaris oryzae</i>	Ppt1
<i>Bipolaris zeicola</i>	PGN1
<i>Blumeria graminis f. sp. hordei</i>	Avrk1, Avra1, Avra13, CSEP0105, BEC1005, BEC1040
<i>Blumeria graminis f. sp. tritici</i>	AvrPm2
<i>Botrytis cinerea</i>	CUTA, Nep1, PGIP2, BcPGA1, BcPG2
<i>Botrytis elliptica</i>	BeNEP2
<i>Candida albicans</i>	BGL2, SAP1, sap1, SAP2, SAP3, sap3, Pga26, RBT4
<i>Candida tropicalis</i>	ACP
<i>Cladosporium fulvum</i>	Avr2, Avr4, Avr4E, Avr5, Avr9, Ecp1, Ecp2, Ecp4, Ecp5, Ecp6
<i>Claviceps purpurea</i>	CPPG1, CPPG2
<i>Colletotrichum graminicola</i>	CgEP1, Cgfl
<i>Fusarium graminearum</i>	FGL1
<i>Fusarium oxysporum</i>	Avr3, PG1, XYL3
<i>Fusarium oxysporum f. sp. lycopersici</i>	Six1, Six2, Six3, Six4, Six5, Six6, Six7, Six8
<i>Fusarium solani</i>	cutA
<i>Histoplasma capsulatum</i>	CBP1
<i>Laccaria bicolor</i>	MiSSP7
<i>Leptosphaeria maculans</i>	AvrLm1, AvrLm4–7, AvrLm6, AvrLm11, AvrLmJ1, SP1
<i>Magnaporthe oryzae</i>	Avr1-CO39, Avr-Pia, AvrPib, Avr-Pita, Avr-Pii, Avr-Pik, AvrPi9, AvrPiz-t, Bas1, Bas2, Bas3, Bas4, Bas107, Bas162, GAS1, GAS2 , lug6, lug9, MC69, MHP1 , MoCDIP1, MoCDIP2, MoCDIP3, MoCDIP4, MoCDIP5, MPG1 , MoHEG13, Msp1, Pwl1, SPD2, SPD4, SPD7, SPD9, SPD10, XYL-6
<i>Melampsora lini</i>	AvrL2-A, AvrL567-A, AvrM, AvrM14, AvrP4, AvrP123,
<i>Metarhizium anisopliae</i>	Pr1

<i>Monilinia fructicola</i>	MfCUT1
<i>Parastagonospora nodorum</i>	SP1
<i>Phakopsora pachyrhizi</i>	PpEC23
<i>Phytophthora cactorum</i>	PcF
<i>Phytophthora capsici</i>	Pc129892, Pcipg2
<i>Phytophthora infestans</i>	EPI10, INF1, INF2A, INF2B
<i>Phytophthora parasitica</i>	CBEL, Ppxyn1
<i>Phytophthora sojae</i>	GIP2
<i>Puccinia graminis f. sp. tritici</i>	AvrSr50, PGTAUSPE-10-1
<i>Puccinia striiformis f. sp. Tritici</i>	Pec6, PstSCR1
<i>Pyrenophora tritici-repentis</i>	ToxB
<i>Rhynchosporium secalis</i>	NIP1, NIP2, NIP3
<i>Sclerotinia sclerotiorum</i>	SsSSVP1
<i>Stagonospora nodorum</i>	ToxA, Tox1, Tox3
<i>Trichoderma virens</i>	Sm1
<i>Uromyces fabae</i>	RTP1
<i>Ustilago hordei</i>	UhAvr1
<i>Ustilago maydis</i>	Cmu1, eff1-1, Mig1, Mig2-1 , Pep1, Pit2, Tin2, See1
<i>Verticillium dahliae</i>	Ave1, PevD1, Vdlsc1, VdSCP7
<i>Zymoseptoria tritici</i>	AvrStb6, Zt6

* One hundred fifty effector proteins were pooled from the list (94) collected by Sperschneider *et al.* [15] and 56 effectors retrieved from the PHI database (labeled in bold). All these effector proteins have been published in peer-reviewed journals.

As negative controls, a subset of 4530 proteins was used, comprising P450 proteins, MFTS and ABC transporters, most of them large, integral membrane proteins. The set of negative controls comprises a large number of proteins to challenge EffHunter, in order to prevent false positive identification as much as possible. Using the same data set with positive and negative controls, we compared our pipeline with EffectorP 2.0, a machine learning classifier for fungal effector prediction.

EffHunter positively identified 105 from 150 effectors (70%); meanwhile, 45 were false negatives (30%). The missing 45 effectors could not be identified because some of them contain transmembrane domains or less than four cysteine residues, two criteria that EffHunter uses for the prediction. From the total subset of proteins evaluated (4680), no false positive was retrieved (Table 3). Sequential analyses with the programs and scripts that make up EffHunter produced the same results as the automatic analysis with EffHunter (Table 3), showing that the pipeline worked as expected.

EffectorP 2.0 identified 166 effectors; 102 of them were true effectors (68%), but 48 true effectors were missing, i.e., 32% of false negatives. The larger difference observed between both predictors was the number of false positives, e.g., 64 for EffectorP 2.0 (41 ABC transporters, 22

cytochrome P450, and 1 MFTS) and none for EffHunter. Furthermore, sensitivities are similar (76% and 75%), but specificity was larger for EffHunter (100%) in comparison with 98% for EffectorP 2.0. Precision of EffHunter was 100% vs. 97% for EffectorP 2.0. Accuracy value was 99% for EffHunter and 97% for EffectorP 2.0 (Table 3).

In addition, following the strategy of Sonah *et al.* [10], i.e., using the pipeline SECRETOOL and then selecting proteins with maximum length of 300 amino acids, 72 true effectors were recognized and 78 true effectors were discarded (false negatives). No false positive was retrieved in this prediction.

Estimation of the F1 score for these predictions was carried out considering their results on the same set of 4680 proteins (containing 150 true effectors and 4530 negative controls). The scores were 0.57 for EffectorP v1.0, 0.64 for EffectorP v2.0, and SECRETOOL/300 amino acids and finally, for EffHunter, 0.82. These results support EffHunter as a good predictor for fungal canonical effectors.

Table 3. Validation of EffHunter for prediction of effector proteins and comparison with EffectorP 2.0.

Step-by-step prediction														
Data*	Proteins in Data set	Total Proteins	Length (30 – 400aas)	≥ 4 Cysteine	Signal peptide by SignalP/Phobius	Proteins without TMD with TMHMM	Total prediction	Results	Sen/Rec	Spe	PPV/Prec	ACC	FPR	F1 score
Set 1	150							105						
Set 2	2329							0						
Set 3	476	4680	765	435	107	105	105	0	70%	100%	100%	99%	0.00%	0.82
Set 4	1725							0						
EffHunter														
Data*	Proteins in data set	Total proteins			Total prediction	Prediction		Sen/Rec	Spe	PPV/Prec	ACC	FPR		F1 score
Set 1	150					105								
Set 2	2329					0								
Set 3	476	4680			105	0		70%	100%	100%	99%	0.00%		0.82
Set 4	1725					0								

Table 3. *Cont.*

EffectorP 2.0										
Data*	Proteins in data set	Total proteins	Total prediction	Prediction	Sen/Rec	Spe	PPV/Prec	ACC	FPR	F1 score
Set 1	150	4680	166	102	68%	98%	61%	97%	1.41%	0.64
Set 2	2329			41						
Set 3	476			22						
Set 4	1725			1						
EffectorP 1.0										
Data*	Proteins in data set	Total proteins	Total prediction	Prediction	Sen/Rec	Spe	PPV/Prec	ACC	FPR	F1 score
Set 1	150	4680	164	91	60%	98%	55%	97%	1.6%	0.57
Set 2	2329			49						
Set 3	476			20						
Set 4	1725			4						
Sonah <i>et al.</i> [10] (SECRETOOL and filter by length <300 amino acids)										
Data*	Proteins in data set	Total proteins	Total prediction	Prediction	Sen/Rec	Spe	PPV/Prec	ACC	FPR	F1 score
Set 1	150	4680	72	72	48%	100%	100%	98%	0.00%	0.64
Set 2	2329			0						
Set 3	476			0						
Set 4	1725			0						

***Set 1:** Positive dataset of true effectors (positive dataset comprises effectors retrieved from pathogen-host interaction (PHI) and the list collected by Sperschneider *et al.* [15]); **Set 2:** ABC transporters; **Set 3:** cytochrome P450; **Set 4:** proteins classified as major facilitator transporter superfamily. Sen/Rec: Sensitivity/Recall; Spe: Specificity; PPV/Prec: Positive Predictive Value/Precision; ACC: Accuracy; FPR: False positive rate; F1 score: Measure of the success of binary classifier (score reaches its best value at 1, and worst score at 0).

EffHunter and EffectorP 2.0 shared 61 candidates. Non-shared predicted candidates were 44 for EffHunter and 41 for EffectorP 2.0 (Figure 2a). All of the 44 EffHunter specific candidates meet the established criteria for effector prediction. From the 41 candidates predicted only by EffectorP 2.0, 9% have no signal peptide, 26% have TMDs and 87% have less than four cysteine residues (Figure 2b). Since effectors are so diverse, it is ambiguous how many of the specific candidates of each predictor are true effectors, but candidates of Effhunter highly meet its own established criteria. The algebraic sum in Figure 1b is greater than 100% because some candidates have two or more characteristics, e.g., no signal peptide and have TMD at the same time.

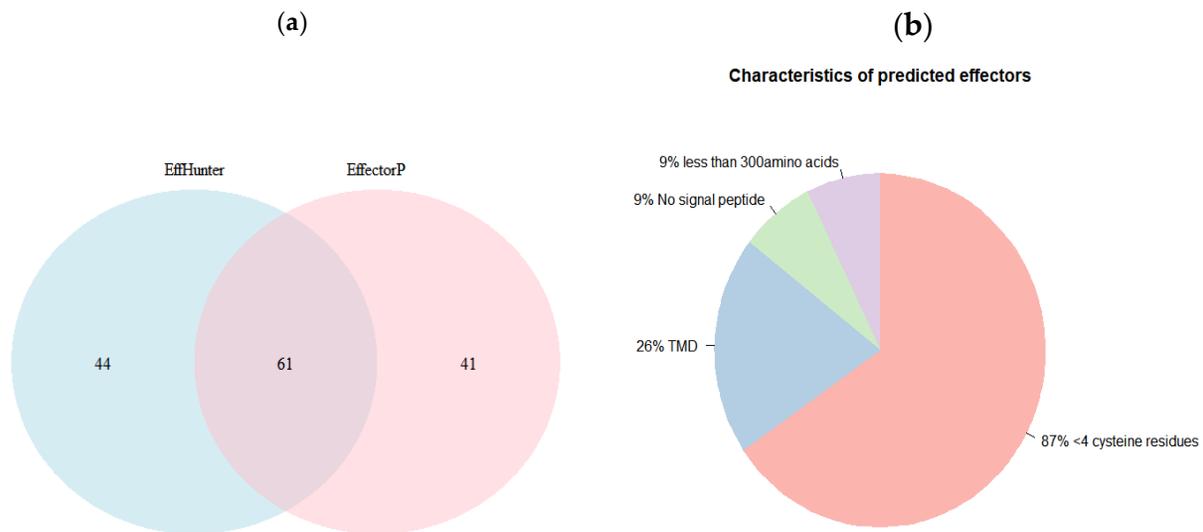


Figure 2. Effector prediction of positive control set. (a) Venn diagram showing the distribution of shared and non-shared predicted effectors by EffHunter and EffectorP 2.0. (b) Pie chart summarizing the characteristics of the 41 non-shared effectors protein predicted by EffectorP 2.0.

3.3. Validation of EffHunter with Fungal Proteomes and Comparison with Other Effector Prediction Tools

EffHunter is versatile and allows the user to set cutoff values for the protein sequence length and number of cysteine residues. To continue the validation of EffHunter in the prediction of effector proteins, additional analysis was carried out on a few economically important plant pathogens whose effectoromes have been previously analyzed. Parameters and software used in each case are described in Table 4. For EffHunter, the length of proteins was set according to the reported criterion in each case. The number of cysteines was fixed at >4 per protein since is difficult to set a proper percentage, as the number of cysteines varies with the size of the protein.

In the plant pathogen *Blumeria graminis* f. sp. *hordei*, EffHunter predicted 490 effector candidates in comparison with the 494 reported by Liang *et al.* [38] using diverse criteria presented in Table 4; 404 of them were common in both predictions. EffHunter identified 82 proteins which were not retrieved by the other study. These 82 proteins were then analyzed with different programs that are not included in EffHunter but were used by other authors (i.e., TargetP 1.1 [45] and big-PI [46]). In addition, the analysis of Liang *et al.* [38] used one additional criterion, searching for secreted proteins that show similarity only with proteins from powdery mildews. These analyses identified 12 false positives in EffHunter candidates because they did not meet this additional criterion, but 70 of them met *all* the author's criteria. On the contrary, Liang *et al.* [38] identified 86 candidates, which were not recognized by EffHunter. Sixty-two of them were larger than 400 amino acids; 19 had no signal peptide and 5 were predicted as GPI-anchored. Therefore, these 86 candidates identified by Liang *et al.* [38] seem to be false positives according to their own criteria (Table 4).

The second comparison was with a list of candidate effectors reported by Chang *et al.* [39] from *Pseudocercospora fijiensis*, the causal agent of black Sigatoka disease in banana and plantain. Chang *et al.* [39] reported 105 candidate effectors for the fungus, while EffHunter predicted 136, with 78 of them shared between both analyses. From the 27 candidates exclusive of the results of Chang *et al.*, 15 appear to be false positives since they do not meet some of their criteria; 12 were ambiguous. Ambiguity arises because both predictions have different settings for some criteria (i.e., they accepted one TMD in the proteins, but EffHunter does not). Since effectors are so diverse, the criterion of having one or no TMD has a similar probability of being acceptable. In the ambiguous candidates, seven have one TMD and five candidates have only two cysteines in their sequence, but they present the 2% cysteine because those are peptides with 60 amino acids or less. EffHunter parameters filtered those proteins. EffHunter predicted 58 effectors that were not enlisted in the reference data; 32 of them may be false negatives for the reference since they meet *all* their parameters. Sixteen EffHunter candidates were probably false positive, and 10 were ambiguous (Table 4).

The third comparison was with a list of deduced effectors from *Mycosphaerella graminicola*, a fungus causing septoria leaf blotch in wheat [40]. In addition to some software previously mentioned (i.e., SignalP 3.0, TargetP 1.1, big-PI, WoLFPSORT), these authors included in their analysis the use of LocDB [47], ProtComp v8.0 and PotLocDB [48] and designed a script to set cysteine at 5%. The list of candidates was filtered to exclude those proteins that have any functional annotation. They reported 171 effector candidates vs. 183 by EffHunter.

One hundred ten were common in both results; among the 61 candidates exclusive for the reference, 60 failed to meet some of the authors' criteria (they may represent false positives), but one was true positive for their result and false negative for EffHunter. Curiously, this protein (ID 82029) is not present in the non-redundant set of *M. graminicola* proteins in the JGI database. For this reason, EffHunter did not analyze it.

EffHunter predicted 74 candidates in addition to the common set. Fifty of these candidates are true positives, according to all criteria from authors, including the search for candidates with no relation with proteins with functional annotation. These 50 candidates are true effector candidates for EffHunter and false negatives for the reference. It is probable that EffHunter had 24 false positives since four have GPI, ProtComp predicts seven non-extracellular proteins, and 13 have homology with proteins with functional annotation in Pfam. EffHunter does not include the search of GPI, the ProtComp program and the Pfam database in the pipeline.

In summary, EffHunter performed well in these comparisons. In the three cases, the number of true positives was higher, and the false positives were lower with EffHunter. The number of false negatives for EffHunter was negligible; meanwhile, in the other reports, it was 14% in the Liang *et al.* [38] effectorome prediction for *B. graminis* f. sp. *hordei*, 30% in Chang *et al.* [39] prediction for *P. fijiensis* and 29% in Morais do Amaral *et al.* [40] prediction for *M. graminicola*, supporting the robust and reliable prediction by EffHunter.

Table 4. Data analysis of the set of fungal effector proteins between EffHunter and other reports.

Species	Criteria in Reference	Genome Size (Mbp)	Total Proteome	Secretome	Effectors Prediction in Reference	EffHunter Prediction	Shared	Difference in: Reference(R) EffHunter(E)	Observations for Effectors Predicted by Reference or by EffHunter	Summary of true or false Positives or Negatives, or Ambiguous in the Specific Sets of Effectors (Considering both Predictions)
<i>Blumeria graminis f. sp. hordei</i> [38]	Secretion signal with SignalP 4.1 and SecretomeP, no TMD with TMHMM 2.0, no hits outside powdery mildews with Blastp; subcellular localization with TargetP 1.1 and GPI anchors by Big-PI.	124.49	7118	726	494	490	408	R = 86	Negatives (86): 62 proteins are larger than 400 amino acids, 5 proteins have GPI binding and 19 proteins have no signal peptide and have TMD. Therefore, all 86 are false positive in this set. TP=0; FP=86	Reference TP = 0 FP = 86 FN = 70 TN = 12
								E=82	Negatives (12): 3 proteins were predicted by TargetP 1.1 as mitochondria target, 6 proteins were predicted with GPI anchors by big-PI and 3 proteins have homologs in fungal species that are not powdery mildew. Positives (70): 70 <i>met al</i> criteria of authors; Blastp retrieved homologs only in powdery mildews. TP=70; FP=12	EffHunter TP = 70 FP = 12 FN = 0 TN = 86
								R = 27	Negatives (15): 2 proteins are larger than 250 amino acids, 2 have no clear localization prediction with TargetP 1.1, 7 have no signal peptide, 4 have no extracellular location. *Ambiguous (12): 7 have one TMD and 5 small proteins (<60 amino acids) with less than 4 cysteine, but ≥2% cysteine. A = 12; FP = 15	Reference TP = 0 FP = 15 FN = 32 TN = 16 A = 12
<i>Pseudocercospora fijiensis</i> [39]	Secretion signal with SignalP 4.1, one or no TM domains with TMHMM 2.0, subcellular localization with TargetP 1.1 and WoLFPSORT, no GPI anchor with PredGPI, length <250aas, >2% cysteine residues	74.1	13107	584	105	136	78	E = 58	Negatives (16): 11 proteins are GPI-anchored, 5 proteins have mitochondria target (by TargetP 1.1). Positives (32): 30 <i>met al</i> criteria and have 2.0-8.2% cysteine; 2 proteins are predicted with signal peptide by SignalP 4.1 and Phobius, and as extracellular with WoLFPSORT, but TargetP 1.1 (reference) fails to predict localization.	EffHunter TP = 32 FP = 16 FN = 0 TN = 15 A = 10

<i>Mycosphaerella graminicola</i> [40]	Size <200 amino acids; secretion signal with SignalP 4.1, one or no TM domain with TMHMM 2.0, secreted by TargetP 1.1, no GPI-anchor with big-PI, subcellular localization with WoLFPSORT and ProtComp, and no functional information	39.7	10933	492	171	183	110	R = 61	<p>*Ambiguous (10): 10 are >200 amino acids and have 4 (but <2%) cysteine. <i>Me et al</i> other criteria. TP = 32; FP = 16; A = 10</p>	Reference TP=1 FP=60 FN=50 TN=24
								E = 74	<p>Negatives (60): 37 proteins are larger than 200 amino acids; 3 do not have signal peptide; 6 are GPI-anchored, 14 are not predicted as secreted by ProtComp. Positives (1): Protein ID 82029 matches all criteria. This protein is not in the nonredundant set of <i>M. graminicola</i> database at JGI, therefore EffHunter could not analyze it. TP = 1; FP = 60</p> <p>Negatives (24): 4 proteins have GPI anchors; 1 has mitochondria target; 6 are predicted cytosolic, mitochondrial or nuclear by ProtComp and 13 proteins have functional annotation in PFAM Positives (50): 50 match all criteria and have no functional annotation or known protein domains. TP = 50; FP = 24</p>	

*Ambiguous: Those candidates that meet criteria from one prediction (positive for this analysis), but do not meet criteria of the other analysis and criteria from one or the other are not definitive for assigning them as positive or negative. Databases analyzed in the references and by EffHunter were the same, except for *M. graminicola*. The authors did not provide that database; the non-redundant protein models from *M. graminicola* at JGI were downloaded in that case.

In another report, Sonah *et al.* [10] used the bioinformatics tool SECRETOOL on 12 proteomes [41] to obtain the respective secretomes and then retrieve the small proteins (<300 amino acids), proteins that they classified as effectors (Table 5). One of these pathogens (*M. graminicola*) was analyzed also by Morais do Amaral *et al.* [40] and compared here with EffHunter (Table 4), which is interesting because this expands the comparison among the different predictive tools for fungal effectors. Comparison among effectoromes predicted by SECRETOOL, EffHunter and EffectorP 2.0 was conducted for these pathogens (Table 5). Sonah *et al.* [10] did not provide a link to the databases that they used; as such, the first attempt in our analyses was conducted on total proteomes of these species. EffHunter works properly on total proteomes, but we observed an elevated number of false positives with EffectorP 2.0. Using a total proteome as input, EffectorP 2.0 identified 1663 candidates for *C. fulvum* vs. 151 candidates with a secretome as input (data not shown). For this reason, the secretome for each pathogen was first obtained submitting the proteomes to the SignalP 4.1 to reduce the number of false positives for the EffectorP 2.0 analysis. Table 5 shows the prediction of effectors by the three programs. The numbers of effectors are similar between Sonah *et al.* [10] predictions and EffHunter; in general, EffectorP 2.0 predicted in some cases a lower number of candidates than EffHunter and SECRETOOL [10].

For example, in the case of *Puccinia graminis* f. sp. *tritici*, 659 effector candidates were predicted by EffHunter, 612 reported by SECRETOOL and 605 predicted by EffectorP 2.0. In the latter, 110 proteins have one or more transmembrane domains and cannot be classified as positive or negative, but as ambiguous. In the list of 94 validated effectors shown in Table 2 [15], 11 have one TMD and 83 have no TMD. In general, in each effectorome, EffectorP 2.0 retrieved few candidates with no signal peptide or possessing transmembrane domains. False positives or false negatives from Sonah *et al.* [10] could not be calculated because of the lack of the datasets of their sequences. Step-by-step analyses of all candidates retrieved by EffHunter showed that they meet all EffHunter criteria, suggesting no false positives in our sets of putative effectors.

Table 5. Effector prediction data from the different predictions tool across 12 proteomes.

Species	Lifestyle	Genome		Total proteins	Effector predictions			Effectors in reference	Reference genome
		Mb	Coverage		EffHunter	EffectorP 2.0	*SECRETOOL		
<i>Alternaria brassicicola</i>	Necrotroph	31.03	120×	10688	227	113	228	139	[49]
<i>Blumeria graminis</i>	Biotroph	158.94	13×	6526	255	109	143	437	[50]
<i>Cladosporium fulvum</i>	Biotroph	61.11	21×	14127	342	151	296	271	[51]

<i>Colletotrichum graminicola</i>	Hemibiotroph	51.6	9×	12006	364	159	352	177	[52]
<i>Fusarium oxysporum</i>	Hemibiotroph	55.72	186.1×	17726	474	256	361	364	[53]
<i>Leptosphaeria maculans</i>	Hemibiotroph	44.81	8.31×	12469	290	162	263	529	[54]
<i>Magnaporthe oryzae</i>	Hemibiotroph	41.7	7×	12755	273	368	528	163	[55]
<i>Mycosphaerella graminicola</i>	Hemibiotroph	39.7	8.9×	10933	286	166	235	NS	[56]
<i>Phytophthora infestans</i>	Hemibiotroph	228.54	7.6×	17787	355	404	343	563	[57]
<i>Puccinia graminis f. sp. Tritici</i>	Biotroph	88.64	6.9×	15979	659	605	612	1106	[26]
<i>Pyrenophora tritici-repentis</i>	Necrotroph	37.84	98×	12169	322	182	328	317	[58]
<i>Ustilago maydis</i>	Biotroph	19.66	10×	6785	113	107	142	426	[59]

*Prediction in Sonah *et al.* [10]; NS: Not specified

To further evaluate EffHunter, its performance on nonconventional known effectors of five species of phytopathogens was analyzed and compared with EffectorP 2.0 (Table 6). PIIN 08944 and AvrSr355 are elusive effectors, and neither EffHunter nor EffectorP 2.0 can recognize them. SAD1 and BEC1054 are not recognized by EffHunter, but they are predicted as effectors by EffectorP 2.0. On the contrary, EffHunter recognizes Mg3LysM, BEC1019 and CSEP0105, which are not recognized by EffectorP 2.0. These results show that EffHunter has strengths and limitations, as does EffectorP 2.0, which is currently the predictive tool for fungal effectors most widely used in the literature. Even with this limitation, EffHunter's performance is acceptable to search for effectors in fungal proteomes and has the characteristic of being able to perform the analyses on total proteomes without prior filtering of the protein set to retrieve the secretome. The results show that EffHunter is a tool that makes the search for effectors friendly, making it a better tool.

Table 6. Comparison of prediction between EffHunter and EffectorP 2.0 on non-canonical effectors.

Species	Effector	Length	No. of Cysteine	Signal Peptide	*TMD	EffHunter	EffectorP 2.0
<i>Sporisorium reilianum</i>	SAD1	626	4	No	0	Non-effector	Effector
<i>Zymoseptoria tritici</i>	Mg3LysM	232	9	Yes	0	Effector	Non-effector
<i>Blumeria graminis f. sp. hordei</i>	BEC1054	118	2	Yes	0	Non-effector	Effector
	BEC1011	118	3	Yes	0	Non-effector	Effector

	BEC1019	316	8	Yes	0	Effector	Non-effector
	CSEP0055	122	3	Yes	0	Non-effector	Effector
	Bcg1	146	2	Yes	0	Non-effector	Effector
	CSEP0105	128	6	Yes	0	Effector	Non-effector
<i>Rhizophagus irregularis</i>	SIS1	149	2	Yes	1	Non-effector	Effector
<i>Fusarium graminearum</i>	Xyla	231	1	Yes	0	Non-effector	Effector
<i>Piriformospora indica</i>	PIIN_08944	120	0	Yes	0	Non-effector	Non-effector
<i>Blumeria graminis f. sp. tritici</i>	AvrPm3	130	2	Yes	0	Non-effector	Effector
<i>Puccinia graminis f. sp. tritici</i>	AvrSr35	577	3	Yes	0	Non-effector	Non-effector

* TMD: Transmembrane Domain.

3.4 Prediction of Effector Proteins in Several Fungal and Oomycetes Proteomes with EffHunter

The prediction of candidate effectors with EffHunter was carried out on 95 proteomes downloaded from the JGI Genome Institute MycoCosm and FungiEsembl platform (Supplementary Table S2).

Since effectors have been defined as pathogenicity-related proteins, effectoromes were compared among 40 phytopathogens with different lifestyles: 9 species of biotrophs (blue bars), 20 species of hemibiotrophs (green bars) and 11 species of necrotrophs (red bars). In general, the highest number of predicted effectors was in the group of hemibiotrophs (close to 400 on average), with the lower number in the necrotrophs (around 200 effectors), followed the biotrophic group (around 300 effectors). In each group, there are fungi with an expanded or contracted set of effectors, such as the necrotroph *Penicillium digitatum*, the hemibiotrophs *Verticillium dahliae* and the oomycete *Phytophthora capsici*, and the biotrophs *Blumeria graminis f. sp. tritici* and *Ustilago maydis*, with smaller effectoromes than the rest.

The performance of EffHunter predicts expanded effectoromes in *Melampsora larici-populina*, *Puccinia graminis f. sp. tritici*, *Colletotrichum higginsianum*, *Fusarium oxysporum f. sp. lycopersici*, *Magnaporthe oryzae* and *Phytophthora sojae* (Figure 3 and Supplementary Table S3).

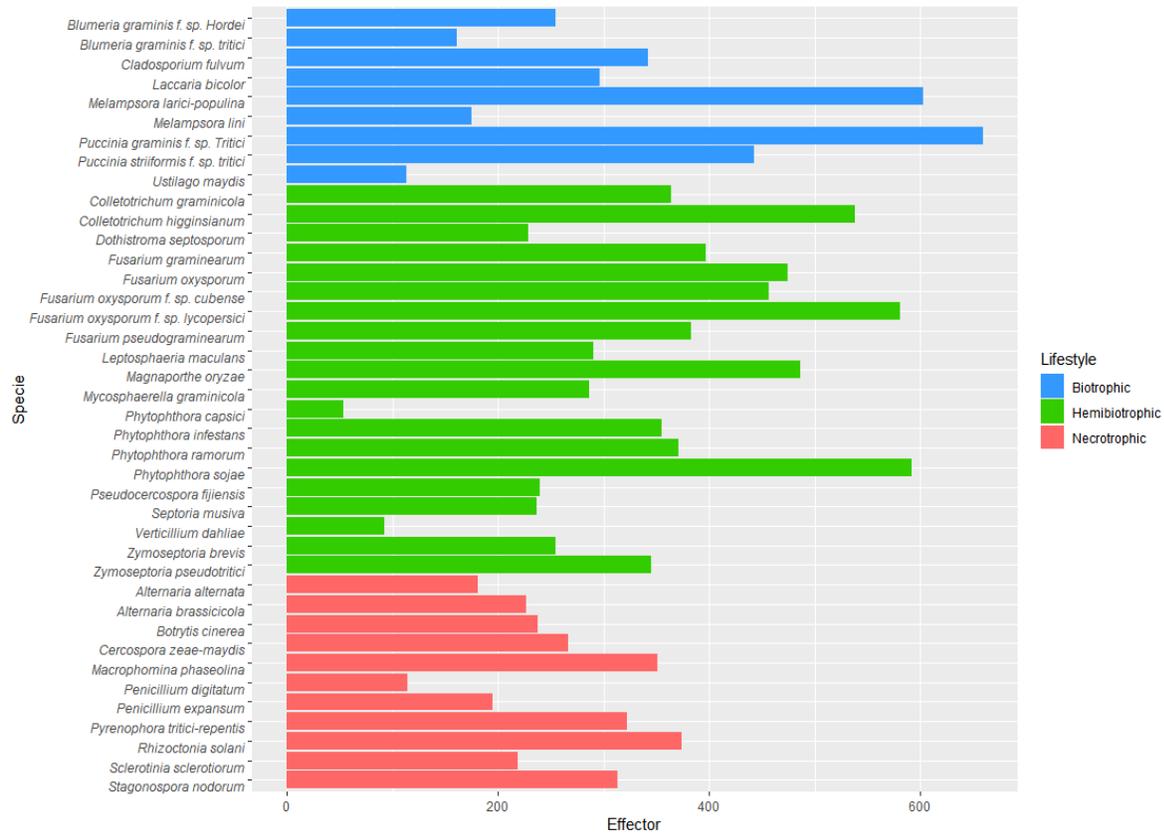


Figure 3. Effectoromes in fungal and oomycetes phytopathogens predicted by EffHunter.

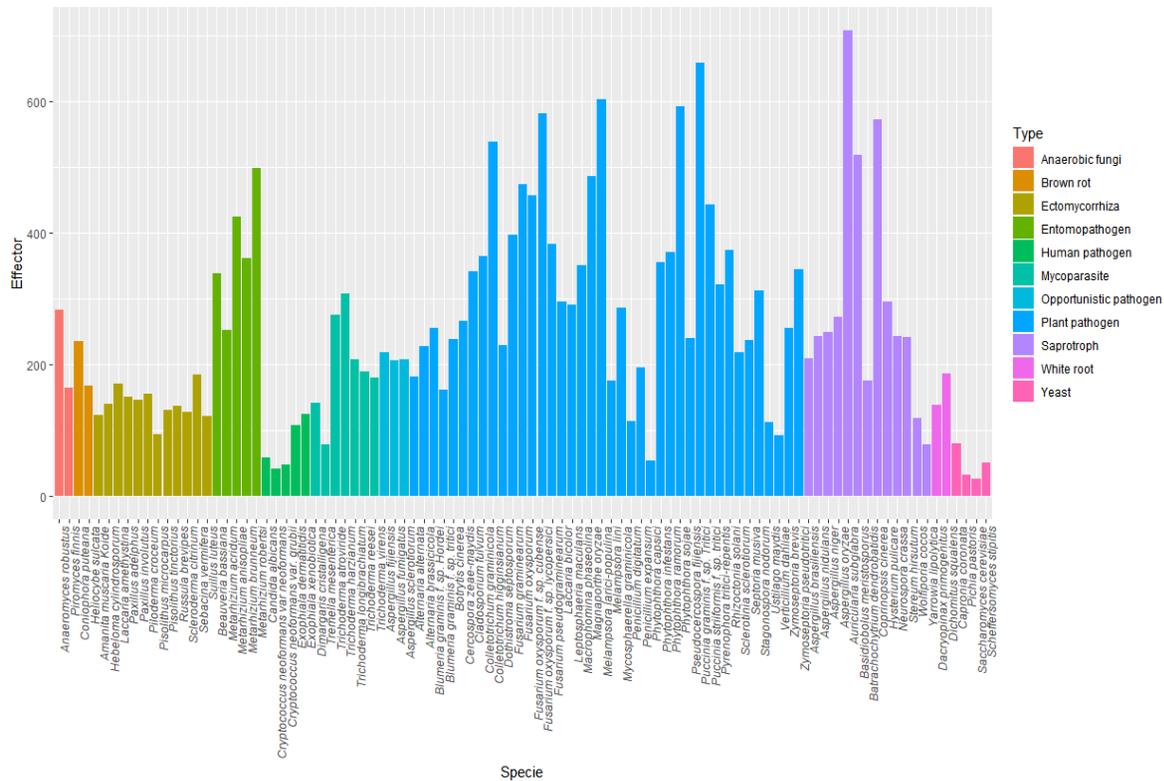


Figure 4. Predicted effectors in fungal and oomycetes proteomes using EffHunter.

In Figure 3, the graphic shows the number of effectors predicted in different types of fungi: yeast, brown and white rot fungi, ectomycorrhiza, opportunistic, mycoparasites, human pathogens, plant pathogens, entomopathogens and saprotrophs (Supplementary Table S4).

The lowest numbers of effectors were predicted in yeast. Curiously, the number of effectors in human pathogens was lower than in the nonpathogenic group, such as ectomycorrhiza and saprotrophs, revealing that pathogens have diversity in the number of predicted effectors. The number of effectors predicted in other pathogens was congruently larger. The groups of fungi with more effectors were plant pathogens and entomopathogens. However, it is not exclusive for pathogens to have the largest effectoromes, because in the case of saprotrophs, EffHunter predicted a similar number of effectors than in phytopathogens and entomopathogens.

4. Discussion

Effectors are key molecules in pathology since they enable the pathogen to modify host cell structure, physiology and metabolism to permit pathogen growth and colonization in the host. Most of the available knowledge has emerged from the study of plant pathogen interactions where

it has been discovered that effectors play diverse functions. Many effectors work as plant immunosuppressors, but others can trigger sugar transport in benefit of the pathogen [60], block or inactivate plant receptors preventing its detection [61] or, on the contrary, activate the immune plant receptors to kill the host [62]. Research on effectors is at its peak, and surely many other roles of effectors remain to be discovered.

In agriculture, fungal diseases cause major losses in several high-value crops such as bananas, coffee, cacao, legumes and spices, provoking ~20% of annual crop yield losses worldwide [63]. Effector proteins are potential biotechnological tools to assist in developing disease control strategies, for example, to select effector-assisted tolerant or resistant plants in breeding programs [64] or to identify plant targets and protein cognates for genetic engineering [65].

High-throughput technologies can currently provide us with high-quality transcriptomes from plant–pathogen interaction and pathogen genomes, meaning powerful data output for effector discovery. However, although progress has been achieved in functional validation of effectors, and multiple candidates can be tested [66], the experimental validation requires experience in genetic engineering. Moreover, this validation is expensive and time-consuming. Currently, less than 200 effectors have been identified and characterized [15]. Hence, robust effector mining from genomic data is key.

EffHunter is a pipeline created in our group by integrating SignalP 4.1, Phobius, WoLFPSORT, TMHMM 2.0 and two scripts for filtering small size (<400 amino acids) and cysteine-rich (> 4) proteins. EffHunter identifies canonical effectors, i.e., secreted, apoplastical, small size and cysteine-rich. To test EffHunter, we used a positive subset of 94 effectors for the initial positive training of EffectorP 2.0 [15] and 56 effectors available in the PHI-database. A large set of negative controls (4530 proteins) was used to challenge EffHunter. No protein in the negative control set is extracellular and, more importantly, none has been described as a fungal effector. In addition, they are highly variable in length, content of cysteine, TMDs and presence or absence of signal peptide, in order to prevent false positive identification by EffHunter as much as possible.

Performance of EffHunter was compared with other effector predictors, running the analysis on the same set of data (4680 proteins, comprising 150 true effectors and 4530 negative controls). F1 score for EffHunter was 0.82, superior to F1 scores for EffectorP v1.0 and effector prediction following the strategy of Sonah *et al.* [10] (Table 3). EffectorP v2.0 showed a lower F1 score than

EffHunter (0.64). The larger difference observed between EffHunter and EffectorP v2.0 was the number of false positives, 64 for EffectorP 2.0 and none for EffHunter.

The positive set of proteins used for developing and training EffectorP 2.0 comprised effectors with transmembrane domains (11 proteins), effectors with no signal peptide (2 proteins), effectors larger than 400 amino acids (2 proteins) and 37 proteins with less than four cysteine residues [15]. This training enables that algorithm to identify effectors with these noncanonical characteristics; however, according to our analysis, it also results in a higher percentage of putative false positives, in contrast to the zero false positives obtained with EffHunter. Another advantage of EffHunter is its suitability for total proteomes or secretomes as input, retrieving the same number of candidate effectors, while EffectorP 2.0 requires a secretome as input. We found that effector prediction using a total proteome in EffectorP 2.0 increases the rate of false positives by almost 10 times (data not shown). EffHunter demonstrated a very good performance as it relates to sensitivity, specificity, precision and accuracy, which were similar to or better than those of EffectorP 2.0.

The ability to predict fungal effectors by the EffHunter pipeline was compared later with reports that used different strategies to predict effectors in *Blumeria graminis* f. sp. *hordei* [38], *Pseudocercospora fijiensis* [39] and *M. graminicola* [40]. We found that similar numbers of effectors were predicted between EffHunter and each of these reports where distinct strategies and criteria were used. For instance, to predict effectors of the different fungi with EffHunter, the length of amino acids was set according to each report. The number of cysteines was not changed because, in the case of the report for *B. graminis* f. sp. *hordei*, the number of cysteines was not defined, and in the case of *Pseudocercospora fijiensis* and *Mycosphaerella graminicola*, authors used, respectively, 2% and 5% cysteine as cut off. However, 2% and 5% would discard many promising candidates. Then, since both extremes are not adequate, the number of cysteines for EffHunter searching was set at >4 as used in other analyses in this manuscript, reinforcing the EffHunter evaluation against other analyses, which use different parameters.

