

# UNIVERSIDAD AUTÓNOMA DEL ESTADO DE HIDALGO



## ESCUELA SUPERIOR DE APAN

### TESIS DE LICENCIATURA

Del aislamiento al análisis genómico: caracterización de *Monographella maydis*, hongo del complejo mancha de asfalto del maíz

### Para obtener el título de

Licenciado en Ingeniería en Biociencias

### PRESENTA

Susana Daniela Arteaga Quiroz

### Directores de tesis

Dr. Akshaya Kumar Biswal

### Codirector de tesis

Dr. Martín Peralta Gil

Apan, Hidalgo.

Fecha: Noviembre 11, 2025

**Lugar donde se realizó la tesis**

Este trabajo se realizó en el Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT), en los laboratorios de Sanidad de Semillas y Bioseguridad, con presupuesto de GRP (Genetic Resources Program), bajo la supervisión del Dr. Akshaya Kumar Biswal y el Dr. Amos Emitati Alakonya, a quienes agradezco ampliamente por su confianza y apoyo a lo largo de este camino.

## Autorización de impresión de tesis



Universidad Autónoma del Estado de Hidalgo  
Escuela Superior de Apan  
Campus Apan

Número de control: ESAP-25/2025  
Asunto: Autorización de impresión de tesis

Susana Daniela Arteaga Quiroz  
P.D.L.I.B.C.  
**PRESENTE**

Por este medio le comunico que, el jurado asignado a la pasante de la licenciatura en Ingeniería en Biociencias Susana Daniela Arteaga Quiroz (356824), a quienes les presentó el trabajo para su titulación en la modalidad de tesis con el título “Del aislamiento al análisis genómico: caracterización de *Monographella maydis*, hongo del complejo mancha de asfalto del maíz”, después de la revisión ha autorizado la impresión.

A continuación, se presentan las firmas de conformidad de los miembros de este jurado:

PRESIDENTE: Dr. Martín Peralta Gil

SECRETARIO: Dr. Víctor Hugo Pérez España

PRIMER VOCAL: Dr. Akshaya Kumar Biswal

SEGUNDO VOCAL: Dr. Jaime Alioscha Cuervo Parra

TERCER VOCAL: Dra. Teresa Romero Cortes

SUPLENTE 1: Dra. María Magdalena Armendáriz Ontiveros

SUPLENTE 2: Dra. Lis Tamayo Rivera

ATENTAMENTE

“Amor, Orden y Progreso”

Chimalpa, Apan, Hgo., México a 11 de noviembre de 2025

Dra. María Magdalena Armendáriz Ontiveros  
Coordinador de Ingeniería en Biociencias



Carretera Apan - Calpulapan Km. 8, Colonia Chimalpa  
Tlalayote, Apan, Hidalgo, México C.P. 43900  
Teléfono: 7717172000 Ext. 50901  
esap@uaeh.edu.mx

“Amor, Orden y Progreso”



2025



uaeh.edu.mx

**Dedicatoria**

A mi madre, Eunice Quiroz Melgarejo, por ser mi más constante aliento, mi incondicional apoyo y mi ejemplo cada día. Sin ti no estaría donde estoy. Con la esperanza de que esto recompense en parte todo tu esfuerzo.

A mi novio y amigo, Joel Jimenez Manzanares, por ser mi apoyo y motivación. Gracias por estos años.

A mis amigas, Celeste, Carolina y Jessica, y a mis hermanos Sofia y Santiago. Gracias por el amor incondicional.

A mi padre, Ángel Fernando Arteaga Ortiz, por su apoyo a lo largo de este camino.

## **Agradecimientos**

Agradezco ampliamente al Dr. Akshaya Kumar Biswal, director de este proyecto, ya que confió en mí desde el primer día, y me dio la guía que necesitaba para ser capaz de desarrollarme plenamente en un laboratorio. Gracias por su paciencia y entendimiento, y por siempre estar pendiente de mí a lo largo de estos dos años.

Al Dr. Amos Emitati Alakonya, gracias por confiar en mí para llevar en mis manos este proyecto, sé que estaba tomando el lugar de alguien muy capaz, y aprecio el voto de confianza.

A todas las personas de CIMMYT que me compartieron sus conocimientos y experiencia, gracias por contribuir en mi aprendizaje, y gracias a los que me proporcionaron su amistad, ya que crearon recuerdos y enseñanzas duraderas.

