

## Breve aproximación a la naturaleza genómica de *Moniliophthora roreri* CPMRT01 aislado de cacao en Tabasco, México

Brief approach to the genomic nature of *Moniliophthora roreri* CPMRT01 isolated from cocoa in Tabasco, Mexico

Enrique Hipólito-Romero<sup>1</sup>, Eliezer Cocoltzi-Vásquez<sup>1</sup>, José M. Ramos-Prado<sup>1</sup>, Cesar Espinoza<sup>2</sup>, Magdiel Torres-de la Cruz<sup>3</sup>, Jorge Ricaño-Rodríguez<sup>1\*</sup>

<sup>1</sup> Centro de EcoAlfabetización y Diálogo de Saberes. Universidad Veracruzana. Av. de las Culturas Veracruzananas s/n, Col. Emiliano Zapata, C.P. 91060. Zona Universitaria, Campus USBI. Xalapa, Veracruz, México.

<sup>2</sup> Centro de Investigación en Micología Aplicada. Universidad Veracruzana. Médicos no. 5, Col. Unidad del Bosque, C.P. 91010. Xalapa, Veracruz, México.

<sup>3</sup> División Académica de Ciencias Biológicas. Universidad Juárez Autónoma de Tabasco. Carretera Villa Hermosa-Cárdenas, Km. 0.5, C.P. 86039. Municipio Centro. Tabasco, México.

### RESUMEN

*Moniliophthora roreri* es un hongo de suma importancia agroeconómica principalmente en el continente americano, ya que es el agente causal de la moniliasis de al menos cuatro especies de cacao. El genoma de este hongo consta de aproximadamente 52,3 Mpb, cuyos genes de interés se reagrupan dependiendo de su naturaleza (e.g. hemibiotróficos, biotróficos y fitopatogénicos). Por otro lado, *Moniliophthora* es capaz de metabolizar proteínas involucradas en procesos de infección, regulación metabólica y mecanismos de defensa. El objetivo principal de este trabajo fue caracterizar un fragmento del genoma de una cepa de *M. roreri*, previamente aislada de cacao en el Estado de Tabasco, México. Para ello, se implementó una estrategia experimental que involucró la generación de bibliotecas genómicas así como una secuenciación de regiones nucleotídicas consenso, predicción heurística-funcional e identificación de hipotéticos dominios conservados, respectivamente. El estudio permitió explorar alrededor del 16 % del genoma de la cepa. Entre los resultados y conclusiones obtenidas más interesantes, se generaron reconstrucciones filogenéticas de máxima identidad respecto a un genoma de referencia y, cuyos transcritos se encontrarían involucrados en procesos de fitopatogenicidad y metabolismo secundario, identificándose aparentes dominios proteicos y sitios catalíticos activos en algunos de los contigios estudiados.

**Palabras clave:** basidiomiceto fitopatógeno; genoma; metabolismo secundario; moniliasis, reconstrucción filogenética

### ABSTRACT

*Moniliophthora roreri* is a fungus of extreme agro-economic importance mainly in the American continent, since it is the causal agent of moniliasis of at least four cocoa species. The genome of this fungus consists of approximately 52,3 Mbp, whose genes of interest may be regrouped depending on their nature (e.g. hemibiotrophics, biotrophics and phytopathogenics). On the other hand, *Moniliophthora* is liable to metabolize proteins involved in infection processes, metabo-

lic regulation and defense mechanisms. The aim of this work was to partially characterize a fragment of the genome of a *M. roreri* strain previously isolated from cocoa in the State of Tabasco, Mexico. To reach this goal, an experimental strategy was implemented that involved the generation of genomic libraries as well as a consensus nucleotide region sequencing, functional-heuristic prediction and identification of hypothetical conserved domains, respectively. The study allowed to explore about 16 % of the strain's genome. Among the most interesting results and conclusions obtained, phylogenetic reconstructions of maximum identity respect to the reference genome were reached, whose transcripts would be involved in phytopathogenic processes and secondary metabolism, identifying apparent protein domains and active catalytic sites in some of the contigs under study.

**Keywords:** genome; moniliasis; phylogenetic reconstruction, phytopathogenic basidiomycete; secondary metabolism

### INTRODUCCIÓN

La naturaleza fitopatógena de los hongos depende en gran medida de su capacidad y modo de interactuar con el organismo huésped. En este sentido, existe una clasificación que delimita tres grupos representativos distintos, que son: fitopatógenos biotróficos, necrotrofoes y hemibiotrofoes (Meinhardt *et al.*, 2014). *Moniliophthora roreri* se considera un hongo basidiomiceto fitopatógeno perteneciente a la familia Marasmiaceae del orden Agaricales (Aime y Phillips-Mora, 2005). El organismo se caracteriza por atacar únicamente frutos de diversas especies de los géneros botánicos *Herrania* y *Theobroma*, por lo que se considera el agente causal de la moniliasis en este último (Figura 1) (Evans, 1981). La infestación en plantas de cacao resulta evidente cuando la especie fitopatógena causa daños necróticos en el tejido interno del fruto, lo que culmina en la mayoría de los casos en la pérdida de éste (Phillips-Mora *et al.*, 2007; Hipólito Romero *et al.*, 2017).

A nivel molecular, existen pocos antecedentes sobre la naturaleza genómica de *M. roreri*. Por mencionar algunos ejemplos importantes; Costa *et al.* (2012) reportaron la se-

\*Autores para correspondencia: Jorge Ricaño-Rodríguez

Correo electrónico: jricano@uv.mx

Recibido: 29 de julio de 2019

Aceptado: 11 de noviembre de 2019



**Figura 1.** Fruto de cacao tipo trinitario infectado por *M. roreri* en una plantación de cacao, Ranchería Galeana segunda sección, Jalpa de Méndez, Tabasco, México. Fotografía tomada por MTC.

**Figure 1.** Fruit of trinitarian type cocoa infected by *M. roreri* in a cocoa plantation, Ranchería Galeana second section, Jalpa de Méndez, Tabasco, Mexico. Photograph taken by MTC.

cuencia del genoma mitocondrial (mt) de esta especie, cuyos resultados se compararon parcialmente con su pariente más cercano *Moniliophthora perniciosa*. Entre muchas características interesantes, el ADNmt de *M. roreri* (94 Kb) posee una topología circular que hipotéticamente codifica 14 genes relacionados con procesos de fosforilación oxidativa. Además de ello, éste transcribe genes tipo ARNr (ARN ribosomal), subunidades proteicas ribosomales, 13 marcos de lectura abierta intrónica (ORFs; Open Reading Frames) y un complemento de 27 genes tipo ARNt (ARN de transferencia). Estudios posteriores permitieron secuenciar y ensamblar 52,3 Mpb representados por 3,298 cóntigos (secuencias consenso de ADN) correspondientes al genoma de una cepa de *M. roreri* proveniente del estado de Los Ríos, Ecuador (Meinhardt *et al.*, 2014). Por su parte, Barbosa *et al.* (2018) realizaron la secuenciación y caracterización molecular de los genomas de seis cepas del género *Moniliophthora* divididas en subpoblaciones empleando plataformas de nueva generación.

Es importante mencionar que parte del genoma de *M. roreri* codifica proteínas relacionadas con fenómenos de retrotransposición, especies reactivas de oxígeno (ROS; Reactive Oxygen Species) y degradación de pared celular. Por otra parte, el considerable número de genes que codifica monooxigenasas citocromo P450, sugiere que el género *Moniliophthora* posee un gran potencial de detoxificación, al igual que síntesis de toxinas y distintos tipos de compuestos indólicos, lo que conferiría una alta capacidad de adaptación ambiental a sus especies fúngicas. El género en mención también es capaz de secretar polipéptidos ricos en cisteína, así como expresar genes involucrados en metilotropía y biosíntesis de hormonas de crecimiento vegetal (*e.g.* giberelinas y auxinas). Los pocos análisis profundos de la familia de genes de *Moniliophthora* sugieren que transcribe cantidades considerables de carboxilesterasas (Mondego *et al.*, 2008).

A la luz de las consideraciones anteriores, este trabajo

tuvo como objetivo principal; caracterizar parte del genoma de una cepa de *M. roreri*, previamente aislada de cacao en el Estado de Tabasco, México. Para lo anterior, se diseñó un conjunto de metodologías que permitieron la construcción de biobancos y su respectiva secuenciación a través de plataformas de nueva generación, incluyendo distintos análisis computacionales que, en conjunto, dieron un vistazo a un fragmento de la naturaleza genómica de la cepa fúngica. De esta forma se develaron parcialmente algunos de sus mecanismos de evolución adaptativa así como de competencia interespecífica, pero principalmente mediante reconstrucciones filogenéticas; la presencia de supuestos genes que estarían relacionados con adaptación al estrés, metabolismo secundario y capacidad necrótica vegetal, cuyos procesos metabólicos son considerados imperativos durante el ciclo de infestación de los hongos hemibiotróficos.