EffHunter performed well on each comparison. In the three cases, their number of true positives was higher and the number of false positives (candidates that do not meet the criteria established by the respective authors) was lower in comparison with these reference works. EffHunter false positives resulted from additional criteria used by the authors, for example, discarding candidates that have homologs in fungi phylogenetically distant from the model under study or that exclude those that have homologous proteins with any functional annotation. These criteria are good, but

we do not recommend including these criteria in automatic analysis, to prevent elimination of many potential true effectors since 18% of true effectors have functional annotation (e.g., hydrophobin, protein with CFEM-domain, cerato-platanin, etc). In the case of false negatives (those candidates proposed only by the other predictor, which largely qualify as potential effectors), the number was negligible for EffHunter; meanwhile, it was 70 in the Liang *et al.* [38] effectorome prediction for *B. graminis f. sp. hordei*, 32 for Chang *et al.* [39] for *P. fijiensis* and 50 for *Morais do Amaral et al.* [40] prediction for *M. graminicola*. Such diversity of approaches used with other fungi presented a great challenge to EffHunter; however, we demonstrated its capacity to perform smoothly with different data.

Another advantage of the program is its versatility since the user can set the cut-off for the length (number of amino acids) and the number of cysteines.

The next evaluation was to compare with the report from Sonah *et al.* [10] since these authors used another bioinformatics tool, SECRETOOL [41] to predict effectoromes in proteomes of 12 fungi; they filtered first by the SECRETOOL pipeline and then selected the proteins < 300 amino acids. The number of effectors predicted by EffHunter was consistent with the number of predicted effectors by the SECRETOOL pipeline used by these authors, probably because both predictors share similarities in their constructions (both comprise analyses by SignalP 4.1, TMHMM 2.0 and WoLFPSORT). However, they are not identical, since SECRETOOL does not analyze the content of cysteine and EffHunter does not integrate TargetP 1.1 and PredGPI as SECRETOOL does. Results obtained with both tools were not similar for all organisms, revealing differences between both predictors. Unfortunately, coincidences, differences, false positives and false negatives in the predictors for both bioinformatics tools could not be checked, because Sonah and colleagues did not provide the sequences or ID of their effector proteins. In the case of the candidates predicted by EffHunter, all of them met EffHunter's criteria, reinforcing its high accuracy and its low false positive rate. EffectorP 2.0 predicted a lower number of effectors in all these cases.

Altogether, EffHunter demonstrated that it is a highly efficient bioinformatics tool for fungal effector prediction, and it can be a suitable tool to search effectoromes in fungal proteomes.

Then, EffHunter was used to predict effectors on different types of fungi. The lowest numbers of effectors were predicted in yeast (Figure 4); consistent with what was reported by Sperschneider *et al.* [16]. These authors proposed that nonpathogenic fungi have less effectors than ectomycorrhiza and saprotrophs. EffHunter predicted a similar number of effectors in

ectomycorrhizal as in white and brown rots. In congruence with these findings, recent literature evidences that small-secreted effectors participate in all types of microbial interactions, and the concept “effector” seems to be rapidly evolving [67–72].

In plant pathogen interactions, effectors can be recognized by the cognate R proteins and trigger a hypersensitive response to prevent the spread of the pathogens. Some virulence factors are shared between plant and human fungal pathogens [73], but mammals have authentic immune systems, and their interactions with pathogens are different in comparison with plants. It is likely that some effectors from animal pathogens have different characteristics than phytopathogen effectors, and as a result, EffHunter could not identify them.

The highest numbers of effectors were predicted for plant pathogens and entomopathogens, suggesting that interactions with host producing toxic, antifungal metabolites demand a larger inventory of effectors. Consistent with this interpretation, *Metharrizium anisoplae*, a generalist entomopathogen, has 68% more effectors than the specialist *Metarrhizium acridum*, enabling *M. anisoplae* to face more divergent challenges. The next group according to the number of effectors was the saprotrophic fungi. Effectors in saprotrophs are probably used for antagonism or in the interaction with microorganisms which inhabit decaying wood. Another explanation is that saprotrophs have effectors because when circumstances change, they could become pathogens [74,75].

Although fungi with small genomes (i.e., yeast) have a smaller number of effectors (30-50 effectors), and the size of effectoromes is not related with the size of the genome. For instance, *Puccinia graminis* f.sp. *tritici* (88Mbp) has 659 effectors, *Blumeria graminis* f.sp. *tritici* (158 Mbp) has 161 effectors, *Magnaporthe oryzae* (41 MbP) has 486 effectors and *Melampsora lini* (189 Mbp) has 175 effectors.

The sizes of effectoromes seem to be related with lifestyles of the fungi: the lower number of effectors was observed in necrotrophs (average ~200 effectors). More complex interactions of biotrophs and hemibiotrophs require larger effectoromes (~300 and ~400 effectors, respectively). Evasion of host perception, suppressing host defense responses and keeping the host alive demand large catalogs of effectors in biotrophic and hemibiotrophic fungi.

As mentioned above, some fungi have unusual large effectoromes (600-700 effectors). The largest effectoromes were predicted for *Auricularia subgrabra* (708), *Puccinia graminis* f. sp. *tritici* (659) and *Melampsora larici-populina* (603) (Supplementary Table S2). Largest sets of effector

candidates in *Puccinia graminis* f.sp. *tritici* and *Melampsora larici-populina* are consistent with predictions by different programs, reported by Sperschneider *et al.* [15]; these authors proposed that these large effectoromes exist because these pathogens require two host species to complete their cycle of life. Recently, Liang *et al.* [38] investigated evolutionary features of the genes in obligate biotrophic fungal pathogens and reported that secreted effectors in powdery mildews of monocots have been subjected to positive selection, which explains the expansion of effectoromes in *P. graminis* f.sp. *tritici* and *M. larici-populina*. On the contrary, the families of secreted effectors in powdery mildews of dicots have been under strong purifying selection, resulting in the contraction in the number of effectors, e.g., *Melampsora lini* (175 effectors). The number of effectors in *Auricularia subgrabra* is large, probably also by expansion of the family of secreted proteins.

On the other hand, it is known that the characteristics of the fungi and oomycete effectors are different, but we took advantage of the fact that both of their effectors have signal peptides for secretion, and we used EffHunter to predict effectors in oomycetes. The number of candidates predicted by EffHunter in *P. infestans* (355) was similar to the prediction reported by Sonah *et al.* [10] (343 candidates). Supporting EffHunter's prediction, 295 candidates contain the motif RxLR, very common in oomycete effectors. This suggests that EffHunter is suitable for searching effectors in oomycete proteomes. Haas *et al.* [57], using Hidden Markov Models to retrieve proteins with oomycete motifs, predicted 563 effectors in *P. infestans*. Restriction by protein length and cysteine content by EffHunter can underestimate the number of effectors in oomycetes, but this pipeline can be used for easy and rapid preliminary searches.

It is important to highlight that any of the available effector predictors is capable to identify all effectors that have been experimentally studied so far [15]. Effector BEC1019, a haustorial protease from *Blumeria graminis* f.sp. *hordei* that suppresses host cell death, and AvrSr35, a 578 amino acids effector from *Puccinia graminis* f.sp. *tritici*, are neither identified by any effector classifier previously created, nor by EffHunter. EffectorP 2.0 does not retrieve the effectors Mg3LysM from *Zymoseptoria tritici* and CSEP0105 from *Bumeria graminis* f.sp. *hordei*; meanwhile, both of them are recognized by EffHunter. On the contrary, BEC1054 and BEC1011 from *Blumeria graminis* f.sp. *hordei* are identified by EffectorP 2.0, but EffHunter is not able to recognize them as effectors. A combination of different tools can increase sensitivity in effector prediction, adding the criteria and the predictive advantages of each tool. For example, a combination of EffectorP 1.0 and EffectorP 2.0 allowed the identification of AvrSr50 effector from *Puccinia graminis* f.sp. *tritici* [15]. However, larger lists make it more difficult to prioritize

candidates for functional validation. The main robustness of EffHunter is its low false positive rate in identifying bona fide canonical effectors. Although EffHunter ignores effectors with atypical characteristics, its high PPV (100%) and accuracy (ACC) (99%) make it a useful tool for the selection of top candidates. This is crucial because the number of fungal effectors per genome is in the order of hundreds [10,16] and false positives slow down the validation and characterization of effectors. After effector mining, high-priority candidates can be selected by filtering with additional criteria (when available) common among many known effectors such as in planta expression data, genomic location (e.g., comprising clusters of putative effectors, or locations in dispensable chromosomes), positive net charge and low content of serine and tryptophan, among others.

5. Conclusions

EffHunter is a pipeline that integrates the software SignalP 4.1, Phobius, TMHMM 2.0 and WoLFPSORT with Perl scripts to filter proteins by length and by cysteine content to search for fungal protein effectors in a single step. This makes EffHunter a user-friendly and amenable tool.

EffHunter is a robust tool that can identify effectors in fungal proteomes, showing higher accuracy and lower false positives than other effector predictors do.

Different types of fungi have varying quantities of effectors. Although exceptions were observed, there are averages in the number of effectors in each type of fungi. The results of our effectoromics study showed that plant pathogens and entomopathogens were the organisms with the largest effectoromes. Within plant pathogens, as it relates to their lifestyle, biotrophic and hemibiotrophic fungi have larger effectoromes than necrotrophic fungi.

6. Patents

The present pipeline was certified at Mexican Public Copyright Registry with the registration number 03-2019-101809310300-01.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1. Table S1: Analysis of the PHI-Database. Table S2: Description of proteomes analyzed by EffHunter. Table S3: List of phytopathogens classified according to their lifestyle. Table S4: List of all fungi and oomycetes species classified according to their lifestyle. Supplementary data set S1: Positive data set of true effector proteins. Supplementary data set S2: Negative control set ABC transport proteins. Supplementary data set S3: Negative control set cytochrome P450 proteins.

Supplementary data set S4: Negative control set major facilitator transporters (MFTS).
Supplementary data set S5: Positive and negative controls pooled in a single database.

Author Contributions: K.G.C.-A. and B.C.-C. conceived, designed the research and wrote the draft; K.G.C.-A. wrote the pipeline and performed the bioinformatics analyses; J.V.-A. contributed to the analysis and interpretation of the manuscript and provided feedback; L.S.-C., I.I.-F. and J.V.-A reviewed and provided feedback of the manuscript. K.G.C.-A. and B.C.-C. contributed equally. All authors have read and approved the final manuscript.

Funding: This research was funded by the National Council of Science and Technology (CONACYT) project FOSEC-SEP 220957, Mexico. We would like to thank the Yucatan Center for Scientific Research (CICY), Mexico, for the support in the realization of this study. Fellowship from CONACyT- 589301 for Carreón-Anguiano K.G. and the support by BioAli-CYTED are acknowledged.

Acknowledgments: English revision by Jewel N. A. Todd is acknowledged.

Conflicts of Interest: The authors declare no conflict of interest.

References

*En el apartado de referencias, números en “[]” representan a las usadas en éste artículo.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

CAPÍTULO III

IDENTIFICACIÓN *IN SILICO* DE PROTEÍNAS EFECTORAS DE *P. FIJIENSIS*.*

Artículo en preparación para publicación

3.1 INTRODUCCIÓN

Pseudocercospora fijiensis es un hongo hemibiotrófico, agente causal de la enfermedad Sigatoka negra en *Musa* spp (banano y plátano) (*Black Sigatoka Disease* “BSD”) (Churchill, 2011). Los patógenos hemibiotróficos, como *P. fijiensis*, *Colletotrichum higginsianum*, *Magnaporthe oryzae*, entre otros, colonizan inicialmente a los huéspedes a través de la invasión biotrófica, seguido de una transición a la etapa necrotrófica, logrando así, aprovechar todos los nutrientes del hospedero, llevándolo en la mayoría de las veces a su muerte (Pandaranayaka *et al.*, 2019).

El genoma de *P. fijiensis* (aislado CIRAD86) ha sido comparado con otros 18 genomas de hongos Dotideomicetos para realizar un análisis global de organización y evolución genómica, concluyendo que junto con *C. fulvum* poseen los genomas más grandes entre los hongos, con 74.1 Mb de *P. fijiensis* y 61.1 Mb de *C. fulvum*, coneniendo 39.5% y 44.4% de secuencias repetidas, respectivamente (Chang *et al.*, 2016; Ohm *et al.*, 2012; de Wit *et al.*, 2012). También se observó que en *P. fijiensis* en conjunto con *Mycosphaerella graminicola*, *Setosphaeria turcica*, *Stagonospora nodorum* y *Cochliobolus heterostrophus* C5, más del 50% de sus genes predicen un dominio PFAM. Una secuenciación posterior y actualización de la base de datos de un nuevo aislado denominado CIRAD186 confirmó a *P. fijiensis* como el hongo con el genoma más grande de todos los integrantes de la clase Dotideomicetos, con un aproximado de 74Mb, con una predicción de 13,107 genes, mientras que el tamaño del genoma de otros hongo es en promedio 40Mb (Arango *et al.*, 2016; Chang *et al.*, 2016).

Muchos fitopatógenos establecen la existencia de enfermedades en las plantas al detectar la secreción o presencia de una serie de moléculas que manipulan los procesos metabólicos vegetales y mitigan las respuestas inmunitarias de las plantas, para crear un entorno favorable para el crecimiento del patógeno (Robin *et al.*, 2018). Estas moléculas se conocen con el término "efectores", un efector se considera una molécula de tipo proteína, carbohidrato, metabolitos secundarios, etc, que están potencialmente involucrados en el proceso de infección (Lu y Edwards, 2016; Sonah *et al.*, 2016).

Las proteínas efectoras son en su mayoría proteínas secretadas; no son conservadas en las secuencias, pero coinciden en algunas características estructurales que se han vuelto criterios para su identificación. Estos criterios incluyen: péptido señal para la secreción, ausencia de dominios transmembranales, en general no similitud con dominios funcionales conocidos, pero actualmente se ha observado en algunos efectores, dominios funcionales asociados a patogenicidad, tamaño pequeño (<400 aminoácidos o <30 kDa) y contenido de cisteína enriquecido (por ejemplo, ≥ 4 residuos de cisteína) (Dörfors *et al.*, 2019; Jones *et al.*, 2018; Gabriëls *et al.*, 2016; Sonah *et al.*, 2016; Bowen *et al.*, 2009).

El análisis de proteomas de fitopatógenos predice una gran cantidad de posibles efectores, llegando a varios cientos en cada genoma fúngico (Robin *et al.*, 2018), sin embargo hasta el momento son pocos los efectores caracterizados, en el caso de *P. fijiensis*, sólo se ha reportado la caracterización de tres efectores, el PfAvr4, PfEcp2 y el PfEcp6 (Stergiopoulos *et al.*, 2010; Bolton *et al.*, 2008). Los efectores Avr4 y Ecp2 de *P. fijiensis*, renombrados PfAvr4 y PfEcp2, son ortólogos funcionales de los efectores CfAvr4 y CfEcp2 de *Cladosporium fulvum*, reportados como factores de virulencia (de Wit *et al.*, 2009).

La proteína PfAvr4, presenta un porcentaje de identidad 42% con la proteína efectora de *C. fulvum* y desencadena una respuesta hipersensible en plantas de tomate transformadas con el receptor Cf-4, así mismo, actúa como factor de virulencia defensivo para proteger las paredes celulares de los hongos contra la hidrólisis de las quitinasas vegetales al unirse a la quitina. Además, se identificaron tres proteínas en el genoma de *P. fijiensis* que muestran homología con el Ecp2 de *C. fulvum*, los efectores PfEcp2, PfEcp2-2 y PfEcp2-3, con porcentajes de identidad de 57%, 28% y 25% respectivamente. Sin embargo, solo PfEcp2 es capaz de desencadenar una respuesta hipersensible en líneas de tomate transformadas con Cf-2 (Stergiopoulos *et al.*, 2010).

Análisis bioinformáticos han permitido predecir la existencia de numerosos efectores en *P. fijiensis*, por ejemplo, Chang y colaboradores (2016) reportaron la identificación de 105 efectores mediante las herramientas bioinformáticas SignalP 4.1, TMHMM 2.0, TargetP 1.1, big-PI, WoLFPSORT, ProtComp, permitiendo la presencia de máximo un dominio transmembranal, definiendo la longitud de proteína en ≤ 250 aminoácidos y el contenido de residuos de cisteínas en $\geq 2\%$.

Un análisis transcriptómico realizado por Noar y Daub, (2016) reportó un total de 231 efectores. Los criterios usados fueron la presencia de péptido señal, una longitud de ≤ 300 aminoácidos y

una presencia de ≥ 4 residuos de cisteínas. Estas secuencias fueron identificadas analizando el transcriptoma de la interacción con *Musa* spp.

La predicción del número de efectores es variable según los predictores que se usen y los límites de tamaño de proteína y contenido de cisteína que se fijen, entre otros criterios. Como se mostró en el capítulo anterior, EffHunter es el algoritmo predictor de efectores fúngicos con mayor robustez actualmente, teniendo el índice F más alto entre los predictores de efectores fúngicos disponibles (Carreón-Anguiano *et al.*, 2020). En el reporte de la construcción de EffHunter se comparó su desempeño en la predicción de los efectoromas en diferentes hongos, comparando con los reportes respectivos de otros autores, en los que se usaron diferentes estrategias en cada caso, ajustándose en el algoritmo EffHunter el tamaño de las proteínas, presencia o no de dominios transmembranales, según las condiciones que hubieran usado los autores del reporte correspondiente. En el caso de *P. fijiensis* se comparó con la predicción de Chang *et al.*, (2016) que seleccionaron proteínas con péptido señal, aceptaron un dominio transmembranal, rechazaron la presencia de dominio de anclaje GPI (*Glycosyl-phosphatidylinositol-anchored*) y definieron la longitud máxima de las proteínas en 250 aminoácidos. En esas condiciones EffHunter predijo 136 candidatos, pero el número pudiera estar subestimado por enfocarse únicamente a proteínas muy pequeñas y con los criterios de esos autores.

En este capítulo se presenta la predicción de efectores canónicos de *P. fijiensis* usando el pipeline bioinformático EffHunter (Carreón-Anguiano *et al.*, 2020). Para fortalecer la predicción y ampliar el efectoroma con efectores no canónicos, se incluyen también la búsqueda con EffectorP 2.0 (Sperschneider *et al.*, 2018) y minería adicional empleando criterios adicionales como la búsqueda de dominios funcionales y análisis de distribución taxonómica de homólogos, esperando obtener el efectoroma completo de *P. fijiensis*.

3.2 MATERIALES Y MÉTODOS.

3.2.1 Identificación y caracterización *in silico* de candidatos a efectores de *Pseudocercospora fijiensis*

Efectores canónicos

Para la predicción e identificación de los efectores canónicos en *P. fijiensis* se utilizó el pipeline bioinformático EffHunter v1.0 (Carreón-Anguiano *et al.*, 2020) compilando la búsqueda

considerando presencia de péptido señal, ≥ 4 residuos de cisteína (Marton *et al.*, 2018; Duplessis *et al.*, 2011), ausencia de dominios transmembranales, longitud de ≤ 400 aminoácidos (Dörfors *et al.*, 2019; Neu *et al.*, 2019; Esna *et al.*, 2018; Krijger *et al.*, 2014; Zou *et al.*, 2013; Bowen *et al.*, 2009), las cuales con las condiciones que fueron usadas en la evaluación de EffHunter en la que su desempeño fue superior a los otros predictores (Carreón-Anguiano *et al.*, 2020). El algoritmo fue ejecutado en un ambiente de SO Linux/Unix (https://github.com/GisCarreon/EffHunter_v.1.0), usando en formato FASTA la base de datos del proteoma total reportado en la página del JGI MycoCosm *Pseudocercospora* (*Mycosphaerella*) *fijiensis* v2.0 (<https://mycocosm.jgi.doe.gov/Mycofi2/Mycofi2.home.html>). El programa EffectorP 2.0 fue ejecutado con los parámetros predeterminados (<http://effectorp.csiro.au/>) (Sperschneider *et al.*, 2018), para su aplicación se requirió de un paso adicional, previo, para predecir el secretoma de *P. fijiensis* con el pipeline bioinformático SECRETOOL (<http://genomics.cicbiogune.es/SECRETOOL/Secretool.php>) (Cortázar *et al.*, 2014), y los resultados en formato FASTA, fueron sometidos al programa EffectorP 2.0.

Efectores no canónicos

Para buscar efectores no canónicos potenciales, el proteoma total deducido es decir, no filtrado ni por EffHunter ni por EffectorP, se analizó con la tubería de InterProScan para identificar los dominios relacionados con la patogénesis. Los dominios funcionales que han sido identificados en los efectores fúngicos comprenden LysM, NPP, Hce2, CBM14, ceratoplatanina, CFEM, hidrofobina, lectina, isocorismatasa, DUF, HsbA, cutinasa, unión de quitinasa y concanavalina A (Liu *et al.*, 2019; Volk *et al.*, 2019; Zhang *et al.*, 2019; Casarrubia *et al.*, 2018; Gui *et al.*, 2018; Uhse y Djamei, 2018; Zhang *et al.*, 2018; Dalio *et al.*, 2017; Chang y Stergiopoulos, 2015; Baccelli, 2015; Guyon *et al.*, 2014; Liu *et al.* al., 2014; Kombrink y Thomma, 2013; Bateman *et al.*, 2010).

3.2.2. Análisis de la longitud de aminoácidos y residuos de cisteínas de los efectores predichos

El conteo de los aminoácidos en longitud total y el conteo de residuos de cisteína de cada secuencia fueron realizados en la base de datos creada en Microsoft Excel versión 15.0.5301.1000 (64-bit). Se emplearon dos fórmulas principales, la fórmula 1) Utilizando la función "Length", función que cuenta letras, números, caracteres y todos los espacios de la celda seleccionada (donde se encuentra depositada la secuencia), formula 2) En esta también se utilizó

la misma función “Length” acoplada con la función “Substitute” (función que sustituye un texto por otro en la cadena de texto especificada) para los residuos de cisteínas en cada secuencia de la base. Las fórmulas usadas fueron las siguientes:

$$\text{Longitud de secuencia} = \text{LENGTH}(\text{celda}) \quad (1)$$

$$\text{Residuos de cisteína} = \text{LENGTH}(\text{celda}) - \text{LENGTH}(\text{SUBSTITUTE}(\text{celda}, "C", "")) \quad (2)$$

3.2.3. Anotación funcional de proteínas efectoras y su localización subcelular.

Para la predicción de dominios funcionales se utilizó el pipeline bioinformático y la base de datos del InterProScan (versión 5.39-77.0) (Quevillon *et al.*, 2005), en modo STANDALONE ejecutándose en ambiente SO LINUX/UNIX. Las secuencias se analizaron contra las bases de datos del PFAM (Fin *et al.*, 2015), CDD (Marchler-Bauer *et al.*, 2015) y SMART (Schultz *et al.*, 2000). Para encontrar coincidencias se estableció como filtro un E_{value} de $\leq 1e-4$.

Para tener información de la probable localización subcelular de cada secuencia, se empleó el programa LOCALIZER (<http://localizer.csiro.au/>) (Sperschneider *et al.*, 2017), programa de aprendizaje automático para la predicción de localización subcelular de proteínas efectoras eucarióticas en células vegetales (ya sea en cloroplasto, mitocondria o núcleo).

3.2.4 Localización genómica (scaffolds).

El análisis de la organización genómica de los candidatos, se llevó a cabo consultando la base de datos del JGI MycoCosm (<https://mycocosm.jgi.doe.gov/Mycfi2/Mycfi2.home.html>), obteniendo el número de “scaffold”. Los datos fueron compilados en una base de datos y analizados, se clasificaron según el número de scaffold.

3.2.5 Búsqueda de homólogos de los efectores de *P. fijiensis*

Para la búsqueda de homólogos, se usó la base de datos específico de hongos (taxid 4751), del NCBI, no redundante, ajustando los parámetros de la búsqueda para un máximo de 50 secuencias blanco, un corte Evalue de $1e-10$ y empleando una matriz BLOSUM62 (Figura 3.1). Para la distribución de la taxonomía, se clasificaron cuatro grupos: distribución ampliamente filogenética, efectores del complejo Sigatoka (compartido con *P. eumusae* y *P. musae*), efectores

compartidos con hongos parientes cercanos (*P. eumusae* o *P. musae*) y efectores específicos de *P. fijiensis*. La Figura 3.1 resume la búsqueda y caracterización de los efectores en *P. fijiensis*.

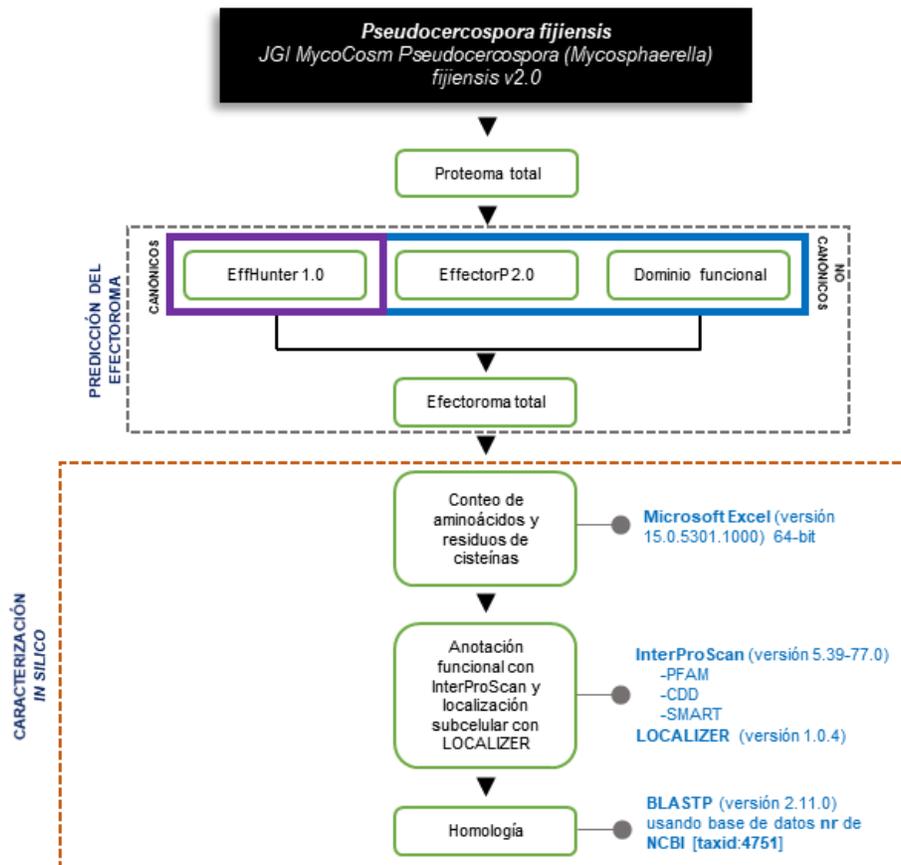


Figura 3.1 Esquemización de la caracterización *in silico* del efectoroma de *P. fijiensis*.

3.3 RESULTADOS

3.3.1 Predicción de efectores en *P. fijiensis*

Efectores canónicos

El algoritmo EffHunter identificó 240 proteínas efectoras, las cuales cumplen con los criterios de longitud de aminoácidos (≤ 400 aas) y residuos de cisteínas (≥ 4) presentes en cada secuencia, así como la presencia de péptido señal y ausencia de dominios transmembranales. La base de datos generada incluye a dos de los efectores ya caracterizados, PfAvr4 (ID 87167) y PfEcp2 (ID

52972), cabe mencionar que el PfEcp6 (ID 202004) no se colectó con EffHunter, pues tiene una longitud de más de 400 residuos de aminoácidos.

Efectores no-canónicos.

En el caso de los efectores no canónicos, primero se recuperaron mediante la búsqueda de proteínas que contienen dominios relacionados con la patogénesis y se refinó la búsqueda mediante la selección de proteínas que presentan una distribución taxonómica discontinua, es decir, están en algunas especies de los géneros en los que se encuentran (Guyon *et al.*, 2014). Este grupo de candidatos (no canónicos) incluye el efector PfEcp6, el cual fue descartado por EffHunter por tener una longitud de 413 aminoácidos.

Se recuperaron 481 proteínas, 73 de las cuales se compartieron con el conjunto de efectores canónicos, es decir, cumplen con sus criterios de tamaño, contenido de cisteínas, etc., y poseen dominio relacionado a patogénesis. Por lo tanto, luego de filtrar estos efectores canónicos, esta búsqueda resultó en 408 candidatos nuevos.

Debido a la dificultad de identificar efectores se recomienda usar más de una herramienta predictiva (Neu y Debener, 2019; Dalio *et al.*, 2018; Sonah *et al.*, 2016; Jones *et al.*, 2018), por lo que en este trabajo, se empleó también el programa EffectorP 2.0 para fortalecer la búsqueda de efectores en *P. fijiensis*, con el que se reconocieron 162 posibles efectores, coincidiendo 91 de ellos con los identificados por EffHunter (candidatos efectores canónicos) y otros 57 que fueron previamente recuperados por búsqueda de dominio (candidatos efectores no canónicos). EffectorP 2.0 identificó 14 nuevos candidatos, que no cumplen con los criterios canónicos, por lo que son candidatos no canónicos, aunque no poseen dominios conocidos relacionados con la patogénesis. Estos 14 nuevos candidatos ampliaron el grupo no canónico, a un total de 422 candidatos. Por lo tanto, se predice un efectoroma de al menos 662 posibles efectores (Cuadro 3.1).

Cuadro 3.1 Predicción y clasificación del efectoroma de *P. fijiensis*.

Efectores predichos en <i>P. fijiensis</i>				
Ánàlisis	No. de proteínas	Rango de longitud	Rango No. de residuos de Cys	Secuencias con Péptido señal
		Canónicos		

EffHunter 1.0	240	70 – 397	4Cys – 24Cys	240
No canónicos				
EffectorP 2.0	14	53 – 269	0Cys – 10Cys	12
*Dominio funcional	408	72 – 3839	0Cys – 52Cys	53

*Dominios funcionales empleados para la búsqueda: LysM (*Lysin motif*), NPP1 (*Necrosis inducing protein*), Hce2 (*Homolog of Cladopsorium fulvum Ecp2 effector*), CBM14 (*Carbohydrate-binding module family 14*), ceratoplatanina, CFEM (*Common in several fungal extracellular membrane proteins*), hidrofobina, lectina, isocorismatasa, cutinasa, DUF, HsbA, unión a quitinasa, CAP (*Cysteine-rich secretory protein*), salicilato hidroxilasa y concanavalina A-like.

**Sólo se incluyen los miembros adicionales, según el criterio de búsqueda. Por ejemplo, EffectorP 2.0 reconoció 162 pero sólo 14 son únicamente reconocidos por EffectorP 2.0.

3.3.2 Caracterización de los efectores predichos de *P. fijiensis*

Los 240 candidatos de efectores de *P. fijiensis* presentan amplio rango en longitud de aminoácidos de 70 a 397 aminoácidos, y los no canónicos de 53 a 3839 aminoácidos.

En cuanto al número de residuos de cisteína, el efectoroma canónico contiene desde 4 hasta 24 cisteínas en la secuencia, mientras que los efectores no canónicos tienen desde 0 hasta 52 residuos de cisteínas. Con base en el contenido de cisteínas, se agruparon 5 subgrupos con los efectores canónicos y 6 subgrupos con los efectores no canónicos (Cuadro 3.2). En ambos subgrupos el mayor número de candidatos contiene menos de 5% de cisteínas, aunque en proteínas efectoras no canónicas de gran tamaño, algunas pueden contener más de 50 residuos de cisteínas.

Cuadro 3.2 Clasificación del efectoroma de *P. fijiensis* de acuerdo al contenido de cisteínas.

Porcentaje de Cys	No. de residuos de Cys	Longitud de los candidatos	No. de proteínas
Canónicos			
1 – 2.99	4 – 11	136 – 397	134
3 – 4.99	4 – 17	87 – 390	66
5 – 6.99	4 – 12	70 – 212	23
7 – 8.99	6 – 14	82 – 185	14
9 – 11.11	11 – 24	108 – 254	3
No canónicos			
0 – 1	0 – 25	72 – 3839	195
1 – 2.99	1 – 52	53 – 3161	190

3 – 4.99	3 – 41	80 – 1040	32
5 – 6.99	4 – 36	72 – 631	3
7.46	10	134	1
11.36	10	88	1

3.3.3 Anotación funcional y localización subcelular del efectoroma de *P. fijiensis*

Para la anotación funcional de la base de datos canónica, se emplearon las bases de datos depositadas en el pipeline bioinformático InterproScan (<http://www.ebi.ac.uk/interpro/search/sequence/>), resultando en la identificación de 74 proteínas con dominio funcional, mientras que 166 candidatos no tuvieron dominios conocidos; entre los dominios funcionales, se pudo observar CFEM, CAP, NPP1, así como también dominios de hidrolasas, glicosil hidrolasas, tirosinasas, entre otros. En el Cuadro 3.3 se muestran los dominios funcionales encontrados en la base de datos de efectores canónicos y en el Cuadro 3.4 los efectores no canónicos.

Cuadro 3.3 Dominios funcionales identificados en la base de datos de efectores canónicos de *P. fijiensis*.

Efector candidato	Efectores candidatos	Residuos de Cys
3-hydroxyacyl-CoA dehydrogenase, NAD binding domain	1	7
Aldose 1-epimerase	1	6
Alpha/beta hydrolase family	1	4
CAP	2	5
Carboxylesterase family	1	5
CFEM	3	4
		9
		11
		10
Common central domain of tyrosinase	3	8
		8
		7
Concanavalin A-like	2	10
lectin/glucanase/Glycosyl hydrolases family		4
16		
Cupin	1	4
Cutinasa	5	4
		4
		7
		4

		4
Deoxyribonuclease NucA/NucB	1	11
Deuterolysin metalloprotease (M35) family	1	7
Dioxygenase	1	4
DUF	4	4
		4
		6
		7
FAD binding domain	1	5
F-box domain	1	10
Hidrofobina	1	6
GDSL-like Lipase/Acylhydrolase	3	4
		7
		5
Glutathione S-transferase, C-terminal domain	1	4
Glucosido hidrolasa (familia 10, 17,18, 28, 43, 61)	9	6
		6
		7
		9
		4
		8
		4
		4
		4
Histidine phosphatase superfamily (branch 1)	1	6
Isocorismatasa	1	5
Lipasa (clase 2)	1	7
LysM	1	9
Lytic polysaccharide mono-oxygenase,	1	4
Cellulose/chitin-binding protein		
Lytic transglycolase	1	6
NAD(P)H-binding	1	5
Necrosis inducing protein (NPP1)	1	5
Palmitoyl protein thioesterase	1	4
PAN	2	12
		7
Pathogen effector	2	10
		5
Pectate lyase	1	8
Peptidase_M43	2	8
Peptidase_S10	1	7
Peroxidase, family 2	1	4
Putative peptidase family	1	6
Rapid ALKalinization Factor (RALF)	2	5
		5
RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	1	5
S1/P1 Nuclease	2	5
		10
Ser-Thr-rich glycosyl-phosphatidyl-inositol-anchored membrane family (GPI-anchored)	2	4
Survival protein SurE	1	7
Tiorredoxina	1	4
Transglycosylase SLT domain	1	4
Tripsina	1	5

Vacuolar protein sorting-associated protein	1	4
62		

*CAP: *Cysteine-rich secretory protein*, CFEM: *common in several fungal extracellular membrane proteins*, DUF: *Domain with unknown function*, Hce2: *Homolog of Cladopsorium fulvum Ecp2 effector*, NPP1: *Necrosis inducing protein 1*, LysM: *Lysin motif*.

En el caso de los efectores no canónicos, primero se recuperaron mediante la búsqueda de proteínas que contienen dominios relacionados con la patogénesis y se trató de afinar seleccionando luego las proteínas con homólogos que muestran distribución taxonómica discontinua (Sędzielewska y Brachmann, 2016; Guzmán-Guzmán *et al.*, 2017; Dalio *et al.*, 2018; An-dong *et al.*, 2020; Kanja y Hammond-Kosack, 2020).

Cuadro 3.4 Dominios funcionales identificados en la base de datos de efectores no canónicos de *P. fijiensis*.

Dominio	Efectores no-canónicos	Residuos de Cys
DUF	319	0-52
Concanavalina A-like lectina	33	0-30
CAP	2	4
Ceratoplatanina	1	4
CFEM	8	8-20
FAD binding	10	4-10
Cutinasa	2	4-13
Isocorismatasa	7	2-14
Unión a quitina	2	13-36
Proteína A de unión a superficie hidrofóbica	15	0-2
LysM	6	1 - 24
NPP1	1	2
Hce2	2	3 -5

*CAP: *Cysteine-rich secretory protein*, CFEM: *common in several fungal extracellular membrane proteins*, DUF: *Domain with unknown function*, Hce2: *Homolog of Cladopsorium fulvum Ecp2 effector*, NPP1: *Necrosis inducing protein 1*, LysM: *Lysin motif*.

Cuadro 3.5 Dominios funcionales compartidos entre las bases de efectores canónica y no canónica de *P. fijiensis*.

Dominio	Efectores canónicos	Efectores no canónicos
CAP	2	2
CFEM	3	8
Cutinasa	5	2

DUF	4	316
Hce2	1	2
Isocorismatasa	1	7
LysM	1	6
NPP1	1	1
Union a quitina	1	3

*CAP: *Cysteine-rich secretory protein*, CFEM: *common in several fungal extracellular membrane proteins*, DUF: *Domain with unknown function*, Hce2: *Homolog of Cladopsorium fulvum Ecp2 effector*, NPP1: *Necrosis inducing protein 1*, LysM: *Lysin motif*.

Aunque los dominios se identificaron solo en 74 candidatos canónicos, en comparación con la base de 408 candidatos en el grupo no canónico que, por su criterio de búsqueda, todos tienen dominios, se identificó un mayor número de dominios diferentes en la base canónica (45 dominios) (Cuadro 3.3), en comparación con el subgrupo no canónico con 13 dominios diferentes (Cuadro 3.4 y 3.5).

Entre la base de datos de efectores canónicos, se identificó una proteína con dominio de unión superficial Hidrofóbica A (HsbA) y una hidrofobina, mientras que en la base de datos de efectores no canónicos se identificaron 15 proteínas con dominio HsbA, de los cuales, 9 fueron altamente expresados en tejido foliar infectado (log₂FC de 2 a 5.58) (datos obtenidos del trabajo reportado por Noar y Daub, 2016).

En cuanto a la localización subcelular, LOCALIZER, predijo que 43 efectores canónicos se localizan en diferentes compartimentos celulares: 13 secuencias mostraron localización en el núcleo (6%), 15 localización en cloroplasto (6%), 7 secuencias con localización en mitocondria (3%) y 8 secuencias tuvieron un resultado ambiguo (3%), mientras que 197 (82%) candidatos no pudieron definirse en cuanto a localización subcelular (Cuadro 4.6). En cuanto a los efectores no canónicos, 192 se predicen en diferentes compartimentos celulares, 20 secuencias con localización en cloroplasto (4.7%), 24 secuencias con localización a mitocondria (5.6%), 98 secuencias con localización en núcleo (23.2%) y 50 secuencias con un resultado ambiguo (11.8%), 230 proteínas no tuvieron localización subcelular definida (54.5%) (Cuadro 3.6).

Cuadro 3.6 Predicción de localización subcelular de los efectores de *P. fijiensis*.

Localización subcelular	Efectores canónicos	Efectores no canónicos
Cloroplasto	15	20
Mitocondria	7	24

Núcleo	13	98
Ambiguo	8	50
No predicción	197	230

La mayoría de los candidatos tienen como descripción “Hypothetical protein”, “Uncharacterized protein” o “unnamed protein product”, las funciones de los efectores son poco conocidas y estos resultados coinciden con las descripciones en otros efectoromas, en los que la mayor parte de los candidatos resultan con función desconocida (Sonah *et al.*, 2016; Sertzielewska y Brachmann, 2016; Neu y Debener, 2019).