Mención especial al Dr. Martin Peralta Gil, codirector de este trabajo, ya que siempre está ahí en los momentos más difíciles, y siempre cuenta con un consejo útil. Gracias por creer en los jóvenes, y por no darse por vencido a pesar de los obstáculos.

Agradezco enormemente a la Dra. Irma Martínez Flores y al Dr. Luis Fernando Lozano Aguirre Beltrán del CCG-UNAM, ya que, sin su apoyo y guía en la parte bioinformática de este proyecto, esta tesis no sería lo que es.

Un agradecimiento a Libia Fernanda Gómez Trejo, por su guía al inicio de este proyecto. Gracias por ser mi amiga en el momento que lo necesitaba y por considerarme como alguien capaz de continuar con esto.

## Tabla de contenido

Resumen.....	vii
Abstract.....	viii
Glosario de abreviaturas .....	ix
Introducción .....	1
Justificación .....	3
Antecedentes .....	5
Complejo Mancha de Asfalto .....	6
Características morfológicas de <i>Monographella maydis</i> .....	7
Marcadores moleculares .....	9
Impacto de la Mancha de Asfalto.....	9
Distribución geográfica .....	10
Factores ambientales y agronómicos .....	11
Aproximaciones genómicas y aplicaciones.....	12
Observaciones finales.....	13
Hipótesis .....	14
Objetivos .....	15
Objetivo general.....	15
Objetivos específicos .....	15
Materiales y métodos .....	16
Aislamiento de <i>Monographella maydis</i> .....	16
Preparaciones permanentes de estructuras de <i>Monographella maydis</i> ....	18
Preparación de cámaras húmedas.....	18
Aislamiento monospórico .....	19
Confirmación de la identidad de <i>Monographella maydis</i> .....	20

Cultivo y subcultivo de <i>Monographella maydis</i> .....	21
Extracción de ADN de <i>Monographella maydis</i> .....	22
Protocolo para extracción de ADN de <i>Monographella maydis</i> .....	23
Estimación de concentración de ADN.....	25
Preservación de cepas .....	26
Resultados y Discusión.....	28
Secuenciación del genoma completo y ensamblaje <i>de novo</i> .....	28
Estrategia de secuenciación.....	28
Control de calidad del ADN .....	29
Ensamblaje del genoma y evaluación de calidad .....	29
BUSCO.....	30
QUAST .....	31
Aplicaciones downstream.....	32
Predicción génica y anotación funcional .....	33
Anotación estructural y funcional.....	33
Análisis de metabolitos secundarios.....	34
Análisis de ortología .....	35
Análisis Pangenómico y Filogenómico .....	37
Conclusiones .....	40
Referencias.....	42
Anexos .....	48
Anexo A. Glosario .....	48
Anexo B. Taxonomía de <i>M. maydis</i> .....	51
Anexo C. Imágenes de estructuras de <i>M. maydis</i> .....	52
Anexo D. Referencia de ladder usados para los geles de agarosa .....	55

Anexo E. Análisis de K-meros usando GenomeScope .....	56
MNG128 S.....	56
MNG133S.....	57
MNG137S.....	58
Anexo F. Resumen de los resultados de la anotación de <i>M. maydis</i> .....	59
Anexo G. Workflow del análisis bioinformático .....	60
Anexo H. Herramientas bioinformáticas y parámetros.....	62
Anexo I. Salidas de KEGG Mapper mostrando vías metabólicas reconstruidas a partir de anotaciones de EggNOG .....	63



## Resumen

El hongo *Monographella maydis* es un agente asociado al complejo de la mancha de asfalto en maíz, una enfermedad que afecta significativamente la productividad del cultivo. En este trabajo se realizó el aislamiento del hongo a partir de muestras infectadas y la posterior extracción de ADN genómico, el cual fue secuenciado mediante la plataforma PacBio Sequel II. El genoma obtenido fue ensamblado *de novo* utilizando Canu, y la calidad del ensamblaje se evaluó con QUAST y BUSCO, mostrando una alta continuidad. La anotación estructural y funcional se llevó a cabo con el pipeline Funannotate, integrando bases de datos como Pfam, CAZy, MEROPS, UniProt, y EggNOG-mapper para la asignación de funciones y ortología. Además, se identificaron clusters biosintéticos mediante antiSMASH, y se reconstruyeron rutas metabólicas con KEGG Mapper. Finalmente, se realizaron análisis filogenómicos y pangenómicos con especies relacionadas del género *Microdochium* y *Phyllachora*, lo que permitió establecer las relaciones evolutivas y genéticas de *M. maydis*. Este estudio constituye la primera caracterización genómica integral de esta especie, proporcionando una base para futuras investigaciones sobre su biología, patogenicidad y potencial metabólico.