## MATERIALES Y MÉTODOS

### Obtención e identificación de *M. roreri* CPMRT01

La cepa de *M. roreri* utilizada en este estudio fue donada por el Laboratorio de Fitopatología del Colegio de Postgraduados, campus Tabasco, y su aislamiento se llevó a cabo de la manera siguiente: el hongo fue obtenido de frutos de cacao en estado inicial de necrosis externa, procedentes del Estado de Tabasco, México. Una vez en el laboratorio, las drupas se sometieron a un proceso de desinfección externa con hipoclorito de sodio al 2.5 % durante 1 min, seguido de tres lavados con agua destilada estéril. Después de remover la dermis, con la ayuda de un bisturí se transfirieron pequeñas fracciones de tejido interno a cajas de Petri que contenían agar papa dextrosa (Evans, 1981). De esta manera se obtuvieron colonias maduras en un periodo de  $\approx 10$  días y se subcultivaron en placas de agar V8 clarificado, a  $25 \pm 2$  °C. Las colonias fueron identificadas morfológicamente con la ayuda de claves taxonómicas (Phillips-Mora *et al.*, 2006; Evans 1981).

Posteriormente, se produjo biomasa (pellets) mediante fermentación líquida en caldo papa dextrosa incubado durante 10 días a 150 rpm y una temperatura de  $25 \pm 2$  °C. Previamente, se llevó a cabo la caracterización molecular de esta cepa a través de la secuenciación de un transcrito de 853 pb mediante PCR (Reacción en Cadena de la Polimerasa), utilizando ADN genómico como plantilla e iniciadores para regiones ITS (Espaciadores Internos Transcritos) (Torres-de la Cruz *et al.*, 2016) (GenBank: GU108605), cuyos resultados de análisis de máxima identidad mostraron el 100 % respecto a distintas accesiones de *M. roreri* previamente depositadas en el repositorio del NCBI (Centro Nacional para la Información Biotecnológica (<https://www.ncbi.nlm.nih.gov>)).

### Extracción de ADN nuclear

Bajo condiciones asépticas, se separó el micelio del medio de cultivo respectivo con la ayuda de asa bacteriológica, eliminando el exceso de agar y congelando el material biológico mediante vertido de nitrógeno líquido. Se pulverizó por maceración en frío con la ayuda de un mortero.

Posteriormente, se partió de 200 mg de tejido en polvo y se adicionaron 900 µL de buffer de Lin (Lin *et al.*, 2001), 12 µL de β-mercaptoetanol y 2 µL de RNasa (20 mg/mL). Se mezcló perfectamente y se incubó a 37 °C durante 1 h y 600 rpm o agitación suave. Se incorporó un volumen de mezcla de fenol-cloroformo-alcohol isoamílico (25:24:1) y se centrifugó a 12,000 rpm por 10 min a 4 °C, repitiendo este paso dos veces más. Se recuperó el sobrenadante adicionando un volumen de isopropanol frío al 100 % y se incubó a -20 °C durante 1 h. Después, se realizó una centrifugación a 12,000 rpm por 10 min a 4 °C y se decantó el sobrenadante. A la pastilla o cúmulo restante se le adicionaron 500 µL de etanol absoluto al 70 % y se centrifugó a 12,000 rpm por 30 s. Cuidadosamente se eliminó el etanol dejando secar el remanente a temperatura ambiente. Por último, se realizó una hidratación de la pastilla adicionando 30 µL de agua libre de nucleasas (AMBION), se cuantificó por espectrofotometría  $A_{260}/A_{280}$  y se corroboró su integridad mediante electroforesis en gel de agarosa. El material genético se almacenó a -70 °C hasta su uso.

### Generación de bibliotecas genómicas y secuenciación por síntesis

Con el material genético de *M. roreri* extraído previamente, se generaron dos bibliotecas tipo TruSeq® con un promedio de 17,5 ng/µL (insertos cortos) y 18,2 ng/µL (insertos largos tipo mate-pairs). Los pasos seguidos fueron los siguientes: (1) fragmentación de ADN con enzimas de restricción (*Bgl* II y *Dde* I) y generación de insertos de ≈550 pb y ≈7Kbp (fragmentos de ADN de doble cadena con extremos 3' y 5'); (2) reparación de extremos y selección del tamaño de la biblioteca y fragmentación de extremos romos; (3) adenilación de extremos 3' y adición de nucleótidos adenina a los extremos romos 3'; (4) ligación de adaptadores a múltiples extremos de los fragmentos de ADN; (5) enriquecimiento de fragmentos de ADN mediante anillamiento a los extremos de los adaptadores mediante PCR; (6) validación de bibliotecas mediante cuantificación fluorométrica por qPCR (PCR cuantitativa) y verificación de la distribución del tamaño de los fragmentos; (7) normalización y reacomodo en placas de secuenciación diluida (normalización de las bibliotecas genómicas a 10 nM). Finalmente, el proceso de secuenciación se realizó en una plataforma NextSeq 500 (Illumina) con 2 lecturas pareadas. Los archivos de lectura de secuencias cortas (SRA) se depositaron en el repositorio Entrez del NCBI, bajo el número de acceso: PRJNA515757.

### Procesamiento de lecturas crudas de secuenciación

Se desarrolló una estrategia técnica para procesar las lecturas de secuenciación provenientes de los insertos genómicos, la cual constó de cinco pasos: (1) análisis de calidad de 65,685,731 lecturas mediante el software MultiQC (Ewels *et al.*, 2016), tomando en cuenta tanto las secuencias ≈26-140 pb como el % de GC; (2) corte de secuencias filtrando adaptadores de secuenciación y, tomando como nivel de calidad Phred ≤30; 1 error en 1,000 bases secuenciadas ( $Q = \log_{10} P$ ); (3) simulación de ensamblaje continuo de secuencias utilizando

el software ALLPATHS-LG: se programó el algoritmo para que empleara cóntigos de 6,000 pb como máximo y posteriormente se ejecutó el corrector de errores ALLPATHS-LG *error corrector* (parámetros de corrección estándar con eliminación de espacios vacíos o brechas y secuencias repetidas). Con el fragmento del genoma previamente terminado, los cóntigos filtrados se calcularon en función del tamaño conocido del genoma de referencia (MacCallum *et al.*, 2009). La corroboración del proceso de corrección de secuencias y revisión de estadísticas de alineación se realizó con el programa *dnadiff* ver 2.0 (Phillippy *et al.*, 2008). Finalmente se partió de lecturas limpias con valores óptimos; (4) filtrado de lecturas cortas inferiores a 200 nt. Para excluir datos provenientes de posibles contaminantes (i.e. material genético de otros organismos) se utilizó la herramienta BlobPlot (<https://blobtools.readme.io>) (Laetsch y Blaxter, 2017) generando líneas de comando de partición del genoma de *M. roreri* a partir del contenido de guanina y citosina, cobertura de secuenciación en las bibliotecas y máxima identidad de secuencias consenso.

### Estimación de máxima identidad de lecturas procesadas de secuenciación

Basado en los resultados obtenidos de la fase de "control de calidad" de los insertos de ADN secuenciados, se procedió a filtrar las secuencias consenso con un valor  $Q \leq 30$  seguido de la distribución de lecturas (calidad vs longitud de lectura), lo que permitió comparar parte del genoma del organismo sujeto a estudio con un fragmento del genoma (shotgun sequence) de *M. roreri* MCA 2952 (LATX01000001.1) que se encuentra disponible en el repositorio del NCBI (datos sin publicar) y es proveniente del Estado de Chiapas, México. El proceso anterior se realizó con el software ACT; Artemis Comparison Tool (Carver *et al.*, 2005).