3.3.4 Localización genómica de los efectores de *P. fijiensis*

Para explorar la organización genómica de los efectores, se clasificaron según los scaffold en los que se encuentran. El gráfico revela que los efectores canónicos predichos por EffHunter y los candidatos no canónicos se distribuyen de manera similar a través de los mismos 12 scaffolds (Figura 3.2). La mayoría de ellos están ubicados en los scaffolds uno y dos. Con la excepción del scaffold 19 que alberga el mismo número de efectores canónicos y no canónicos (4 de cada conjunto), todos los demás scaffolds contienen más efectores no canónicos que canónicos. Curiosamente, la proporción de ambas clases de efectores fue similar en todos los scaffolds, excepto en el scaffold 19. Por ejemplo, el scaffold 1 contiene el 17,9% de los candidatos canónicos y el 19,4% de los candidatos no canónicos; en el scaffold 2, su proporción fue de 16,2 y 14,9% respectivamente. No se produce concentración de efectores de alguna longitud particular o contenido de cisteína en ningún scaffold, sino que en todos los scaffold se distribuyen candidatos con rangos similares tanto de longitudes (200 a 300 aminoácidos) como de riqueza de cisteína (de 4 a 20 residuos de Cys).

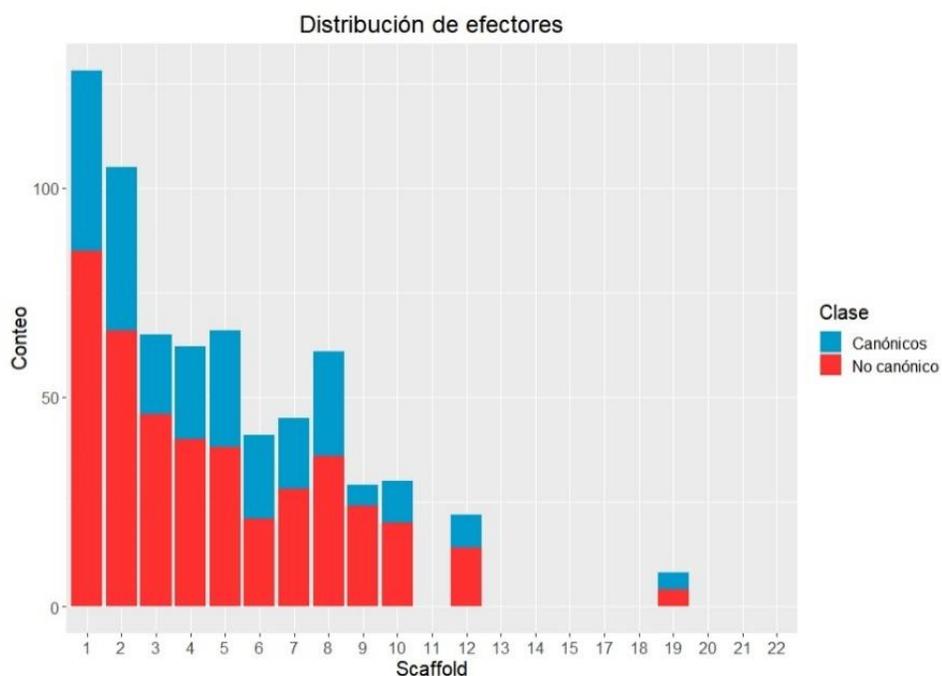


Figura 3.2 Distribución genómica del efectograma predicho de *P. fijiensis*.

3.3.5 Homólogos de los efectores de *P. fijiensis*

Cerca del 10% de los posibles efectores encontrados en *P. fijiensis* tuvieron homólogos en *P. eumusae* y/o *P. musae*, mientras que el 24% parecen ser efectores específicos de *P. fijiensis*, al no encontrarse homólogos en otras especies. La mayoría de los efectores candidatos (40%) tienen homólogos en algunas especies de los géneros estrechamente relacionados, como *Cercospora*, *Dothistroma* y *Zymoseptoria*, además de *P. eumusae* y *P. musae*; 213 (26%) candidatos tuvieron una distribución taxonómica amplia, incluyendo en hongos no relacionados y con una distribución filogenética discontinua de los homólogos; es decir, algunos posibles efectores se observaron en varios géneros, pero solo en unas pocas especies de ellos, lo cual ya se había observado en la distribución taxonómica de otros efectores (Sędzielewska-Toro y Brachmann, 2016; Guyon *et al.*, 2014). El Cuadro 3.7 muestra la distribución taxonómica de los homólogos de los efectores de *P. fijiensis*.

Cuadro 3.7 Distribución de efectores candidatos de *P. fijiensis* con homología con otras especies.

Homología	Hits	Efectores canónicos	Efectores no canónicos
Ampliamente distribuido filogenéticamente	Algunos ejemplos: Especies de <i>Aspegillus</i> , <i>Fusarium</i> , <i>Verticillium</i> , <i>Ramularia</i> , <i>Bipolaris</i> , <i>Alternaria</i> , <i>Friedmanniomyces</i> , <i>Hortae</i> , etc. Hits con pocos o muchos de ellos por candidato.	145	68
Género de hongos relacionados	<i>Pseudocercospora eumusae</i> , <i>Pseudocercospora musae</i> , <i>Cercospora</i> , <i>Dothistroma</i> , <i>Zymoseptoria</i> , etc.	37	351
Complejo Sigatoka	Solo con <i>P. eumusae</i> y <i>P. musae</i>	14	0
Especies de hongos cercanos	<i>Pseudocercospora eumusae</i>	4	0
	<i>Pseudocercospora musae</i>	6	0
Específicos	Solo con <i>P. fijiensis</i>	34	3

3.4 DISCUSIÓN.

Es bien sabido que la identificación de efectores es complicada, ya que cada candidato puede cumplir uno o más de los criterios usados en la predicción de los efectores, la mayoría no cumple con todos los criterios establecidos para considerarse como tal. La predicción de efectores fúngicos es una tarea desafiante, debido a que los efectores están sujetos a una rápida coevolución con el hospedero y carecen de conservación significativa entre ellos (Neu y Debener, 2019; Sonah *et al.*, 2016; Sperschneider *et al.*, 2015). Por lo tanto, para la identificación de efectores fúngicos se emplean métodos de búsqueda basados en características estructurales que incluyen una longitud corta (menos de 300-400 aminoácidos), ser ricos en cisteína (4% o al menos cuatro residuos de cisteína), y ser secretados al entorno extracelular (Dörfors *et al.*, 2019; Neu y Debener, 2019; Esna *et al.*, 2018; Marton *et al.*, 2018; Krijger *et al.*, 2014; Zou *et al.*, 2013; Duplessis *et al.*, 2011; Bowen *et al.*, 2009). Sin embargo, no todas las proteínas efectoras comparten estas características (Dalio *et al.*, 2018; Sonah *et al.*, 2016; Jones *et al.*, 2018), por lo

que en este trabajo se emplearon dos enfoques, una búsqueda de efectores canónicos, y otro enfoque para identificar efectores no canónicos, para así, obtener el efectograma de *P. fijiensis*.

Para la búsqueda de efectores canónicos, utilizamos como predictor principal el pipeline EffHunter, que actualmente es el predictor de efectores fúngicos con el mayor índice F1. Sin embargo, EffHunter tiene una puntuación de falsos negativos similar a otros predictores (Carreón-Anguiano *et al.*, 2020), por lo que se combinó el uso de este con el predictor EffectorP 2.0 (Sperschneider *et al.*, 2018) para obtener una mayor cantidad de predicción de posibles efectores. Con EffHunter se predijeron 240 efectores canónicos, lo cual es un mayor número al encontrado en *P. fijiensis* por Chang *et al.*, (2016), una de las posibles razones sería el hecho de que en este trabajo se consideró como límite de longitud 300 aminoácidos mientras que nosotros establecimos el límite en 400 aminoácidos con el programa EffHunter. El tamaño del efectograma canónico predicho en este estudio para *P. fijiensis* es similar al tamaño encontrado en la secuencias para *Blumeria graminis* (255), *Magnaporthe oryzae* (273) y *Lesptosphaeria maculans* (290) (Carreón-Anguiano *et al.*, 2020).

EffHunter ha demostrado que puede ser utilizado sobre proteomas completos, mientras que EffectorP 2.0 necesita analizar el secretoma previamente identificado. Los análisis comparativos con EffectorP 2.0 arrojaron que de las 240 proteínas efectoras identificadas por EffHunter, 91 son predichas como efector por EffectorP 2.0. Dentro de la base de datos resultante, los efectores conocidos PfAvr4 (ID 87167) y PfEcp2 (ID 52972) de *P. fijiensis*, ortólogos funcionales identificados de los efectores de *Cladosporium fulvum* (Stergiopoulos *et al.*, 2010), fueron recuperados por EffHunter en el conjunto de candidatos efectores canónicos, también fueron identificados como efectores por EffectorP 2.0. El análisis para identificar efectores no canónicos con EffectorP 2.0 arrojó 14 nuevos hits, que no cumplen con los criterios canónicos y no predicen dominios conocidos relacionados a la patogénesis, siendo estos candidatos no canónicos. Este conjunto tiene un rango de longitud de 53 a 269 aminoácidos, y de cero a 10 residuos de Cys.

Como se mencionó anteriormente, los candidatos efectores no canónicos se recuperaron del proteoma de *P. fijiensis* mediante la búsqueda de dominios relacionados con la patogenicidad. Otros reportes también han usado este criterio para caracterizar a sus candidatos a efectores (Mukhi *et al.*, 2020; Xu *et al.*, 2019; Franceschetti *et al.*, 2017; Guzmán-Guzmán *et al.*, 2017; Pennington *et al.*, 2016). En el presente trabajo esta estrategia se utilizó primero para buscar efectores no canónicos y luego para caracterizarlos, lo que difiere a lo previamente realizado en otros trabajos similares (Neu *et al.*, 2019; Guo *et al.*, 2014; Morais do Amaral *et al.*, 2012;

O'Connell *et al.*, 2012). Esto expande el efectograma de *P. fijiensis* en 408 candidatos no canónicos, mientras que solo 74 candidatos canónicos tienen dominios relacionados con la patogenicidad. Con esta última búsqueda se recuperó el efector PfEcp6 (ID 202004) de *P. fijiensis*, que había sido excluido en la búsqueda de candidatos canónicos pues es una proteína de 413 aminoácidos. La suma de los candidatos no canónicos predichos en *P. fijiensis* es al menos de 422 candidatos (408 identificados por dominios + 14 identificados por EffectorP 2.0).

El análisis de la distribución de los candidatos de efectores en los scaffold reveló que la mayoría están ubicados en los scaffolds uno y dos, pero se distribuyen en dos scaffolds en proporciones similares de efectores canónicos y efectores no canónicos (Figura 3.2). La coexistencia de ambos tipos de efectores en los mismos scaffolds y la observación de equilibrio entre los porcentajes de ambos tipos de efectores no parece un hecho casual y apoya la identificación *in silico* de efectores realizada en el presente trabajo.

Uno de los criterios más frecuentemente para identificar efectores es el conteo de cisteína, pero no se pueden descartar los candidatos que no lo cumplen, ya que existen efectores verdaderos que no son ricos en residuos de Cys como en *Melampsora lini* con el efector AvrM, que tiene una longitud de secuencia de 314 aminoácidos y solo un residuo de Cys (Catanzariti *et al.*, 2004), BAS1 de *Magnaporthe oryzae* con 115 aminoácidos de longitud y ningún residuo de Cys (Yang *et al.*, 2019) y Cmu1 de *Ustilago maydis* con 290 aminoácidos de longitud y 2 residuos de Cys (Han *et al.*, 2019). Los resultados de las predicciones de efectores en *P. fijiensis*, hasta ahora encontrados va de 105 proteínas (Chang *et al.*, 2016) hasta 231 proteínas (Noar y Daub, 2016), por lo que el efectograma predicho en este trabajo es más grande. La distribución taxonómica discontinua de los homólogos de los candidatos a efectores que encontramos refuerza la efectividad del número predicho por nuestro análisis.

Se detectaron 74 efectores canónicos de *P. fijiensis* que predicen dominios funcionales variados, relacionados con patogenicidad. También se identificó una proteína con unión superficial hidrofóbica A (HsbA) y una hidrofobina entre los candidatos efectores. Las proteínas que predicen dominio HsbA son capaces de reclutar enzimas líticas a la superficie y promover la degradación de la pared celular (Feldman *et al.*, 2020; Valette *et al.*, 2017; Ohtaki *et al.*, 2006). En la base de datos de efectores no canónicos se identificaron 15 HsbA, de los cuales nueve de ellos se vieron altamente expresados (Log₂FC de 2 a 5,58) en tejido foliar infectado, según lo reportado por Noar y Daub (2016), lo que sugiere un papel importante de estas proteínas durante la interacción.

Entre los efectores con funciones de hidrolasas se encontraron carboxilesterasas, palmitoil proteína tioesterasa, lipasa/acilhidrolasa similar a GDSL, pectato liasa, celulasas, familias de glicosil hidrolasa (10, 61, 16,17,18,28, 43) y cutinasa, así como proteasas como peptidasa A4, peptidasa M43, peptidasa S10 y tripsina. Todas estas funciones han sido descritas en la interacción patógeno-hospedador (Medina *et al.*, 2020; Franceschetti *et al.*, 2017; Pennington *et al.*, 2016; Sędzielewska *et al.*, 2016).

Curiosamente, 8 candidatos (uno en el grupo canónico y siete en el grupo no canónico) predicen hidrolasas de isocorismato y 10 candidatos no canónicos predicen hidroxilasas de salicilato. Estos efectores tienen sus dianas en el metabolismo del salicilato en la célula hospedera; el primero hidroliza el isocorismato, el precursor del salicilato, y el segundo cataliza la hidroxilación descarboxilativa del salicilato en catecol. Liu *et al.*, (2014) reportaron dos efectores no convencionales, Pslsc1 de *Phytophthora sojae* y VdIsc1 de *Verticillium dahliae*, correspondientes a isocorismatasas secretadas, las cuales suprimen la acumulación de ácido salicílico (SA) en las células del hospedero y evitan la señalización que resultaría en una reacción de hipersensibilidad (RH) (Zhu *et al.*, 2017; Toruño *et al.*, 2016; Liu *et al.*, 2014).

Dos candidatos canónicos predicen factores de alcalinización rápida (RALF). Estos pequeños péptidos secretados ricos en cisteína se describieron por primera vez en plantas, en las que participan en muchos procesos diferentes. Se han identificado homólogos de RALF en fitopatógenos fúngicos, en los que desempeñan un papel en la alcalinización de la célula hospedera, la activación de factores de virulencia y la infección del hospedero (Wood *et al.*, 2020; Thynne *et al.*, 2017). Se identificaron dos homólogos de ceratoplatanina en la base de datos de efectores no canónicos. La ceratoplatanina se describió por primera vez en el ascomyceto *Ceratocystis fimbriata* como una proteína fitotóxica secretada (Carresi *et al.*, 2006). Las ceratoplataninas son proteínas ricas en cisteína, ampliamente distribuidas en los hongos Ascomycota y Basidiomycota, secretadas extracelularmente en interacciones hongo-hospedero (Chen *et al.*, 2013). Otros efectores se relacionan con la necrotrofia, por ejemplo, dos candidatos homólogos de las proteínas inductoras de necrosis (NPP) (Li *et al.*, 2020; Zhao *et al.*, 2020; Marton *et al.*, 2018) y 11 proteínas que contienen el dominio CFEM “*Common in fungal extracellular membrane*”, tres candidatos en la base canónica y ocho en la base de no canónicos. Este dominio contiene ocho residuos de Cys espaciadas de forma característica y es un dominio único de los hongos. Desempeña un papel importante en la infección y el desarrollo fúngico, como receptores de la superficie celular, transductores de señales o como moléculas de adhesión (Li *et al.*, 2020; Kulkarni *et al.*, 2003). En *P. fijiensis*, siete candidatos efectores CFEM

muestran una alta expresión en hojas de banano infectadas (Noar y Daub, 2016), y se predicen que se dirigen al cloroplasto, las mitocondrias, el núcleo y el apoplasto. Estas distribuciones coinciden con lo reportado para este dominio pues los efectores CFEM tienen localización celular diversa, dirigiéndose a diferentes membranas del huésped, incluida la membrana celular y el núcleo (Gong *et al.*, 2020).

Otro dominio que fue encontrado es el dominio DUF, las proteínas con dominios DUF son comunes en las interacciones planta-patógeno fúngico (Sarris *et al.*, 2016). Uno de los miembros DUF de *P. fijiensis* (ID 144082, DUF3328) está en la lista de candidatos con mayor expresión en planta (\log_2FC 7,64) (Noar y Daub, 2016). Algunos estudios han asociado cambios en la expresión del gen DUF con el ciclo sexual de los hongos (Grognet *et al.*, 2014). Dado que la reproducción sexual y la producción de ascosporas ocurren en la etapa necrotrófica de la enfermedad de Sigatoka negra, DUF3328 puede ser importante para el desarrollo de pseudotecios en *P. fijiensis*. Otros 11 miembros muestran alta expresión en planta (Noar y Daub, 2016), evidenciando el papel clave de las proteínas DUF en la patogénesis de *P. fijiensis*.

Como fue descrito, muchos candidatos no presentaron predicción subcelular (197), pero otros predicen objetivos diferentes en la célula hospedera, incluyendo localización en mitocondria (7), cloroplasto (15) y núcleo (13). La identificación de péptidos de translocación de hongos aún está en sus inicios, pero se hace evidente que algunos efectores de hongos se translocan a la célula hospedera y se dirigen a diferentes compartimentos subcelulares (Gong *et al.*, 2020; Uhse y Djamei, 2018; Petre *et al.*, 2015).

Durante mucho tiempo se creyó que la mayoría de los efectores fúngicos no muestran similitudes significativas con secuencias conocidas en otro organismo (Hacquard *et al.*, 2011). Curiosamente, una gran parte (75,41%) de los 240 efectores canónicos de *P. fijiensis* tienen homólogos en diferentes hongos. La idea de que los efectores son exclusivos de un organismo y que por lo tanto no tiene homólogos está cambiando. Sędziewska-Toro y Brachmann, (2016) informaron que el 95% de los candidatos del hongo micorrízico arbuscular *Rhizophagus clarus* tienen homólogos en otros hongos en el filo de Glomeromycota. Asimismo, Guyon *et al.* (2014) informaron que en *Sclerotinia sclerotiorum* más del 50% de los efectores muestran conservación en los ascomicetos.

En la actualidad, los estudios ómicos están cambiando el concepto de efectores y están identificando potenciales efectores “core” o comunes entre los hongos. Los efectores comunes

pueden desempeñar funciones generales fundamentales durante la interacción del patógeno con sus diversas plantas hospedadoras. En este capítulo se identificó por primera vez el pan-eftectoroma de *P. fijiensis*, el cual resultó grande y sofisticado. Hay familias efectoras involucradas en la adhesión celular, la señalización, la interferencia con la percepción y la defensa del huésped. Esta estrategia se puede utilizar para predecir efectectoromas en otros hongos y ayudarnos a ampliar nuestro conocimiento sobre los patógenos.

CAPÍTULO IV

SELECCIÓN Y CLONACIÓN DE LOS CANDIDATOS DE EFECTORES DE *P. FIJIENSIS*

4.1 INTRODUCCIÓN

El estudio de efectores continúa siendo un reto y con un lento avance (Arroyo-Velez *et al.*, 2020; Dalio *et al.*, 2018; Jones *et al.*, 2018). La identificación *in silico* ha estado en continua mejora, evidenciando que los efectomas de los patógenos son más grandes de lo supuesto anteriormente (Carreón-Anguiano *et al.*, 2020; Marton *et al.*, 2018; Sonah *et al.*, 2016). Una vez identificados los posibles efectores, usualmente se seleccionan algunos miembros para confirmar su papel como efector y entender su función (Zhang *et al.*, 2020; Marton *et al.*, 2018; Liu *et al.*, 2014). Los análisis pueden incluir la generación de mutantes por pérdida de función, como el uso de mutantes knockout (Marton *et al.*, 2018; Assis *et al.*, 2017; Gawehns *et al.*, 2015), o la generación de genes quiméricos, ligando el gen de interés con una región codificante de etiquetas como la proteína verde fluorescente (Zhang *et al.*, 2020; Liu *et al.*, 2018), expresando la proteína para evaluarla *in vitro* (See *et al.*, 2019; Robin *et al.*, 2018; Kohler *et al.*, 2016), o agroinfiltrando en la planta hospedera o en una planta modelo (Zhang *et al.*, 2020; Kettles *et al.*, 2017; Petre *et al.*, 2016).

En el caso de *P. fijiensis* la transformación es altamente recalcitrante. Existen pocos reportes de transformación en este hongo (Escobar-Tovar *et al.*, 2015b; Portal *et al.*, 2012; Donzelli y Churchill, 2003). Las metodologías como polietilenglicol (PEG) y transformación con *Agrobacterium tumefaciens*, que son las más comunes en la literatura para otros hongos, son poco eficientes en *P. fijiensis*. El uso de ondas de choque ha sido más efectivo (Escobar-Tovar *et al.*, 2015b), pero requiere de un equipo sofisticado; debido a las limitantes técnicas, los análisis knockout o silenciamiento de genes no son análisis rutinarios en este hongo.

Se realizan escrutinios efectorómicos para adquirir conocimientos sobre los candidatos como primer paso, para luego priorizar algunos de ellos para su análisis y formular hipótesis sobre su función. Existen varios ensayos eficaces en cuanto al tiempo y rentabilidad (Petre *et al.*, 2017; Ma *et al.*, 2012; Win *et al.*, 2011; Alfano *et al.*, 2009). Estos ensayos a menudo se basan en la expresión heteróloga de efectores marcados (sin sus señales de secreción) directamente dentro de las células de las hojas de especies de plantas modelo, como *Arabidopsis thaliana* o *Nicotiana benthamiana* (Vyacheslavova *et al.*, 2012; Yesilirmak y Sayers, 2009). Para los efectores de oomicetos, el análisis de su comportamiento en sistemas vegetales heterólogos ha sido una

herramienta muy útil y ha permitido obtener información relevante sobre su localización específica en las células del hospedero e identificar moléculas que interactúan en el hospedero con el efector (King *et al.*, 2014; Petre *et al.*, 2014; Stam *et al.*, 2013). En su conjunto, estos estudios han mostrado cómo los efectores atacan estructuras o bloquean o manipulan funciones del hospedero, y han permitido avanzar significativamente en la comprensión de las interacciones moleculares entre patógenos y plantas.

Dado que la predicción de efectores de *P. fijiensis* resultó en una lista amplia de candidatos (Capítulo 3), se tuvo interés en la implementación de un ensayo que permitiera el análisis de un gran número de candidatos de efectores, para ello el primer filtro fue la selección y amplificación de éstos. Como parte de la estrategia para priorizar los candidatos, se comparó la base generada en el presente trabajo contra aquellas bases de datos reportadas por Chang *et al.*, (2016), Noar y Daub, (2016) y Escobar-Tovar *et al.*, (2015a).

En el trabajo reportado por Chang *et al.*, (2016) se realizaron la comparación genómica del complejo Sigatoka e identificaron 105 candidatos de efectores para *P. fijiensis* usando diversas herramientas bioinformáticas como SignalP 4.1, TMHMM 2.0, TargetP 1.1, big-PI, WoLFPSORT, ProtComp, permitiendo en su identificación la presencia de un dominio transmembranal, y definiendo la longitud de proteína en ≤ 250 aminoácidos y un contenido de $\geq 2\%$ de residuos de Cys. Por otro lado, Noar y Daub, (2016), reportaron el análisis transcriptómico de la asociación de *P. fijiensis* y *Musa* spp, donde identificaron un total de 231 efectores al analizar el transcriptoma de su interacción. En este estudio también se generaron una amplia lista con la información de expresión de los genes identificados (“log2Fold change”). Los criterios que emplearon para la identificación de los efectores fueron la predicción del péptido señal, una longitud de ≤ 300 aminoácidos y presencia de ≥ 4 residuos de cisteínas. Escobar-Tovar y colaboradores (2015a) realizaron el análisis proteómico de los secretomas obtenidos de *P. fijiensis* cultivado *in vitro* y en interacción con *Musa* spp, comparando, en ambos casos, dos cepas que difieren en patogenicidad. Estos autores reportaron una lista con 90 proteínas identificadas, las cuales codifican 42 proteínas únicas y 35 “spots” diferenciales entre ellas.

Existen numerosos reportes que se han enfocado en la aplicación de dos técnicas primordiales para el análisis funcional de efectores, una de ellas es la expresión transitoria mediada por *Agrobacterium tumefaciens*, conocida comúnmente como agroinfiltración, la cual se ha empleado para el análisis masivo de candidatos en *Puccinia triticina* (Zhang *et al.*, 2020), *Colletotrichum*

higginsianum (Robin *et al.*, 2018), *Zymoseptoria tritici* (Kettles *et al.*, 2017), *Phytophthora infestans* (Yin *et al.*, 2017), entre otros.

En la transformación mediada por *Agrobacterium* la bacteria transfiere el segmento de ADN (llamado T-ADN) a las células hospederas. El T-ADN contiene los genes requeridos para la transferencia y la región codificante del efector en estudio. Es una técnica rápida y simple que es ampliamente utilizada para la expresión transitoria, siendo exitosa en plantas modelos como tomate, lechuga, *A. thaliana*, *Nicotiana* spp, entre otros, siendo los dos últimos los más usados (Ma *et al.*, 2012; Goodin *et al.*, 2008; Wroblewski *et al.*, 2005).

Se ha explotado la agroinfiltración en muchos tipos de experimentos, por ejemplo para estudiar la función de los genes (Kettles *et al.*, 2017; Zhu *et al.*, 2017; Mesarich *et al.*, 2016), la producción de proteínas (Liu *et al.*, 2014), la interacción patógeno-hospedero (Deb *et al.*, 2018; Ramachandran *et al.*, 2016; Stam *et al.*, 2013; Fabro *et al.*, 2011), interacciones proteína-proteína (Kohler *et al.*, 2016; King *et al.*, 2014;) y la localización subcelular de proteínas (Robin *et al.*, 2018, Stam *et al.*, 2013).

La otra estrategia empleada es la expresión heteróloga de las proteínas efectoras en bacterias o levaduras y su posterior infiltración, tal es el caso del reporte en *Phytophthora cactorum* (Huang *et al.*, 2020), *Pyrenophora tritici-repentis* (See *et al.*, 2019) y *Puccinia striiformis* f. sp *tritici* (Petre *et al.*, 2016). En el caso de *P. fijiensis* en la cepa GS115 de *Pichia pastoris* se expresó la proteína madura del efector PfAvr4 de *P. fijiensis*, el extracto de la proteína secretada por la levadura fue infiltrado en hojas de banano de la variedad Calcutta4, resistente a la Sigatoka negra, observándose un disparo en la reacción de hipersensibilidad (Arango *et al.*, 2016).

Independientemente de la estrategia que se elija, ambas requieren de la clonación de la secuencia codificante del gen de interés, ya sea analizando la región codificante completa, la región sin el péptido señal o secuencias con mutaciones puntuales, para la evaluación del reconocimiento en el hospedero.

En este trabajo se seleccionaron y clonaron 12 de los posibles efectores de *P. fijiensis* predichos anteriormente para verificar que efectivamente son proteínas efectoras.

4.2 MATERIALES Y MÉTODOS

4.2.1 Selección de genes candidatos de efectores de *P. fijiensis*

De la base total del efectoroma predicho de *P. fijiensis* de 662 candidatos (240 efectores canónicos y 422 efectores no canónicos), se seleccionaron 12 candidatos de efectores canónicos para los primeros análisis funcionales. Cabe aclarar que, hasta el último semestre, y con la experiencia bioinformática que se ha adquirido, se identificaron los candidatos de efectores no canónicos. Una gran cantidad de trabajos de identificación *in silico* de efectores continúan centrándose en los efectores canónicos, pues realiza el escrutinio de efectores basándose en la selección de proteínas de tamaño pequeño que son ricas en cisteínas (Hao *et al.*, 2020; Queiroz y Santana, 2020; Neu y Debener, 2019; Chang *et al.*, 2016; Noar y Daub, 2016; Sonah *et al.*, 2016; Wicker *et al.*, 2013; Bowen *et al.*, 2009). En el presente trabajo se optó por ir más allá, pero la identificación de los efectores no canónicos se realizó recientemente, como se mencionó al inicio de este capítulo. Por esa razón, en el presente capítulo, para los análisis funcionales se presenta la selección de candidatos del grupo canónico únicamente, es decir, todos cumplen con las características de tamaño de proteína menor a 400 aminoácidos, presencia de péptido señal, poseen ≥ 4 residuos de Cys y carecen de dominio transmembranal.

Para la selección de los candidatos, se usaron diferentes criterios alternativos para seleccionar unos y otros, es decir, no fueron criterios homogéneos para todos los candidatos, ya que en la actualidad numerosos reportes manejan diferentes criterios de identificación de efectores (Sonah *et al.*, 2016; de Wit, 2016; Sperschneider *et al.*, 2015; de Wit *et al.*, 2009; Jones *et al.*, 2018). Hoy en día han surgido otros enfoques para la predicción de efectores que no se basan únicamente en los criterios canónicos.

Por lo tanto, ya que ningún criterio es universal para la identificación de efectores (Kanja *et al.*, 2020; Jones *et al.*, 2018; Sonah *et al.*, 2016; Sperschneider *et al.*, 2015), en este trabajo se usaron diferentes criterios alternativos para la selección de los candidatos que se continuarían estudiando, con el fin de ampliar las oportunidades de seleccionar algún(os) efectores verdaderos. Uno de los filtros de priorización fue el empleo del programa EffectorP 2.0, herramienta usada para identificación de efectores con base en aprendizaje de máquina. Otro filtro de selección fue la homología con efectores verdaderos conocidos en otros hongos, evaluando si pertenecían a alguno de los grupos establecidos como “Core fúngico” (homología con diferentes especies de hongos), “Core Sigatoka” (homología solo con *P. fijiensis*, *P. eumusae*

y *P. musae*) o “específico” (candidatos presentes solo en *P. fijiensis* sin homología con otras especies); estos grupos se establecieron siguiendo la nomenclatura propuesta por Chang *et al.* (2016). Se realizaron los análisis con BLASTP con la base de datos no redundante del NCBI, y también se incluyó la información disponible en los portales públicos de los genomas de *P. fijiensis*, *P. eumusae* y *P. musae*. Otra consideración fue usar la información de la literatura disponible sobre la expresión de genes de *P. fijiensis* durante su interacción con banano. Estos son los trabajos de Escobar-Tovar *et al.*, (2015a) que realizaron análisis proteómicos de los secretomas de *P. fijiensis in vitro* y de muestras de hojas de banano infectados con *P. fijiensis*. Otro reporte es el de Noar y Daub, (2016) que generaron el transcriptoma de la interacción de *P. fijiensis* y *Musa acuminata*, y generaron una amplia lista con la información de expresión (“log2Fold change”) de todos los genes de *P. fijiensis* que identificaron en la interacción con el hospedero. Como criterio adicional, se tomó en cuenta a aquellos candidatos de efectores que contengan algún dominio funcional conocido, información obtenida con el programa InterProScan. Cabe mencionar que los criterios no se aplicaron necesariamente de manera conjunta, sino paralela. La estrategia implementada se resume en la Figura 4.1.

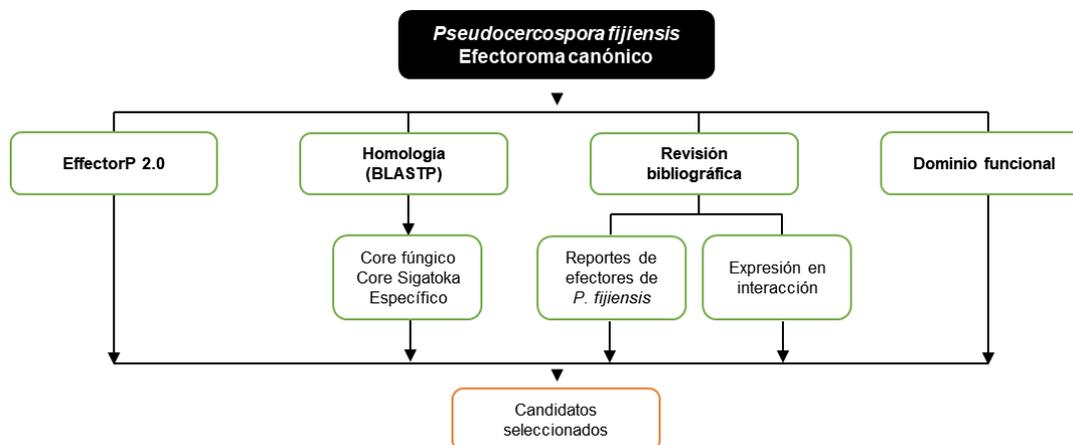


Figura 4.1 Criterios seguidos para priorizar la selección de candidatos de efectores de *P. fijiensis* para continuar con los análisis funcionales.

En consecuencia, dado que la mayoría de los efectores no cumplen con todos los criterios y que difícilmente se unifican, se realizó una selección “por fracciones”, teniendo los siguientes cuatro conjuntos de candidatos seleccionados; conjunto 1: que coincidiera con los tres reportes de *P. fijiensis* mencionados arriba y sea predicho como efector por EffectorP 2.0; se seleccionó un candidato. Conjunto 2: que tuvieran coincidencia con al menos un reporte de *P. fijiensis* y sea predicho como efector por EffectorP 2.0; se seleccionaron 8 candidatos, 4 de ellos coinciden con

uno de los 3 reportes y los otros 4 coinciden en dos de los 3 reportes. Conjunto 3: no tuviera coincidencias con reportes de *P. fijiensis* pero sea predicho como efector por EffectorP 2.0; este grupo es de mayor riesgo por lo que se seleccionó un candidato y conjunto 4: que no tuviera coincidencia con reportes de *P. fijiensis* y no sea predicho como efector por EffectorP 2.0; se seleccionaron 2 candidatos.

4.2.2 Condiciones de crecimiento de microorganismos.

Para los análisis funcionales y moleculares, se utilizó la cepa C1233 de *P. fijiensis* (procedente de Uxmal, Yucatán, colectado por el Dr. Andrew James Kay). De cada colonia reactivada se cortó 1 cm² y se maceró en un mortero con pistilo, en 1 mL de medio PDB (4g/L almidón de papa, 20g/L dextrosa) (Difco™ BD Sparks, MD, USA). El micelio macerado se inoculó en matraces Erlenmeyer de 250 mL con 50 mL de PDB. El cultivo líquido se mantuvo en agitación durante 12 días a 100 rpm a temperatura ambiente. Este material fue utilizado para los inóculos en los experimentos posteriores.

Las células de *E. coli* Top10 fueron cultivadas en medio LB “Luria-Bertani” (Invitrogen, CA, USA) (10 g/L peptona, 5 g/L extracto levadura, 10 g/L de NaCl) adicionado con 100 µg/mL de ampicilina.

4.2.3 Obtención de material de interacción *Musa spp-P. fijiensis*

La expresión de los candidatos fue evaluada analizando los diferentes ADNc obtenidos de la interacción de *P. fijiensis* con el cultivar de banano *M. acuminata* cv. Calcutta 4. Se infectaron las hojas 1 y 2 de plantas con cuatro meses de edad; para ello se usó un volumen de 150 µL de una solución de 1x10⁶ conidios/mL. Las plantas infectadas se mantuvieron bajo condiciones de invernadero en el CICY. Para los análisis de expresión se evaluaron muestras de diferentes “tiempos de interacción” (días después de la infección o dpi por “*days post-inoculation*”), siendo los días 3dpi, 9 dpi, 10 dpi, 11 dpi, 12 dpi, 13 dpi, 14 dpi, 15 dpi, 16 dpi, 17 dpi, 18 dpi, 19 dpi, 20 dpi, 21 dpi, 22 dpi y 23 dpi). Para facilitar el análisis las muestras se mezclaron en pools; pool 1: 3dpi, 9dpi y 10dpi; pool 2: 11dpi, 12 dpi y 13 dpi; pool 3: 14dpi, 15dpi y 16dpi; pool 4: 17dpi, 18dpi y 19dpi; pool 5: 20dpi, 21dpi, 22dpi y 23 dpi (Figura 4.2). Se usó 1 µL de ADNc de cada pool como templado explorar la expresión de los efectores candidatos.

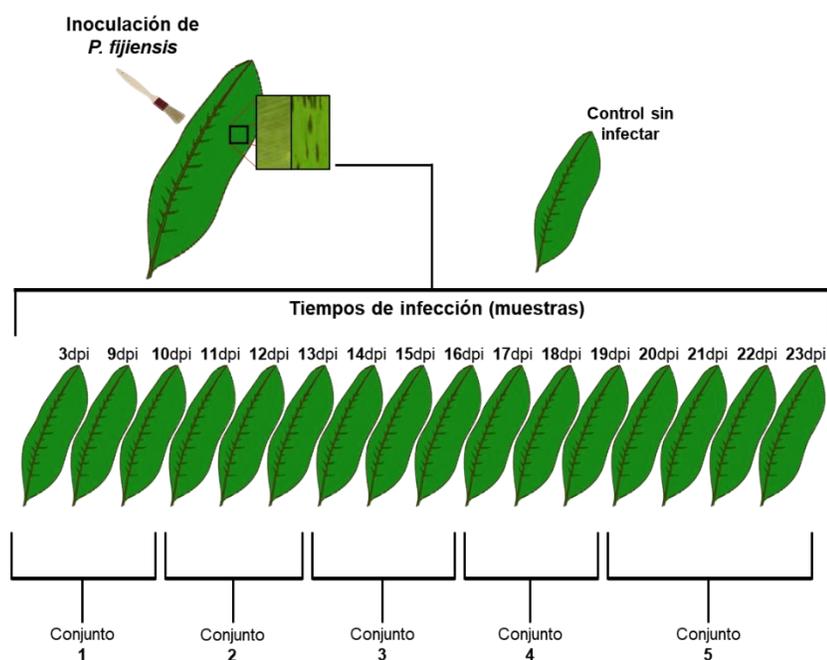


Figura 4.2 Diagrama de los tiempos de infección de *P. fijiensis* y *Musa spp.*

4.2.4 Extracción de ARN y síntesis de ADNc

Las muestras de tejido infectado y tejido control almacenadas a -80°C se pulverizaron con nitrógeno líquido usando mortero y pistilo estériles. El ARN total se aisló utilizando el reactivo PureLink™ Plant RNA Reagent (ambion®). Una vez obtenido el ARN, éste se trató con ADNasa (Invitrogen™) siguiendo las indicaciones del fabricante. El ADNc se sintetizó con la transcriptasa reversa SuperScript™III (Invitrogen™) siguiendo las recomendaciones del fabricante. El ADNc resultante se utilizó para explorar la expresión génica mediante PCR punto final.

4.2.5 Amplificación y clonación de los candidatos a efectores de *P. fijiensis*

La amplificación y clonación de los genes candidatos se llevó a cabo a partir de ADNc de la interacción de *Musa spp.*- *P. fijiensis* y en los casos que no se detectó expresión en los ADNc, se empleó ADNc de *P. fijiensis* crecido *in vitro*. Los cebadores específicos de cada candidato se diseñaron con el programa OligoAnalyzer Tool del IDT (<https://www.idtdna.com/calc/analyzer>), de tal forma que permitieran la amplificación del CDS (“Coding Sequence”). Para la amplificación se

utilizó la polimerasa de alta fidelidad, Herculase II Fusion DNA Polymerase (Agilent Technologies ©, CA, US).

Para la purificación de los productos amplificados, los amplicones fueron visualizados sobre un transiluminador de UV (Benchtop UV Daigger® Scientific) y las bandas fueron cortadas del gel de agarosa con ayuda de un bisturí. Posteriormente las bandas cortadas se colocaron en un tubo Eppendorf de 1.5 mL y se siguió el protocolo del kit de extracción *Qiaquick Gel Extraction Kit* (Qiagen). Para la clonación de los fragmentos, se utilizaron los vectores pGEM®-T Easy (Promega), para el cual en un tubo Eppendorf estéril se agregaron 5 µL del amortiguador de ligación 2X y 1 µL de ADN Ligasa T4, 1 µL de vector pGEM-Teasy (50ng), 2 µL de ADN Ligasa y agua libre de nucleasas. El volumen final de la mezcla fue de 10 µL y se incubaron por 1 hora a temperatura ambiente, siguiendo las instrucciones del fabricante (Promega).

4.2.6 Transformación de *Escherichia coli* TOP10 y confirmación de los genes de interés en el vector pGEM-Teasy.

Para las transformaciones se usaron bacterias quimio-competentes *Escherichia coli* cepa TOP10. La introducción del plásmido en *E. coli* se realizó mediante choque térmico; 2 µL de la reacción de ligación se adicionaron a las células competentes en un tubo Eppendorf de 1.5 mL. La mezcla se incubó por 30 minutos en hielo y enseguida se colocaron a 42 °C por 30 segundos, y se adicionaron a 1 mL de medio LB líquido (10g/L peptona, 5g/L extracto levadura, 10g/L de NaCl) a la alícuota de las células. Las células transformadas se sembraron en medio LB sólido (10g/L peptona, 5g/L extracto levadura, 10g/L de NaCl, Agar 20g/L), adicionando ampicilina (Amp) 100µg/µL como agente de selección, X-gal (80 µg/mL) como sustrato cromogénico e IPTG 0.5 mM (Isopropil-β-D-1-tiogalactopiranosido) como inductor de la expresión del gen *lacZ*. Se seleccionaron las clonas blancas y estas se recuperaron en cajas de Petri con medio LB suplementado con agar y ampicilina. Se tomaron muestras de las clonas para confirmar la presencia de los fragmentos de interés mediante PCR en colonia con la polimerasa DreamTaq PCR Master Mix (2x) (Thermo Scientific™).

4.2.7 Purificación y secuenciación de los genes clonados en pGEM-Teasy.

Las colonias recombinantes de *E. coli* TOP10 se inocularon en medio LB líquido (10g/L peptona, 5g/L extracto levadura, 10g/L de NaCl), adicionado con 100 µg/mL de Amp. La purificación de los plásmidos se realizó mediante el kit de purificación “QIAprep Spin Miniprep Kit (50)” (QIAGEN, QIAprep® Spin Kit), siguiendo las indicaciones del fabricante. Para la cuantificación del ADN

plasmídico, se utilizó el espectrofotómetro NanoDrop™ 2000c (Thermo Scientific™). Se ajustaron las concentraciones a 50 ng/μL en un volumen final de 50 μL.

Para corroborar la secuencia de los insertos clonados se solicitó el servicio de secuenciación al Laboratorio Nacional de Biotecnología Agrícola, Médica y Ambiental, del Instituto Potosino de Investigación Científica y Tecnológica (LAMBAMA, IPICT). Las secuencias fueron editadas y analizadas mediante el software “GenomeNet,” disponible en la página <http://www.genome.jp>, usando la herramienta ClustalW; para la obtención de las cadenas complementarias se usó el software “Reverse Complement-Bioinformatics.Org”, disponible en la página http://www.bioinformatics.org/sms/rev_comp.html y la página; <http://web.expasy.org/translate/>, para el software ExPASy: SIB Bioinformatics Resource Portal”, usando la herramienta *Translate tool* que permite obtener el marco de lectura y la traducción del mismo. Para la predicción de estructura tridimensional se usó la herramienta SWISS-MODEL (<https://swissmodel.expasy.org/>). Para la búsqueda de homólogos y evaluación en la patogénesis en especies caracterizadas, se usó la base de datos depositada en el PHI-Base (<http://www.phi-base.org/>) (Urban *et al.*, 2017; Winnenburg *et al.*, 2006) empleando la herramienta BLASTP.

4.3 RESULTADOS

4.3.1 Selección de efectores candidatos en *P. fijiensis*

Como se mencionó antes, de la base total del efectoroma predicho de *P. fijiensis* de 240 efectores canónicos, se seleccionaron para los primeros análisis funcionales una pequeña población de la base de datos de efectores canónicos.

Para priorizar los candidatos, se comparó la lista de secuencias de la base generada de efectores de *P. fijiensis* predichas por EffHunter en el presente trabajo, contra las bases de datos reportadas por Chang *et al.*, (2016), Noar y Daub, (2016) y Escobar-Tovar *et al.*, (2015). En el trabajo de Escobar-Tovar *et al.*, (2015). Para manejar la información se dividieron en diferentes bases, Base A: Escobar-Tovar *et al.*, (2015); Base B: Chang *et al.*, (2016); Base C: Noar y Daub, (2016).

Debido a que la mayoría de los efectores no cumplieron con los criterios convencionales y que a su vez difícilmente se unifican, se optó por una selección basados en criterios alternativos. Se seleccionó un candidato que coincidió con los tres reportes de *P. fijiensis* y es predicho como efector por EffectorP 2.0 (candidato con mayor probabilidad de ser efector); cuatro candidatos

coincidieron con dos de los reportes de *P. fijiensis* y son predichos como efectores por EffectorP 2.0; otros cuatro candidatos predichos como efectores por EffectorP 2.0, tuvieron coincidencia con alguno de los reportes de *P. fijiensis*. Se seleccionó un candidato predicho por EffectorP 2.0 que no tuvo coincidencias con ninguno de los reportes de *P. fijiensis*, y dos candidatos que no fueron identificados en ninguno de los reportes de *P. fijiensis* y tampoco son predichos como efectores por EffectorP 2.0. Este último grupo es muy interesante, aunque arriesgado, ya que la predicción *in silico* no apoya fuertemente que sean efectores, sin embargo, los análisis funcionales permitirán evaluar si efectivamente estos candidatos identificados por EffHunter, pero no por EffectorP 2.0, son efectores. De observarse respuesta en el hospedero, podrían considerarse como falsos negativos de la predicción por EffectorP 2.0 y confirmar la identificación de efectores por EffHunter. Por lo tanto, para continuar con los análisis funcionales, se seleccionaron un total de 12 candidatos de la base de efectores canónicos.

Con base al criterio de selección por coincidencia con los reportes de *P. fijiensis*, 9 candidatos estuvieron presentes en al menos una de las bases de datos referidas. La proteína con el ID210733 tuvo coincidencia con la Base A, ID77734 tuvo coincidencias con la Base A y B, los candidatos ID181415, ID212258 tuvieron coincidencias con las bases A y C, el candidato ID85220 con las tres bases A, B y C, los candidatos ID191819, ID39450, ID212087 tuvieron con la Base B y C y el candidato ID60298 con la base C. Con el criterio de selección de no estar previamente reportado, tres candidatos (ID70410, ID80609, ID189452) no fueron identificados previamente, por lo que son nuevos.

Los análisis con BlastP mostraron que nueve candidatos tienen homólogos en otros hongos, por ejemplo especies del complejo Sigatoka, en los géneros *Cercospora*, *Fusarium*, *Aspergillus*, entre otros. De los otros tres candidatos, los análisis mostraron que uno tiene homólogos en las especies del complejo Sigatoka (ID191819), otro tiene homología con especies cercanas como *Zasmidium cellare*, *Septoria musiva*, *Cercospora berteroae*, *C. beticola*, *Dothistroma septosporum*, *Ramularia collo-cygni* más no se comparte con el complejo Sigatoka (ID80609), y otro no tiene homología con ningún otro microorganismo (ID191819), siendo específico de *P. fijiensis*. Cuatro de los candidatos seleccionados tienen predichos un dominio funcional (ID77734, ID70410, ID80609, ID189452), de acuerdo al análisis *in silico* por InterProScan.

En el Cuadro 4.1 se muestran los candidatos de efectores seleccionados con la descripción detallada de sus características.

Cuadro 4.1 Candidatos a efectores de *P. fijiensis* seleccionados para los análisis funcionales.

Candidato ID	Longitud		Caracterización in silico				Búsqueda de homólogos		
	Proteína	Transcrito	Residuos de Cys	Dominio InterProScan	EffectorP 2.0	Homología en PHI: Database	*Homología en otras bases efectorómicas de <i>P. fijiensis</i>	Homología en la base nr NCBI	Descripción
210733	195	588	4	-	Efector	BEC1040	Base A	Varias especies	Proteína hipotética, proteína extracelular 32-2, proteína no caracterizada, producto proteico sin nombre
191819	114	345	6	-	Efector	-	Base B y C	-	-
77734	131	396	7	PAN	Efector	-	Base A y B	Varias especies	Proteína hipotética, relacionada con el precursor de la alfa-amilasa A, proteína que contiene el dominio WSC, metabolito fungistático putativo, lectina / glucanasa similar a la concanavalina A
181415	193	582	4	-	Efector	-	Base A y C	Varias especies	Proteína hipotética, proteína no caracterizada, producto proteico sin nombre
212258	254	765	24	-	Efector	-	Base A y C	Varias especies	Proteína hipotética, proteína no caracterizada, producto proteico sin nombre
39450	181	546	8	-	Efector	-	Base B y C	Varias especies	Proteína hipotética, proteína no caracterizada, supuesta proteína integral de membrana, transportador no caracterizado
60298	260	783	8	-	Efector	-	Base C	Varias especies	Proteína hipotética, producto proteico no identificado, proteína no caracterizada
85220	188	567	8	-	Efector	-	Base A, B y C	Varias especies	Proteína hipotética, proteína no caracterizada
212087	162	489	5	-	Efector	-	Base B y C	Complejo Sigatoka	Proteína hipotética
70410	243	732	5	Isocorismatasa	-	-	-	Varias especies	Proteína hipotética, producto proteico no identificado, proteína no caracterizada, isocorismatasa hidrolasa, probable ycaC, hidrolasa de especificidad desconocida
80609	198	597	4	Carboxylesterasa	Efector	LIP1, FGSG_03243	-	Varias especies	Proteína no caracterizada, proteína hipotética, alfa / beta-hidrolasa, relacionada con para-nitrobenzil esterasa, lipasa putativa, acetilcolinesterasa
189452	356	1071	7	Deuterolisina metaloproteasa (M35)	-	AVR-Pita1	-	Varias especies	Proteína hipotética, proteína no caracterizada, proteasa neutra 2

*Base A: Escobar-Tovar *et al.*, 2015; Base B: Chang *et al.*, 2016; Base C: Noar y Daub, 2016.

4.3.2 Amplificación y clonación de los candidatos a efectores de *P. fijiensis*

Se evaluó la expresión de los candidatos en diferentes tiempos de interacción de *P. fijiensis* con el cultivar de banano *M. acuminata* cv. Calcutta4, mencionados anteriormente en el apartado de metodología. La Figura 4.3 corresponde a las plantas de banano *M. acuminata* cv. Calcutta 4 infectadas con *P. fijiensis*.



Figura 4.3 Plantas de banano *M. acuminata* cv. Calcutta4 infectados con *P. fijiensis* C1233.

Se confirmó la síntesis correcta de ADNc obtenido de *P. fijiensis* cepa C1233 cultivado *in vitro*, usando los cebadores diseñados para amplificar el gen de la actina de *P. fijiensis*. En la Figura 4.4 se observa la amplificación del fragmento de actina a partir de ADNc (carril 4) y ADN genómico (carril 2). Se obtuvo un inserto aproximado de 400pb que es menor al amplicón genómico (~500pb), debido a que los cebadores flanquean un intrón de 50pb. Estos resultados indicaron que no existe contaminación de ADN genómico en la muestra de ADNc.

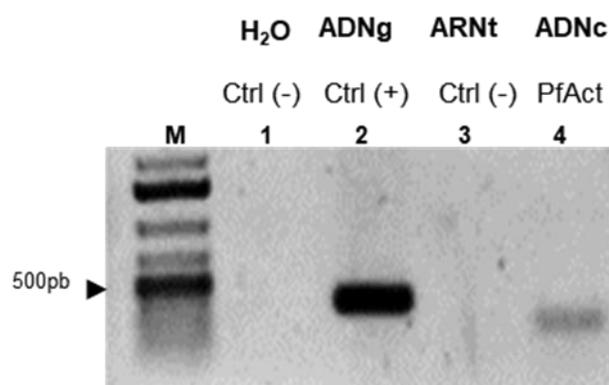


Figura 4.4 Fotografía gel de agarosa al 0.8% de control Actina de *P. fijiensis*. Controles negativos H₂O (carril 1) y ARN tratado con DNaseI (carril 3); control positivo (+) ADN genómico (carril 2) y evaluación de ADNc (carril 4).

Para la amplificación de los candidatos, se evaluaron diferentes templados, ADNc de *P. fijiensis* cultivado *in vitro*, ADNc de material de infección *Musa spp - P. fijiensis* y ADNg de *P. fijiensis*. Los cebadores específicos utilizados para la amplificación de cada región codificante se presentan en el Cuadro 4.2.

Cuadro 4.2 Cebadores utilizados para la amplificación de los candidatos a efectores de *P. fijiensis*.

Cebador	Secuencia 5' → 3'	Tm °C	Tamaño inserto
Fw_191819	ATGCAACTTGCTACAATTTTTGCCTCCCTCTCGCCTTAGTTGCCCTC	69.0	345
Rv_191819	TCAGTGGTGATGGTGATGGTGTGCTGGGGCCTTGG	70.2	
Fw_70410	ATGAAGAACTTCTACTCCATTCTTTGCTCTGTTGCC	62.4	732
Rv_70410	TTGCTCCGGCAACAGGACATCTTCACCAAG	65.4	
Fw_181415	ATGTACGCTTTCACTACCATCTTGCCACTGCCTCCCTCGCT	70.5	582
Rv_181415	TCAGTGGTGATGGTGATGGTGTGCTGGGCAGTTGGC	70.3	
Fw_39450	ATGCTCTCCATCACTGCCACCCTGGG	66.2	546
Rv_39450	TCAGTGGTGATGGTGATGGTGGTTAGCATTGGGGTA	66.7	
Fw_210733	ATGAAAGCCTTCACACTTGCCTCTTCGCG	65.7	588
Rv_210733	TCAGTGGTGATGGTGATGGTGTACGTACACATAAGCCGTA	66.3	
Fw_80609	ATGAGGTCTCTCGATCTTGGAACTTCTT	60.4	597
Rv_80609	CAGCCAGCAGGAACCTTTCGATATTGT	60.6	
Fw_60298	ATGAAGTCCTATAGCATCTTCTCTGTCTCGCTGCT	64.7	783
Rv_60298	TCAGTGGTGATGGTGATGGTGGACTATGTAGTCCAT	64.8	
Fw_77734	ATGTCCAAGATGCTTCTTCTCGCAGCCACCATCCTTGGTCTC	68.8	396
Rv_77734	TCAGTGGTGATGGTGATGGTGAGCAGCCTGACCACA	69.3	
Fw_85220	ATGTTCCGACTCACCTTCGAGCCCTCAGCCT	69.5	567
Rv_85220	TCAGTGGTGATGGTGATGGTGTGGCTCAGCCGCAAT	69.9	
Fw_189452	ATGAAGTTTGCAATCTTCTCCATTGCAGCC	61.4	1071

Rv_189452	ACTCAACTCACCATTTCGCATAAAGCGCATA	61.7	
Fw_212087	ATGAAGCCCACTACACCAACCATCCTCGCCC	67.6	489
Rv_212087	TCAGTGGTGATGGTGATGGTGCGCAACACAACCTCTG	67.6	
Fw_212258	ATGAGGCTCACTTTCATTGCTGCAGCTCTAGC	64.6	765
Rv_212258	TCAGTGGTGATGGTGATGGTGAGTGAGTGCAGATCTCATT	65.4	

Las reacciones de PCR se llevaron a cabo utilizando ADNc de la interacción *Musa spp-P. fijiensis*, siendo los candidatos ID85220 e ID189452 los únicos amplificados en esta condición. Por lo tanto, se optó por la amplificación de los candidatos a partir de la muestra de ADNc de *P. fijiensis* cultivado *in vitro*, pero ninguno tuvo amplificación en esta condición.

Dado a lo anterior, se siguió una estrategia diferente para la amplificación de los candidatos usando como templado ADNg de *P. fijiensis*. Aunque no es lo más recomendado, esta estrategia es válida, ya que los efectores ID210733, ID77734, ID212258 e ID80609 no predicen tener intrones. En la Figura 4.5 se muestran los candidatos amplificados a partir de ADNg (ID210733, ID77734, ID212258 e ID80609) y los amplificados a partir de ADNc (ID85220, ID212087 e ID189452).

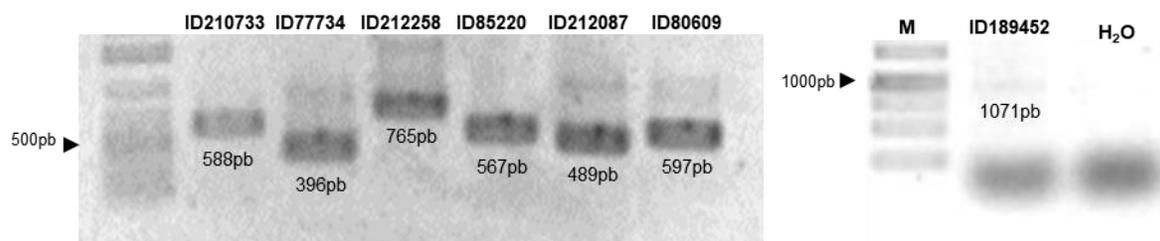


Figura 4.5 Amplificación por PCR de los efectores candidatos de *P. fijiensis*, usando como templado ADNc y ADNg, como se describió en metodología. ADNg (ID210733, ID77734, ID212258 e ID80609) y los amplificados a partir de ADNc (ID85220, ID212087 e ID189452).

Una vez amplificados los fragmentos, se procedió con la purificación de los productos, donde se cortaron las bandas del gel de agarosa y se siguió el protocolo correspondiente de QuiAquick Gel Extraction Kit (Qiagen). Para la clonación se usó el vector pGEM-Teasy® (Promega).

4.3.3 Confirmación *in silico* de los candidatos de efectores de *P. fijiensis*

De las colonias transformantes de los candidatos ID210733, ID77734, ID212258, ID85220, ID212087, ID80609 e ID189452, se seleccionaron dos clones de cada uno, las cuales se enviaron para secuenciar en el IPICyT (servicios de secuenciación antes mencionados) con el uso de los cebadores T7 y SP6, confirmando los CDS de cada efector. En la Figura 4.6 se muestran los marcos de lectura de las secuencias obtenidas, las cuales corresponden a las secuencias reportadas.

Cabe mencionar que las secuencias correspondientes a los candidatos de efectores ID191819, ID181415, ID39450, ID60298 e ID70410 no amplificaron en los diferentes templados usados, tampoco hubo amplificación usando ADNg como templado, por lo que no pudieron clonarse.

Se clonaron siete efectores, entre los cuales están los candidatos ID210733 e ID80609 que tienen cuatro residuos de Cys con una longitud de 195 y 198 aminoácidos respectivamente, los candidatos ID212087 e ID189452 con siete residuos de Cys y con longitud de 131 y 356 aminoácidos respectivamente, el candidato ID85220 con una longitud de 188 aminoácidos con ocho residuos de Cys, y el candidato ID212258 con una longitud de 254 aminoácidos y con 24 residuos de Cys. Los candidatos ID210733, ID80609, ID77734 e ID212087 están localizados en los scaffolds 3,4, 5 y 7 respectivamente. Los candidatos ID189452 e ID85220 se encuentran localizados en el scaffold 6, mientras que el candidato ID212258 está localizado en el scaffold 8. De los siete posibles efectores clonados, tres tienen dominio funcional predicho, ID80609 con dominio de carboxilesterasa (IPR002018/PF00135), ID77734 con dominio PAN (IPR003609/PF14295) e ID189452 con dominio de Deuterolisina metaloproteasa M35 (IPR001384/PF02102).

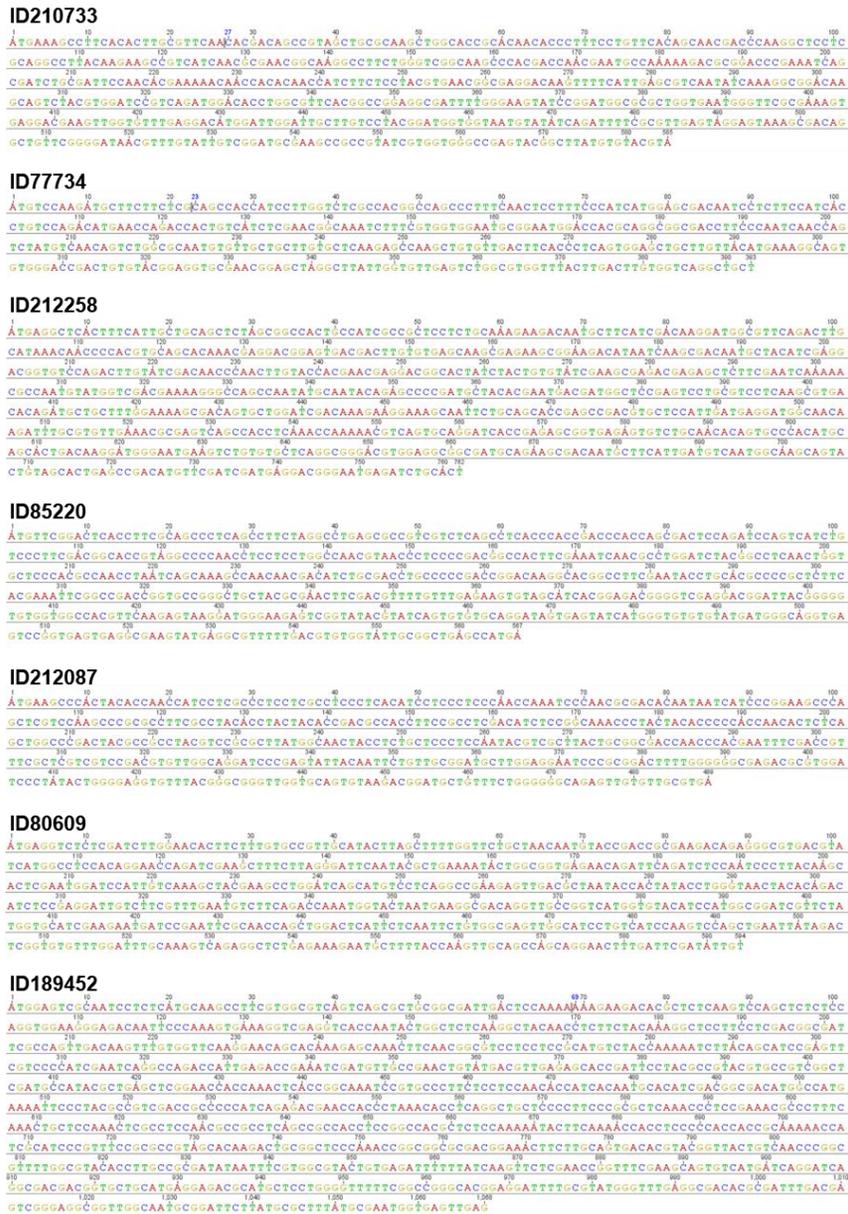


Figura 4.6 Marcos de lectura abierto de los efectores candidatos que fueron amplificados y clonados.

4.3.4 Modelamiento *in silico* del efector candidato PfID77734 de *P. fijiensis*

Entre los candidatos clonados, el efector PfID77734 con el dominio predicho de PAN/Apple domain (IPR003609/PF14295), tuvo homólogos en varios géneros de hongos (especies de *Pseudocercospora*, especie de *Hortea*, *Ramularia* y *Aspergillus*). La descripción de las proteínas con las que tuvo homología es de “proteínas hipotéticas” y “proteínas relacionadas a metabolitos

fungistáticos”. Debido a que el candidato PfID77734 correspondió a un efector con posible función conservada, se optó por realizar el *modelamiento in silico* del candidato, encontrando homología estructural con la proteína MICRONEMAL PROTEIN 4 de *Toxoplasma gondii* (Templado 4a5v.1.A) de la base de datos de SWISS-MODEL. Esta proteína también cuenta con dominio PAN/Apple, teniendo un porcentaje de identidad de 30.77%. En la figura 4.7 se observa, en color azul, el modelado construido para el candidato PfID77734 y en color marrón la proteína de *Toxoplasma gondii* *Microneme protein 4* (Templado 4a5v.1.A) (Figura 4.7), en las cuales es notorio el traslape entre algunas regiones del dominio PAN de ambas proteínas, lo que apoya que pertenece a la misma familia de proteínas.

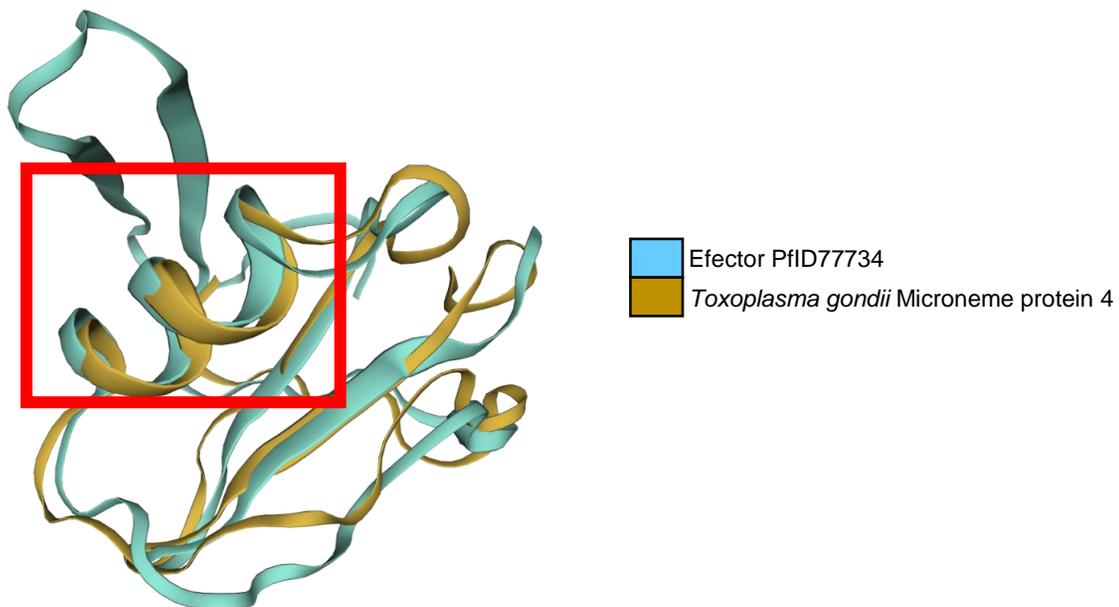


Figura 4.7 Modelamiento tridimensional del traslape del dominio PAN/Apple de las estructuras de PfID77734 de *P. fijiensis* (azul) y Microneme protein 4 de *Toxoplasma gondii* (marrón). El cuadro rojo señala las α hélices traslapadas.

El alineamiento entre PfID77734 de *P. fijiensis* y *Microneme protein 4* de *Toxoplasma gondii*, mostró correlación entre las α hélice de ambas proteínas.

Se ha reportado, que este dominio está en proteínas de varias especies de *Phytophthora*, *Toxoplasma*, *Fusarium* (Li *et al.*, 2020; McGowan *et al.*, 2020; Gong *et al.*, 2012) y su función es

la unión con otras proteínas o con carbohidratos. En *T. gondii*, parásito que causa toxoplasmosis en humano y en animales, se ha reportado que pueden mediar interacciones interespecíficas, proporcionando un vínculo entre el hospedero y el parásito (Gong *et al.*, 2012).

Los candidatos (PfIF85220, Pf212087 y PfID210733) tienen expresión diferencial en la interacción de *P. fijiensis* con *Musa* spp, teniendo valores de Log2FC de 4.35, 2.95 y 0.74 respectivamente en lo reportado por Noar y Daub, (2016).

De los anteriores, según los resultados de BLASTP con la base del PHI (*Pathogen-Host Interaction*), el candidato PfID210733 tuvo hit con el efector BEC1040 de *Blumeria graminis* (PHI:2901), con un porcentaje de identidad del 27.46%, el candidato PfID189452 tuvo hit con una proteína de *Magnaporthe oryzae*, el efector Avr-Pita1 (PHI:2850), con 29.63% de identidad, mientras que el candidato PfID80609 tuvo hit con una proteína de *Botrytis cinerea*, LIP1 (PHI:541), con un 21.52% de identidad y con la proteína FGSG_03243 (PHI:6121) de *Fusarium graminearum* con un 28.42% de identidad. En este último caso se le ha reportado como un gen que se expresa significativamente en la interacción del patógeno con maíz y la eliminación de ese gen provocó 40% de disminución de la patogenicidad en la mutante de *F. graminearum* (Zhang *et al.*, 2016).

4.4 DISCUSIÓN

El repertorio de efectores de *Pseudocercospora fijiensis* es poco conocido. La identificación de efectores y su caracterización puede proveer de un mejor entendimiento del mecanismo de acción de *P. fijiensis* y pueden dirigir a la generación de nuevos métodos de control.

En este trabajo, se exploró la expresión de los efectores de *P. fijiensis* en su interacción con diferentes cultivares de banano. Resultados obtenidos de la estudiante de Licenciatura la Ing. Celina Ayil realizó la evaluación de los diferentes tiempos de infección con Enano Gigante, reportando que no hubo amplificación de los candidatos en las diferentes muestras. Sin embargo, sólo dos candidatos, el PfID85220 y PfID189452, se detectaron en el pool 1 de la interacción de *Musa* spp – *P. fijiensis* (Calcutta 4).

De acuerdo con los resultados que se han obtenido, se logró clonar en el vector pGEM®-Teasy los genes que codifican para los efectores candidatos PfID210733, PfID77734, PfID212258, PfID85220, PfID212087, PfID80609 y PfID189452.

Sin embargo, cinco de los candidatos seleccionados no lograron clonarse (PflD191819, PflD181415, PflD39450, PflD60298 y PflD70410), esto pudiera deberse a que estos candidatos se encuentran en regiones condensadas del ADN, siendo regiones no accesibles para la amplificación de estos; como se reportó en el hongo hemibiotrófico *Verticillium dahliae* en el que mediante análisis con aprendizaje de máquina se identificó metilación incompleta del ADN de elementos repetitivos, asociados con compartimentos genómicos específicos originalmente definidos como regiones específicas de linaje (LS) que contienen genes involucrados en la adaptación del huésped (Cook *et al.*, 2020). Basado en el enriquecimiento de genes expresados en la interacción con el hospedero se observó que los presuntos efectores identificados en *V. dahliae* tienen un alto índice de PAV (variación de la Presencia/Ausencia) entre cepas de la misma especie (Cook *et al.*, 2020; de Jonge *et al.*, 2013).

Otra posibilidad sería que existan diferencias entre la cepa de referencia y la cepa usada en este trabajo en la región del ADN donde se diseñaron los oligos para la amplificación de los efectores, imposibilitando la amplificación de estos. Se sabe que estos genes a menudo se pierden o mutan rápidamente (Plissonneau *et al.*, 2017).

Por otro lado, también se tiene la posibilidad de que estos candidatos no hayan podido ser amplificados por los rearrreglos cromosómicos que se llevan a cabo en el genoma. Por ejemplo, el efector Avr-Pita de *Magnaporthe oryzae* es muy variable en su ubicación cromosómica y se asocia con polimorfismos frecuentes de presencia-ausencia dentro de los diferentes linajes del patógeno, siendo localizado cerca de los telómeros y, a veces, en minicromosomas (Peng *et al.*, 2019; Chuma *et al.*, 2011). Similarmente el caso del efector Zt_8_609 de *Zymoseptoria tritici*, donde reportan que segrega un polimorfismo de presencia/ausencia dentro de la especie, congruente con ello, se ha podido anotar este efector en el genoma de referencia IPO323 pero no en los nuevos genomas ensamblados de *Z. tritici* (Plissonneau *et al.*, 2018; Hartmann *et al.*, 2017; Marshall *et al.*, 2011).

Tres de los efectores clonados tienen dominio funcional predicho; PflD80609 con dominio de carboxilesterasa (IPR002018/PF00135), PflD77734 dominio PAN (IPR003609/PF14295) e ID189452 con dominio de Deuterolisina metaloproteasa M35 (IPR001384/PF02102). En el caso del efector PflD77734 con el dominio predicho de PAN/Apple domain (IPR003609/PF14295), tiene homología de acuerdo al BLASTP con varias especies de los géneros de hongos *Pseudocercospora*, *Hortea*, *Ramularia* y *Aspergillus*. En el análisis tridimensional de la proteína arrojó homología estructural con la proteína MICRONEMAL PROTEIN 4 de *Toxoplasma gondii*,

parásito que afecta humanos y animales, porque es la proteína que se ha estudiado y está depositada en esa base de datos. Se ha reportado que pueden mediar interacciones interespecíficas, favoreciendo la interacción de *T. gondii* con su hospedero (Gong *et al.*, 2012). También se ha reportado este dominio en varias especies de *Phytophthora* y *Fusarium* (Li *et al.*, 2020; McGowan *et al.*, 2020; Gong *et al.*, 2012), donde reportan la unión a otras proteínas o carbohidratos. En el reporte de Li *et al.*, (2020), donde trabajaron con el secretoma de *F. oxysporum* f. sp. *lycopersici* usando LC-MS/MS, reportaron una lista de 32 proteínas como potenciales a efectores, donde 24 de ellas fueron anotadas funcionalmente y entre las cuales, una proteína fue predicha con el dominio funcional de PAN/Apple (FOXG_10672). McGowan *et al.*, (2020), realizaron recientemente una revisión de este dominio y encontraron una amplia lista de proteínas efectoras con dominio funcional PAN/Apple en especies de *Phytophthora*; reportaron 25, 32 y 20 proteínas efectoras anotadas con PAN/Apple en *P. chlamydospora*, *P. gonapodyides* y *P. pseudosyringae* respectivamente. Más del 90% de estas proteínas de *Phytophthora* son similares a las proteínas CBEL (*Cellulose Binding Elcitor Lectin*) en el que CMB1 (*Carbohydrate-binding module family 1*) está asociado con un dominio PAN/Apple no catalítico, conocido por interactuar con carbohidratos o proteínas específicas (Larroque *et al.*, 2012). Es posible que el candidato PfID77734 con dominio PAN/Apple participe en el proceso de infección de *Musa* spp por *P. fijiensis* mediando la interacción y secuestro de posibles oligosacáridos generados por el patógeno en el proceso de infección, evitando así el reconocimiento por el hospedero y evitando desencadenar la respuesta hipersensible. Por otra parte, por su probable interacción con polisacáridos y proteínas, este candidato puede facilitar de la adhesión a la pared celular del hospedero.

Otro dominio también relacionado con el proceso de interacción, es el de Deuterolisina metaloproteasa M35, que fue predicho en el efector PfID189452. Se ha reportado que este dominio es importante en proteínas de especies de *Metarrizium* (A. Huang *et al.*, 2020), en enzimas secretadas que muestran actividad proteolítica y facilitan la penetración del hongo en las cutículas de los insectos. En *M. robertsii*, la sobreexpresión de este gen puede inhibir la síntesis de cutícula y de melanización de hemocitos en los insectos. Por otra parte Pan *et al.*, (2020) reportaron que la proteína RcMEP2 de *Rhizoctonia cereales*, hongo necrotrófico del trigo, actuó como factor de virulencia desencadenando una acumulación excesiva de especies reactivas de oxígeno (“*Reactive oxygen species*” ROS), inducción de la muerte celular e inhibición de la expresión de quitinasas en trigo. Lo anterior apoya la hipótesis de que el candidato PfID189452 podría estar involucrado en el proceso de infección participando en el cambio de la

etapa biotrófica a necrotrófica, desencadenando una producción de ROS la inhibición de la expresión de quitinasas de *Musa* spp. y/o induciendo posiblemente la muerte de las células del hospedero.

En el caso del efector PfID80609, con dominio de carboxilesterasa (IPR002018/PF00135), en *Zymoseptoria tritici*, se reportó una alta expresión diferencial (con un Log2FC de 8.9) en respuesta a diferentes condiciones de luz de una proteína con este dominio (ZtritIPO323_04g04132) (McCorison y Goodwin, 2020). Por otro lado, en un análisis comparativo de los secretomas predichos *in silico* de los hongos necrotróficos *Sclerotinia sclerotiorum* y *Botrytis cinerea*, se encontró que el dominio de carboxilesterasa se encuentra más abundante en el secretoma de *B. cinerea* con 18 proteínas anotadas con este dominio, mientras que en *S. sclerotiorum* ocho proteínas anotaron para este dominio. Para el caso de *B. cinerea*, se identificó un candidato de efector con el motivo Y/FWxC y anotado con el dominio de carboxilesterasa, dominio relacionado con la degradación de lípidos (Heard *et al.*, 2015). Esto podría sugerir que en el caso de *P. fijiensis*, el candidato PfID80609 estuviera involucrado en la degradación de lípidos durante el desarrollo de la infección de *Musa* spp. Se sabe que los lípidos tienen un papel fundamental en la interacción, aunque es un área en la que falta estudiar más. El candidato PfID80609 podría participar en la transición de las etapas de infección, pasando de una etapa biotrófica a una necrotrófica, ya que durante la interacción, existe una producción de oxilipinas por parte del hospedero. Estas moléculas podrían ser aprovechadas por *P. fijiensis* para facilitar su propia virulencia y desarrollo, promoviendo la regulación de la esporulación y la producción de micotoxinas (Dal Molin *et al.*, 2018; Christensen y Kolomiets, 2011).

Al momento ya se han clonado y caracterizado una variedad de efectores en hongos biotróficos y hemibiotróficos, como por ejemplo los efectores biotróficos AvrA10, AvrK1 de *Blumeria graminis* f. sp. *hordei* (Shen *et al.*, 2007; Ridout *et al.*, 2006), los efectores AvrL567, AvrM, AvrP123 y AvrP4 de *Melampsora lini* (Kale *et al.*, 2010; Lawrence *et al.*, 2010; Catanzariti *et al.*, 2006; Dodds *et al.*, 2004); el efector UhAvr1 de *Ustilago hordei* (Grewal *et al.*, 2008); mientras que en hongos hemibiotróficos como *Fusarium oxysporum* f. sp. *lycopersici* los efectores Avr1 (SIX4), Avr2 (SIX3) y Avr3 (SIX1) (Houterman *et al.*, 2008; Rep *et al.*, 2004); los efectores Avr-Pita1, PWL1, PWL2, ACE1, Avr-CO39 de *Magnaporthe oryzae* (Khang *et al.*, 2010; Li *et al.*, 2009; Bohener *et al.*, 2004; Orbach *et al.*, 2000); y los efectores Avr2, Avr4, Avr4E y Avr9 de *Cladosporium fulvum* (van den Burg *et al.*, 2006; Westerink *et al.*, 2004; van den Hooven *et al.*, 2001; Dixon *et al.*, 1996). Esto da pauta a que las bases de datos de efectores vayan aumentando, sin embargo, la clonación de estos y los análisis funcionales siguen siendo los

principales cuellos de botella; ejemplo de ello es *P. fijiensis*, donde en la actualidad únicamente tres efectores han sido caracterizados, el PfAvr4, PfEcp2 y PfEcp6 (Stergiopoulos *et al.*, 2010; Bolton *et al.*, 2008).