### Reconstrucción filogenética de regiones consenso de ADN

Para realizar el estudio de reconstrucción filogenética se emplearon 35 cóntigos que mostraron máxima similitud respecto a proteínas transcritas de los géneros *Moniliophthora* y *Theobroma* previamente depositadas en el repositorio del NCBI y, cuya predicción funcional se correlacionaría con diversos procesos metabólicos de etapas hemi y biotróficas por parte del hongo. Se utilizó en primera instancia el método de unión de secuencias vecinas (Neighbor-Joining; NJ) (Saitou y Nei, 1987) complementado por predicciones de probabilidad máxima compuesta (Tamura *et al.*, 2004) que eliminaron todas las posiciones que contenían brechas (gaps) y datos vacíos, generando un árbol radial con parámetros de re-muestreo (bootstrapping) de 500 réplicas. Por otra parte, a través de una búsqueda heurística y aplicando algoritmos de unión-vecindad y BioNJ a una matriz de distancias estimadas en pares, se realizaron análisis estadísticos de máxima verosimilitud (MLE) bajo el modelo Tamura-Nei (Tamura y Nei, 1993).

Se recurrió también al método de máxima probabilidad de composición (MCL) para seleccionar topologías con un va-



lor de probabilidad de registro superior. Asimismo se generó un agrupamiento jerárquico aglomerado simple con media aritmética (UPGMA; Unweighted Pair Group Method with Arithmetic Mean) (Sneath y Sokal, 1973). Las distancias evolutivas se calcularon utilizando el método de probabilidad máxima compuesta (Tamura *et al.*, 2004) y se representaron como el número de sustituciones de bases nucleotídicas por sitio. Todos los análisis evolutivos anteriores se realizaron con MEGA7 versión 7.0 (Kumar *et al.*, 2016).

### Caracterización de dominios conservados y predicción funcional de regiones consenso de ADN

La búsqueda de hipotéticos dominios conservados de las secuencias consenso de ADN se llevó a cabo con el algoritmo CDD/SPARCLE a nivel de subfamilia y se programó para que éste realizara predicciones aleatorias de marcos de lectura abierta (predicted sequences) (Marchler-Bauer *et al.*, 2017) (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>). Las secuencias de aminoácidos de los 35 cóntigos transcritos se obtuvieron mediante el software EXPASy Translate Tool (<https://web.expasy.org/translate/>) y se identificaron marcos de lectura abierta en orientación 5'-3' así como sitios activos y de enlace, al igual que triadas y residuos catalíticos. De manera complementaria se realizó una alineación múltiple de secuencias de algunos cóntigos basada en restricciones por pares. La predicción heurística funcional se generó mediante comparaciones genómicas de máxima identidad utilizando BLAST (<https://blast.ncbi.nlm.nih.gov>) y tomando como referencia una base de datos con 20,128 secuencias de proteínas reportadas en NCBI para *M. royeri*. Se emplearon de manera complementaria los algoritmos MAKER-GMOD (<http://www.yandell-lab.org/software/maker.html>) así como BUSCO v3 (<https://busco.exlab.org>) (Waterhouse *et al.*, 2017) para analizar las secuencias respecto a su predicción funcional y evolución molecular a partir de copias únicas de genes ortólogos.

## RESULTADOS

### Máxima identidad genómica de *M. royeri* CPMRT01

Mediante el análisis correspondiente, se logró correlacionar el material genético de *M. royeri* CPMRT01 con un 93,11 % de identidad genómica respecto a la cepa de referencia. Aunado a lo anterior, se realizó un análisis BLAST para identificar al resto de los organismos que coincidieron con la muestra y que fueron descartados para este estudio, por lo que se filtraron las secuencias correspondientes a *M. royeri* para seleccionar al resto de posibles organismos homólogos durante el análisis (e.g. *Theobroma cacao*). En la tabla 1 se observa la descripción de cóntigos identificados por su máxima similitud y características contenidas por columna.

### Reconstrucción filogenética

Mediante reconstrucciones filogenéticas se generó un árbol radial óptimo empleando el método NJ (Figura 2) que involucró la alineación de secuencias nucleotídicas de 35 cóntigos que mostraron máxima identidad en los análisis

correspondientes. La suma total de longitud de las ramas fue de 28.6254 sustituciones por sitio, incluyendo las distancias evolutivas para cada una de éstas y tomando en cuenta su probabilidad máxima compuesta. Al eliminar todas las posiciones que contenían espacios faltantes, se generó un conjunto final de datos de 219 nucleótidos distintos. La Figura 2 muestra una reconstrucción que sugiere la presencia de seis grupos principales, que a su vez albergan subgrupos adyacentes conformados por accesiones genéticas previamente anotadas y predichas. La mayoría de las predicciones respectivas para *Moniliophthora* y *Theobroma* sugieren conjunción en subgrupos, inclusive para el caso de las secuencias hipotéticas.

Las figuras 3 (MLE) y 4 (UPGMA) muestran los resultados de los análisis evolutivos moleculares complementarios, respectivamente. La suma total de la longitud de las ramas fueron las siguientes: MLE= -19,207,954.26 y UPGMA= 0.00925867 sustituciones por sitio (probabilidad de registro más alta). Se calculó la divergencia evolutiva con estimadores mínimos para cada caso y en virtud de ello, se observó una reagrupación distinta de las muestras. La filogenia evolutiva obtenida con los métodos de unión de secuencias vecinas (Figura 2) y agrupamiento jerárquico aglomerado (Figura 4) sugiere que los grupos se organizan por pares o subramas primarias dependiendo de los valores obtenidos de los parámetros de re-muestreo replicado. No obstante, mediante el análisis de máxima verosimilitud (Figura 3), se observó un aglomerado en pares de cuatro secuencias consenso (i.e. cóntigos 4478 y 2879; 4961 y 8564) que representa la rama primaria de la filogenia (out group) que da origen a los clados restantes.

### Caracterización de dominios conservados

Respecto a la caracterización de dominios conservados, se observaron similitudes secuenciales en al menos cuatro cóntigos distintos (i.e. 1295, 1609, 1961, 8353) (Figura 5). Los análisis predictivos sugieren la presencia de posibles marcos de lectura abierta pertenecientes a superfamilias de proteínas ricas en cisteína (SCP), peptidasas tipo S8 y transposasas tipo Tnp4, entre otras. También se detectaron dominios de proteínas relacionadas con procesos de fitopatogenicidad, sitios de hibridación y enlace, y ORFs relacionados con el metabolismo secundario de *Moniliophthora*, incluyendo procesos de transcripción y retrotransposición. De manera complementaria, se observaron distintos sitios activos y catalíticos en un rango de 250 a 1000 pb mediante la alineación de 11 secuencias de cóntigos relacionados con *M. perniciososa* y proteínas hipotéticas de *T. cacao* (considerándolo como organismo homólogo). Las principales alineaciones de nucleótidos se predicen entre las 610 y 710 pb a lo largo de los cóntigos (Figura 5).

### Predicción funcional de regiones consenso

La predicción funcional de regiones consenso se realizó preparando una base de datos que constó de 20,128 péptidos reportados para *M. royeri*. El algoritmo Maker-GMOD