La identificación de los repertorios de efectores de los patógenos está facilitando una mayor comprensión de los mecanismos moleculares subyacentes a la virulencia, así como guiando el desarrollo de estrategias de control de enfermedades. Los estudios moleculares han caracterizado más de 60 efectores fúngicos en múltiples especies; sin embargo, esto es muy poco en comparación con el repertorio de efectores candidatos para cada especie patógena (Sperschneider *et al.*, 2015). Por ejemplo, en *Blumeria graminis* f. sp. *hordei* se predice que el 7% de su genoma codifica proteínas efectoras secretadas candidatas (CSEP) (Pedersen *et al.*, 2012). Los efectores funcionan de múltiples formas, incluida la inhibición de las enzimas del huésped, la modulación de las respuestas inmunitarias de las plantas y los mecanismos de silenciamiento de genes del huésped. Por ejemplo el Avr4 de *Cladosporium fulvum*, además de proteger las paredes del hongo del ataque de quitinasas producidas por el hospedero, inhibe varias proteasas, incluida la proteasa de tomate Rcr3, que es necesaria para la defensa basal de la planta e induce la RH en especies de tomate que portan el receptor Cf-4 (Song *et al.*, 2009; van Esse *et al.*, 2008; Rooney *et al.*, 2005) o el efector AvrLm1 de *Leptosphaeria maculans* que interactúa con la proteína MAP cinasa 9 (BnMPK9) del huésped *Brassica napus*, resultando en un incremento en la acumulación de las proteínas del hospedero, así como también potenciando su fosforilación, lo que provoca la inducción de la muerte celular (Fouché *et al.*, 2018; Ma *et al.*, 2018; Soyer *et al.*, 2014).

Con los rápidos avances de la genómica comparativa y nuevas plataformas de secuenciación, muchos genomas fúngicos se secuenciarán en un futuro y, con avances paralelos en herramientas bioinformáticas, esto permitirá acelerar el descubrimiento de nuevos efectores. Sin embargo, hasta ahora, la función y/o participación de candidatos a efectores en la interacción con el hospedero se puede investigar experimentalmente mediante la alteración de genes, el silenciamiento de genes o la sobreexpresión. Por lo tanto, la elucidación de los mecanismos para la identificación de efectores fúngicos sigue siendo un importante desafío.

CAPÍTULO V

EXPRESIÓN HETERÓLOGA DE LAS PROTEÍNAS EFECTORAS PfAVR4 Y PfECP2 DE *P. FIJIENSIS*

5.1 INTRODUCCIÓN

Las enfermedades de las plantas causadas por hongos fitopatógenos son una gran amenaza para la producción de cultivos y la seguridad alimentaria mundial. Para poder colonizar las plantas, los patógenos utilizan diversas estrategias para vencer la inmunidad vegetal y lograr colonizar las plantas hospederas. En el caso de *Pseudocercospora fijiensis*, hongo patógeno causante de la enfermedad de la Sigatoka negra en bananos y plátanos (Friesen, 2016; Churchill, 2011; Stergiopoulos *et al.*, 2010), analizando el genoma de *P. fijiensis* reportaron dos ortólogos funcionales de *Cladosporium fulvum*, los efectores Avr4 y Ecp2, que se sabe son factores de virulencia (De Wit *et al.*, 2009). El ortólogo funcional PfAvr4 de *P. fijiensis* actúa como un factor de virulencia en defensa que protege la pared celular del hongo contra la hidrólisis por las quitinasas del hospedero, PfAvr4, y es también capaz de desencadenar una respuesta hipersensible (RH) mediada por Cf-4 en tomate (Stergiopoulos *et al.*, 2010).

Además del PfAvr4, también se identificaron tres proteínas en el genoma de *P. fijiensis* que muestran homología con CfEcp2 de *C. fulvum* (PfEcp2, PfEcp2-2, PfEcp2-3), de los cuales, la que presenta más alta identidad (57%) promueve la virulencia al interactuar con un posible receptor del hospedero, induciendo una RH en una línea de tomate transformada con el receptor Cf-2. Curiosamente, a diferencia de CfEcp2 que no provoca reacción en *Nicotiana benthamiana* en ausencia de Cf-2, PfEcp2 por sí solo provoca necrosis celular, por lo que CfEcp2 y PfEcp2 pueden ser similares en función, o al menos son reconocidos por el mismo cognado (Stergiopoulos *et al.*, 2010), pero puede haber diferencias entre ellos.

El PfEcp2 es una proteína secretada de 161 aminoácidos que se identificó originalmente como un factor de virulencia en *C. fulvum*, ya que la eliminación del gen reduce su virulencia en interacción con plantas de tomate (Laugé *et al.*, 1997). Aunque aún se desconoce la función intrínseca de esta proteína efectora durante la patogénesis, se ha planteado la hipótesis de que Ecp2 interactúa con un objetivo de virulencia del hospedero para inducir necrosis y la liberación de nutrientes durante la infección (Stergiopoulos *et al.*, 2010).

En otro reporte Stergiopoulos *et al.*, (2012), realizaron un análisis de búsqueda de homólogos del efector Ecp2 en la base de datos del GenBank e identificaron presuntos ortólogos de este efector en otras especies de hongos de la clase Dotideomicetos, como lo son *Dothistroma septosporum*, *Septoria musiva*, *Setosphaeria turcica*, *Cochliobolus heterostrophus*, *Pyrenophora tritici-repentis* y *Stagonospora nodorum* (Cuadro 5.1).

Cuadro 5.1 Distribución del efector Ecp2 en patógenos de clase de Dotideomicetos.

Especie	Efector	Referencia
<i>Cladosporium fulvum</i>	CfEcp2, CfEcp2-2, CfEcp2-3	Lauge <i>et al.</i> , 1997
<i>Dothistroma septosporum</i>	Dotse1_158381, Dotse1_127671, Dotse1_23431	Stergiopoulos <i>et al.</i> , 2012
<i>Pseudocercospora</i> (<i>Mycosphaerella</i>) <i>fijiensis</i>	CfEcp2, CfEcp2-2, CfEcp2-3	Stergiopoulos <i>et al.</i> , 2010; Stergiopoulos <i>et al.</i> , 2012
<i>Mycosphaerella</i> <i>graminicola</i>	Mgra_104404, Mgra_107904, Mgra_111636, Mgra_106176	Stergiopoulos <i>et al.</i> , 2010; Stergiopoulos <i>et al.</i> , 2012
<i>Septoria musiva</i>	Sepmu1_146583, Sepmu1_149875	Stergiopoulos <i>et al.</i> , 2012
<i>Setosphaeria turcica</i>	Settu_38372	Stergiopoulos <i>et al.</i> , 2012
<i>Cochliobolus</i> <i>heterostrophus</i>	Chet_100623	Stergiopoulos <i>et al.</i> , 2012
<i>Pyrenophora tritici-repentis</i>	Ptri_00203, Ptri_07799	Stergiopoulos <i>et al.</i> , 2012
<i>Stagonospora nodorum</i>	Snod_04278.3	Stergiopoulos <i>et al.</i> , 2012

Sin embargo, la función intrínseca del efector Ecp2 hasta el momento no ha sido elucidada en *C. fulvum* ni en ninguno de los hongos que poseen homólogos.

Actualmente hay diferentes enfoques para el análisis de los efectores. En el caso de *P. fijiensis*, Arango *et al.* (2016), reportaron la expresión heteróloga de PfAvr4 en el sistema de *Pichia pastoris* cepa GS115, infiltrando en hojas de las variedades de banano Calcuta IV y Enano

Gigante los extractos de proteínas secretadas por la levadura. La infiltración activó la respuesta hipersensible en la variedad Calcuta IV, sugiriendo que la variedad de banano Calcuta IV posee proteínas de resistencia (proteínas R), como es de esperarse por ser un cultivar resistente. Dicha proteína cognada aún no ha sido identificada y que podría ser ortóloga del Cf-4 que está presente en las especies resistentes de tomate (Arango *et al.*, 2016; Stergiopoulos *et al.*, 2010).

Los enfoques bioquímicos y estructurales reportados por Kohler *et al.*, (2016), evidenciaron que el CfAvr4 de *C. fulvum* y PfAvr4 de *P. fuligena* comparten especificidad funcional en cuanto a la unión de carbohidratos, también, se encontró que estos efectores confieren protección contra las quitinasas provenientes de la célula vegetal, mientras que el análisis por cristalografía mostró que con la mutagénesis dirigida, los residuos importantes para la unión con la quitina, no influyen directamente en el reconocimiento mediado por Cf-4. Mientras que Mesarich *et al.*, (2016), reportaron para Avr4 que el residuo de prolina 87 (Pro87) es importante para que se lleve a cabo el reconocimiento por Cf-4 y se desencadene la RH.

Las proteínas recombinantes pueden ser expresadas usando sistemas procariotas (*Escherichia coli* y *Bacillus subtilis*), sistemas eucariotas (levadura, células de insectos y células de mamíferos) o sistemas *in vitro* (Jia y Jeon, 2016).

La tecnología de producción de ADN y proteínas recombinantes ha permitido avanzar en muchos estudios bioquímicos y proyectos que requieren obtener o aumentar la cantidad de la proteína de interés (enzimas, anticuerpos, hormonas, etc.), para su análisis, logrando alcanzar una producción a gran escala (Rosano y Ceccarelli, 2014).

E. coli se ha usado como sistema de producción de efectores dado que es un microorganismo de manipulación sencilla, de crecimiento rápido y además de que su medio de crecimiento es menos costoso en comparación al medio de crecimiento de otro tipo de células. Tripathi, (2016), describió que *E. coli* se ha utilizado para la producción de farmacéuticos, nutracéuticos, proteínas terapéuticas, intermediarios para el diagnóstico, vacunas, anticuerpos, enzimas y proteínas complejas, todos ellos de alto valor comercial. En *E. coli*, las proteínas heterólogas pueden alcanzar del 5 al 50 % de la proteína celular total (Rosano y Ceccarelli, 2014). En estudios recientes, la cepa de expresión *E. coli* Rosetta Gami B se ha usado para expresar la proteína efectora PfAvr4 de *P. fijiensis* (Muñoz-Pérez, 2018).

El PfAvr4 en su forma madura e inmadura, fue expresado usando el vector pLATE, purificadas por cromatografía de afinidad con resina de níquel afín para la etiqueta de histidina presente en

las recombinantes y secuenciadas por LC-MS/MS. Una vez que confirmaron la presencia del efector PfAvr4, realizaron ensayos de infiltración en hojas de banano resistente y susceptible, observando que PfAvr4 en su forma madura e inmadura, ambas generaron respuesta hipersensible (Muñoz-Pérez, 2018).

Por otra parte, el sistema de expresión por levaduras ha sido relevante en el desarrollo de metodologías con alto rendimiento en la purificación de proteínas. *Pichia pastoris* es un huésped muy utilizado para expresar proteínas recombinantes debido a que puede alcanzar densidades de células muy elevadas, y debido a que naturalmente secreta pocas proteínas propias, lo que resulta muy conveniente. *P. pastoris* es un productor eficiente de proteínas secretadas. Su uso como alternativa a *S. cerevisiae* se ha ido consolidando por la variedad de proteínas heterólogas que se han producido con altos rendimientos, que van desde la toxina del tétanos y el factor de crecimiento epidérmico del ratón (Clare *et al.*, 1991) hasta proteínas de membrana, incluidos los transportadores ABC humanos, acuaporinas y tetraspaninas (Rosenberg *et al.*, 2010; Jamshad *et al.*, 2008; Urbatsh *et al.*, 2001).

Entre los ensayos principales utilizados para demostrar la entrada de efectores en las células huésped, el "ensayo de reentrada celular", se basa en la expresión heteróloga de la región codificante completa, incluido su péptido señal de secreción; puede estar fusionada a una etiqueta fluorescente, permitiendo la evaluación en la planta hospedera (Catanzariti *et al.*, 2006). Este método se ha utilizado para evaluar la entrada celular de varios efectores de hongos y oomicetos y a su vez ha permitido identificar los dominios de reconocimiento, necesarios para la entrada (Kale *et al.*, 2010; Rafiqi *et al.*, 2010; Dou *et al.*, 2008).

Kettles *et al.*, (2017) reportaron los análisis funcionales de 63 proteínas efectoras de *Z. tritici*, clonadas con el péptido señal nativo; los constructos fueron agroinfiltrados en plantas de *Nicotiana benthamiana*, lo que permitió la identificación de 14 efectores que generan muerte celular. Estos 14 efectores fueron clonados en otro experimento, pero ahora sin el péptido señal de secreción, y fueron agroinfiltrados de nuevo en hojas de *N. benthamiana*, resultando en la falta de reconocimiento y ausencia de RH, evidenciando que los péptidos señal, al menos los de éstos efectores, son requeridos para la redirección y reconocimiento de la proteína efectora dentro de la célula vegetal.

En otro trabajo, reportado por See *et al.*, (2019), realizaron la expresión en *E. coli* cepa SHuffle y en *P. pastoris* de dos efectores ya conocidos de *Pyrenophora tritici-repentis*, ToxA y ToxB (los

cuales confieren susceptibilidad a la enfermedad conocida como mancha amarilla); estos efectores se expresaron en los sistemas procariota y eucariota para hacer análisis comparativos de la expresión. En el caso de la expresión de ToxA, los rendimientos de la proteína obtenida fueron bajos tanto en su versión madura e inmadura en los dos sistemas de expresión utilizados (entre 1.2µg y 3.3 µg). Mientras que en el caso de la expresión del efector ToxB, para la versión inmadura se tuvo un rendimiento bajo en *E. coli* y *P. pastoris* (6.9 µg y 13.1 µg respectivamente), pero en su versión madura se tuvo un alto rendimiento en los dos sistemas, obteniendo 79.6 µg de proteína purificada de *E. coli* y 69.8 µg de proteína obtenida de *P. pastoris*. Al infiltrar las proteínas purificadas de ToxA y ToxB en hojas de trigo, observaron que ToxB confiere mayor sensibilidad que ToxA, generando también clorosis en plantas de cebada. Estos trabajos muestran que se puede evaluar a los efectores mediante infiltración de la proteína (See *et al.*, 2019). Similarmente, en el caso de Muñoz-Pérez (2018), la infiltración de la proteína madura generó la respuesta esperada.

Por otra parte, la eliminación del gen del efector VmE02 en el hongo necrotrófico *Valsa mali* atenuó en gran medida la conidiación del patógeno, pero no su virulencia, y la inoculación de VmE02 de *N. benthamiana* mejora la resistencia de las plantas ante la infección por *Sclerotinia sclerotiorum* y *Phytophthora capsici*. El efector VmE02 codifica para una proteína secretada y en los análisis funcionales se observó que la proteína infiltrada indujo la muerte celular en *N. benthamiana*, tanto cuando contiene como cuando no contiene el péptido señal (Nie *et al.*, 2019).

Los efectores sin el péptido señal nativo pueden desencadenar una RH, por ejemplo, Wang y colaboradores, (2019) reportaron que el efector gsr1 (smut_2965) de *Tilletia horrida* no requería del péptido señal para desencadenar la muerte celular agroinfiltrando el constructo en *N. benthamiana*. Este hallazgo es similar al del efector SsCP1, una proteína extracelular secretada por *Sclerotinia sclerotiorum*, que también desencadena la muerte celular sin el péptido señal (Yang *et al.*, 2018). Con base en todos estos reportes se puede esperar que la expresión de efectores sin péptido señal de *P. fijiensis* debieran ser capaces de disparar respuesta de RH.

Por lo tanto, con base en lo descrito, en el presente trabajo se siguió la metodología establecida por Muñoz-Pérez (2018) para la expresión de los candidatos en el sistema de *E. coli* Rosetta gami B. Por ello, como primer enfoque se decidió reproducir la técnica para la expresión de PfAvr4 y poder luego aplicarla para la expresión de los candidatos de efectores de *P. fijiensis* seleccionados en el presente trabajo. Asimismo, se decidió realizar la expresión del efector PfEcp2 con la finalidad de estudiarlo y caracterizarlo, ya que como se mencionó anteriormente

su función es completamente desconocida. Por estas razones los ensayos de expresión heteróloga de efectores se iniciaron con estos dos efectores, por lo que en este capítulo se presenta la expresión heteróloga de los efectores PfAvr4 y PfEcp2 en el sistema de *E. coli* Rosetta gami B, evaluando los constructos de los efectores con péptido señal y sin péptido señal de secreción.

Cabe resaltar, que uno de los intereses en éste capítulo, era comparar la función de los efectores PfAvr4 y PfEcp2 expresados en sistema procariota contra el sistema eucariota, sin embargo, por la situación de la pandemia, estos experimentos no fueron posibles de concluirse. Sin embargo, es importante mencionar que se confirmaron las clonas transformantes en células de *Pichia pastoris* cepa GS115, transformada con los vectores confirmados pPICZ-PfAvr4 (+SP), pPICZ-PfAvr4 (-SP), pPICZ-PfEcp2 (+SP), pPICZ-PfEcp2 (-SP). Este material podrá dar pauta para el seguimiento del análisis funcional de estos efectores de *P. fijiensis* cuando pueda continuarse en el grupo de investigación. Sin embargo, debido a que estas clonas no fueron evaluadas en el análisis de expresión, a continuación se presentan únicamente los resultados de clonación y expresión en *E. coli* Rosetta gami B.

5.2 MATERIALES Y MÉTODOS

5.2.1 Secuencias de los efectores reportados de *P. fijiensis*: PfAvr4 y PfEcp2

Para los análisis, se emplearon las secuencias codificantes de los efectores PfAvr4 y Ecp2, tanto incluyendo como excluyendo la región que codifica el péptido señal. Para el caso de PfAvr4, se recibieron dos clonas del grupo de investigación del Dr. Ignacio Islas Flores clonados en vector pLATE por Muñoz-Pérez, (2018).

En el caso de PfEcp2, los fragmentos fueron amplificados en nuestro grupo de investigación por Barahona (2017), clonado en el vector pGEM-Teasy®, por lo que en el presente trabajo se subclonó en pLATE. Los constructos se denominaron PfAvr4 (+SP), PfAvr4 (-SP), PfEcp2 (+SP) y PfEcp2 (-SP), según contuvieran o no la región correspondiente al péptido señal.

5.2.2 Amplificación y clonación de los fragmentos del efector PfEcp2

La amplificación de la región codificante correspondientes a la proteína PfEcp2 en su versión madura (sin péptido señal) e inmadura (con péptido señal), se realizó usando cebadores conteniendo las secuencias adaptadoras para la ligación en pLATE. Los cebadores (Cuadro 5.1) se diseñaron con el programa OligoAnalyzer Tool del IDT (<https://www.idtdna.com/calc/analyzer>).

A los productos de PCR obtenidos con estos cebadores se les adicionó el sufijo “LIC” (por aLICatos Ligation, al que pertenece el sistema pLATE).

Cuadro 5.2 Cebadores utilizados para la amplificación y clonación de los efectores PfAvr4 y PfEcp2.

Cebador	Secuencia 5' → 3'	Tm °C
	Específicos	
*FcAvr4	ATGCTTTCAACTACGAGAATAACACTGCTCG	60.1
*FpAvr4	GGAACGATGCAGGTGCGAGAC	60.7
*RcAvr4	GACGAAACAGGATGCGACAACGTGTAA	60.9
FcEcp2	ATGCATTTCTCCCGCGCTGCTG	63.0
FpEcp2	ACATGCACTCGATGCAAGAGCAAG	61.2
RcEcp2	TGGAGCATCGTCCACCCTTGA	60.0
pLATE		
*FcEcAvr4	GGTTGGGAATTGCAAATGCTTTCAACTACGAGAATAACACTGCTCG	60.1
*FpEcAvr4	GGTTGGGAATTGCAAGGAACGATGCAGGTGCGAGAC	60.7
*RcEcAvr4	GGAGATGGGAAGTCATTATTACAGTTGTGCGCATCCTGTTTCGTC	60.9
FcEcEcp2	AGAAGGAGATATAACTATGCATTTCTCCCGCGCTGC	61.6
FpEcEcp2	AGAAGGAGATATAACTATGACATGCACTCGATGCAAGAGCAAG	60.0
RcEcEcp2	GTGGTGGTGATGGTGATGGCCAGGGTGGACGATGCTCCAGTC	61.3

*Cebadores proporcionados por el grupo de investigación del Dr. Ignacio Islas Flores, producto del trabajo de Muñoz-Pérez, 2018.

La ligación de los productos de PCR, PfEcp2 (+SP) y PfEcp2 (-SP) se realizó siguiendo el protocolo descrito en el manual del paquete comercial “aLICator Ligation Independent Cloning and Expression System” (Thermo Scientific). La reacción de la ligación de cada candidato en el plásmido pLATE, contenía amortiguador LIC 5X, el producto de PCR purificado (correspondiente al efector), agua libre de nucleasas y la ADN polimerasa T4 (1 U/μL), todo en un volumen final de 10 μL. La mezcla de reacción se incubó a temperatura ambiente por 5 minutos. Enseguida, se añadió a la mezcla 1 μL (60 ng= 0.02 pmol ADN) del plásmido pLATE. Una vez pasado el tiempo de reacción, se procedió con la transformación de células quimiocompetentes de *E. coli* TOP10, usando el método por choque térmico, 2 μL de la reacción de ligación se adicionaron a las células competentes en un tubo Eppendorf de 1.5 mL. La mezcla se incubó por 30 minutos en hielo y enseguida se colocaron a 42 °C por 30 segundos, y se adicionó a las células 1 mL de medio LB líquido (10g/L peptona, 5g/L extracto levadura, 10g/L de NaCl). Las células transformadas se sembraron en medio LB sólido (10g/L peptona, 5g/L extracto levadura, 10g/L de NaCl, Agar 20g/L), adicionando ampicilina (Amp) 100μg/μL. Se seleccionaron las clonas transformantes y estas se recuperaron en cajas de Petri con medio LB suplementado con agar y ampicilina. Se tomaron muestras de las clonas para confirmar la presencia de los insertos mediante PCR en colonia; se empleó la polimerasa DreamTaq PCR Master Mix (2x) (Thermo Scientific™) (Figura 5.1).

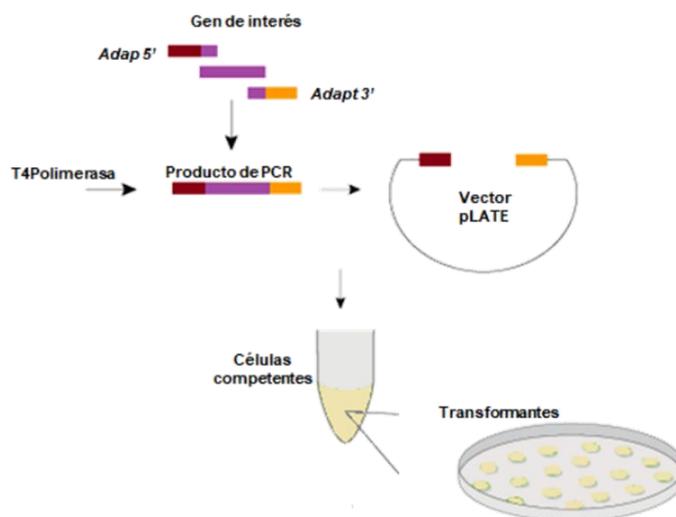


Figura 5.1 Diagrama experimental aplicado para la ligación y clonación de los fragmentos de PfEcp2 en vector pLATE.

5.2.3 Secuenciación de los plásmidos pLATE-PfEcp2 (+SP) y pLATE-PfEcp2 (-SP)

Una vez ligadas las secuencias codificantes correspondientes para el efector PfEcp2 de *P. fijiensis* (con y sin péptido señal), los plásmidos resultantes se secuenciaron para confirmar que los insertos estuvieran en marco de lectura abierto. Para ellos se usaron los cebadores LICFw (5'- TAATACGACTCACTATAGGG -3') y LICRv (5'- GAGCGGATAACAATTTTCACACAGG-3'), indicados por el kit “aLICator Ligation Independent Cloning and Expression System” (Thermo Scientific), los cuales flanquean la región donde se une el fragmento de interés en el vector. La secuenciación se realizó en el Instituto Potosino de Investigación Científica y Tecnológica (IPYCIT, México).

5.2.4 Transformación de *E. coli* cepa Rosetta gami B con los constructos pLATE-PfEcp2 (+SP) y pLATE-PfEcp2 (-SP)

Una vez confirmada la correcta construcción de los vectores pLATE-PfEcp2 (+SP) y pLATE-PfEcp2 (-SP), se procedió con la transformación de la cepa *E. coli* Rosetta Gami B. Esta cepa contiene los elementos regulatorios que codifican para la ARN polimerasa T7, misma que está bajo el control transcripcional del promotor lacUV5, el cual es inducido por IPTG. La

transformación de la cepa se llevó a cabo según el protocolo de choque térmico, plaqueando las células transformadas en medio LB “Luria-Bertani” (10g/L peptona, 5g/L extracto levadura, 10g/L de NaC, Agar bacteriológico 20g/L), adicionado con 100µg/mL de ampicilina (Amp), 15µg/mL de kanamicina (Kn) y 35µg/mL de cloranfenicol (Cm). Una vez crecidas las clonas transformadas, se resembraron en medio de cultivo fresco. Estas colonias, fueron evaluadas empleando la técnica de PCR en colonia usando la polimerasa DreamTaq PCR Master Mix (2x) (Thermo Scientific™) y los oligos que flanquean el inserto, LICFw y LICRv, mencionadas anteriormente (apartado 5.2.3).

5.2.5 Expresión heteróloga de los efectores PfEcp2 (+SP) y PfEcp2 (-SP)

Las cepas confirmadas de *E. coli* Rosetta gami B de cada constructo se cultivaron e indujeron para la expresión de la proteína efectora PfEcp2 (+SP) y PfEcp2 (-SP); también se cultivo el control sin transormar (vector pLATE solo) y PfAvr4 (-SP), como control de técnica. Las células se cultivaron en medio LB “Luria-Bertani” liquido (Invitrogen, CA, USA) (10g/L peptona, 5g/L extracto levadura, 10g/L de NaC) adicionado con 100µg/mL Amp, 15µg/mL Kn y 35µg/mL Cm, se dejó saturar 1L de medio. Una vez alcanzada la densidad óptica de 0.5 unidades de absorbancia (U. A.) a 600nm, se añadió IPTG (concentración final de 1 mM), incubando los cultivos a 15°C con agitación de 120 rpm durante 16 horas (Figura 5.2).

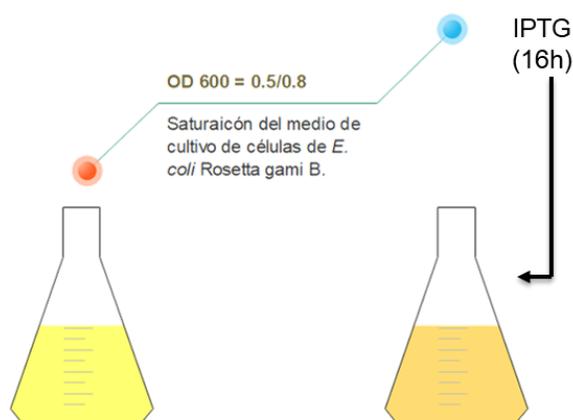


Figura 5.2 Proceso de inducción con IPTG 1 mM en *E. coli* Rosetta gami B.

5.2.6 Purificación de PfAvr4 y PfEcp2 de *P. fijiensis*

La purificación de los efectores recombinantes PfEcp2 (+SP), Ecp2 (-SP) y el control PfAvr4M (-SP) se realizó mediante cromatografía de afinidad a partir de la fracción soluble de la bacteria, utilizando la resina "ProBond™" de Invitrogen. Después de transcurridas las 16 horas de inducción, el cultivo se centrifugó a 6,000 rpm por 10 min a 4°C. Se separó la fase del medio de cultivo de las pastillas celulares, las pastillas se resuspendieron en amortiguador STET (8% sacarosa, 50mM Tris-HCl pH8, 1mM EDTA y 0.1% de Tritón X100, 1 mM de PMSF) y se lisaron mediante 5 ciclos de sonicación. El lisado se centrifugó a 10,000 rpm por 30 minutos a 4°C, el sobrenadante se filtró en membranas Millex-GS de 0.45µm (MF-Millipore™membrane). El filtrado se aplicó a la columna de níquel (resina) equilibrada previamente con el amortiguador 1 (50 mM Tris-HCl pH 7.5 y 350 mM NaCl). La columna se lavó con el amortiguador 1 y los siguientes lavados se hicieron con el amortiguador 1 adicionado con imidazol a diferentes concentraciones (10, 50, 200 y 250 mM). Las fracciones eluidas se analizaron por SDS-PAGE en gel al 12%.

5.2.7 Solubilización de cuerpos de inclusión

Después de la purificación de los efectores recombinantes PfEcp2 (+SP), Ecp2 (-SP) y el control PfAvr4M (-SP) en la fase soluble de la bacteria inducida, las partículas de las células lisadas fueron usadas para la solubilización de los cuerpos de inclusión. Las pastillas se resuspendieron en 1 mL de amortiguador Tris-HCl 50 mM pH8, sonicando 3 veces. Las muestras se transfirieron

en tubos Eppendorf de 1.5mL, se centrifugaron a 12,000 rpm por 10 min a 4°C. Se descartó el sobrenadante y se hicieron 2 lavados con 1 mL de amortiguador Tris-HCl 50 mM pH8 cada uno, centrifugando cada vez a 12,000 rpm por 5 min a 4°C y descartando cada vez el sobrenadante. Se resuspendió la pastilla final en 500 µL de amortiguador Tris-HCl 50 mM pH8 para luego pasar las pastillas resuspendidas a un tubo falcon de 50 mL, adicionando 3 mL de amortiguador Tris-HCl 50 mM pH8 con 6 M de guanidina. Se incubaron a 4°C toda la noche. Posteriormente se centrifugó a 12,000 rpm por 10 minutos a 4°C, se recuperó la fase soluble; a 100 µL de la fase solubilizada se adicionó 900 µL de etanol absoluto, dejándose a -20°C toda la noche. Se centrifugó como se ha descrito y la pastilla se lavó dos veces con etanol al 90% (centrifugando cada vez a 12,000 rpm por 5 min a 4°C), y se dejó secar la pastilla a 45°C por 5 min. Al final, la pastilla se resuspendió en 30 µL de amortiguador Tris-HCl 50 mM pH8.

5.2.8 Cuantificación de proteínas por método Bradford

Una alícuota del dializado (5 µL) se colocó en tubos de ensaye y se adicionó con agua bidestilada, hasta alcanzar un volumen de 200 µL. A la muestra se le adicionó 1 mL de reactivo de Bradford (1976) y se incubó por 15 minutos a temperatura ambiente. A la mezcla resultante se le determinó la absorbancia a 595 nm (Ernst y Zor, 2010; Kielkopf *et al.*, 2020). La curva de proteína se construyó usando 1, 2.5, 5, 7.5, 10, 15, 20 y 25 µg de albúmina sérica bovina. La concentración de proteína en la muestra se determinó por interpolación del valor de la absorbancia en el gráfico de referencia.

5.3 RESULTADOS

5.3.1 Amplificación de los CDSs del efector PfEcp2 (+SP) y PfEcp2 (-SP) con los adaptadores LIC y generación de los constructos pLATE-efector

Los CDSs pertenecientes a PfEcp2 (+SP) y pLATE-PfEcp2 (-SP) clonados en pGEM-T-easy fueron usados como templado para llevar a cabo la PCR con los cebadores FcEcp2 (sentido) y RcEcp2 (reverso) para la amplificación del CDS versión inmadura de la proteína (Figura 5.3a), mientras que los cebadores FpEcp2 (sentido) y RcEcp2 (reverso) fueron usados para la amplificación del CDS versión madura de la proteína. La Figura 5.3 presenta los resultados de las amplificaciones con estos cebadores.

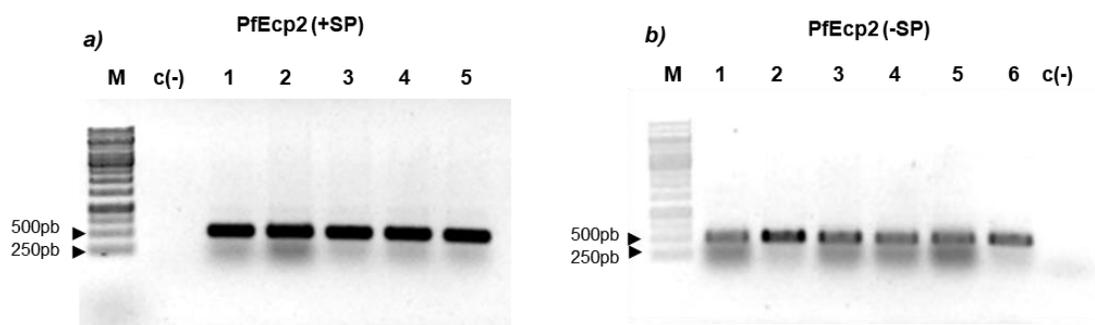


Figura 5.3 Amplificación de la región codificante de PfEcp2 (+SP) y PfEcp2 (-SP) de transformantes de *E. coli* TOP10. a) PfEcp2 (+SP) 1-5: Amplificones correspondientes al fragmento de PfEcp2 (+SP); b) PfEcp2 (-SP) 1-6: Amplificones correspondientes al fragmento de PfEcp2 (-SP). M: Marcador molecular; c(-): Control negativo sin templado.

5.3.2 Secuenciación de los vectores de expresión pLATE-PfEcp2 (+SP) y pLATE-PfEcp2 (-SP)

Los constructos de expresión pLATE-PfEcp2 (+SP) y pLATE-PfEcp2 (-SP) fueron secuenciados usando los cebadores LIC del kit de clonación. Las clonas que tuvieron 100% de identidad con la secuencia reportada de PfEcp2 fueron seleccionadas para los experimentos de expresión heteróloga. El análisis de las secuencias confirmó la presencia de la etiqueta de histidina y el sitio de unión para el ribosoma (Figura 5.4)



Figura 5.4 Secuencias nucleotídicas de los constructos pLATE-PfEcp2 (+SP) y pLATE-PfEcp2 (-SP). Los colores representados en las secuencias indican los componentes de la construcción. Gris, secuencia que codifica para los seis residuos de histidinas; magenta, secuencia de unión a ribosoma; rojo, secuencia codificante correspondiente al efector; rosa, codón de inicio; guinda, codones de paro.

5.3.3 Transformación de *E. coli* Rosetta gami B con los vectores pLATE-PfEcp2 (+SP) y pLATE-PfEcp2 (-SP).

Las colonias transformantes fueron analizadas por medio de PCR en colonia usando los cebadores específicos para PfEcp2 (+SP) y PfEcp2 (-SP). Se esperaban amplificaciones de ~530pb y ~480pb. La Figura 5.5 presenta los resultados del análisis de los amplicones.

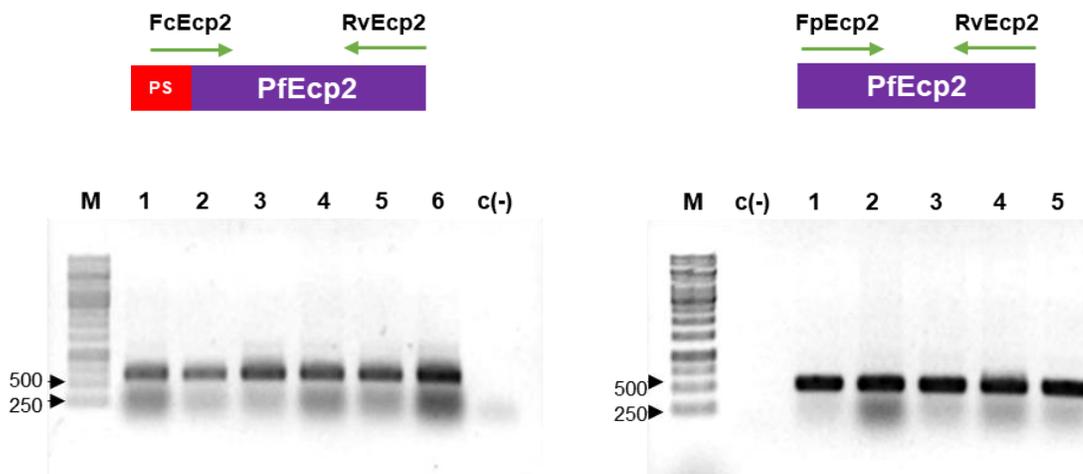


Figura 5.5 Confirmación por PCR en colonia de las cepas transformadas de *E. coli* Rosetta gami B con los constructos de expresión pLATE-PfEcp2 (+SP) y pLATE-PfEcp2 (-SP).

5.3.4 Expresión de PfEcp2 (+SP) y PfEcp2 (-SP) en *E. coli* Rosetta gami B.

Se colectaron las fracciones eluídas con las distintas concentraciones de imidazol (~ 500 μ L por alícuota). El Cuadro 5.3 corresponde a los resultados de cuantificación de la proteína

Cuadro 5.3 Cuantificación de proteína en las alícuotas recuperadas durante la purificación de PfEcp2 (+SP) y PfEcp2 (-SP). Cuadro 6 a, valores de absorbancia de las diferentes concentraciones de BSA. Cuadro 6 b, cantidad de proteína en las muestras colectadas durante la purificación de proteínas.

a)		b)			
CURVA BSA		Concentración de proteína μ g/ μ L			
μ g/mL	Valores	Lavado	Avr4M	Ecp2I	Ecp2M
1	0.032	FT	12.17	9.16	14.16
2.5	0.071	L	0.75	1.58	2.22
5	0.114	L10	0.35	0.42	0.565
7.5	0.164	L50	0.67	0.84	0.76
10	0.203	L200	0.37	0.69	1.06
15	0.254	L250	0.78	0.024	0.75
20	0.302				
25	0.34				

Para la curva de calibración se obtuvo una R de 0.98 y una R^2 de 0.96, (Figura 5.6), lo que indica que la linealidad permite usarla para calcular la concentración de las muestras en la purificación. Para cada muestra se analizó 12 μg de proteína por SDS-PAGE en gel al 15%. En la Figura 5.7 en las fracciones L50, L200 y L250 (con 50, 200 y 250 mM respectivamente), se observan perfiles proteicos complejos. Se señalan en rectángulo color negro la banda probable correspondiente a PfEcp2 (+SP) de ~ 19 kDa, mientras que para PfEcp2 (-SP) se indica una banda de proteína de ~ 17.8 kDa.

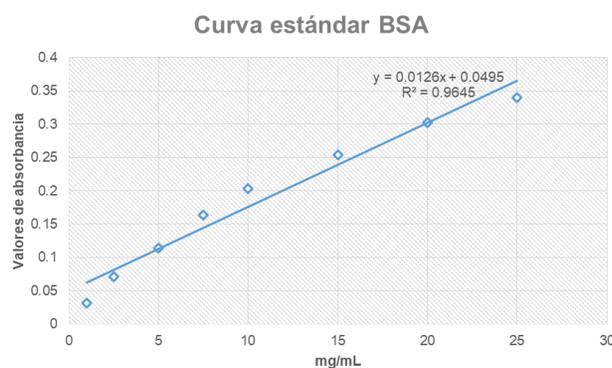


Figura 5.6 Curva estándar de cuantificación de proteínas por el método de Bradford, usando la proteína BSA como estándar.

En el análisis electroforético en el carril de la proteína solubilizada de los cuerpos de inclusión, se pudo observar la presencia de una banda correspondiente a PfEcp2 (+SP) de ~ 19 kDa, mientras que para PfEcp2 (-SP) se indica una banda de proteína de ~ 17.8 kDa, ambos señalados con el símbolo '*' en color rojo, indicando la posibilidad de que la proteína se esté yendo a cuerpos de inclusión.

Cabe mencionar que este gel fue el resultado de varios intentos de expresión en los que se analizó las proteínas secretada, soluble, y proteína obtenida de cuerpos de inclusión, dónde se puede sugerir que se distingue la proteína esperada, sin embargo, hacen falta los análisis de secuenciación para confirmar este resultado.

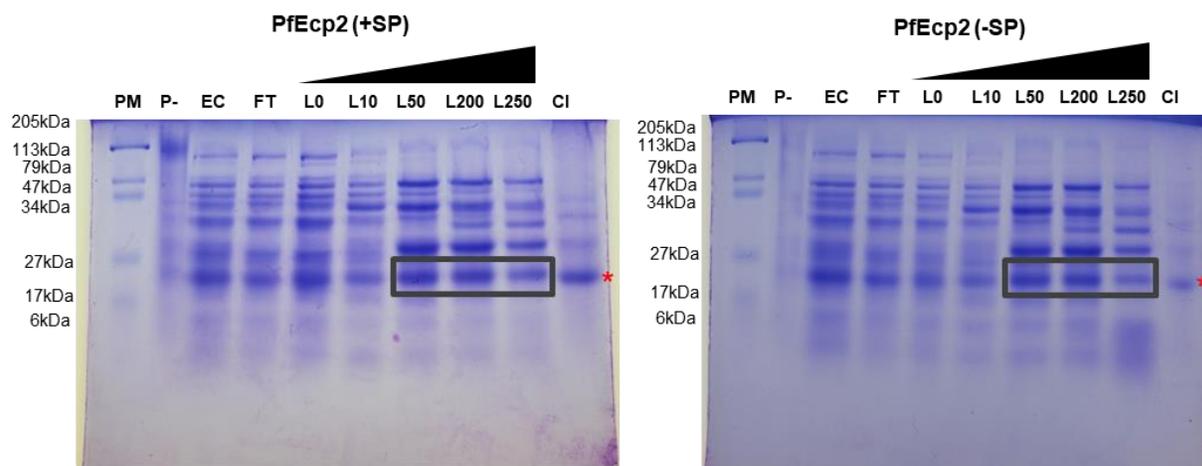


Figura 5.7 Análisis SDS-PAGE del proceso de purificación de PfEcp2 (+SP) y de PfEcp2 (-SP). Se cargó una cantidad de proteína de 12 μ g de cada muestra; al final de la electroforesis el gel fue teñido con Coomassie Blue. Carril 1, (PM) Marcador de masa molecular de proteínas; carril 2, (P-) pastilla proteica obtenida de la bacteria sin inducir; carril 3, (EC) extracto proteico del lisado de *E. coli* Rosetta gami B inducido; carril 4, (FT) flujo de extracto inducido pasado por la columna; carril 5-9, (L0, L10, L50, L200 y L250) eluatos con gradiente de imidazol (10, 50, 200, 250 mM); carril 9, (CI) proteína solubilizada de cuerpos de inclusión.

La expresión de PfAvr4 se realizó con dos propósitos. Por un lado, para estudiar la respuesta molecular del hospedero ante su infiltración, pues hasta el momento no ha sido caracterizada la respuesta de la planta, y por otro lado para usar el proceso de Muñoz-Pérez, (2018) como referencia, es decir, para corroborar si se estaba realizando adecuadamente la expresión de las proteínas; por ello se le llamó “control de la expresión”. Las fracciones de proteína se fraccionaron en gel SDS-PAGE 15% y en la Figura 5.8 se observó la banda de proteína esperada para PfAvr4 (-SP), con peso molecular de ~13.83 kDa. Observándose la proteína en los carriles L200 y L250, fracciones eluidas con 200mM y 250 mM de imidazol (indicadas con el símbolo “*” en color negro).

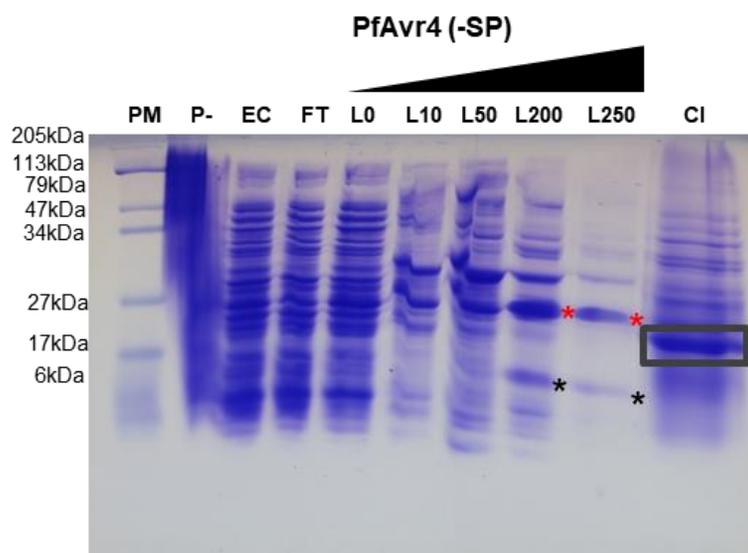


Figura 5.8 Análisis electroforético de las muestras colectadas en la purificación de la proteína recombinante PfAvr4 (-SP) en gel de 15% SDS-PAGE. Se cargó una cantidad de proteína de 12 μ g. Carril 1, (PM) Marcador de masa molecular de proteínas; carril 2, (P-) pastilla proteica obtenida de la bacteria sin inducir; carril 3, (EC) extracto proteico del lisado de *E. coli* Rosetta gami B inducido; carril 4, (FT) flujo de extracto inducido pasado por la columna; carril 5-9, (L0, L10, L50, L200 y L250) eluatos con gradiente de imidazol (10, 50, 200, 250 mM); carril 9, (CI) proteína solubilizada de cuerpos de inclusión.

En los carriles L200 y L250, se observan múltiples bandas entre las que está una de \sim 27 kDa, indicada en el reporte de Muñoz-Pérez, 2018 (fracciones eluidas con 200 y 250 mM de imidazol indicadas con el símbolo "*" en color rojo). Se pudo observar también en la proteína solubilizada de los cuerpos de inclusión, la presencia de una banda que podría ser PfAvr4 (-SP) con un peso arriba de \sim 17 kDa (indicado en rectángulo color negro).

Durante el proceso de la purificación del PfAvr4 (-SP), se observó una banda con un peso de \sim 28 kDa (Figura 5.8), se plantea que este péptido es un producto de la dimerización de los monómeros de PfAvr4 (-SP), similar a lo reportado por Muñoz-Pérez, (2018), quien también reportó que, a pesar de realizar los ensayos de reducción y alquilación, no se logró tener la disociación de la proteína. Es posible que en el presente trabajo, con el proceso de resolubilización pudo resuspenderse el monómero de PfAvr4, aunque faltaría corroborarse por secuenciación.

A continuación, el Cuadro 5.4 presenta la proteína total recuperada en las fracciones obtenidas durante la purificación de los efectores PfEcp2 (+SP), PfEcp2 (-SP) y PfAvr4 (-SP) expresados en *E. coli* Rosetta gami B.

Cuadro 5.4 Proteína total obtenida para los efectores PfEcp2 (+SP), PfEcp2 (-SP) y PfAvr4 (-SP).

Efactor	*Proteína total (µg)
PfEcp2 (+SP)	
FT	6085
L	375
L10	175
L50	335
L200	185
L250	390
PfEcp2 (-SP)	
FT	4580
L	790
L10	210
L50	420
L200	345
L250	12
PfAvr4 (-SP)	
FT	7080
L	1110
L10	282
L50	380
L200	530
L250	375

*Proteína total calculada de las diferentes fracciones obtenidas de la columna de purificación de la fracción soluble de las pastillas inducidas.

En la caracterización de efectores, se han usado diferentes cantidades de proteína en los análisis funcionales. Para los análisis funcionales, en el trabajo de Muñoz-Pérez, (2018) usaron 150 y 300 ng de la proteína PfAvr4, infiltrados en hojas del banano resistente Calcuta 4. Por otra parte, See *et al.*, (2019) usaron diferentes concentraciones de proteína a evaluar, siendo 25, 50, 100, 200 y 500 ng para los efectores ToxA y ToxB infiltrados en hoja de trigo, resultando que 200 ng

de ToxA y 500ng de ToxB son requeridos para inducir una respuesta en las hojas de trigo; mientras que Assis *et al.*, (2017), en el estudio de siete efectores de bacterias de la familia Xanthomonadaceae usaron 90µg de proteína (0.3µg/µl) de cada uno para su infiltración en plantas de *N. benthamiana*.

Con base en los resultados de cuantificación de proteína, se obtuvo proteína suficiente para los análisis funcionales futuros de los efectores PfEcp2 y PfAvr4.

Para tener elementos para facilitar la interpretación de los resultados, se analizó la proteína obtenida de la fracción soluble de la cepa *E. coli* Rosetta gami B sin transformar e inducida con IPTG (1mM). En la Figura 5.9 se observan las proteínas eluidas en los diferentes lavados con imidazol. En la figura se señalan las zonas del gel donde se muestra una mayor abundancia en los péptidos de un peso aproximado de ~17 kDa, ~30 kDa y ~113 kDa respectivamente (Figura 5.9).

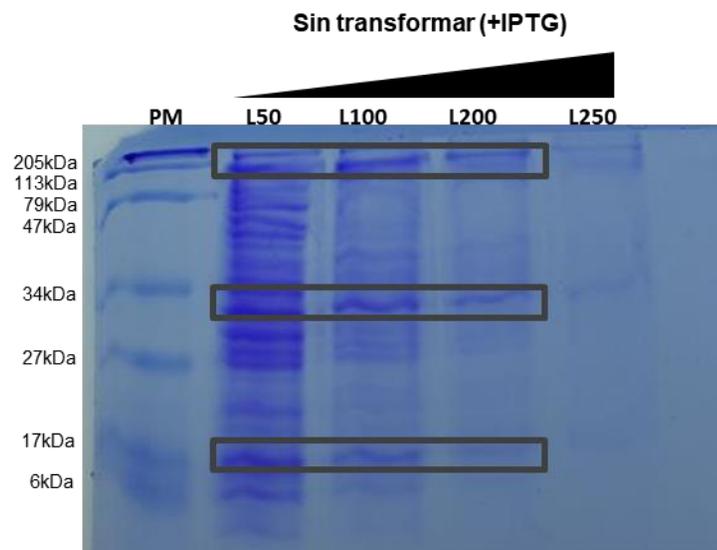


Figura 5.9 Perfil de la fracción soluble de *E. coli* Rosetta gami B sin transformar inducida con IPTG de 12% SDS-PAGE. Se cargó una cantidad de proteína de ~12µg de cada muestra. El gel fue teñido con Coomassie Blue. Carril 1, (PM) Marcador de masa molecular de proteínas; carril 2: fracción obtenida con 50mM de imidazol (L50); carril 3: fracción obtenida con 100mM de imidazol (L100); carril 4: fracción obtenida con 200mM de imidazol (L200) y carril 5: fracción obtenida con 250mM de imidazol (L250).

Sin embargo, los resultados observados muestran que la mayoría de las proteínas presentes en la fracción soluble del cultivo son eluidas con una concentración de 50 mM de imidazol, mientras

que en el perfil de proteínas obtenido de las fracciones con una concentración de 200 mM y 250 mM de imidazol se observa poca cantidad de proteína eluída. Esto apoya que las proteínas señaladas en la Figura 5.7 y Figura 5.8, corresponden a las proteínas efectoras PfEcp2 y PfAvr4, sin embargo, de la Figura 5.7 es necesario realizar análisis por secuenciación para confirmar que efectivamente sea la proteína PfEcp2.

Algunas proteínas de *E. coli* se unen a resinas quelantes de metales, debido a su contenido natural de histidinas. Eso las hace unirse a la resina junto con las proteínas recombinantes de interés etiquetadas con poli-histidina; esto se ha reportado en otros trabajos, siendo relativamente común (Mokhonov *et al.*, 2018; Bolanos-García y Daves, 2006). La cantidad de proteínas de *E. coli* “contaminantes” que se unen a la resina de metal a través de sus residuos de histidina expuestos en la superficie con numerosas, entre ellas se han identificado la *Glucosamine-6-phosphate synthase* (GlmS, 66.8 kDa), la *Chloramphenicol-O-acetyl transferase* (CAT, 25.5 kDa), la *cAMP-regulatory protein* (CRP, 23.6 kDa), *Peptidoylproline cis–trans isomerase* (SlyD, 20.8 kDa), *Host factor-I protein* (Hfq, 11.1 kDa) y *Regulatory ribosomal protein* (S15, 10.2 kDa) (Bolanos-García y Daves, 2006). Esto explica que proteínas de la cepa hospedera puedan co-purificar con la proteína de interés, en este caso PfEcp2 y PfAvr4, como se observa en la Figura 5.9. En el trabajo de Muñoz-Pérez, (2018) secuenciaron la banda de la proteína PfAvr4, y reportaron que también se co-purificó con la proteína de *E. coli* *Peptidylproline cis–trans isomerase* (SlyD) de 20.8 kDa. En nuestro caso, hacen falta secuenciar las fracciones obtenidas del efector PfEcp2. Así como también, de las proteínas solubilizadas de los cuerpos de inclusión tanto para PfEcp2 y PfAvr4 para evaluar si están homogéneas o existe la presencia de proteínas nativas de *E. coli* en estas fracciones.

5.4 DISCUSIÓN

En la actualidad, la expresión heteróloga se ha utilizado para producir diversas proteínas como lo son las celulasas, lipasas, complejos de coagulación, entre otros (Contesini *et al.*, 2020; Vieira *et al.*, 2018; Fathi-Roudsari *et al.*, 2016; Lambertz *et al.*, 2014; Rosano *et al.*, 2014). Sin embargo, también se ha empleado el uso de esta tecnología para producir proteínas relacionadas a la patogénesis.

Existen diversos reportes donde aplican la expresión heteróloga de proteínas efectoras para el análisis estructural y bioquímico, y obtener información del patosistema.

En este capítulo, se realizó la expresión heteróloga del efector PfEcp2 (+SP) y PfEcp2 (-SP), así como PfAvr4 (-SP) en *E. coli* Rosetta gami B. Se sabe, que estos efectores, ortólogos funcionales de los efectores de *C. fulvum* CfAvr4 y CfEcp2 (Stergiopoulos *et al.*, 2010), han sido expresados de forma heteróloga y se han caracterizado de manera parcial. Numerosos trabajos han tratado de obtener mayor información del patosistema *P. fijiensis-Musa spp.* Para el caso de PfAvr4, Arango *et al.*, (2016) la expresaron en *P. pastoris* y la infiltración produjo respuesta hipersensible en la variedad Calcuta IV. El análisis estructural reportada por Kohler *et al.*, (2016) mostró que el CfAvr4 de *C. fulvum* y PfAvr4 de *P. fuligena* comparten especificidad funcional en la unión de carbohidratos y ambos efectores confieren protección contra las quitinasas.

Existen ensayos de precipitación, co-expresión transitoria, sobreexpresión de genes, expresión de proteínas de fusión, así como numerosos análisis *in silico* que se enfocan en el entendimiento de estos efectores (Onyilo *et al.*, 2018; Arango *et al.*, 2016; Kohler *et al.*, 2016; Mesarich *et al.*, 2016; Stergiopoulos *et al.*, 2010)

En el presente capítulo se realizó la expresión heteróloga de las proteínas PfEcp2 (+SP), PfEcp2 (-SP) y PfAvr4 (-SP). Al evaluar los perfiles de proteínas obtenidas de la fracción soluble e insoluble de cada experimento, se observaron las presuntas bandas esperadas para PfEcp2 (+SP), PfEcp2 (-SP) y PfAvr4 (-SP). El sistema de expresión empleado fue el de Rosetta gami B, la cual ha tenido reportes de uso exitoso para la expresión de la proteína efectora CfAvr4 de *Cladosporium fulvum* (Hurlburt *et al.*, 2018), así como la expresión de PfAvr4 (+SP) y PfAvr4 (-SP), ésta última empleada como control de expresión en este capítulo (Muñoz-Pérez, 2018).

La expresión del efector PfAvr4 (-SP) como control reprodujo lo reportado por Muñoz-Pérez, (2018), lo que indica que la metodología se empleó de manera adecuada.

En los análisis de expresión de PfEcp2 (+SP) y PfEcp2 (-SP), se observaron bandas que coinciden con el rango aproximado de ~27 kDa, y se observan en los carriles varias otras proteínas. Resultados similares se observan en el reporte de See *et al.*, (2019), que expresaron y purificaron a ToxA y ToxB (13.2 kDa y 6.5 kDa respectivamente), y en los geles se observan bandas por arriba del peso mencionado, tanto en la expresión en *E. coli* SHuffle como en *E. coli* BL21 (DE3).

Tan *et al.* (2012), mencionaron la recuperación de la proteína ToxA a partir de la proteína obtenida en cuerpos de inclusión (CI). Es posible que la problemática se deba a que las proteínas son ricas en residuos de cisteínas, y no estén teniendo un correcto plegamiento estructural y

sean dirigidas para su degradación al no pasar el control de calidad celular (Moghadam *et al.*, 2015; Skaar *et al.*, 2015; Tyo *et al.*, 2012). Por lo tanto, se decidió solubilizar los cuerpos de inclusión que pudieran generarse, observándose que una cantidad importante de proteína está en esa fracción. Probablemente, este perfil de proteína sea el resultado de la dimerización de los monómeros del PfAvr4 o por mal plegamiento que se está llevando a cabo en el sistema. También se realizó la solubilización de los cuerpos de inclusión de PfEcp2 (+SP) y PfEcp2 (-SP). Sin embargo, se observa una cantidad igual o menor que lo recuperado en las eluciones con 200mM y 250mM de imidazol, al contrario de lo que se observó en la purificación de PfAvr4, en la que se observó más de la proteína de interés en los cuerpos de inclusión en comparación con lo eluido con 200nM y 250nM de imidazol. Esta diferencia en el patrón de purificación de estos dos efectores se ha observado también en otros trabajos. Guillen *et al.*, (2019), expresaron 11 proteínas efectoras de *Melampsora larici populina*, y encontraron sólo tres de ellas en la fracción soluble, aunque evaluaron diferentes cepas de *E. coli* y protocolos de inducción. Algunos de los otros efectores candidatos no se expresaron mientras otros se expresaron y se fueron a cuerpos de inclusión. Esto evidencia que la expresión heteróloga de efectores es un área que está aún en infancia y que quedan muchos retos técnicos por resolver. En el presente trabajo, se obtuvo un bajo rendimiento en la expresión de las proteínas efectoras PfAvr4 y PfEcp2 de *P. fijiensis*; en el caso de PfAvr4 la cantidad de proteína en los cuerpos de inclusión fue abundante, y aparentemente se recuperó como monómero y como dímero, aunque falta el análisis de secuenciación para la confirmación de los resultados.

Aunque la producción de proteínas recombinantes en *E. coli* ha sido una estrategia útil para realizar análisis biofísicos y bioquímicos de efectores candidatos, la expresión de efectores no ha sido siempre fácil posiblemente, porque muchas de ellas son proteínas pequeñas ricas en residuos de Cys. Es posible purificar proteínas recombinantes de cuerpos de inclusión mediante el uso de condiciones de extracción desnaturalizantes y realizar renaturalización. Sin embargo, no es lo más recomendable, ya que el replegamiento de las proteínas *in vitro* puede ser diferente del plegamiento nativo (de Guillen *et al.*, 2019). Otra limitante de los sistemas procariontes para producir proteínas eucariotas es la falta de modificaciones postraduccionales como las metilaciones y glucosilaciones (Gifre-Renom *et al.*, 2020; Huang *et al.*, 2020; de Guillen *et al.*, 2019; Singh *et al.*, 2015).

Se sabe que las proteínas del hospedero que poseen residuos de histidina pueden unirse a la resina de la misma forma que la proteína de interés que está siendo expresada. Estos resultados ya han sido reportados por Mokhonov *et al.*, (2018) y Bolanos-García y Daves, (2006) que

mencionan que las proteínas Glms, CAT, CRP, SlyD, Hfq y S15 son las más comunes en co-purificar en las columnas de resina con níquel. Tal es el caso reportado por Muñoz-Pérez, (2018) donde al secuenciar la banda correspondiente a PfAvr4, confirmó su identidad, pero también la co-purificación con la proteína SlyD de *E. coli* de tamaño aproximado de 20.8 kDa. Bolanos-García y Daves, (2006) reportaron la clasificación de las proteínas de *E. coli* que co-purifican de acuerdo a las diferentes concentraciones de imidazol, siendo 3 clases principales, Clase I comprendido por proteínas que requieren ≥ 80 mM de imidazol (Fur, Crp, SlyD, ArgE, Cu/Zn-SODM y YodA), Clase II proteínas que requieren de 55 a 80 mM de imidazol (Glms, ODO2, YadF, CAT, GlgA, YfbG y G6-PD) y Clase III comprendido por proteínas que requieren una concentración baja de imidazol, que va entre 30 y 50 mM (Hsp60 y ODO1).

Sin embargo, del presente trabajo falta la secuenciación de las fracciones obtenidas del efector PfEcp2 para la confirmar su expresión. Así como también, secuenciar las proteínas obtenidas de la solubilización de los cuerpos de inclusión tanto para PfEcp2 y PfAvr4, ya que esta metodología de re-solubilización fue aplicada por primera vez para estos efectores. También es importante evaluar si las proteínas renaturalizadas de los cuerpos de inclusión son capaces de desencadenar una respuesta hipersensible. De ser el caso, este procedimiento podría usarse para estudiar estos y otros efectores, ya que permitirá obtener un mejor rendimiento de la proteína de interés, en comparación con la recuperación de los extractos solubles.

Cabe resaltar, como se mencionó en la introducción de este capítulo, que el interés inicial fue la comparación de la función de PfAvr4 y PfEcp2 expresados en sistema procariota vs. sistema eucariota, por lo tanto, se decidió por el sistema de *Pichia pastoris*, usando el vector pPICZ, sin embargo, por la situación de la pandemia, estos experimentos no fueron posibles de concluir. Se confirmaron las clonas transformantes de las células de *P. pastoris* cepa GS115, con los constructos denominados pPICZ-PfAvr4 (+SP), pPICZ-PfAvr4 (-SP), pPICZ-PfEcp2 (+SP), pPICZ-PfEcp2 (-SP). Las clonas se encuentran almacenadas a -20°C para sus posteriores análisis.

La metodología empleada en este capítulo podrá usarse para expresar los demás genes seleccionados para su análisis funcional.

CAPÍTULO VI

DISCUSIÓN GENERAL

Se sabe que los fitopatógenos poseen una batería de moléculas llamadas “efectoras”, generalmente de naturaleza proteica, importantes para iniciar la enfermedad y evitar que el huésped desarrolle una respuesta de defensa, ocultando al patógeno o matando directamente a la célula hospedera (Jones *et al.*, 2018).

La predicción computacional de efectores en hongos es un desafío por las escasas características comunes entre ellos y el bajo número aún de efectores caracterizados. Estas moléculas muestran un bajo número de características determinantes, las cuales van desde la longitud de la proteína, presencia de un péptido señal de secreción, ≥ 4 residuos de cisteínas y/o ausencia de TMD (Dörfors *et al.*, 2019; Jones *et al.*, 2018; Gabriëls *et al.*, 2016; Sonah *et al.*, 2016; Bowen *et al.*, 2009).

Existen numerosos reportes donde se han dedicado a la identificación y caracterización *in silico* de efectores en proteomas fúngicos con ayuda de varias herramientas bioinformáticas (Cortázar *et al.*, 2016; Sonah *et al.*, 2016; Petersen *et al.*, 2011; Pierleoni *et al.*, 2008; Horton *et al.*, 2007; Kall *et al.*, 2007; Krogh *et al.*, 2001). En los últimos cinco años se ha tenido un gran auge en el desarrollo y establecimiento de nuevas herramientas bioinformáticas para la predicción e identificación de efectores en proteomas de hongos y oomicetos, que van desde pipelines bioinformáticos hasta algoritmos de aprendizaje de máquina (Carreón-Anguiano *et al.*, 2020; Wang *et al.*, 2020; Tabima y Grünwald, 2019; Sperschneider *et al.*, 2018, 2015). En muchos casos, estos análisis han revelado que los genomas fúngicos contienen una gran cantidad de efectores que evolucionan rápidamente en comparación con el resto del genoma. Esta rápida evolución puede deberse a la fuerte presión de selección ejercida por el sistema de reconocimiento de la planta y la adaptación específica del hospedero frente a diferentes especies patogénicas.

Por lo tanto, para realizar avances en la investigación de efectores, en el presente trabajo se reporta el establecimiento de un pipeline bioinformático para la identificación de proteínas efectoras de las bases de datos de proteínas fúngicas. La generación y estructuración del pipeline bioinformático EffHunter, corresponde a una herramienta generada para la identificación de efectores en proteomas fúngicos, integrado por los programas SignalP 4.1 (Petersen *et al.*, 2011), Phobius (Kall *et al.*, 2007), TMHMM 2.0 (Krogh *et al.*, 2001) y WoLFPSORT (Horton *et al.*,

2007), pipeline estructurado con un conjunto de scripts escritos en lenguaje Perl para filtrar proteínas por longitud y por contenido de cisteínas. Los análisis realizados con EffHunter demostraron que es un pipeline bioinformático amigable y accesible, robusto para identificar efectores en proteomas de hongos, ya que muestra una mayor precisión y menos falsos positivos que otros predictores de efectores. El número de efectores predichos fue muy variante en las diferentes especies de hongos analizadas. Aunque se observaron excepciones, existen promedios en el número de efectores en cada tipo de hongo. Entre ellos, se identificó que los hongos patógenos y entomopatógenos de plantas fueron aquellos que mostraron tener un mayor número de efectores. A su vez, en los análisis y comparación de los hongos con diferente estilo de vida, resultó que en los hongos biotróficos y hemibiotróficos se predicen efectomas más grandes que en los hongos necrotrofos. EffHunter es un pipeline bioinformático muy astringente, todos los candidatos que recluta son canónicos y pueden confirmarse en análisis manuales (Carreón-Anguiano *et al.*, 2020).

El pipeline EffHunter fue usado para la predicción de los efectores canónicos de *P. fijiensis*, resultando en la identificación de 240 candidatos. En los análisis de verificación manuales de los resultados no se encontraron proteínas en las que se identifique *in silico* la presencia de regiones transmembranales, concluyendo que la lista cumple con todos los criterios que se usaron para la búsqueda (sugiriendo que no hubo escapes). La búsqueda se repitió al menos tres veces de manera independiente para la identificación reproducible y confiable. Cabe mencionar que el número de efectores predichos depende de los criterios tomados para la búsqueda, puesto que difícilmente se unifican; y eso puede provocar que las bases de datos finales sean diferentes, por ejemplo, Sonah *et al.*, (2016) reportaron 528 efectores para *Magnaporthe oryzae*, mientras que EffHunter identificó 273 (Carreón-Anguiano *et al.*, 2020). En la actualidad la diversidad de criterios implementados para la búsqueda de efectores, aumenta la probabilidad de identificar efectores verdaderos.

Durante mucho tiempo no se encontró similitud entre los efectores fúngicos y las secuencias caracterizadas en otros organismos (Hacquard *et al.*, 2011). Sin embargo, el 75.41% de los 240 efectores canónicos de *P. fijiensis* tienen homólogos en diferentes hongos, de los cuales el 97.8% de ellos tienen evidencia de expresión (Noar y Daub, 2016). Actualmente ya se han identificado efectores de algunos hongos que tienen homología con proteínas de otros hongos de diferente filo, por ejemplo, en el hongo *Rhizophagus clarus* el 95% de los candidatos identificados tienen homólogos en otros hongos en el filo de Glomeromycota, y el 50% de los efectores de *Sclerotinia*

sclerotiorum muestran conservación a través de ascomicetos (Sędziewska-Toro y Brachmann, 2016; Guyon *et al.*, 2014).

En comparación con las diversas bases de efectores de *P. fijiensis* reportadas por otros autores, se ha encontrado variabilidad en el número de candidatos identificados y escapes en las predicciones, ya que en el trabajo reportado por Chang *et al.*, 2016, reportaron 105 proteínas efectoras en *P. fijiensis*, pero al hacer la comparación con nuestros resultados se pudo detectar variabilidad en los resultados publicados (Carreón-Anguiano *et al.*, 2020). Estos autores reportaron 105 efectores que tienen ≤ 200 aminoácidos y son presuntamente secretados, pero 11 presentan al menos un dominio transmembranal, que evidencian que la predicción tuvo errores. Éste grupo realizó la búsqueda con base en las características generales, como tamaño (≤ 300), presencia de péptido señal de secreción y presencia de al menos 4 cisteínas. Sin embargo, un análisis manual de sus resultados, mostró que el 40% de su base no cumplían con los criterios que establecieron para la búsqueda Noar y Daub, (2016), reportaron un total de 231 efectores identificados en la interacción de *P. fijiensis* con *Musa spp*, sin embargo, la base completa no se encuentra disponible, sólo un conjunto de 40 proteínas que fueron las que tuvieron mayor expresión en el análisis transcriptómico de la asociación de *P. fijiensis* con *M. acuminata* (base conformada por 10 proteínas identificadas como efectores en hongo crecido *in vitro* y 30 proteínas altamente expresadas en interacción), resultando que de las 40 proteínas, 31 se compartían con nuestra base (9 efectores identificados en hongo *in vitro* y 22 efectores identificados en interacción). Sin embargo, al realizar un análisis manual de sus candidatos de efectores, se identificaron dos proteínas que no cumplían con el criterio de tener ≥ 300 aminoácidos de longitud (una proteína de cada condición). Cabe mencionar que todas estas bases de datos de efectores de *P. fijiensis* cumplen con el criterio de distribución discontinua (Carreón-Anguiano *et al.*, 2020; Chang *et al.*, 2016; Noar y Daub, 2016).

Para ampliar la predicción de efectores de *P. fijiensis*, se decidió buscar efectores por medio de la predicción de dominios funcionales, para identificar a los efectores no canónicos. Esta metodología ha sido utilizada en otros reportes (Gong *et al.*, 2020; Sonah *et al.*, 2016; Levin *et al.*, 2017; Sperschneider *et al.*, 2015; Morais do Amaral *et al.*, 2012). Levin *et al.*, 2017 reportaron la identificación de 18 efectores con dominio LysM en el hongo *Penicillium expansum*, usando las bases del InterproScan (Jones *et al.*, 2014). En el caso de *Hyaloperonospora arabidopsidis*, *Phytophthora ramorum* y *Phytophthora sojae* se identificaron 63, 282 y 285 efectores con búsqueda de motivos RXLR (Ai *et al.*, 2019). Sin embargo, existen efectores no canónicos que

no tienen dominios funcionales conocidos, por lo que deben de haber efectores que escapan de ser reconocidos.

En el presente trabajo se realizó la predicción del efectoroma completo de *P. fijiensis* compuesto por 662 proteínas, reportando una base de efectores canónicos de 240 proteínas y otra base de datos de efectores no canónicos con 422 proteínas. Dado a lo anterior, en este trabajo se seleccionaron 12 candidatos de la base de datos de efectores canónicos, cabe mencionar que no se seleccionaron de la base de datos de efectores no canónicos ya que no se podía esperar hasta que se tuviera el efectoroma terminado, por ello se tuvo que comenzar con la base de efectores canónicos, que se identificó mucho tiempo antes que los efectores no canónicos.

Los candidatos de efectores seleccionados fueron el PfID210733, PfID191819, PfID77734, PfID181415, PfID212258, PfID39450, PfID60298, PfID85220, PfID212087, PfID70410, PfID80609, PfID189452. Sin embargo los únicos que se clonaron fueron los candidatos PfID210733, PfID77734, PfID212258, PfID85220, PfID212087, PfID80609 y PfID189452, sin embargo, cinco (PfID191819, PfID181415, PfID39450, PfID60298 e PfID70410) no pudieron ser clonados. En el análisis de la expresión de los candidatos, el candidato PfID191819 muestra amplificación débil en el pool 1, que comprende los días 3dpi, 9dpi y 10dpi. Esto llama la atención ya que en el trabajo de Noar y Daub (2016), lo identifican durante la fase necrotrófica, mientras que en éste trabajo, se encontró expresado dentro de la fase biotrófica en la interacción *P. fijiensis* – *Musa spp.* Siendo la posibilidad de la existencia de efectores que estén participando de manera latente en las diferentes etapas de la infección.

Así mismo, se tenía el objetivo en este trabajo de analizar mediante Agroinfiltración las construcciones de los candidatos de efectores canónicos de *P. fijiensis*, aunque desafortunadamente por la pandemia provocada por SARS-CoV-2 (COVID-19), no se pudo concluir, se avanzó en clonar los candidatos en el vector pDONR221, vector de entrada Gateway, para que en el grupo se continúe trabajando para realizar la Agroinfiltración (*Agrobacterium transient transformation assay*, ATTA) de los candidatos en hojas de *N. benthamiana*. Cabe mencionar que las plantas ya se tenían. Recientemente, el M.C. Bartolomé Chí Manzanero reprodujo la metodología de ATTA, usando un constructo que contiene la región codificante de la proteína verde fluorescente (GFP), logrando expresar la proteína en las hojas de *N. benthamiana* (datos no publicados). Con ello, ya se tiene la plataforma metodológica para que el equipo de investigación realice los análisis funcionales de los candidatos una vez que la contingencia sanitaria termine.

Las interacciones huésped-patógeno se han estudiado intensamente durante las últimas décadas (Jaswal *et al.*, 2020; Sperschneider *et al.*, 2015). En *Pseudocercospora fijiensis* se conocen a los efectores PfAvr4, PfEcp2 y PfEcp6 como ortólogos funcionales de efectores de *C. fulvum*, los cuales han sido utilizados en diversos análisis de expresión (Stergiopoulos *et al.*, 2010). Arango *et al.*, (2016), reportaron la expresión heteróloga en células de *P. pastoris*, infiltrando proteína en plantas de *Musa spp*, lo que provocó respuesta de hipersensibilidad.

Por su parte Muñoz-Pérez, (2018), reportó la expresión del efector PfAvr4 de *P. fijiensis*, obteniendo la proteína en su versión inmadura (con péptido señal) y madura (sin péptido señal), proteínas confirmadas por secuenciación. En el presente trabajo se realizó la expresión heteróloga de las proteínas efectoras PfEcp2 y PfAvr4, en su versión madura (sin péptido señal) o inmadura (con péptido señal); por tal motivo, para los análisis de expresión, se reprodujo la expresión del efector PfAvr4 (-SP) como control de experimento, de acuerdo a lo reportado por Muñoz-Pérez, (2018). En los análisis de expresión heteróloga de la proteína PfEcp2 (+SP) y PfEcp2 (-SP), se observaron bandas que coinciden con el rango esperado de ~27 kDa. Sin embargo, se observan varias bandas de proteína en los carriles. Éstos resultados fueron similares a lo reportado por See *et al.*, (2019), quienes expresaron y purificaron a los efectores ToxA y ToxB (13.2 kDa y 6.5 kDa respectivamente) de *Pyrenophora tritici-repentis*, y en los geles de SDS-PAGE se observan bandas por arriba de la masa molecular esperada. Resultados similares obtuvieron tanto en la expresión en *E. coli* SHuffle como en *E. coli* BL21 (DE3). Éstos resultados sugieren que las proteínas ricas en residuos de Cys no estén siguiendo un plegamiento estructural correcto, por lo que son dirigidas para su degradación (Tyo *et al.*, 2012; Moghadam *et al.*, 2015; Skaar *et al.*, 2015). Por lo anterior, en el presente trabajo se optó por la solubilización de los cuerpos de inclusión de las células inducidas de *E. coli* Rosetta gami B para recuperar las proteínas PfAvr4 (-SP), PfEcp2 (+SP) y PfEcp2 (-SP). Estos resultados son nuevos, pues en los reportes anteriores no se habían solubilizado los cuerpos de inclusión, lo que permitió descubrir que la mayoría de la proteína se dirige a los cuerpos de inclusión. Será importante evaluar si una vez renaturalizadas, estas proteínas son capaces de inducir una respuesta hipersensible en la planta hospedera resistente. En caso de tener resultados positivos, representará una oportunidad para obtener más proteína, mejorando el rendimiento de la proteína de interés. Cabe mencionar que falta corroborar mediante secuenciación de las bandas proteicas promisorias, así como realizar los análisis funcionales inoculando las proteínas en hojas de banano resistente cv Calcutta 4.

En efecto, se tienen listos en nuestro grupo de investigación los efectores clonados en los vectores de expresión pLATE y pDONR221, para que el grupo pueda continuar los análisis pendientes, que por medidas institucionales de restricción de entrada por la contingencia sanitaria, no se terminaron en la presente tesis. Sin embargo, contribuí de manera intelectual con el diseño experimental y con el material biológico usado en las tesis de dos estudiantes de licenciatura del grupo, y en conjunto con el M.C. Bartolomé Chí Manzanero, en la tesis de la estudiante de Lic. Mercedes Rodríguez (tesis en proceso) se logró la agroinfiltración y expresión transitoria de la proteína GFP en hojas de *N. benthamiana*. Mientras que en la tesis de la Ing. Celina Ayil Naal se realizaron análisis por qRT-PCR para evaluar la expresión de efectores que se predijeron para la etapa biotrófica. Los resultados de la Ing. Ayil apoyan que los análisis predictivos desarrollados en la presente tesis de Doctorado son correctos (Ayil-Naal, 2021), y sin ser resultados directos de la presente tesis, sí son contribuciones de mi trabajo de Doctorado para el avance del grupo de investigación y finalizan el proceso que se planteó establecer en la presente tesis.

CAPÍTULO VII

CONCLUSIONES Y PERSPECTIVAS GENERALES

7.1 CONCLUSIONES

Se logró establecer un pipeline que permite la identificación masiva de proteínas efectoras a partir de las bases de datos de proteínas en hongos (análisis de minería de datos), el cual está registrado ante el INDAUTOR (ANEXO) y fue publicado en la revista *Biomolecules* del MDPI (Capítulo 2).

Los análisis de la predicción de proteínas efectoras mostraron un efectoroma amplio de *P. fijiensis*. Los resultados predictivos con EffHunter permitió obtener una base de efectores canónicos de 240 proteínas, efectoroma que se amplió después con 422 proteínas y que se predicen como efectores no canónicos.

Se identificó que el candidato de efector PfID77734 tiene dominio PAN, observándose homología estructural con el dominio PAN de la proteína *Microneme protein 4* de *Toxoplasma gondii*.

Se expresó de manera heteróloga a las proteínas efectoras PfAvr4 y PfEcp2 de *P. fijiensis*, PfEcp2 en su versión madura (sin péptido señal) e inmadura (con péptido señal), y PfAvr4 en su versión madura en *E. coli* Rosetta gami B. La confirmación de PfAvr4 fue reportada por Muñoz-Pérez (2018), mientras que para PfEcp2 aún queda pendiente la secuenciación de la proteína recuperada.