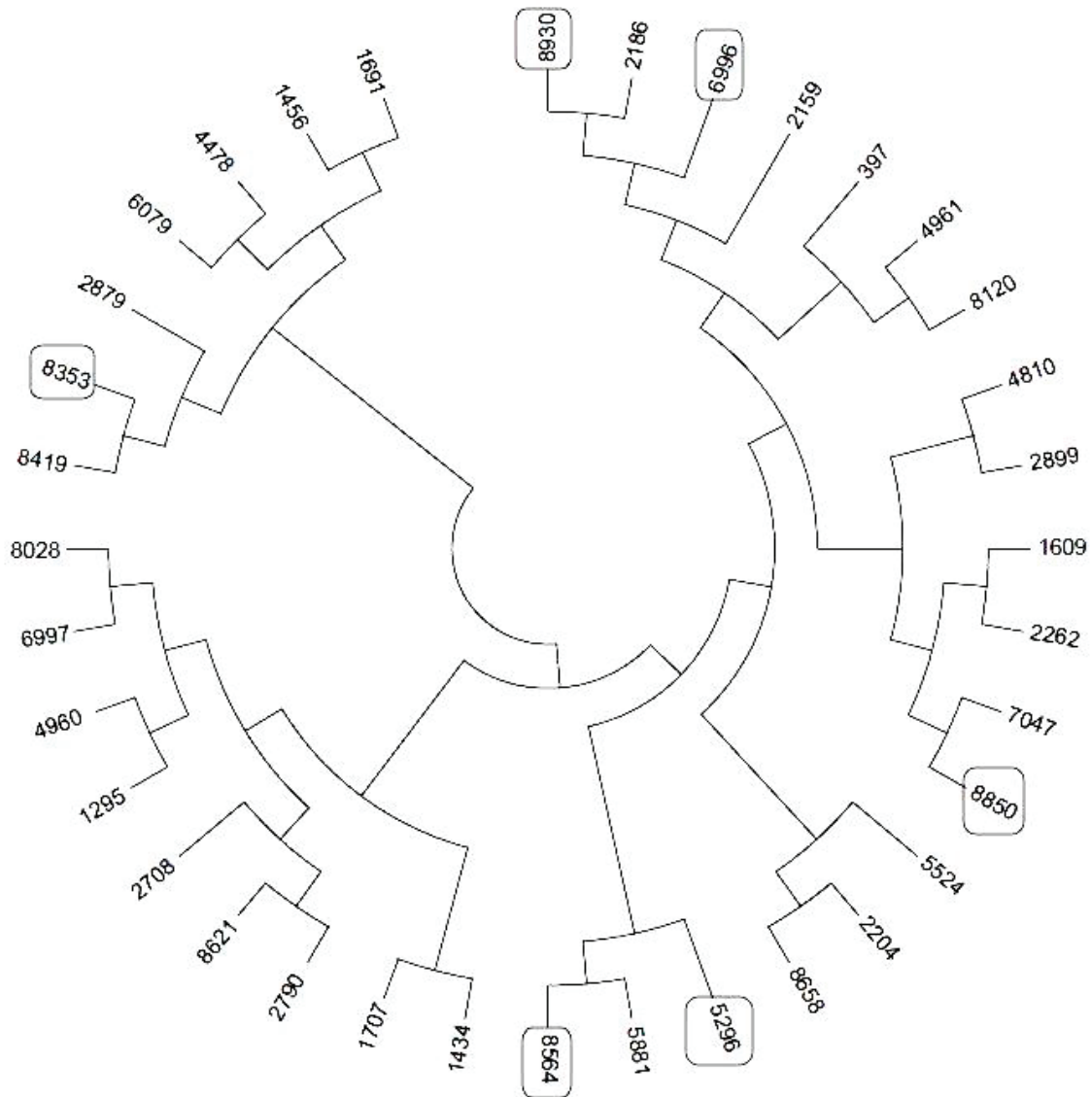
**Tabla 1.** Máximas similitudes identificadas en el genoma de *M. roreri* CPMRT01.

**Table 1.** Maximum similarities identified in the *M. roreri* CPMRT01 genome.

Cóntigo	Número accesión NCBI	% lectura alineada	Longitud nucleótidos alineados	Posición inicial lectura	Posición final lectura	Organismo/predicción
397	gi 397648737 gb JX024739.1	87,8	3,516	7,421	10,910	<i>M. pernicioso</i> methanol oxidase gene
1295	gi 374683156 gb JN620344.1	84,34	645	4,263	4,888	<i>M. pernicioso</i> CP02 PR-1 protein (PR-1e) gene
1434	gi 599126472 gb KJ526207.1	89,68	862	2,144	3,005	<i>M. pernicioso</i> MpSaci26 retrotransposon hypothetical protein
1609	gi 599126451 gb KJ526187.1	80,54	6,012	1	5,955	<i>M. pernicioso</i> MpSaci6 retrotransposon unautonomous hypothetical protein
1456	gi 159786346 gb EU218539.1	85,22	846	1,073	1,917	<i>M. pernicioso</i> transposon Boto hypothetical protein
2204	gi 1042907538 emb LT594789.1	100	28	1,229	1256	<i>T. cacao</i> genome assembly, chromosome: II
4810	gi 1042907684 emb LT594796.1	96,86	542	1	542	<i>T. cacao</i> genome assembly, chromosome: IX
5881	gi 1184802641 gb KY085907.1	100	380	1	380	<i>T. cacao</i> plastid, complete genome
8419	gi 337730156 gb JN127762.1	99,56	228	1	228	<i>T. cacao</i> clone TCC_BA003D04, complete sequence hypothetical protein
8658	gi 337730159 gb JN127765.1	79,06	191	27	215	<i>T. cacao</i> clone TCC_BA049P20, complete sequence hypothetical protein
2790	gi 5991266476 gb kj526211.1	88,24	68	27	94	<i>M. pernicioso</i> Mpsaci30 retrotransposon
2899	gi 599126468 gb kj526203.1	86,67	75	1,469	1543	<i>M. pernicioso</i> CP01 gene (defense protein)
4960	gi 599126463 gb kj526198.1	95,12	41	268	308	<i>M. pernicioso</i> MPsaci29 retrotransposon hypothetical protein
5524	gi 599126475 gb kj5262210.1	100	28	216	243	<i>M. pernicioso</i> MPsaci30 retrotransposon non-autonomous
2262	gi 599126466 gb kj526201.1	74,14	321	1,578	1894	<i>M. pernicioso</i> genome assembly, partial cds
2159	gi 599126448 gb kj526184.1	96,97	33	2,391	2,423	<i>M. pernicioso</i> MPsaci13 retrotransposon
8621	gi 397648737 gb JX024739.1	97,8	1,490	6,425	7,921	<i>M. pernicioso</i> MO hypothetical protein
2708	gi 599126472 gb KJ526207.1	89,56	930	5,144	6,078	<i>M. pernicioso</i> MpSaci27 retrotransposon hypothetical protein
6997	gi 374683156 gb JN620344.1	88,57	2,560	4,321	6,887	<i>M. pernicioso</i> protein (PR-1e) gene
8028	gi 599126463 gb kj526198.1	91,23	230	560	790	<i>M. pernicioso</i> MPsaci25 retrotransposon hypothetical protein
8120	gi 134118907 ref XM_766864.1	98,56	342	115	457	<i>M. pernicioso</i> Mpsaci33 retrotransposon hypothetical protein
1707	gi 159786346 gb EU218539.1	89,21	1,220	1,067	2,300	<i>M. pernicioso</i> Boto hypothetical protein
2186	gi 599126453 gb kj526189.1	95,43	553	89	642	<i>M. roreri</i> hypothetical protein
1691	gi 397648737 gb JX024739.1	89,89	1,427	5,491	6,918	<i>M. pernicioso</i> methanol oxidase hypothetical protein
2879	gi 599126466 gb kj526201.1	86,4	1,314	2,576	3,895	<i>M. pernicioso</i> genome assembly, partial cds
7047	gi 599126473 gb kj526208.1	97,53	283	1	283	<i>M. pernicioso</i> hypothetical protein
4478	gi 599126453 gb kj526189.1	100	28	573	600	<i>M. roreri</i> hypothetical protein
4961	gi 134118907 ref XM_766864.1	97,37	38	88	125	<i>M. pernicioso</i> Mpsaci30 retrotransposon hypothetical protein
6079	gi 599126464 gb kj526203.1	82,73	110	195	303	<i>M. pernicioso</i> Mpsaci36 retrotransposon hypothetical protein
6996	gi 1063457698 ref XM_018129897.1	94,78	115	3	117	<sup>1</sup> PREDICTED: <i>Theobroma cacao</i> vasodilator-stimulated phosphoprotein-like (LOC108663971), partial mRNA
8353	gi 1063465904 ref XM_018114244.1	99,57	230	1	230	<sup>1</sup> PREDICTED: <i>Theobroma cacao</i> subtilisin-like protease SBT1.1 (LOC18614246), mRNA
8850	gi 1063461799 ref XM_007050642.2	100	213	1	213	<sup>1</sup> PREDICTED: <i>Theobroma cacao</i> transcription termination factor
8564	gi 1063509784 ref XM_018123082.1	100	91	77	167	MTERF4, chloroplastic (LOC18613422), mRNA
8930	gi 1063502088 ref XM_018120943.1	92,45	106	4	109	<sup>1</sup> PREDICTED: <i>Theobroma cacao</i> subtilisin-like protease SBT1.1 (LOC18614246), mRNA
5296	gi 1063509680 ref XR_001928432.1	91,04	134	158	290	<sup>1</sup> PREDICTED: <i>Theobroma cacao</i> uncharacterized LOC108661946 (LOC108661946), mRNA
						<sup>1</sup> PREDICTED: <i>Theobroma cacao</i> uncharacterized LOC108662550 (LOC108662550), ncRNA

Nota: Número de accesión NCBI; código del repositorio para futuras consultas, % lectura alineada; se tomó como referencia secuencias del repositorio NCBI.

<sup>1</sup>PREDICTED: referencia de cóntigos anotados funcionalmente que no se encuentran reportados en el NCBI (cóntigos predichos).