Los resultados sugieren que gran parte de las proteínas efectoras expresadas en *E. coli* se van a cuerpos de inclusión. Será interesante evaluar este punto en *Pichia pastoris*. Así como comparar la efectividad de las proteínas expresadas en uno u otros sistemas para inducir una respuesta en el hospedero.

7.2 PERSPECTIVAS

Con el establecimiento del pipeline bioinformático EffHunter para la identificación de efectores canónicos, se espera realizar una aplicación de manera interactiva, la cual se espera ser más amigable con el usuario.

En este trabajo se generaron dos listas de efectores de *P. fijiensis*, una de proteínas de candidatos de efectores canónicos y no canónicos. Por ende, quedan preguntas aún sin responder, ¿son novedosos estos nuevos candidatos? ¿Serán capaces de inducir RH? o ¿podrán evadir la respuesta hipersensible para desencadenar el desarrollo de la enfermedad?

Los candidatos PfID210733, PfID77734, PfID212258, PfID85220, PfID212087, PfID80609 e PfID189452 se encuentran clonados en los diferentes vectores de expresión, lo que permitirá realizar los análisis funcionales para profundizar en el entendimiento del patosistema *P. fijiensis*-*Musa* spp.

Los efectores que se encuentren indispensables para la patogenicidad y virulencia de *P. fijiensis*, podrán ser nuevos blancos moleculares para el diseño de estrategias de control de la Sigatoka negra.

ANEXOS

CERTIFICADO

Registro Público del Derecho de Autor

Para los efectos de los artículos 13, 162, 163 fracción I, 164 fracción I, 168, 169, 209 fracción III y demás relativos de la Ley Federal del Derecho de Autor, se hace constar que la **OBRA** cuyas especificaciones aparecen a continuación, ha quedado inscrita en el Registro Público del Derecho de Autor, con los siguientes datos:

AUTORES: CANTO CANCHE BLONDY BEATRIZ
CARREÓN ANGUIANO KARLA GISEL

TÍTULO: EFFHUNTER VERSION 1.0

RAMA: PROGRAMAS DE COMPUTACION

TITULAR: CENTRO DE INVESTIGACION CIENTIFICA DE YUCATAN, A.C. (CON FUNDAMENTO EN EL ARTICULO 83 DE LA L.F.D.A.)

Con fundamento en lo establecido por el artículo 168 de la Ley Federal del Derecho de Autor, las inscripciones en el registro establecen la presunción de ser ciertos los hechos y actos que en ellas consten, salvo prueba en contrario. Toda inscripción deja a salvo los derechos de terceros. Si surge controversia, los efectos de la inscripción quedarán suspendidos en tanto se pronuncie resolución firme por autoridad competente.

Con fundamento en los artículos 2, 208, 209 fracción III y 211 de la Ley Federal del Derecho de Autor; artículos 64, 103 fracción IV y 104 del Reglamento de la Ley Federal del Derecho de Autor; artículos 1, 3 fracción I, 4, 6 fracción I y 9 del Reglamento Interior del Instituto Nacional del Derecho de Autor, se expide el presente certificado.

Número de Registro: 03-2019-101809310300-01

Ciudad de México, a 23 de octubre de 2019

EL DIRECTOR DEL REGISTRO PÚBLICO DEL DERECHO DE AUTOR

JESUS PARETS GOMEZ



SECRETARÍA DE CULTURA
INSTITUTO NACIONAL DEL
DERECHO DE AUTOR
REGISTRO PÚBLICO DEL
DERECHO DE AUTOR

BIBLIOGRAFÍA

*Números en “[]” representan a las referencias usadas en el artículo publicado en *Biomoleculas*.

- Ayil-Naal, C. (2021). Análisis de expresión de posibles efectores durante la fase biotrófica de la Sigatoka negra. Tesis de Licenciatura en opción al título de Ingeniería en Bioquímica. Instituto Tecnológico Superior de Calkiní (ITESCAM, Campeche). 19 de enero 2021. 95pag.
- Ai, G., Yang, K., Ye, W., Tian, Y., Du, Y., Zhu, H., Li, T., Xia, Q., Shen, D., Peng, H., Jing, M., Xia, A. y Dou, D. (2020). Prediction and characterization of RXLR effectors in *Pythium* species. *Molecular Plant-Microbe Interactions*, 33(8), 1046–1058.
- Arroyo-Velez, N., González-Fuente, M., Peeters, N., Lauber, E. y Noël, L.D. (2020). From effectors to effectoromes: Are functional studies of individual effectors enough to decipher plant pathogen infectious strategies?. *PLoS pathogens*, 16(12), e1009059.
- Almeida, D.S., Menezes-Blackburn, D., Zhang, H., Haygarth, P.M. y Rosolem, C.A. (2019). Phosphorus availability and dynamics in soil affected by long-term ruzigrass cover crop. *Geoderma*, 337, 434–443 ^[63].
- Armitage, A.D., Taylor, A., Sobczyk, M.K., Baxter, L., Greenfield, B., Bates, H.J., Wilson, F., Jackson, A.C., Ott, S., Harrison, R.J. y Clarkson, J.P. (2018). Characterisation of pathogen-specific regions and novel effector candidates in *Fusarium oxysporum* f. sp. *cepae*. *Scientific reports*, 8(1), 13530 ^[53].
- Assis, R., Polloni, L.C., Patané, J., Thakur, S., Felestrino, É.B., Diaz-Caballero, J., Digiampietri, L.A., Goulart, L.R., Almeida, N.F., Nascimento, R., Dandekar, A.M., Zaini, P.A., Setubal, J.C., Guttman, D.S. y Moreira, L.M. (2017). Identification and analysis of seven effector protein families with different adaptive and evolutionary histories in plant-associated members of the Xanthomonadaceae. *Scientific Reports*, 7(1), 16133
- Arango Isaza, R.E., Diaz-Trujillo, C., Dhillon, B., Aerts, A., Carlier, J., Crane, C.F., V de Jong, T., de Vries, I., Dietrich, R., Farmer, A.D., Fortes Ferreira, C., Garcia, S., Guzman, M., Hamelin, R.C., Lindquist, E.A., Mehrabi, R., Quiros, O., Schmutz, J., Shapiro, H., Reynolds, E., ... Kema, G.H. (2016). Combating a Global Threat to a Clonal Crop: Banana Black Sigatoka

- Pathogen *Pseudocercospora fijiensis* (Synonym *Mycosphaerella fijiensis*) Genomes Reveal Clues for Disease Control. *PLoS genetics*, 12(8), e1005876
- Ali, S. y Bakkeren, G. (2011). Fungal and oomycete effectors- strategies to subdue a host, *Plant Pathology*, 425-446.
- Alfano J. R. (2009). Roadmap for future research on plant pathogen effectors. *Molecular Plant Pathology*, 10(6), 805–813.
- Ahn, J.O., Jang, H.W., Lee, H.W., Choi, E.S., Haam, S., Oh, T.K. y Jung, J.K. (2003). Overexpression of thermoalkalophilic lipase from *Bacillus stearothermophilus* L1 in *Saccharomyces cerevisiae*. *Journal of microbiology and biotechnology*, 13(3), 451-456.
- Bacelli, I. (2015). Cerato-platanin family proteins: one function for multiple biological roles? *Frontiers in Plant Science*, 5:796
- Bignell, D.R.D., Fyans, J.K., Cheng, Z. (2014). Phytotoxins produced by plant pathogenic *Streptomyces* species. *J Appl Microbiol*, 116, 223–235 ^[74].
- Bateman, A., Coggill, P., Finn, R. (2010). DUF: families in search of function. *Acta Cryst. F66*: 1148-1152.
- Boller, T., Felix, G. (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annual Review of Plant Biology*, 60, 379–406.
- Bornacelly, H. (2009). Estudio del Ciclo de Vida de *Mycosphaerella fijiensis* en tres clones de banano (*Musa AAA*) en tres regiones de la zona bananera del Magdalena. Tesis de Maestría. Universidad Nacional de Colombia, Colombia. pp 23-28.
- Bowen, J.K., Mesarich, C.H., Rees-George, J., Cui, W., Fitzgerald, A., Win, J., Plummer, K. M. y Templeton, M.D. (2009). Candidate effector gene identification in the ascomycete fungal phytopathogen *Venturia inaequalis* by expressed sequence tag analysis. *Molecular Plant Pathology*, 10(3), 431–448 ^[21].
- Bolanos-Garcia, V.M. y Davies, O.R. (2006). Structural analysis and classification of native proteins from *E. coli* commonly co-purified by immobilised metal affinity chromatography. *Biochimica et Biophysica Acta*, 1760(9), 1304–1313.

- van den Burg, H.A., Harrison, S.J., Joosten, M.H., Vervoort, J. y de Wit, P.J. (2006). *Cladosporium fulvum* Avr4 protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection. *Molecular Plant-Microbe Interactions: MPMI*, 19(12), 1420–1430.
- Böhnert, H.U., Fudal, I., Diouf, W., Tharreau, D., Notteghem, J.L. y Lebrun, M.H. (2004). A putative polyketide synthase/peptide synthetase from *Magnaporthe grisea* signals pathogen attack to resistant rice. *The Plant Cell*, 16(9), 2499–2513.
- Carreón-Anguiano, K.G., Islas-Flores, I., Vega-Arreguín, J., Sáenz-Carbonell, L. y Canto-Canché, B. (2020). EffHunter: a tool for prediction of effector protein candidates in fungal proteomic databases. *Biomolecules*, 10(5), 712.
- Contesini, F.J., Davanço, M.G., Borin, G.P., Vanegas, K.G., Cirino, J.P.G., Melo, R.R. de, Mortensen, U.H., Hildén, K., Campos, D.R. y Carvalho, P.O. de (2020). Advances in recombinant lipases: production, engineering, immobilization and application in the pharmaceutical industry. *Catalysts*, 10(9), 1032. MDPI AG.
- Cook, D.E., Kramer, H.M., Torres, D.E., Seidl, M.F. y Thomma, B. (2020). A unique chromatin profile defines adaptive genomic regions in a fungal plant pathogen. *eLife*, 9, e62208.
- Casarrubia, S., Daghino, S., Kohler, A., Morin, E., Khouja, H.R., Daguerre, Y., Veneault-Fourrey, C., Martin, F.M., Perotto, S. y Martino, E. (2018). The hydrophobin-like OmSSP1 may be an effector in the ericoid mycorrhizal symbiosis. *Frontiers in Plant Science*, 9, 546 ^[67].
- Chang, T.C., Salvucci, A., Crous, P.W. y Stergiopoulos, I. (2016). Comparative genomics of the sigatoka disease complex on banana suggests a link between parallel evolutionary changes in *Pseudocercospora fijiensis* and *Pseudocercospora eumusae* and increased virulence on the banana host. *PLoS Genetics*, 12(8), e1005904 ^[39].
- Crous, P. W., Groenewald, J. Z., Slippers, B., & Wingfield, M. J. (2016). Global food and fibre security threatened by current inefficiencies in fungal identification. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 371(1709), 20160024.
- Chang, T.C. y Stergiopoulos, I. (2015). Evolutionary analysis of the global landscape of protein domain types and domain architectures associated with family 14 carbohydrate-binding modules. *FEBS Letters*, 589(15), 1813–1818.

- Cortázar, A.R., Aransay, A.M., Alfaro, M., Oguiza, J.A., y Lavín, J.L. (2014). SECRETOOL: integrated secretome analysis tool for fungi. *Amino acids*, 46(2), 471–473 ^[41].
- Chen, H., Kovalchuk, A., Keriö, S. y Asiegbu, F.O. (2013). Distribution and bioinformatic analysis of the cerato-platanin protein family in Dikarya. *Mycologia*, 105(6), 1479–1488.
- Crous, P. W., Braun, U., Hunter, G. C., Wingfield, M. J., Verkley, G. J., Shin, H. D., Nakashima, C., & Groenewald, J. Z. (2013). Phylogenetic lineages in Pseudocercospora. *Studies in mycology*, 75(1), 37–114.
- Chen, Y., Liu, Z. y Halterman, D.A. (2012). Molecular determinants of resistance activation and suppression by *Phytophthora infestans* effector IPI-O. *PLoS Pathogens*, 8(3), e1002595 ^[60].
- Chettri, P., Calvo, A.M., Cary, J.W., Dhingra, S., Guo, Y., McDougal, R.L. y Bradshaw, R.E. (2012). The veA gene of the pine needle pathogen *Dothistroma septosporum* regulates sporulation and secondary metabolism. *Fungal Genetics and Biology: FG y B*, 49(2), 141–151.
- Christensen, S.A. y Kolomiets, M.V. (2011). The lipid language of plant-fungal interactions. *Fungal Genetics and Biology: FG y B*, 48(1), 4–14.
- Chuma, I., Isobe, C., Hotta, Y., Ibaragi, K., Futamata, N., Kusaba, M., Yoshida, K., Terauchi, R., Fujita, Y., Nakayashiki, H., Valent, B. y Tosa, Y. (2011). Multiple translocation of the AVR-Pita effector gene among chromosomes of the rice blast fungus *Magnaporthe oryzae* and related species. *PLoS Pathogens*, 7(7), e1002147.
- Churchill A.C. (2011). *Mycosphaerella fijiensis*, the black leaf streak pathogen of banana: progress towards understanding pathogen biology and detection, disease development, and the challenges of control. *Molecular Plant Pathology*, 12(4), 307–328.
- Coll, N.S., Epple, P. y Dangl, J.L. (2011). Programmed cell death in the plant immune system. *Cell Death and Differentiation*, 18(8), 1247–1256.
- Crous, P.W. (2009) Taxonomy and phylogeny of the genus *Mycosphaerella* and its anamorphs. *Fungal Divers*, 38, 1-24.
- Carresi, L., Pantera, B., Zoppi, C., Cappugi, G., Oliveira, A.L., Pertinhez, T.A., Spisni, A., Scala, A. y Pazzagli, L. (2006). Cerato-platanin, a phytotoxic protein from *Ceratocystis fimbriata*:

- expression in *Pichia pastoris*, purification and characterization. *Protein Expression and Purification*, 49(2), 159–167.
- Catanzariti, A.M., Dodds, P.N., Lawrence, G.J., Ayliffe, M.A. y Ellis, J.G. (2006). Haustorially expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *The Plant Cell*, 18(1), 243–256
- Chisholm, S.T., Coaker, G., Day, B. y Staskawicz, B.J. (2006). Host-microbe interactions: shaping the evolution of the plant immune response. *Cell*, 124(4), 803–814.
- Dörfors, F., Holmquist, L., Dixelius, C. y Tzelepis, G. (2019). A LysM effector protein from the basidiomycete *Rhizoctonia solani* contributes to virulence through suppression of chitin-triggered immunity. *Molecular Genetics and Genomics*, 294(5), 1211–1218 ^[22].
- Dal Molin, A., Minio, A., Griggio, F., Delledonne, M., Infantino, A. y Aragona, M. (2018). The genome assembly of the fungal pathogen *Pyrenochaeta lycopersici* from single-molecule real-time sequencing sheds new light on its biological complexity. *PloS One*, 13(7), e0200217.
- Dalio, R., Herlihy, J., Oliveira, T.S., McDowell, J.M. y Machado, M. (2018). Effector biology in focus: a primer for computational prediction and functional characterization. *Molecular Plant-Microbe Interactions*, 31(1), 22–33^[1].
- Deb, D., Anderson, R.G., How-Yew-Kin, T., Tyler, B.M. y McDowell, J.M. (2018). Conserved RxLR effectors from oomycetes *Hyaloperonospora arabidopsidis* and *Phytophthora sojae* suppress PAMP- and effector-triggered immunity in diverse plants. *Molecular Plant-Microbe Interactions*, 31(3), 374–385.
- Du, J., Rietman, H. y Vleeshouwers, V.G. (2014). Agroinfiltration and PVX agroinfection in potato and *Nicotiana benthamiana*. *Journal of Visualized Experiments: JoVE*, (83), e50971.
- Duplessis, S., Cuomo, C. A., Lin, Y.C., Aerts, A., Tisserant, E., Veneault-Fourrey, C., Joly, D.L., Hacquard, S., Amselem, J., Cantarel, B.L., Chiu, R., Coutinho, P.M., Feau, N., Field, M., Frey, P., Gelhaye, E., Goldberg, J., Grabherr, M.G., Kodira, C.D., Kohler, A., ... Martin, F. (2011). Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proceedings of the National Academy of Sciences of the United States of America*, 108(22), 9166–9171 ^[26].

- Dou, D., Kale, S.D., Wang, X., Jiang, R.H., Bruce, N.A., Arredondo, F.D., Zhang, X. y Tyler, B.M. (2008). RXLR-mediated entry of *Phytophthora sojae* effector Avr1b into soybean cells does not require pathogen-encoded machinery. *The Plant Cell*, 20(7), 1930–1947.
- Dean, R.A., Talbot, N.J., Ebbole, D.J., Farman, M.L., Mitchell, T.K., Orbach, M.J., Thon, M., Kulkarni, R., Xu, J.R., Pan, H., Read, N.D., Lee, Y.H., Carbone, I., Brown, D., Oh, Y. Y., Donofrio, N., Jeong, J.S., Soanes, D.M., Djonovic, S., Kolomiets, E., ... Birren, B.W. (2005). The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature*, 434(7036), 980–986 ^[55].
- Dodds, P.N., Lawrence, G.J., Catanzariti, A.M., Ayliffe, M.A. y Ellis, J.G. (2004). The *Melampsora lini* AvrL567 avirulence genes are expressed in haustoria and their products are recognized inside plant cells. *The Plant Cell*, 16(3), 755–768.
- Durrant, W.E. y Dong, X. (2004). Systemic acquired resistance. *Annual Review of Phytopathology*, 42, 185–209.
- Donzelli, B.G.G. y Churchill, A.C.L. (2003). *Agrobacterium tumefaciens* mediated transformation of *Mycosphaerella fijiensis*, the black sigatoka pathogen of bananas. *Fungal Genet. Newsl.* 50.
- Dixon, M.S., Hatzixanthis, K., Jones, D.A., Harrison, K. y Jones, J.D. (1998). The tomato Cf-5 disease resistance gene and six homologs show pronounced allelic variation in leucine-rich repeat copy number. *The Plant Cell*, 10(11), 1915–1925.
- Esna-Ashari, Z., Dasgupta, N., Brayton, K.A. y Broschat, S.L. (2018). An optimal set of features for predicting type IV secretion system effector proteins for a subset of species based on a multi-level feature selection approach. *PLoS One*, 13(5), e0197041 ^[23].
- Escobar-Tovar, L., Guzmán-Quesada, M., Sandoval-Fernández, J.A. y Gómez-Lim, M.A. (2015a). Comparative analysis of the *in vitro* and *in planta* secretomes from *Mycosphaerella fijiensis* isolates. *Fungal Biology*, 119(6), 447–470.
- Escobar-Tovar, L., Magaña-Ortíz, D., Fernández, F., Guzmán-Quesada, M., Sandoval-Fernández, J.A., Ortiz-Vázquez, E., Loske, A.M. y Gómez-Lim, M.A. (2015b). Efficient transformation of *Mycosphaerella fijiensis* by underwater shock waves. *Journal of Microbiological Methods*, 119, 98–105.

- van Esse, H.P., Van't Klooster, J.W., Bolton, M.D., yadeta, K.A., van Baarlen, P., Boeren, S., Vervoort, J., de Wit, P.J. y Thomma, B.P. (2008). The *Cladosporium fulvum* virulence protein Avr2 inhibits host proteases required for basal defense. *The Plant Cell*, 20(7), 1948–1963.
- Emanuelsson, O., Nielsen, H., Brunak, S. y von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *Journal of Molecular Biology*, 300(4), 1005–1016 ^[45].
- Eisenhaber, B., Bork, P. y Eisenhaber, F. (1999). Prediction of potential GPI-modification sites in proprotein sequences. *Journal of Molecular Biology*, 292(3), 741–758 ^[46].
- Feldman, D., Yarden, O. y Hadar, Y. (2020). Seeking the Roles for Fungal Small-Secreted Proteins in Affecting Saprophytic Lifestyles. *Frontiers in Microbiology*, 11, 455.
- Fouché, S., Plissonneau, C. y Croll, D. (2018). The birth and death of effectors in rapidly evolving filamentous pathogen genomes. *Current Opinion in Microbiology*, 46, 34–42
- Frantzeskakis, L., Kracher, B., Kusch, S., yoshikawa-Maekawa, M., Bauer, S., Pedersen, C., Spanu, P.D., Maekawa, T., Schulze-Lefert, P. y Panstruga, R. (2018). Signatures of host specialization and a recent transposable element burst in the dynamic one-speed genome of the fungal barley powdery mildew pathogen. *BMC Genomics*, 19(1), 381 ^[50].
- Franceschetti, M., Maqbool, A., Jiménez-Dalmaroni, M.J., Pennington, H.G., Kamoun, S. y Banfield, M. J. (2017). Effectors of filamentous plant pathogens: commonalities amid diversity. *Microbiology and Molecular Biology Reviews*, 81(2), e00066-16.
- Fathi-Roudsari, M., Akhavian-Tehrani, A. y Maghsoudi, N. (2016). Comparison of three *Escherichia coli* strains in recombinant production of reteplase. *Avicenna Journal of Medical Biotechnology*, 8(1), 16–22.
- Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J. y Bateman, A. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research*, 44(D1), D279–D285.

- Friesen T. L. (2016). Combating the sigatoka disease complex on banana. *PLoS Genetics*, 12(8), e1006234.
- Fabro, G., Steinbrenner, J., Coates, M., Ishaque, N., Baxter, L., Studholme, D. J., Körner, E., Allen, R.L., Piquerez, S.J., Rougon-Cardoso, A., Greenshields, D., Lei, R., Badel, J.L., Caillaud, M.C., Sohn, K.H., Van den Ackerveken, G., Parker, J.E., Beynon, J. y Jones, J.D. (2011). Multiple candidate effectors from the oomycete pathogen *Hyaloperonospora arabidopsidis* suppress host plant immunity. *PLoS Pathogens*, 7(11), e1002348.
- Gifre-Renom, L., Seras-Franzoso, J., Rafael, D., Andrade, F., Cano-Garrido, O., Martínez-Trucharte, F., Ugarte-Berzal, E., Martens, E., Boon, L., Villaverde, A., Opdenakker, G., Schwartz, S., Jr, Arís, A. y García-Fruitós, E. (2020). The biological potential hidden in inclusion bodies. *Pharmaceutics*, 12(2), 157.
- Gong, A.D., Jing, Z.Y., Zhang, K., Tan, Q.Q., Wang, G.G.L., Liu, W.D. (2020). Bioinformatic analysis and functional characterization of the CFEM proteins in maize anthracnose fungus *Colletotrichum graminicola*. *Journal of Integrative Agriculture*, 19, 541–550.
- de Guillen, K., Lorrain, C., Tsan, P., Barthe, P., Petre, B., Saveleva, N., Rouhier, N., Duplessis, S., Padilla, A. y Hecker, A. (2019). Structural genomics applied to the rust fungus *Melampsora larici-populina* reveals two candidate effector proteins adopting cystine knot and NTF2-like protein folds. *Scientific Reports*, 9(1), 18084.
- Gui, Y.J., Zhang, W.Q., Zhang, D.D., Zhou, L., Short, D., Wang, J., Ma, X.F., Li, T.G., Kong, Z.Q., Wang, B.L., Wang, D., Li, N.Y., Subbarao, K.V., Chen, J.Y. y Dai, X.F. (2018). A *Verticillium dahliae* extracellular cutinase modulates plant immune responses. *Molecular Plant-Microbe Interactions*, 31(2), 260–273.
- Guzmán-Guzmán, P., Alemán-Duarte, M.I., Delaye, L., Herrera-Estrella, A. y Olmedo-Monfil, V. (2017). Identification of effector-like proteins in *Trichoderma* spp. and role of a hydrophobin in the plant-fungus interaction and mycoparasitism. *BMC Genetics*, 18(1), 16^[68].
- Gabriëls, S., Kroj, T. (2016). Workshop “pathogen-informed crop improvement”, April 8-10, 2015, Wageningen, The Netherlands. *Euphytica*. 210:461-464.

-
- Gawehns, F., Ma, L., Bruning, O., Houterman, P. M., Boeren, S., Cornelissen, B. J., Rep, M. y Takken, F. L. (2015). The effector repertoire of *Fusarium oxysporum* determines the tomato xylem proteome composition following infection. *Frontiers in Plant Science*, 6, 967.
- Grigoriev, I.V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otilar, R., Riley, R., Salamov, A., Zhao, X., Korzeniewski, F., Smirnova, T., Nordberg, H., Dubchak, I. y Shabalov, I. (2014). MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Research*, 42(Database issue), D699–D704 ^[42].
- Grognet, P., Bidard, F., Kuchly, C., Tong, L.C., Coppin, E., Benkhali, J.A., Couloux, A., Wincker, P., Debuchy, R. y Silar, P. (2014). Maintaining two mating types: structure of the mating type locus and its role in heterokaryosis in *Podospira anserina*. *Genetics*, 197(1), 421–432.
- Guo, L., Han, L., Yang, L., Zeng, H., Fan, D., Zhu, Y., Feng, Y., Wang, G., Peng, C., Jiang, X., Zhou, D., Ni, P., Liang, C., Liu, L., Wang, J., Mao, C., Fang, X., Peng, M. y Huang, J. (2014). Genome and transcriptome analysis of the fungal pathogen *Fusarium oxysporum* f. sp. *ubense* causing banana vascular wilt disease. *PLoS One*, 9(4), e95543.
- Guyon, K., Balagué, C., Roby, D. y Raffaele, S. (2014). Secretome analysis reveals effector candidates associated with broad host range necrotrophy in the fungal plant pathogen *Sclerotinia sclerotiorum*. *BMC Genomics*, 15(1), 336.
- Giraldo, M.C. y Valent, B. (2013). Filamentous plant pathogen effectors in action. *Microbiology*, 11(11), 800–814 ^[3].
- Gong, H., Kobayashi, K., Sugi, T., Takemae, H., Kurokawa, H., Horimoto, T., Akashi, H. y Kato, K. (2012). A novel PAN/apple domain-containing protein from *Toxoplasma gondii*: characterization and receptor identification. *PLoS One*, 7(1), e30169.
- Gómez, D. y Zavaleta, E. (2011). Respuesta de hipersensibilidad, una muerte celular programada para defenderse de ataque por fitopatógenos. *Revista Mexicana de Fitopatología*, 29(2), 154-164.
- Goodwin, S.B., M'barek, S.B., Dhillon, B., Wittenberg, A.H., Crane, C.F., Hane, J.K., Foster, A.J., Van der Lee, T.A., Grimwood, J., Aerts, A., Antoniw, J., Bailey, A., Bluhm, B., Bowler, J., Bristow, J., van der Burg, A., Canto-Canché, B., Churchill, A.C., Conde-Ferràez, L., Cools, H.J., ... Kema, G.H. (2011). Finished genome of the fungal wheat pathogen *Mycosphaerella*
-

- graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. *PLoS Genetics*, 7(6), e1002070 ^[56].
- Gehl, C., Waadt, R., Kudla, J., Mendel, R. y Hänsch, R. (2009). New Gateway vectors for high throughput analysis of protein-protein interactions by biomolecular fluorescence complementation. *Mol. Plant*, 2: 1051-1058.
- Grewal, T.S., Rossnagel, B.G., Bakkeren, G. y Scoles, G.J. (2008). Identification of resistance genes to barley covered smut and mapping of the *Ruh1* gene using *Ustilago hordei* strains with defined avirulence genes. *Can. J. Plant Pathol*, 30:277–284.
- Goodin, M.M., Zaitlin, D., Naidu, R.A. y Lommel, S.A. (2008). *Nicotiana benthamiana*: its history and future as a model for plant-pathogen interactions. *Molecular Plant-Microbe Interactions*, 21(8), 1015–1026.
- Gudesblat, G. (2007). Señalización en la interacción entre bacterias fitopatógenas y su hospedador. *Química Viva*. 36-47.
- Hao, G., McCormick, S., Usgaard, T., Tiley, H. y Vaughan, M.M. (2020). Characterization of three *Fusarium graminearum* effectors and their roles during *Fusarium* head blight. *Frontiers in Plant Science*, 11, 579553.
- Huang, A., Lu, M., Ling, E., Li, P. y Wang, C. (2020). A M35 family metalloprotease is required for fungal virulence against insects by inactivating host prophenoloxidasas and beyond. *Virulence*, 11(1), 222–237.
- Huang, S.X., Zhang, Z. H., Liu, W., Tao, H., Zhang, Y., Shi, N.X., Zhu, F., Ji, Z.L. y Chen, X.R. (2020). Expression of the small cysteine-rich protein SCR96 from *Phytophthora cactorum* in mammalian cells: phytotoxicity and exploitation of its polyclonal antibody. *Biotechnology Letters*, 42(1), 125–133.
- Han, X., Altegoer, F., Steinchen, W., Binnebesel, L., Schuhmacher, J., Glatter, T., Giammarinaro, P.I., Djamei, A., Rensing, S.A., Reissmann, S., Kahmann, R. y Bange, G. (2019). A kiwellin disarms the metabolic activity of a secreted fungal virulence factor. *Nature*, 565(7741), 650–653.

- Hurlburt, N.K., Chen, L.H., Stergiopoulos, I. y Fisher, A.J. (2018). Structure of the *Cladosporium fulvum* Avr4 effector in complex with (GlcNAc)₆ reveals the ligand-binding mechanism and uncouples its intrinsic function from recognition by the Cf-4 resistance protein. *PLoS Pathogens*, 14(8), e1007263.
- Hartmann, F.E., Sánchez-Vallet, A., McDonald, B.A. y Croll, D. (2017). A fungal wheat pathogen evolved host specialization by extensive chromosomal rearrangements. *The ISME Journal*, 11(5), 1189–1204.
- Helliwell, E.E., Vega-Arreguín, J., Shi, Z., Bailey, B., Xiao, S., Maximova, S.N., Tyler, B.M. y Guiltinan, M.J. (2016). Enhanced resistance in *Theobroma cacao* against oomycete and fungal pathogens by secretion of phosphatidylinositol-3-phosphate-binding proteins. *Plant Biotechnology Journal*, 14(3), 875–886.
- Heard, S., Brown, N.A. y Hammond-Kosack, K. (2015). An Interspecies Comparative Analysis of the Predicted Secretomes of the Necrotrophic Plant Pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. *PloS One*, 10(6), e0130534.
- Hacquard, S., Petre, B., Frey, P., Hecker, A., Rouhier, N. y Duplessis, S. (2011). The poplar-poplar rust interaction: insights from genomics and transcriptomics. *Journal of Pathogens*, 2011, 716041.
- Haas, B.J., Kamoun, S., Zody, M.C., Jiang, R.H., Handsaker, R.E., Cano, L.M., Grabherr, M., Kodira, C.D., Raffaele, S., Torto-Alalibo, T., Bozkurt, T.O., Ah-Fong, A.M., Alvarado, L., Anderson, V.L., Armstrong, M.R., Avrova, A., Baxter, L., Beynon, J., Boevink, P.C., Bollmann, S.R., ... Nusbaum, C. (2009). Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature*, 461(7262), 393–398 ^[57].
- Houterman, P.M., Cornelissen, B.J. y Rep, M. (2008). Suppression of plant resistance gene-based immunity by a fungal effector. *PLoS Pathogens*, 4(5), e1000061.
- Huang, D., Gore, P.R. y Shusta, E.V. (2008). Increasing yeast secretion of heterologous proteins by regulating expression rates and post-secretory loss. *Biotechnology and Bioengineering*, 101(6), 1264–1275.
- Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., Nakai, K. (2007) WoLF PSORT: protein localization predictor. *Nucleic Acids Research*. 35, W585–W587 ^[13].

- van den Hooven, H.W., van den Burg, H.A., Vossen, P., Boeren, S., de Wit, P. J. y Vervoort, J. (2001). Disulfide bond structure of the AVR9 elicitor of the fungal tomato pathogen *Cladosporium fulvum*: evidence for a cystine knot. *Biochemistry*, 40(12), 3458–3466.
- Jaswal, R., Rajarammohan, S., Dubey, H. y Sharma, T.R. (2020). Smut fungi as a stratagem to characterize rust effectors: opportunities and challenges. *World Journal of Microbiology and Biotechnology*, 36(10), 150.
- Jones, D.A., Bertazzoni, S., Turo, C.J., Syme, R.A. y Hane, J.K. (2018). Bioinformatic prediction of plant-pathogenicity effector proteins of fungi. *Current Opinion in Microbiology*, 46, 43–49.
- Jia, B. y Jeon, C.O. (2016). High-throughput recombinant protein expression in *Escherichia coli*: current status and future perspectives. *Open Biology*, 6(8), 160196.
- Johansson, O. (2015). Plant pathogen defense: Signalling, resistance and death. Tesis de Doctorado. Universidad de Gotemburgo. Suecia. 21 de enero 2015. 71pag.
- de Jonge, R., Bolton, M.D., Kombrink, A., van den Berg, G.C. yadeta, K.A. y Thomma, B.P. (2013). Extensive chromosomal reshuffling drives evolution of virulence in an asexual pathogen. *Genome Research*, 23(8), 1271–1282.
- de Jonge, R., Bolton, M.D., Thoma, B.P. (2011) How filamentous pathogens co-opt plants: The ins and outs of fungal effectors. *Curr. Opin. Plant Biology*, 14, 400-4006 ^[2].
- Jones, J.D. y Dangl, J.L. (2006). The plant immune system. *Nature*, 444(7117), 323–329.
- Kanja, C. y Hammond-Kosack, K.E. (2020). Proteinaceous effector discovery and characterization in filamentous plant pathogens. *Molecular Plant Pathology*, 21(10), 1353–1376.
- Kersey, P.J., Allen, J.E., Allot, A., Barba, M., Boddu, S., Bolt, B.J., Carvalho-Silva, D., Christensen, M., Davis, P., Grabmueller, C., Kumar, N., Liu, Z., Maurel, T., Moore, B., McDowall, M. D., Maheswari, U., Naamati, G., Newman, V., Ong, C.K., Paulini, M., ... Yates, A. (2018). Ensembl Genomes 2018: an integrated omics infrastructure for non-vertebrate species. *Nucleic Acids Research*, 46(D1), D802–D808 ^[43].

- Kettles, G.J., Bayon, C., Canning, G., Rudd, J.J. y Kanyuka, K. (2017). Apoplastic recognition of multiple candidate effectors from the wheat pathogen *Zymoseptoria tritici* in the nonhost plant *Nicotiana benthamiana*. *The New Phytologist*, 213(1), 338–350 ^[66].
- Kohler, A.C., Chen, L.H., Hurlburt, N., Salvucci, A., Schwessinger, B., Fisher, A.J. y Stergiopoulos, I. (2016). Structural Analysis of an Avr4 Effector Ortholog Offers Insight into Chitin Binding and Recognition by the Cf-4 Receptor. *The Plant Cell*, 28(8), 1945–1965.
- King, S.R., McLellan, H., Boevink, P.C., Armstrong, M.R., Bukharova, T., Sukarta, O., Win, J., Kamoun, S., Birch, P.R. y Banfield, M.J. (2014). *Phytophthora Infestans* RXLR effector PexRD2 interacts with host MAPKKK ϵ to suppress plant immune signaling. *The Plant Cell*, 26(3), 1345–1359.
- Krijger, J.J., Thon, M.R., Deising, H.B. y Wiersel, S.G. (2014). Compositions of fungal secretomes indicate a greater impact of phylogenetic history than lifestyle adaptation. *BMC Genomics*, 15(1), 722 ^[24].
- Kuo, H.C., Hui, S., Choi, J., Asiegbu, F.O., Valkonen, J.P. y Lee, Y.H. (2014). Secret lifestyles of *Neurospora crassa*. *Scientific Reports*, 4, 5135 ^[75].
- Kombrink, A. y Thomma, B.P. (2013). LysM effectors: secreted proteins supporting fungal life. *PLoS Pathogens*, 9(12), e1003769.
- Koeck, M., Hardham, A.R. y Dodds, P.N. (2011). The role of effectors of biotrophic and hemibiotrophic fungi in infection. *Cellular Microbiology*, 13(12), 1849–1857.
- Kale, S.D., Gu, B., Capelluto, D.G., Dou, D., Feldman, E., Rumore, A., Arredondo, F.D., Hanlon, R., Fudal, I., Rouxel, T., Lawrence, C.B., Shan, W. y Tyler, B.M. (2010). External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells. *Cell*, 142(2), 284–295.
- Khang, C.H., Berruyer, R., Giraldo, M.C., Kankanala, P., Park, S.Y., Czymmek, K., Kang, S. y Valent, B. (2010). Translocation of *Magnaporthe oryzae* effectors into rice cells and their subsequent cell-to-cell movement. *The Plant Cell*, 22(4), 1388–1403.
- Kema, G. (2009) Sequencing the major *Mycosphaerella* pathogens of wheat and banana. *Acta Horti*, 828, 147-152.