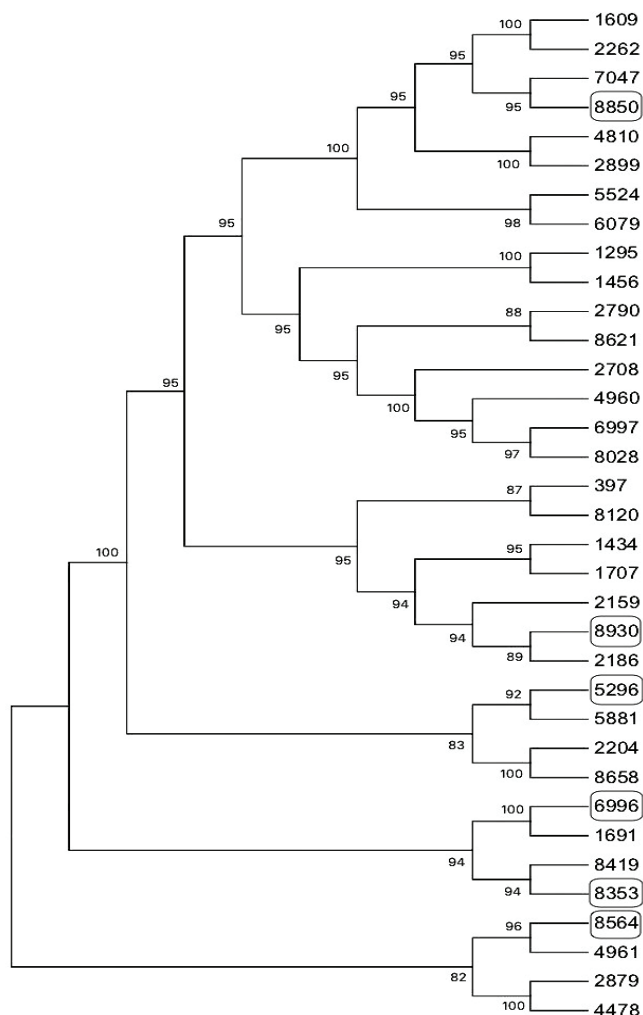


**Figura 2.** Reconstrucción filogenética radial óptima de vecinos cercanos (NJ) (Saitou y Nei, 1987) que muestra las máximas similitudes de secuencias de nucleótidos identificadas en parte del genoma de *M. roreri* CPMRT01. Cada número (etiqueta) corresponde al cóntigo identificado. Los cóntigos predichos se encuentran en recuadros.

**Figure 2.** Optimal radial phylogenetic reconstruction of Neighbor Joining (NJ) (Saitou and Nei, 1987) showing the maximum similarities of nucleotide sequences identified within the *M. roreri* CPMRT01 genome. Each number (label) corresponds to the identified contig. Predicted contigs are found in boxes.

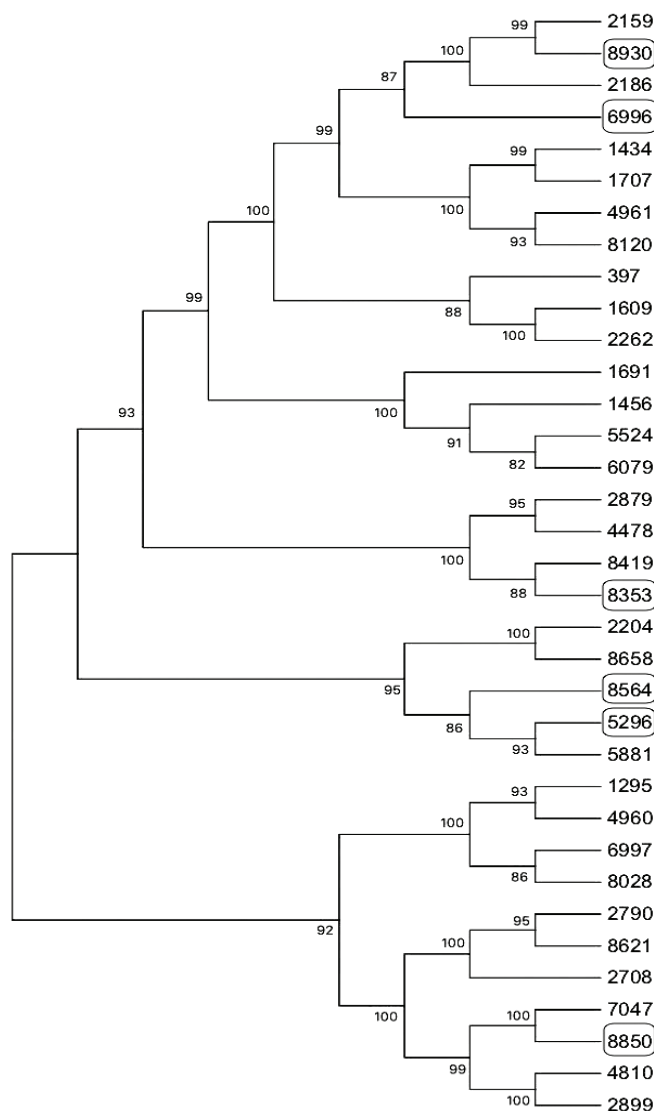
permitió identificar regiones con identidades homólogas respecto a secuencias proteicas descargadas del repositorio correspondiente, así como regiones con repeticiones y predicciones *de novo* sin referencia reportada en el NCBI (Tabla 1). El análisis mostró 424,276 predicciones; equivalentes a

9,142 cóntigos analizados de los cuales 5,420 presentaron cierta evidencia respecto a su predicción heurística y, al menos 35 de ellos sugieren máxima identidad en relación con proteínas tipo PR-1, metanol oxidasas, retrotransposones, fosfoproteínas, proteasas y genes CP01, entre otros (Tabla 1).



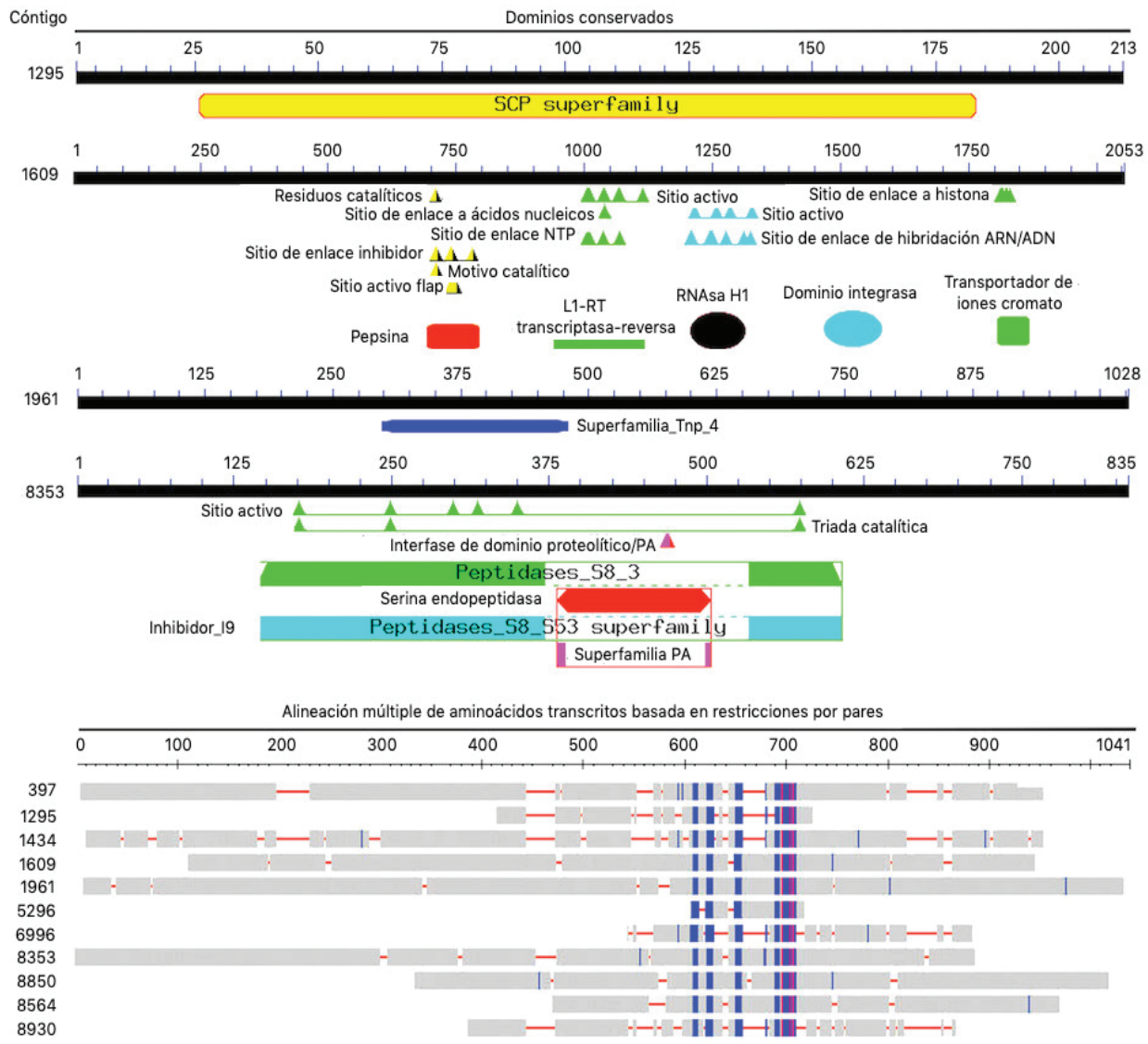
**Figura 3.** Reconstrucción filogenética de máxima verosimilitud (MLE) de secuencias de nucleótidos identificadas en parte del genoma de *M. roleri* CPMRT01 bajo el modelo Tamura-Nei (Tamura y Nei, 1993). La suma total de la longitud de las ramas fue de -19,207,954.26 sustituciones. Los números entre las coyunturas de las ramas representan la supuesta estabilidad de los grupos. Cada número (etiqueta) corresponde al cóntigo identificado. Los cóntigos predichos se encuentran en recuadros.

**Figure 3.** Maximum likelihood phylogenetic reconstruction (MLE) of nucleotide sequences identified within the *M. roleri* CPMRT01 genome under the Tamura-Nei model (Tamura and Nei, 1993). The total sum of the length of the branches was -19,207,954.26 substitutions. Numbers between the joints of the branches represent the supposed stability of the groups. Each number (label) corresponds to the identified contig. Predicted contigs are found in boxes.



**Figura 4.** Reconstrucción filogenética de agrupamiento jerárquico aglomerado simple (UPGMA) (Sneath y Sokal, 1973) de secuencias de nucleótidos identificadas en parte del genoma de *M. roleri* CPMRT01. Se muestra el árbol óptimo con la suma de la longitud de rama = 0.00725912. El porcentaje de árboles replicados en los taxones asociados, agrupados en la prueba de arranque (500 repeticiones) se muestran junto a las ramas. Cada número (etiqueta) corresponde al cóntigo identificado. Los cóntigos predichos se encuentran en recuadros.

**Figure 4.** Phylogenetic reconstruction of single agglomerated hierarchical clustering (UPGMA) (Sneath and Sokal, 1973) of nucleotide sequences identified within the *M. roleri* CPMRT01 genome. The optimal tree is shown with the sum of the branch length = 0.00725912. The percentage of trees replicated in the associated taxa, grouped in the boot test (500 repetitions) are shown next to the branches. Each number (label) corresponds to the identified contig. Predicted contigs are found in boxes.



**Figura 5.** Representación gráfica de la predicción de dominios conservados hipotéticos con máximas similitudes identificados en parte del genoma de *M. roreri* CPMRT01. La alineación inferior múltiple de aminoácidos basada en restricciones por pares muestra 11 transcritos que sugieren similitud con sus secuencias residuales en donde el color azul representa regiones similares y el rojo regiones altamente conservadas.

**Figure 5.** Graphical representation of the prediction of hypothetical conserved domains with maximum similarities identified in part of the genome of *M. roreri* CPMRT01. The multiple lower alignment of amino acids based on peer restrictions shows 11 transcripts that suggest similarity to their residual sequences where the blue color represents similar regions and the red highly conserved regions.

## DISCUSIÓN

En el caso del reino de los hongos, es posible estudiar mediante técnicas moleculares la naturaleza de interacciones fitopatógeno-hospedador, las bases bioquímicas de conidiación, así como rutas metabólicas que conducen a la reproducción sexual, entre otros conceptos fisiológicos. En relación con aspectos filogenéticos de *T. cacao*, el genoma del clon Matina 1-6 se ha anotado en un 99 % respecto a su variabilidad secuencial, observándose un tamaño aproximado de 445 Mpb, lo que resulta considerablemente más grande

que el genoma de la variedad criolla (430 Mpb) (Motamayor *et al.*, 2013). Estos antecedentes permitieron identificar y caracterizar distintos genes de interés biotecnológico y agrícola (Ricaño-Rodríguez *et al.*, 2018) incluyendo al campo de la fitopatología, pues el correcto entendimiento de la organización genómica del cacao develaría el origen de respuestas epigenéticas desencadenadas por la interacción con



su principal antagonista, *Moniliophthora roreri*. Lo anterior se fundamenta en el comportamiento biotrófico de los géneros *Theobroma* y *Moniliophthora*, y en virtud de ello, se abre la posibilidad de que el fenómeno sea estudiado inclusive a nivel transcriptómico (Jones y Dangl, 2006).

Respecto al género *Moniliophthora*, tales antecedentes se complementan con el estudio de su mecanismo de fitopatogenicidad para con el cacao a nivel molecular (Teixeira *et al.*, 2014). Es importante mencionar que Barbosa *et al.* (2018) realizaron la secuenciación genómica de seis subpoblaciones de *M. perniciosa* y *M. roreri*. Los genomas de estas cepas oscilaron entre las 44 y 47 Mpb, lo que representa en promedio el 87.5 % del genoma reportado por Meinhardt *et al.* (2014) y un 13.5 % del genoma de *M. roreri* CPMRT01. A la luz de las consideraciones anteriores y de acuerdo con los resultados de esta investigación referentes a la reconstrucción filogenética de secuencias consenso de ADN la cepa CPMRT01, se concluye que los índices de máxima identidad encontrados corresponden a los géneros *Moniliophthora* y *Theobroma*, pues las predicciones muestran porcentajes >85 % en relación con accesiones de genes tipo PR-1e (estrictamente involucrados en procesos de fitopatogenicidad) (Aime y Phillips-Mora, 2005), metanol oxidasas y retrotransposones, genes CP01 relativos a mecanismos de defensa al igual que fosfoproteínas y proteasas, entre otras (Teixeira *et al.*, 2014).

Asimismo, se observaron similitudes con los géneros *Trichoderma* y *Tremella* del reino Fungi, al sugerir la presencia de transcritos hipotéticos. Así, se sabe que durante su interacción biotrófica con el cacao, el género *Moniliophthora* expresa distintos alelos que codifican transcritos de interés biotecnológico (e.g. citoquininas deshidrogenasas, glicósido hidrolasas, proteínas I3, proteínas ricas en cisteína, peptidasas, transposasas y thaumatinas). Algunos de estos metabolitos se encuentran relacionados con procesos de fitopatogenicidad al igual que transposición replicativa y transcripción, y se involucran en rutas metabólicas que desencadenan respuestas sistémicas en la planta huésped (Jones y Dangl, 2006; Teixeira *et al.*, 2014).

Hoy en día, aún con la atención prestada al estudio de las interacciones moleculares entre los géneros *Theobroma* y *Moniliophthora*, existe relativamente poca información respecto a la naturaleza del genoma y secretoma de este último fitopatógeno fúngico. Meinhardt *et al.* (2014) ensamblaron 52,3 Mb en 3,298 cóntigos que representan el genoma completo de *M. roreri*, prediciendo 17,920 ORF's y validando 13,760 secuencias mediante RNA-seq. De los 1,535 genes que transcriben teóricamente proteínas, 1,355 de estos podrían transcribirse en fases biotróficas y necrotróficas. Igualmente, el análisis de datos revelados por el secretoma de *M. roreri* sugiere correlación con procesos de infección y crecimiento intercelular en la fase biotrófica, al igual que con el crecimiento invasivo y la correspondiente fase necrótica por parte del hospedador.

Mediante nuestro trabajo se predijeron distintos cóntigos en *M. roreri* con altos porcentajes de identidad respecto a secuencias de *T. cacao* (>95 %) (Tabla 1) previamente anotadas y, cuya predicción funcional refiere proteínas aun

sin caracterizar, pero que podrían clasificarse como fosfoproteínas vasodilatadoras, proteasas y factores de transcripción cloroplásticos. Todos ellos involucrados en el metabolismo secundario del hongo. Tiburcio *et al.* (2010) sugieren que el género *Moniliophthora* posee al menos tres genes que han sido adquiridos mediante transferencia horizontal (HGT; horizontal gene transfer) (i.e. (1) proteínas inductoras de necrosis (NIPs; necrosis-inducing proteins), (2) hidrolasas dependientes de metal (MDH; metallo-dependent hydrolases) y (3) deshidrogenasas manitol fosfato (MPDH; mannitol phosphate dehydrogenase).