-
- Käll, L., Krogh, A. y Sonnhammer, E.L. (2007). Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. *Nucleic Acids Research*, 35(Web Server issue), W429–W432 ^[18].
- Kämper, J., Kahmann, R., Bölker, M., Ma, L.J., Brefort, T., Saville, B.J., Banuett, F., Kronstad, J.W., Gold, S.E., Müller, O., Perlin, M.H., Wösten, H.A., de Vries, R., Ruiz-Herrera, J., Reynaga-Peña, C.G., Snetselaar, K., McCann, M., Pérez-Martín, J., Feldbrügge, M., Basse, C.W., ... Birren, B.W. (2006). Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature*, 444(7115), 97–101 ^[59].
- Mysore, K.S. y Ryu, C.M. (2004). Nonhost resistance: how much do we know?. *Trends in Plant Science*, 9(2), 97–104.
- Kulkarni, R.D., Kelkar, H.S. y Dean, R.A. (2003). An eight-cysteine-containing CFEM domain unique to a group of fungal membrane proteins. *Trends in Biochemical Sciences*, 28(3), 118–121.
- Krogh, A., Larsson, B., von Heijne, G. y Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology*, 305(3), 567–580 ^[14].
- Li, J., Gao, M., Gabriel, D. W., Liang, W. y Song, L. (2020). Secretome-wide analysis of lysine acetylation in *Fusarium oxysporum* f. sp. *lycopersici* provides novel insights into infection-related proteins. *Frontiers in Microbiology*, 11, 559440.
- Laur, J., Ramakrishnan, G.B., Labbé, C., Lefebvre, F., Spanu, P.D. y Bélanger, R.R. (2018). Effectors involved in fungal-fungal interaction lead to a rare phenomenon of hyperbiotrophy in the tritrophic system biocontrol agent-powdery mildew-plant. *The New Phytologist*, 217(2), 713–725 ^[69].
- Liang, P., Liu, S., Xu, F., Jiang, S., Yan, J., He, Q., Liu, W., Lin, C., Zheng, F., Wang, X. y Miao, W. (2018). Powdery mildews are characterized by contracted carbohydrate metabolism and diverse effectors to adapt to obligate biotrophic lifestyle. *Frontiers in Microbiology*, 9, 3160 ^[38].
- Lin, H.C., Yu, P.L., Chen, L.H., Tsai, H.C., y Chung, K.R. (2018). A major facilitator superfamily transporter regulated by the stress-responsive transcription factor Yap1 Is required for

- resistance to fungicides, xenobiotics, and oxidants and full virulence in *Alternaria alternata*. *Frontiers in microbiology*, 9, 2229 ^[32].
- Levin, E., Ballester, A.R., Raphael, G., Feigenberg, O., Liu, Y., Norelli, J., Gonzalez-Candelas, L., Ma, J., Dardick, C., Wisniewski, M. y Droby, S. (2017). Identification and characterization of LysM effectors in *Penicillium expansum*. *PLoS One*, 12(10), e0186023.
- Lu, S. y Edwards, M.C. (2016). Genome-wide analysis of small secreted cysteine-rich proteins identifies candidate effector proteins potentially involved in *Fusarium graminearum*-wheat interactions. *Phytopathology*, 106(2), 166–176 ^[4].
- Lambertz, C., Garvey, M., Klinger, J., Heesel, D., Klose, H., Fischer, R. y Commandeur, U. (2014). Challenges and advances in the heterologous expression of cellulolytic enzymes: a review. *Biotechnology for Biofuels*, 7(1), 135.
- Leiva, M., Portal, O., Alvarado-Capó, Y., Acosta-Suárez, M., Cruz-Martín, M., Sánchez-García, C., Rojas, L., Roque, B. y Carrera, K. (2014). Caracterización molecular de aislados cubanos de *Mycosphaerella fijiensis* Morelet en base a grupos de apareamiento sexual. *Protección Vegetal*, 70-74.
- Liu, T., Song, T., Zhang, X., Yuan, H., Su, L., Li, W., Xu, J., Liu, S., Chen, L., Chen, T., Zhang, M., Gu, L., Zhang, B. y Dou, D. (2014). Unconventionally secreted effectors of two filamentous pathogens target plant salicylate biosynthesis. *Nature Communications*, 5, 4686 ^[9].
- Larroque, M., Barriot, R., Bottin, A., Barre, A., Rougé, P., Dumas, B. y Gaulin, E. (2012). The unique architecture and function of cellulose-interacting proteins in oomycetes revealed by genomic and structural analyses. *BMC Genomics*, 13, 605.
- Lawrence, G.J., Dodds, P.N. y Ellis, J.G. (2010). Transformation of the flax rust fungus, *Melampsora lini*: selection via silencing of an avirulence gene. *The Plant Journal: for Cell and Molecular Biology*, 61(2), 364–369.
- Leiva, M., Alvarado, Y., Acosta, M., Cruz, M., Sánchez, C. y Roque, B. (2010). Protocolo para la inoculación artificial de plantas de *Musa* spp. con *Mycosphaerella fijiensis* y evaluación de su respuesta mediante variables epifitológicas y componentes de la resistencia. *Bioteología Vegetal*, 79-88.

- Lehti-Shiu, M.D., Zou, C., Hanada, K. y Shiu, S.H. (2009). Evolutionary history and stress regulation of plant receptor-like kinase/pelle genes. *Plant Physiology*, 150(1), 12–26.
- Li, W., Wang, B., Wu, J., Lu, G., Hu, Y., Zhang, X., Zhang, Z., Zhao, Q., Feng, Q., Zhang, H., Wang, Z., Wang, G., Han, B., Wang, Z. y Zhou, B. (2009). The *Magnaporthe oryzae* avirulence gene AvrPiz-t encodes a predicted secreted protein that triggers the immunity in rice mediated by the blast resistance gene Piz-t. *Molecular Plant-Microbe Interactions*, 22(4), 411–420.
- Lorang, J.M., Sweat, T.A. y Wolpert, T.J. (2007). Plant disease susceptibility conferred by a "resistance" gene. *Proceedings of the National Academy of Sciences of the United States of America*, 104(37), 14861–14866 ^[62].
- Madina, M.H., Rahman, M.S., Huang, X., Zhang, Y., Zheng, H. y Germain, H. (2020). A poplar rust effector protein associates with protein disulfide isomerase and enhances plant susceptibility. *Biology*, 9(9), 294.
- McCorison, C.B. y Goodwin, S.B. (2020). The wheat pathogen *Zymoseptoria tritici* senses and responds to different wavelengths of light. *BMC Genomics*, 21(1), 513.
- McGowan, J., O'Hanlon, R., Owens, R.A. y Fitzpatrick, D.A. (2020). Comparative genomic and proteomic analyses of three widespread *Phytophthora* Species: *Phytophthora chlamydospora*, *Phytophthora gonapodyides* and *Phytophthora pseudosyringae*. *Microorganisms*, 8(5), 653.
- Mukhi, N., Gorenkin, D. y Banfield, M.J. (2020). Exploring folds, evolution and host interactions: understanding effector structure/function in disease and immunity. *The New Phytologist*, 227(2), 326–333.
- Ma, L., Djavaheri, M., Wang, H., Larkan, N.J., Haddadi, P., Beynon, E., Gropp, G. y Borhan, M.H. (2018). *Leptosphaeria maculans* effector protein AvrLm1 modulates plant immunity by enhancing MAP kinase 9 phosphorylation. *iScience*, 3, 177–191.
- Marton, K., Flajšman, M., Radišek, S., Košmelj, K., Jakše, J., Javornik, B. y Berne, S. (2018). Comprehensive analysis of *Verticillium nonalfalfae* *in silico* secretome uncovers putative effector proteins expressed during hop invasion. *PloS One*, 13(6), e0198971 ^[27].

- Mokhonov, V.V., Vasilenko, E.A., Gorshkova, E.N., Astrakhantseva, I.V., Novikov, D.V. y Novikov, V.V. (2018). SlyD-deficient *Escherichia coli* strains: A highway to contaminant-free protein extraction. *Biochemical and Biophysical Research Communications*, 499(4), 967–972.
- Muñoz-Pérez, G. (2018). Expresión heteróloga de la proteína PfAvr4 de *Pseudocercospora fijiensis* y análisis de su efector in planta. Tesis de Maestría, Centro de Investigación Científica de Yucatán, México. 21 de noviembre 2018. 121pag.
- Mesarich, C.H., Stergiopoulos, I., Beenen, H.G., Cordovez, V., Guo, Y., Karimi Jashni, M., Bradshaw, R.E. y de Wit, P.J. (2016). A conserved proline residue in Dothideomycete Avr4 effector proteins is required to trigger a Cf-4-dependent hypersensitive response. *Molecular Plant Pathology*, 17(1), 84–95.
- Marchler-Bauer, D. (2015). CDD: NCBI's conserved domain database. *Nucleic Acids Research*, 43(Database issue), D222–D226.
- Moghadam, M., Ganji, A., Varasteh, A., Falak, R. y Sankian, M. (2015). Refolding process of cysteine-rich proteins: Chitinase as a model. *Reports of Biochemistry and Molecular Biology*, 4(1), 19–24.
- Manning, V.A., Pandelova, I., Dhillon, B., Wilhelm, L.J., Goodwin, S.B., Berlin, A.M., Figueroa, M., Freitag, M., Hane, J.K., Henrissat, B., Holman, W.H., Kodira, C.D., Martin, J., Oliver, R.P., Robbertse, B., Schackwitz, W., Schwartz, D.C., Spatafora, J.W., Turgeon, B.G., Yandava, C., ... Ciuffetti, L.M. (2013). Comparative genomics of a plant-pathogenic fungus, *Pyrenophora tritici-repentis*, reveals transduplication and the impact of repeat elements on pathogenicity and population divergence. *G3 (Bethesda, Md.)*, 3(1), 41–63 ^[58].
- Ma, L., Lukasik, E., Gawehns, F. y Takken, F.L. (2012). The use of agroinfiltration for transient expression of plant resistance and fungal effector proteins in *Nicotiana benthamiana* leaves. *Methods in Molecular Biology (Clifton, N.J.)*, 835, 61–74.
- Martínez, L. (2012). Epidemiología y manejo de la Sigatoka Negra (*Mycosphaerella fijiensis* Morelet) del Banano (*Musa AAA*). Tesis de Doctorado. Colegio de Postgraduados. México. 20 de noviembre 2012. 61pag.

- Morais do Amaral, A., Antoniw, J., Rudd, J.J. y Hammond-Kosack, K.E. (2012). Defining the predicted protein secretome of the fungal wheat leaf pathogen *Mycosphaerella graminicola*. *PLoS One*, 7(12), e49904 ^[40].
- Marshall, R., Kombrink, A., Motteram, J., Loza-Reyes, E., Lucas, J., Hammond-Kosack, K.E., Thomma, B.P. y Rudd, J.J. (2011). Analysis of two *in planta* expressed LysM effector homologs from the fungus *Mycosphaerella graminicola* reveals novel functional properties and varying contributions to virulence on wheat. *Plant Physiology*, 156(2), 756–769.
- Neu, E. y Debener, T. (2019). Prediction of the *Diplocarpon rosae* secretome reveals candidate genes for effectors and virulence factors. *Fungal Biology*, 123(3), 231–239 ^[5].
- Nagarathinam, K., Nakada-Nakura, Y., Parthier, C., Terada, T., Juge, N., Jaenecke, F., Liu, K., Hotta, Y., Miyaji, T., Omote, H., Iwata, S., Nomura, N., Stubbs, M.T. y Tanabe, M. (2018). Outward open conformation of a major facilitator superfamily multidrug/H⁺ antiporter provides insights into switching mechanism. *Nature Communications*, 9(1), 4005 ^[33].
- Noar, R.D. y Daub, M.E. (2016). Transcriptome sequencing of *Mycosphaerella fijiensis* during association with *Musa acuminata* reveals candidate pathogenicity genes. *BMC Genomics*, 17(1), 690.
- Nürnberg, T. y Lipka, V. (2005). Non-host resistance in plants: new insights into an old phenomenon. *Molecular Plant Pathology*, 6(3), 335–345.
- Onyilo, F., Tusiime, G., Tripathi, J.N., Chen, L.H., Falk, B., Stergiopoulos, I., Tushemereirwe, W., Kubiriba, J. y Tripathi, L. (2018). Silencing of the mitogen-activated protein kinases (MAPK) Fus3 and SlT2 in *Pseudocercospora fijiensis* reduces growth and virulence on host plants. *Frontiers in Plant Science*, 9, 291.
- O'Connell, R.J., Thon, M.R., Hacquard, S., Amyotte, S.G., Kleemann, J., Torres, M.F., Damm, U., Buiate, E.A., Epstein, L., Alkan, N., Altmüller, J., Alvarado-Balderrama, L., Bauser, C.A., Becker, C., Birren, B.W., Chen, Z., Choi, J., Crouch, J.A., Duvick, J.P., Farman, M.A., ... Vaillancourt, L. J. (2012). Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. *Nature Genetics*, 44(9), 1060–1065 ^[52].
- Ohm, R.A., Feu, N., Henrissat, B., Schoch, C.L., Horwitz, B.A., Barry, K.W., Condon, B.J., Copeland, A.C., Dhillon, B., Glaser, F., Hesse, C.N., Kosti, I., LaButti, K., Lindquist, E. A.,

- Lucas, S., Salamov, A.A., Bradshaw, R.E., Ciuffetti, L., Hamelin, R.C., Kema, G.H., ... Grigoriev, I.V. (2012). Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen Dothideomycetes fungi. *PLoS Pathogens*, 8(12), e1003037 ^[49].
- Ojito-Ramos, K. y Portal, O. (2010). Introducción al sistema inmune en plantas. *Bioteología Vegetal*, 3-19.
- Oliver, R. (2009) Plant breeding for disease resistance in the age of effectors. *Phytoparasitica*. 37, 1–5 ^[65].
- Ohtaki, S., Maeda, H., Takahashi, T., Yamagata, Y., Hasegawa, F., Gomi, K., Nakajima, T. y Abe, K. (2006). Novel hydrophobic surface binding protein, HsbA, produced by *Aspergillus oryzae*. *Applied and Environmental Microbiology*, 72(4), 2407–2413.
- Orbach, M.J., Farrall, L., Sweigard, J.A., Chumley, F.G. y Valent, B. (2000). A telomeric avirulence gene determines efficacy for the rice blast resistance gene Pi-ta. *The Plant Cell*, 12(11), 2019–2032.
- Osuji, J., Okoli, B., Vuylsteke, D. y Ortiz, R. (1997). Multivariate pattern of quantitative trait variation in triploid banana and plantain cultivars. *Scientia Horticulturae*, 197-202.
- Pan, L., Wen, S., Yu, J., Lu, L., Zhu, X. y Zhang, Z. (2020). Genome-wide identification of M35 family metalloproteases in *Rhizoctonia cerealis* and functional analysis of RcMEP2 as a virulence factor during the fungal infection to wheat. *International Journal of Molecular Sciences*, 21(8).
- Pandaranayaka, E.P., Frenkel, O., Elad, Y., Prusky, D. y Harel, A. (2019). Network analysis exposes core functions in major lifestyles of fungal and oomycete plant pathogens. *BMC Genomics*, 20(1), 1020.
- Peng, Z., Oliveira-Garcia, E., Lin, G., Hu, Y., Dalby, M., Migeon, P., Tang, H., Farman, M., Cook, D., White, F.F., Valent, B. y Liu, S. (2019). Effector gene reshuffling involves dispensable mini-chromosomes in the wheat blast fungus. *PLoS Genetics*, 15(9), e1008272.
- Plissonneau, C., Hartmann, F.E. y Croll, D. (2018). Pangenome analyses of the wheat pathogen *Zymoseptoria tritici* reveal the structural basis of a highly plastic eukaryotic genome. *BMC Biology*, 16(1), 5.

- Petre, B., Win, J., Menke, F. y Kamoun, S. (2017). Protein-protein interaction assays with effector-GFP fusions in *Nicotiana benthamiana*. *Methods in Molecular Biology* (Clifton, N.J.), 1659, 85–98
- Pennington, H.G., Gheorghe, D.M., Damerum, A., Pliego, C., Spanu, P.D., Cramer, R. y Bindschedler, L.V. (2016). Interactions between the Powdery Mildew Effector BEC1054 and Barley Proteins Identify Candidate Host Targets. *Journal of Proteome Research*, 15(3), 826–839.
- Petre, B., Saunders, D.G., Sklenar, J., Lorrain, C., Krasileva, K.V., Win, J., Duplessis, S. y Kamoun, S. (2016). Heterologous expression screens in *Nicotiana Benthamiana* identify a candidate effector of the wheat yellow rust pathogen that associates with processing odies. *PloS One*, 11(2), e0149035.
- Petre, B., Saunders, D.G., Sklenar, J., Lorrain, C., Win, J., Duplessis, S. y Kamoun, S. (2015). Candidate effector proteins of the rust pathogen *Melampsora larici-populina* target diverse plant cell compartments. *Molecular Plant-Microbe Interactions*, 28(6), 689–700 ^[6].
- Petre, B., Joly, D.L. y Duplessis, S. (2014). Effector proteins of rust fungi. *Frontiers in Plant Science*, 5, 416.
- Pedersen, C., Ver Loren van Themaat, E., McGuffin, L.J., Abbott, J.C., Burgis, T.A., Barton, G., Bindschedler, L.V., Lu, X., Maekawa, T., Wessling, R., Cramer, R., Thordal-Christensen, H., Panstruga, R. y Spanu, P.D. (2012). Structure and evolution of barley powdery mildew effector candidates. *BMC Genomics*, 13, 694.
- Portal, O., Acosta-Suárez, M., Ocaña, B., Schäfer, W., Jiménez, E. y Höfte, M. (2012). A green fluorescent protein-transformed *Mycosphaerella fijiensis* strain shows increased aggressiveness on banana. *Plant Pathology*, 41, 645–647
- Petersen, T.N., Brunak, S., von Heijne, G. y Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, 8(10), 785–786 ^[12].
- Patiño, L., Bustamante, E. y Salazar, L. (2006). Efecto de sustratos foliares sobre la Sigatoka Negra (*Mycosphaerella fijiensis* Morolet) en banano (*Musa X paradisiaca* L.) y plátano (*Musa acuminata* Colla). *Agricultura Técnica*, 437-455.

- Queiroz, C.B. y Santana, M.F. (2020). Prediction of the secretomes of endophytic and nonendophytic fungi reveals similarities in host plant infection and colonization strategies. *Mycologia*, 112(3), 491–503.
- Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R. y Lopez, R. (2005). InterProScan: protein domains identifier. *Nucleic Acids Research*, 33(Web Server issue), W116–W120.
- Ramírez-Valdespino, C.A., Casas-Flores, S. y Olmedo-Monfil, V. (2019). *Trichoderma* as a model to study effector-like molecules. *Frontiers in Microbiology*, 10, 1030 ^[70].
- Robin, G.P., Kleemann, J., Neumann, U., Cabre, L., Dallery, J.F., Lapalu, N. y O'Connell, R.J. (2018). Subcellular localization screening of *Colletotrichum higginsianum* effector candidates identifies fungal proteins targeted to plant peroxisomes, golgi bodies, and microtubules. *Frontiers in Plant Science*, 9, 562.
- Ramachandran, S.R., Yin, C., Kud, J., Tanaka, K., Mahoney, A.K., Xiao, F. y Hulbert, S.H. (2017). Effectors from wheat rust fungi suppress multiple plant defense responses. *Phytopathology*, 107(1), 75–83.
- Rosano, G. L. y Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in Microbiology*, 5, 172
- Rovenich, H., Boshoven, J.C. y Thomma, B.P. (2014). Filamentous pathogen effector functions: of pathogens, hosts and microbiomes. *Current Opinion in Plant Biology*, 20, 96–103.
- Raffaele, S., Kamoun, S. (2012) Genome evolution in filamentous plant pathogens: Why bigger can be better. *Nature Reviews Microbiology*, 10, 417-430 ^[11].
- Rastogi, S. y Rost, B. (2011). LocDB: experimental annotations of localization for *Homo sapiens* and *Arabidopsis thaliana*. *Nucleic Acids Research*, 39(Database issue), D230–D234 ^[47].
- Rouxel, T., Grandaubert, J., Hane, J.K., Hoede, C., van de Wouw, A.P., Couloux, A., Dominguez, V., Anthouard, V., Bally, P., Bourras, S., Cozijnsen, A.J., Ciuffetti, L.M., Degrave, A., Dilmaghani, A., Duret, L., Fudal, I., Goodwin, S.B., Gout, L., Glaser, N., Linglin, J., ... Howlett, B.J. (2011). Effector diversification within compartments of the *Leptosphaeria*

- maculans* genome affected by repeat-induced point mutations. *Nature Communications*, 2, 202 [54].
- Rafiqi, M., Gan, P.H., Ravensdale, M., Lawrence, G.J., Ellis, J.G., Jones, D.A., Hardham, A.R. y Dodds, P.N. (2010). Internalization of flax rust avirulence proteins into flax and tobacco cells can occur in the absence of the pathogen. *The Plant Cell*, 22(6), 2017–2032.
- Rees, D.C., Johnson, E. y Lewinson, O. (2009). ABC transporters: the power to change. *Molecular Cell Biology*, 10(3), 218–227 [29].
- Rodríguez, P. y Cayón, G. (2008). Efecto de *Mycosphaerella fijiensis* sobre la fisiología de la hoja de banano. *Agronomía Colombiana*, 256-265.
- Ridout, C.J., Skamnioti, P., Porritt, O., Sacristan, S., Jones, J.D. y Brown, J.K. (2006). Multiple avirulence paralogues in cereal powdery mildew fungi may contribute to parasite fitness and defeat of plant resistance. *The Plant Cell*, 18(9), 2402–2414.
- Rooney, H.C., Van't Klooster, J.W., van der Hoorn, R.A., Joosten, M.H., Jones, J.D. y de Wit, P.J. (2005). *Cladosporium Avr2* inhibits tomato Rcr3 protease required for Cf-2-dependent disease resistance. *Science* (New York, N.Y.), 308(5729), 1783–1786.
- Rep, M. (2005). Small proteins of plant-pathogenic fungi secreted during host colonization. *Fems Microbiology*, 253, 19-27 [7].
- Rep, M., van der Does, H.C., Meijer, M., van Wijk, R., Houterman, P.M., Dekker, H.L., de Koster, C.G. y Cornelissen, B.J. (2004). A small, cysteine-rich protein secreted by *Fusarium oxysporum* during colonization of xylem vessels is required for I-3-mediated resistance in tomato. *Molecular Microbiology*, 53(5), 1373–1383.
- Sarkar, D., Rovenich, H., Jeena, G., Nizam, S., Tissier, A., Balcke, G.U., Mahdi, L.K., Bonkowski, M., Langen, G. y Zuccaro, A. (2019). The inconspicuous gatekeeper: endophytic *Serendipita vermifera* acts as extended plant protection barrier in the rhizosphere. *The New Phytologist*, 224(2), 886–901 [71].
- See, P.T., Iagallo, E.M., Oliver, R.P. y Moffat, C.S. (2019). Heterologous expression of the *Pyrenophora tritici-repentis* effector proteins ToxA and ToxB, and the prevalence of effector sensitivity in Australian cereal crops. *Frontiers in Microbiology*, 10, 182.

- Sperschneider, J., Dodds, P.N., Gardiner, D.M., Singh, K.B. and Taylor, J.M. (2018). Improved prediction of fungal effector proteins from secretomes with EffectorP 2.0. *Molecular Plant Pathology*, 19, 2094-2110 ^[15].
- SAGARPA (2017). Comunicado de prensa: Producción de plátano “Hecho en México” aumenta siete por ciento. SAGARPA.
- Sperschneider, J., Catanzariti, A.M., DeBoer, K., Petre, B., Gardiner, D.M., Singh, K.B., Dodds, P.N. y Taylor, J.M. (2017). LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. *Scientific Reports*, 7, 44598 ^[35].
- Sarris, P.F., Cevik, V., Dagdas, G., Jones, J.D. y Krasileva, K.V. (2016). Comparative analysis of plant immune receptor architectures uncovers host proteins likely targeted by pathogens. *BMC Biology*, 14, 8
- Sędziewska Toro, K. y Brachmann, A. (2016). The effector candidate repertoire of the arbuscular mycorrhizal fungus *Rhizophagus clarus*. *BMC Genomics*, 17, 101.
- Sonah, H., Deshmukh, R.K. y Bélanger, R.R. (2016). Computational prediction of effector proteins in fungi: opportunities and challenges. *Frontiers in Plant Science*, 7, 126 ^[10].
- Sperschneider, J., Gardiner, D. M., Dodds, P. N., Tini, F., Covarelli, L., Singh, K. B., Manners, J. M., y Taylor, J. M. (2016). EffectorP: predicting fungal effector proteins from secretomes using machine learning. *The New phytologist*, 210(2), 743–761 ^[16].
- Singh, A., Upadhyay, V., Upadhyay, A.K., Singh, S.M. y Panda, A.K. (2015). Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. *Microbial Cell Factories*, 14, 41.
- Skaar, K., Korza, H.J., Tarry, M., Sekyrova, P. y Högbom, M. (2015). Expression and subcellular distribution of GFP-Tagged human tetraspanin proteins in *Saccharomyces cerevisiae*. *PloS One*, 10(7), e0134041.
- Sperschneider, J., Dodds, P., Gardiner, D., Manners, J., Singh, B., Taylor, J. (2015) Advances and challenges in computational prediction of effectors from plant pathogenic fungi, *PLOS Pathogens*, 1-7.

- Sperschneider, J., Williams, A. H., Hane, J. K., Singh, K. B., y Taylor, J. M. (2015). Evaluation of secretion prediction highlights differing approaches needed for oomycete and fungal effectors. *Frontiers in plant science*, 6, 1168 [28].
- Soyer, J.L., El Ghalid, M., Glaser, N., Ollivier, B., Linglin, J., Grandaubert, J., Balesdent, M. H., Connolly, L.R., Freitag, M., Rouxel, T. y Fudal, I. (2014). Epigenetic control of effector gene expression in the plant pathogenic fungus *Leptosphaeria maculans*. *PLoS Genetics*, 10(3), e1004227.
- Sánchez-Vallet, A., Saleem-Batcha, R., Kombrink, A., Hansen, G., Valkenburg, D.J., Thomma, B.P. y Mesters, J.R. (2013). Fungal effector Ecp6 outcompetes host immune receptor for chitin binding through intrachain LysM dimerization. *eLife*, 2, e00790 [61].
- Stam, R., Jupe, J., Howden, A.J., Morris, J.A., Boevink, P.C., Hedley, P.E. y Huitema, E. (2013). Identification and characterisation CRN effectors in *Phytophthora capsici* shows modularity and functional diversity. *PLoS One*, 8(3), e59517.
- Sumit, R. (2013). Investigation of the *Arabidopsis* non-host resistance mechanism against the soybean pathogen, *Phytophthora sojae*. Universidad Estatal de Iowa. Estados Unidos. Tesis de Doctorado. 121pag.
- Santana, M.F., Silva, J.C., Batista, A.D., Ribeiro, L.E., da Silva, G.F., de Araújo, E.F. y de Queiroz, M.V. (2012). Abundance, distribution and potential impact of transposable elements in the genome of *Mycosphaerella fijiensis*. *BMC Genomics*, 13, 720.
- Stergiopoulos, I., Kourmpetis, Y.A., Slot, J.C., Bakker, F.T., De Wit, P.J. y Rokas, A. (2012). In silico characterization and molecular evolutionary analysis of a novel superfamily of fungal effector proteins. *Molecular Biology and Evolution*, 29(11), 3371–3384.
- Sanzón, D. y Zavaleta, E. (2011). Respuesta de Hipersensibilidad, una Muerte Celular Programada para Defenderse del Ataque por Fitopatógenos. *Revista Mexicana de Fitopatología*, 29, 154–164.
- Stergiopoulos, I., van den Burg, H.A., Okmen, B., Beenen, H.G., van Liere, S., Kema, G.H. y de Wit, P.J. (2010). Tomato Cf resistance proteins mediate recognition of cognate homologous effectors from fungi pathogenic on dicots and monocots. *Proceedings of the National Academy of Sciences of the United States of America*, 107(16), 7610–7615.

- Song, J., Win, J., Tian, M., Schornack, S., Kaschani, F., Ilyas, M., van der Hoorn, R.A. y Kamoun, S. (2009). Apoplastic effectors secreted by two unrelated eukaryotic plant pathogens target the tomato defense protease Rcr3. *Proceedings of the National Academy of Sciences of the United States of America*, 106(5), 1654–1659.
- Sels, J., Mathys, J., De Coninck, B.M., Cammue, B.P. y De Bolle, M.F. (2008). Plant pathogenesis-related (PR) proteins: a focus on PR peptides. *Plant Physiology and Biochemistry: PPB*, 46(11), 941–950.
- Shen, Q.H., Saijo, Y., Mauch, S., Biskup, C., Bieri, S., Keller, B., Seki, H., Ulker, B., Somssich, I.E. y Schulze-Lefert, P. (2007). Nuclear activity of MLA immune receptors links isolate-specific and basal disease-resistance responses. *Science*, 315(5815), 1098–1103.
- Schultz, C. (2000). SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Research*, 28(1), 231–234.
- Solovyev, V. y Sagitov, V. (1997) Softberry, ProtComp; *Softberry*. Available online: <http://www.softberry.com/> (accessed on 15 August 2019) ^[48].
- Simmonds, N. y Shepherd, K. (1995). Taxonomy and origins of cultivated bananas. *Journal of the Linnean Society of London Botany*, 55(359), 302-312.
- Tabima, J.F. y Grünwald, N.J. (2019). *effectR*: An expandable R package to predict candidate RxLR and CRN effectors in oomycetes using motif searches. *Molecular Plant-Microbe Interactions*, 32(9), 1067–1076.
- Thynne, E., Saur, I., Simbaqueba, J., Ogilvie, H.A., Gonzalez-Cendales, Y., Mead, O., Taranto, A., Catanzariti, A.M., McDonald, M.C., Schwessinger, B., Jones, D.A., Rathjen, J.P. y Solomon, P.S. (2017). Fungal phytopathogens encode functional homologues of plant rapid alkalization factor (RALF) peptides. *Molecular Plant Pathology*, 18(6), 811–824.
- Toruño, T.Y., Stergiopoulos, I. y Coaker, G. (2016). Plant-pathogen effectors: cellular probes interfering with plant defenses in spatial and temporal manners. *Annual Review of Phytopathology*, 54, 419–441.
- Tripathi, N.K. (2016). Production and purification of recombinant proteins from *Escherichia coli*. *ChemBioEng Reviews*, 3: 116-133.

- Tan, K.C., Ferguson-Hunt, M., Rybak, K., Waters, O.D., Stanley, W.A., Bond, C.S., Stukenbrock, E.H., Friesen, T.L., Faris, J.D., McDonald, B.A. y Oliver, R.P. (2012). Quantitative variation in effector activity of ToxA isoforms from *Stagonospora nodorum* and *Pyrenophora tritici-repentis*. *Molecular Plant-Microbe Interactions*, 25(4), 515–522.
- Tyo, K.E., Liu, Z., Petranovic, D. y Nielsen, J. (2012). Imbalance of heterologous protein folding and disulfide bond formation rates yields runaway oxidative stress. *BMC Biology*, 10, 16.
- Torrado, M. y Castaño, J. (2008). Incidencia y severidad de las Sigatoka negra (*Mycosphaerella fijiensis* Morelet) y amarilla (*Mycosphaerella musicola* Leach et Mulder) del plátano según los estadios fenológicos. *Agronomía Colombiana*, 435-442.
- Uhse, S. y Djamei, A. (2018). Effectors of plant-colonizing fungi and beyond. *PLoS Pathogens*, 14(6), e1006992.
- Urban, M., Cuzick, A., Rutherford, K., Irvine, A., Pedro, H., Pant, R., Sadanadan, V., Khamari, L., Billal, S., Mohanty, S. y Hammond-Kosack, K.E. (2017). PHI-base: a new interface and further additions for the multi-species pathogen-host interactions database. *Nucleic Acids Research*, 45(D1), D604–D610 ^[19].
- Volk, H., Marton, K., Flajšman, M., Radišek, S., Tian, H., Hein, I., Podlipnik, Č., Thomma, B., Košmelj, K., Javornik, B. y Berne, S. (2019). Chitin-binding protein of verticillium nonalfalae disguises fungus from plant chitinases and suppresses chitin-triggered host immunity. *Molecular Plant-Microbe Interactions*, 32(10), 1378–1390.
- Vieira Gomes, A.M., Souza Carmo, T., Silva Carvalho, L., Mendonça Bahia, F. y Parachin, N.S. (2018). Comparison of yeasts as hosts for recombinant protein production. *Microorganisms*, 6(2), 38.
- Voß, S., Betz, R., Heidt, S., Corradi, N. y Requena, N. (2018). RiCRN1, a crinkler effector from the arbuscular mycorrhizal fungus *Rhizophagus irregularis*, functions in arbuscule development. *Frontiers in Microbiology*, 9, 2068 ^[72].
- Valette, N., Benoit-Gelber, I., Falco, M.D., Wiebenga, A., de Vries, R.P., Gelhaye, E. y Morel-Rouhier, M. (2017). Secretion of small proteins is species-specific within *Aspergillus* sp. *Microbial Biotechnology*, 10(2), 323–329.

- Vleeshouwers, V.G. y Oliver, R.P. (2014). Effectors as tools in disease resistance breeding against biotrophic, hemibiotrophic, and necrotrophic plant pathogens. *Molecular Plant-Microbe Interactions*, 27(3), 196–206 ^[64].
- Wang, C., Wang, P., Han, S., Wang, L, Zhao, Y. y Juan, J. (2020). FunEffector-Pred: Identification of fungi effector by activate learning and genetic algorithm sampling of imbalanced data. *IEEE Access*, 8. 57674-57683.
- Wood, A., Walker, C., Lee, W.S., Urban, M. y Hammond-Kosack, K.E. (2020). Functional evaluation of a homologue of plant rapid alkalisation factor (RALF) peptides in *Fusarium graminearum*. *Fungal Biology*, 124(9), 753–765.
- White, F.F., y Jones, J.B. (2018). One effector at a time. *Nature plants*, 4(3), 134–135 ^[17].
- Wilkens S. (2015). Structure and mechanism of ABC transporters. *F1000prime Reports*, 7, 14 ^[30].
- Whittaker, M.M. y Whittaker, J.W. (2014). Expression and purification of recombinant *Saccharomyces cerevisiae* mitochondrial carrier protein YGR257Cp (Mtm1p). *Protein Expression and Purification*, 93, 77–86.
- Wicker, T., Oberhaensli, S., Parlange, F., Buchmann, J.P., Shatalina, M., Roffler, S., Ben-David, R., Doležel, J., Šimková, H., Schulze-Lefert, P., Spanu, P.D., Bruggmann, R., Amselem, J., Quesneville, H., Ver Loren van Themaat, E., Paape, T., Shimizu, K.K. y Keller, B. (2013). The wheat powdery mildew genome shows the unique evolution of an obligate biotroph. *Nature Genetics*, 45(9), 1092–1096.
- de Wit, P. J., van der Burgt, A., Ökmen, B., Stergiopoulos, I., Abd-Elsalam, K.A., Aerts, A.L., Bahkali, A.H., Beenen, H.G., Chettri, P., Cox, M.P., Datema, E., de Vries, R.P., Dhillon, B., Ganley, A.R., Griffiths, S.A., Guo, Y., Hamelin, R.C., Henrissat, B., Kabir, M.S., Jashni, M.K., ... Bradshaw, R.E. (2012). The genomes of the fungal plant pathogens *Cladosporium fulvum* and *Dothistroma septosporum* reveal adaptation to different hosts and lifestyles but also signatures of common ancestry. *PLoS Genetics*, 8(11), e1003088 ^[51].
- Win, J., Kamoun, S. y Jones, A.M. (2011). Purification of effector-target protein complexes via transient expression in *Nicotiana benthamiana*. *Methods in Molecular Biology* (Clifton, N.J.), 712, 181–194.

- de Wit, P.J., Mehrabi, R., Van den Burg, H.A. y Stergiopoulos, I. (2009). Fungal effector proteins: past, present and future. *Molecular Plant Pathology*, 10(6), 735–747 ^[8].
- Winnenburg, R., Baldwin, T.K., Urban, M., Rawlings, C., Köhler, J. y Hammond-Kosack, K.E. (2006). PHI-base: a new database for pathogen host interactions. *Nucleic Acids Research*, 34(Database issue), D459–D464 ^[20].
- Wroblewski, T., Tomczak, A. y Michelmore, R. (2005). Optimization of *Agrobacterium*-mediated transient assays of gene expression in lettuce, tomato and *Arabidopsis*. *Plant Biotechnology Journal*, 3(2), 259–273.
- Westerink, N., Brandwagt, B.F., de Wit, P.J. y Joosten, M.H. (2004). *Cladosporium fulvum* circumvents the second functional resistance gene homologue at the Cf-4 locus (Hcr9-4E) by secretion of a stable avr4E isoform. *Molecular Microbiology*, 54(2), 533–545.
- Werck-Reichhart, D. y Feyereisen, R. (2000). Cytochromes P450: a success story. *Genome Biology*, 1(6), REVIEWS3003 ^[31].
- Wall, L. y Schwartz, R. (1991). Programming PERL; *Nutshell handbooks*; 1st ed.; O'Reilly y Associates, Inc.: 101 Morris Street, Sebastopol, CA 95472; ISBN 978-0-937175-64-4 ^[44].
- Xu, Q., Tang, C., Wang, X., Sun, S., Zhao, J., Kang, Z. y Wang, X. (2019). An effector protein of the wheat stripe rust fungus targets chloroplasts and suppresses chloroplast function. *Nature Communications*, 10(1), 5571.
- Xiong, Y., Wang, Q., Yang, J., Zhu, X. y Wei, D.Q. (2018). PredT4SE-Stack: prediction of bacterial type IV secreted effectors from protein sequences using a stacked ensemble method. *Frontiers in Microbiology*, 9, 2571 ^[36].
- Xu, Y., Zhan, J., Wijeratne, E.M., Burns, A.M., Gunatilaka, A.A. y Molnár, I. (2007). Cytotoxic and antihaptotactic beauvericin analogues from precursor-directed biosynthesis with the insect pathogen *Beauveria bassiana* ATCC 7159. *Journal of Natural Products*, 70(9), 1467–1471 ^[73].
- Yang, J., Wang, Y., Liu, L., Liu, L., Wang, C., Wang, C. y Li, C. (2019). Effects of exogenous salicylic acid and pH on pathogenicity of biotrophy-associated secreted protein 1 (BAS1)-

-
- overexpressing strain, *Magnaporthe oryzae*. *Environmental Science and Pollution Research International*, 26(14), 13725–13737.
- Yin, J., Gu, B., Huang, G., Tian, Y., Quan, J., Lindqvist-Kreuze, H. y Shan, W. (2017). Conserved RXLR effector genes of *Phytophthora infestans* expressed at the early stage of potato infection are suppressive to host defense. *Frontiers in Plant Science*, 8, 2155.
- Yan, N. (2015). Structural biology of the major facilitator superfamily transporters. *Annual Review of Biophysics*, 44, 257–283 ^[34].
- Yesilirmak, F. y Sayers, Z. (2009). Heterologous expression of plant genes. *International Journal of Plant Genomics*, 2009, 296482.
- Zhao, H., Zhou, T., Xie, J., Cheng, J., Jiang, D., y Fu, Y. (2020). Host transcriptional response of *Sclerotinia sclerotiorum* induced by the mycoparasite *Coniothyrium minitans*. *Frontiers in Microbiology*, 11, 183.
- Zhang, Y., Wei, J., Qi, Y., Li, J., Amin, R., Yang, W., y Liu, D. (2020). Predicating the effector proteins secreted by *Puccinia triticina* through transcriptomic analysis and multiple prediction approaches. *Frontiers in Microbiology*, 11, 538032.
- Zeng, C. y Zou, L. (2019). An account of *in silico* identification tools of secreted effector proteins in bacteria and future challenges. *Briefings in Bioinformatics*, 20(1), 110–129 ^[37].
- Zhang, M., Xie, S., Zhao, Y., Meng, X., Song, L., Feng, H., y Huang, L. (2019). Hce2 domain-containing effectors contribute to the full virulence of *Valsa mali* in a redundant manner. *Molecular Plant Pathology*, 20(6), 843–856.
- Zhang, L., Huang, X., He, C., Zhang, Q. Y., Zou, X., Duan, K., y Gao, Q. (2018). Novel fungal pathogenicity and leaf defense strategies are revealed by simultaneous transcriptome analysis of *Colletotrichum fructicola* and strawberry infected by this fungus. *Frontiers in Plant Science*, 9, 434.
- Zhu, X., Soliman, A., Islam, M.R., Adam, L.R., y Daayf, F. (2017). *Verticillium dahliae*'s isochorismatase hydrolase is a virulence factor that contributes to interference with potato's salicylate and jasmonate defense signaling. *Frontiers in Plant Science*, 8, 399.
-

- Zhang, Y., He, J., Jia, L.J., Yuan, T.L., Zhang, D., Guo, Y., Wang, Y., y Tang, W.H. (2016). Cellular tracking and gene profiling of *Fusarium graminearum* during maize stalk rot disease development elucidates its strategies in confronting phosphorus limitation in the host apoplast. *PLoS Pathogens*, 12(3), e1005485.
- Zhang, Z.N., Wu, Q. Y., Zhang, G. Z., Zhu, Y.Y., Murphy, R.W., Liu, Z., y Zou, C.G. (2015). Systematic analyses reveal uniqueness and origin of the CFEM domain in fungi. *Scientific Reports*, 5, 13032.
- Zou, L., Nan, C., y Hu, F. (2013). Accurate prediction of bacterial type IV secreted effectors using amino acid composition and PSSM profiles. *Bioinformatics* (Oxford, England), 29(24), 3135–3142 ^[25].