De acuerdo con esta información y analizando los resultados obtenidos de las respectivas reconstrucciones filogenéticas a partir de los cóntigos ensamblados de *M. roreri* CPMRT01 (Figuras 2, 3 y 4), aquí proponemos "hipotéticamente" que existe la posibilidad de que algunas de las regiones conservadas de los genes estudiados se heredarían de oomicetes, actinobacterias y firmicutes (i.e. con base a las supuestas anotaciones funcionales observadas que se relacionarían con especies diferentes y que son el resultado del proceso evolutivo de genomas que provienen de organismos antecesores comunes). A partir de este antecedente es que se propone la idea de una posible divergencia evolutiva interespecífica. Esto explicaría (aunque de manera muy limitada) la predicción hipotética de proteínas en nuestro trabajo que aun no han sido anotadas y que mostraron un alto porcentaje de similitud respecto a secuencias de aminoácidos de otros géneros, e inclusive, especies de otros reinos. Por otra parte, los aglomerados y subgrupos de taxones observados en los análisis correspondientes sugieren variación en su estructura nucleotídica y, la proximidad de cada componente dependería en gran medida de la naturaleza de dichas regiones nucleotídicas. De ser correcto lo anterior, un solo taxón representaría dentro de la población a los componentes con mayor rastro evolutivo, pues en principio; cada grupo de secuencias consenso forman inclusive clústeres subconjuntos (Ricaño-Rodríguez *et al.*, 2018). Tomando en cuenta dichos resultados, los aglomerados restantes se encuentran directamente emparentadas con los cóntigos 4478, 2879, 4961 y 8564, aunque como se ha mencionado en repetidas ocasiones, las predicciones al ser heurísticas sólo sugieren una aproximación a la naturaleza genómica de la especie (Motamayor *et al.*, 2013).

De los 12 genes tipo PR-1 identificados en el genoma de *M. roreri*, 10 de estos poseen similitud con secuencias descritas en *M. perniciosa* (Teixeira *et al.*, 2012). Estas proteínas se regulan durante la fase biotrófica, aunque también es posible observarlas en la fase necrotrófica. Todas las proteínas tipo MrPR-1 tienen al menos un dominio SCP conservado. El C-terminal de la proteína MrPR-1g que se encuentra tanto en *M. perniciosa* como en *M. roreri* tendrían una función específica en su fase fitopatogénica (Meinhardt *et al.*, 2014), resultado que es consistente con la predicción de ciertos dominios conservados en el genoma de nuestra cepa (Figura 5). Otro dato importante es que; durante la interacción con plantas de cacao, el micelio de *M. roreri* tiene acceso a estructuras plasmodesmáticas y por consiguiente, a los nutrimentos que

pasan a través de ellas. En este sentido, las proteínas tipo PR-1 son importantes para la virulencia, y se ha sugerido que intervienen en la diseminación sistémica del hongo, posiblemente limitando la susceptibilidad de defensa del huésped, actuando como efectoras que suprimen su mecanismo de defensa (Prados-Rosales et al., 2012).

Las proteínas PR-1 interactúan también de manera independiente en las fases biotróficas y necrotróficas. Por otra parte, estos péptidos alteran las paredes celulares de *Theobroma* como si fueran inhibidores competitivos, minimizando la respuesta de defensa de la planta e inclusive mostrarían actividad antimicrobiana, para prevenir infecciones subsiguientes y el desarrollo de microorganismos competidores. En principio, este fenómeno sugeriría el planteamiento de una hipótesis que explique parte de las causas por las cuales *Moniliophthora* ha sido un agente causal de moniliasis en el estado de Tabasco, México, lo cual evidencia su alta capacidad de adaptación epigenética al medio ambiente, apoyando el viejo modelo propuesto de fitopatogénesis (García et al., 2007).

Las proteínas secretadas en la fase biotrófica se asocian principalmente con la descomposición de la matriz intercelular del fruto de cacao, así como con el proceso de modificación de micelios fúngicos, posiblemente para enmascarar al hongo de la defensa de la planta. Igualmente, las proteínas que se hacen presentes en la fase necrotrófica atacan las paredes celulares de su huésped (Ricaño-Rodríguez et al., 2018). En virtud de lo anterior, a través del estudio de predicción de dominios conservados en el genoma de la cepa CPMRT01 se propone la presencia de sitios catalíticos relacionados con enlaces a histonas, ácidos nucleicos y NTP e hibridación ADN/ARN. Todos ellos serían factores moleculares importantes en la fase necrotrófica del hongo (Texeira et al. 2012; Meinhardt et al., 2014) y que soportarían una explicación plausible sobre el origen de la naturaleza virulenta del hongo al menos contra el género *Theobroma*.

Por último, cabe mencionar que el estudio genómico de esta especie fúngica permite desarrollar una mejor comprensión sobre cómo la interacción *Moniliophthora-Theobroma* concluye en una enfermedad vegetal progresiva, lo que se traduce en la necesidad de implementar nuevas herramientas que reduzcan su impacto agrícola en ambos sentidos; tanto económico como sociocultural.

## CONCLUSIONES

Hasta donde los autores sabemos, éste es uno de los primeros trabajos realizados en México donde se propone una aproximación heurística a la naturaleza genómica de una cepa de *M. roreri* aislada de frutos de cacao nativos del Estado de Tabasco, México. La evaluación de calidad de las lecturas de secuenciación generadas mostró distribuciones muy variadas y, los resultados de clasificación para las muestras guardaron relación con el organismo estudiado. Tanto los estudios de reconstrucción filogenética como las hipotéticas predicciones funcionales de regiones conservadas, permitirían identificar posibles alelos que aún no han sido reportados en el repositorio del NCBI, incluyendo proteínas

transcritas tipo RP-1 asociadas a procesos de fitopatogenicidad. Lo anterior aportaría valioso conocimiento sobre la capacidad de adaptación de *Moniliophthora* a distintas condiciones ambientales. Los resultados de caracterización de dominios conservados y su consiguiente evolución molecular taxonómica, sugieren la reagrupación de cóntigos que teóricamente evolucionaron de manera conjunta a partir de un clado primario. Por otra parte, los marcos de lectura abierta de los cóntigos estudiados codificarían en principio; superfamilias de proteínas ricas en cisteína, peptidasas tipo S8 y transposasas tipo Tnp4, por citar a las más relevantes.

## AGRADECIMIENTOS

A la Dirección General de Superación Universitaria del Gobierno de México a través del Programa para el Desarrollo Profesional Docente para el tipo Superior (PRODEP).

## REFERENCIAS

- Aime, M.C., Phillips-Mora, W. 2005. The causal agents of witches' broom and frosty pod rot of cacao (chocolate, *Theobroma cacao*) form a new lineage of Marasmiaceae. *Mycologia*. 97: 1012-1022.
- Barbosa, C.S., Fonseca, R.R. da, Batista, T.M., Barreto, M.A., Argolo, C.S., Carvalho, M.R. de, Amaral, D.O.J., Silvia, E.M. de A., Arévalo-Gardini, E., Hidalgo, K.S., Franco, G.R., Pirovani, C.P., Micheli, F., Gramacho, K.P. 2018. Genome sequence and effectorome of *Moniliophthora perniciosa* and *Moniliophthora roreri* subpopulations. *BMC Genomics*. 19: 509.
- Carver, T.J., Rutherford, K.M., Berriman, M., Rajandream, M.-A., Barrell, B.G., Parkhill, J. 2005. ACT: the Artemis comparison tool. *Bioinformatics*. 16: 3422-3423.
- Costa, G.G.L., Cabrera, O.G., Tiburcio, R.A., Medrano, F.J., Carazzolle, M.F., Thomazella, D.P.T., Schuster S.C., Carlson, J.E., Guiltinan, M.J., Bailey, B.A., Mieczkowski, P., Pereira, G.A.G., Meinhardt, L.W. 2012. The mitochondrial genome of *Moniliophthora roreri*, the frosty pod rot pathogen of cacao. *Fungal Biology*. 116: 551-562.
- Evans, H.C. 1981. Pod rot of cacao caused by *Moniliophthora* (Monilia) roreri. 1<sup>st</sup> Edition. *Phytopathological papers*. Commonwealth Mycological Institute. London, UK.
- Ewels, P., Magnusson, M., Lundin, S., Käller, M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*. 32: 3047-3048.
- García, O., Macedo, J.A., Tiburcio, R., Zaparoli, G., Rincones, J., Bittencourt, L.M., Ceita, G.O., Micheli, F., Gesteira, A., Mariano, A.C., Schiavinato, M.A., Medrano, F.J., Meinhardt, L.W., Pereira, G.A.G., Cascardo, J.C. 2007. Characterization of necrosis and ethylene inducing proteins (NEP) in the basidiomycete *Moniliophthora perniciosa*, the causal agent of witches' broom in *Theobroma cacao*. *Mycological Research*. 111: 443-455.
- Hipólito-Romero, E., Carcaño-Montiel, M.G., Ramos-Prado, J.M., Vázquez-Cabañas, E.A., López-Reyes, L., Ricaño-Rodríguez, J. 2017. Efecto de inoculantes bacterianos edáficos mixtos en el desarrollo temprano de cultivares mejorados de cacao (*Theobroma cacao* L) en un sistema agroforestal tradicional del norte de Oaxaca, México. *Revista Argentina de Microbiología*. 49: 356-365.
- Jones, J.D., Dangl J.L. 2006. The plant immune system. *Nature*. 444: 323-329.

- Kumar, S., Stecher, G., Tamura, K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*. 33: 1870-1874.
- Laetsch, D.R., Blaxter, M.L. 2017. BlobTools: Interrogation of genome assemblies [version 1; peer review: 2 approved with reservations]. *F1000Research*. 6: 1287.
- Lin, R.-C., Ding, Z.-S., Li, L.-B., Kuang, T.-Y. 2001. A rapid and efficient DNA miniprep suitable for screening transgenic plants. *Plant Molecular Biology Reporter*. 19: 379a-379e.
- Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C.J., Lu, S., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwladz, M., Hurwitz, D.I., Lu, F., Marchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C., Geer, L.Y., Bryant, S.H. 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Research*. 45: D200-D203.
- MacCallum, I., Przybylski, D., Gnerre, S., Burton, J., Shlyakhter, I., Gnirke, A., Malek, J., McKernan, K., Ranade, S., Shea, T.P., Williams, L., Young, S., Nusbaum, C., Jaffe, D.B. 2009. ALLPATHS 2: small genomes assembled accurately and with high continuity from short paired reads. *Genome Biology*. 10: R103.
- Meinhardt, L.W., Gilson Costa, G., Thomazella, D.P.T., Teixeira, P.J., Falsarella Carazzolle, M., Schuster, S.C., Carlson, J.E., Guiltinan, M.J., Mieczkowski, P., Farmer, A., Ramaraj, T., Crozier, J., Davis, R.E., Shao, J., Melnick, R.L., Pereira, G.A.G., Bailey, B.A. 2014. Genome and secretome analysis of the hemibiotrophic fungal pathogen, *Moniliophthora roreri*, which causes frosty pod rot disease of cocoa: mechanisms of the biotrophic and necrotrophic phases. *BMC Genomics*. 15: 1-25.
- Mondego, J.M.C., Carazzolle, M.F., Costa, G.G.L., Formighieri, E.F., Parizzi, L.P., Rincones, J., Cotomacci, C., Carraro, D.M., Cunha, A.F., Carrer, H., Vidal, R.O., Estrela, R.C., García, O., Thomazella, D.P.T., Oliveira, B.V., Pires, A.B.L., Rio, M.C.S., Araújo, M.R.R., de Moraes, M.H., Castro, L.A.B., Gramacho, K.P., Gonçalves, M.S., Moura Neto, J.P., Góes Neto, A., Barbosa, L.V., Guiltinan, M.J., Bailey, B.A., Meinhardt, L.W., Cascardo, J.C.M., Pereira, G.A.G. 2008. A genome survey of *Moniliophthora pernicioso* gives new insights into Witches' Broom Disease of cocoa. *BMC Genomics*. 9: 548.
- Motamayor, J.C., Mockaitis, K., Schmutz, J., Haiminen, N., Livingstone, D., Cornejo, O., Findley, S.D., Zheng, P., Utro, F., Royaert, S., Saski, C., Jenkins, J., Podicheti, R., Zhao, M., Scheffler, B.E., Stack, J.C., Feltus, F.A., Mustiga, G.M., Amores, F., Phillips, W., Marelli, J.P., May, G.D., Shapiro, H., Ma, J., Bustamante, C.D., Schnell, R.J., Main, D., Gilbert, D., Parida, L., Kuhn, D.N. 2013. The genome sequence of the most widely cultivated cacao type and its use to identify candidate genes regulating pod color. *Genome Biology*. 14(6): R53.
- Phillippy, A.M., Schatz, M.C., Pop, M. 2008. Genome assembly forensics: Finding the elusive mis-assembly. *Genome Biology*. 9: R55.
- Phillips-Mora, W., Aime, M.C., Wilkinson, M.J. 2007. Biodiversity and biogeography of the cacao (*Theobroma cacao*) pathogen *Moniliophthora roreri* in tropical America. *Plant Pathology*. 56: 911-922.
- Phillips-Mora, W., Coutiño, A., Ortiz, C.F., López, A.P., Hernández, J., Aime, M.C. 2006. First report of *Moniliophthora roreri* causing frosty pod rot (moniliasis disease) of cocoa in Mexico. *Plant Pathology*. 55: 584.
- Prados-Rosales, R.C., Roldan-Rodriguez, R., Serena, C., Lopez-Berges, M.S., Guarro, J., Martinez-del-Pozo, A., Di Pietro, A. 2012. A PR-1-like protein of *Fusarium oxysporum* functions in virulence on mammalian hosts. *The Journal of Biological Chemistry*. 287: 21970-21979.
- Ricaño-Rodríguez, J., Ramos-Prado, J.M., Cocolletzi-Vásquez, E., Hipólito-Romero, E. 2018. El estudio genómico del cacao (*Theobroma cacao* L.); breve recopilación de sus bases conceptuales. *Agroproductividad*. 11(9): 29-35.
- Saitou, N., Nei, M. 1987. The Neighbor-Joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*. 4: 406-425.
- Sneath P.H.A., Sokal R.R. 1973. *Numerical Taxonomy*. 1<sup>st</sup> ed. Freeman, San Francisco.
- Tamura, K., Nei, M., Kumar, S. 2004. Prospects for inferring very large phylogenies by using the Neighbor-Joining method. *Proceedings of the National Academy of Sciences (USA)*. 101: 11030-11035.
- Tamura, K., Nei, M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*. 10: 512-526.
- Teixeira, P.J.P.L., Thomazella, D.P., Reis, O., do Prado, P.F., do Rio, M.C., Fiorin, G.L., José, J., Costa, G.G., Negri, V.A., Mondego, J.M., Mieczkowski, P., Pereira, G.A. 2014. High-resolution transcript profiling of the atypical biotrophic interaction between *Theobroma cacao* and the fungal pathogen *Moniliophthora pernicioso*. *Plant Cell*. 11: 4245-4269.
- Teixeira, P.J.P.L., Thomazella, D.P.T., Vidal, R.O., Do Prado, P.F.V., Reis, O., Baroni, R.M., Franco, S.F., Mieczkowski, P., Pereira, G.A.G., Mondego, J.C.M. 2012. The fungal pathogen *Moniliophthora pernicioso* has genes similar to plant PR-1 that are highly expressed during its interaction with cacao. *Plos One*. 7(9): e45929.
- Tiburcio, R.A., Lacerda Costa, G.G., Falsarella Carazzolle, M., Costa Mondego, J.M., Schuster, S.C., Carlson, J.E., Guiltinan, M.J., Bailey, B.A., Mieczkowski, P., Meinhardt, L.W., Guimarães Pereira, G.A. 2010. Genes acquired by horizontal transfer are potentially involved in the evolution of phytopathogenicity in *Moniliophthora pernicioso* and *Moniliophthora roreri*, two of the major pathogens of cacao. *Journal of Molecular Evolution*. 70: 85-97.
- Torres-de la Cruz, M., Ortiz-García, F.F., Nava-Díaz, C., de la Cruz-Pérez, A. 2016. Confirmación morfológica y molecular de *Moniliophthora roreri*, agente causal de la moniliasis del cacao (*Theobroma cacao*) en Tabasco, México. En: *Perspectiva Científica desde la UJAT*. W.M. Contreras Sánchez (ed.), pp 407-412. Universidad Juárez Autónoma de Tabasco. México.
- Waterhouse, R.M., Seppey, M., Simao, F.A., Manni, M., Ioannidis, P., Klioutchnikov, G., Kriventseva, E.V., Zdobnov, E.M. 2017. BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics. *Molecular Biology and Evolution*. 35: 543-548.