## Abstract

*Monographella maydis* is a fungal species associated with the maize tar spot complex, a disease that significantly impacts crop productivity. In this study, the fungus was isolated from infected maize tissues, and genomic DNA was extracted and sequenced using the PacBio Sequel II platform. The genome was assembled *de novo* with Canu, and assembly quality was assessed with QUAST and BUSCO, revealing high completeness. Structural and functional annotations were performed using the Funannotate pipeline, integrating databases such as Pfam, CAZy, MEROPS, UniProt, and EggNOG-mapper for functional and orthology assignments. Biosynthetic gene clusters were identified using antiSMASH, and metabolic pathways were reconstructed via KEGG Mapper. Comparative pangenomic and phylogenomic analyses with related species from *Microdochium* and *Phyllachora* genera provided insights into the evolutionary relationships and genomic diversity of *M. maydis*. This study represents the first comprehensive genomic characterization of this fungus, offering a valuable foundation for future research on its biology, pathogenicity, and metabolic potential.

## Glosario de abreviaturas

1. TSC. Tar Spot Complex. Complejo Mancha de Asfalto
2. *M. maydis*. *Monographella maydis*.
3. *P. maydis*. *Phyllachora maydis*
4. ITS. Internal Transcribed Spacer. Espaciador Interno Transcrito
5. PCR. Polymerase Chain Reaction. Reacción en Cadena de la Polimerasa
6. Tm. Melting Temperature. Temperatura de Fusión
7. WGS. Whole Genome Sequencing. Secuenciación de Genoma Completo
8. SNVs. Single Nucleotide Variations. Variaciones de Nucleótido Único
9. QUAST. Quality Assessment Tool for Genome Assemblies. Herramienta de Evaluación de Calidad para Ensamblados de Genomas
10. GC. Guanina–Citosina
11. SMS. Single Molecule Sequencing. Secuenciación de Molécula Única
12. LTRs. Long Terminal Repeats. Repetición Terminal Larga
13. CLR. Continuous Long Reads. Lecturas Largas Continuas
14. BUSCO. Benchmarking Universal Single-Copy Orthologs. Ortólogos Universales de Copia Única para Benchmarking
15. NCBI. National Center for Biotechnology Information. Centro Nacional para la Información Biotecnológica
16. BLASTn. Basic Local Alignment Search Tool (nucleotide). Herramienta Básica de Búsqueda de Alineamientos Locales (nucleótidos)
17. CTAB. Bromuro de Cetiltrimetilamonio
18. DMSO. Dimetilsulfóxido
19. RT. Room Temperature. Temperatura Ambiente
20. TE buffer. Buffer Tris–EDTA
21. KEGG. Kyoto Encyclopedia of Genes and Genomes. Enciclopedia de Genes y Genomas de Kioto
22. BGC. Biosynthetic Gene Cluster. Clúster de Genes Biosintéticos
23. KOs. KEGG Orthology. Ortología KEGG
24. MLE. Maximum Likelihood Estimation. Estimación de Máxima Verosimilitud

## Introducción

El maíz (*Zea mays L.*) desempeña un papel fundamental en la seguridad alimentaria mundial, la sostenibilidad económica y la innovación agrícola. Más allá de su importancia agronómica, el maíz tiene una profunda relevancia sociocultural, reflejando una relación milenaria entre el cultivo y los pueblos originarios de las Américas. Su diversidad genética y fenotípica deriva de una larga historia de domesticación, cultivo y selección artificial realizada por más de 80 grupos indígenas en México y Centroamérica (Secretaría de Agricultura y Desarrollo Rural, 2023). Tan solo en México se han registrado 64 razas de maíz, 59 de las cuales se consideran nativas del país (CONABIO, 2022).

Según FAOSTAT, el maíz es el cultivo más producido en el mundo, con una producción que supera los 1.2 mil millones de toneladas y una tasa de crecimiento del 46% desde 2010. Su capacidad de adaptación a una amplia gama de climas, altitudes y tipos de suelo ha permitido su cultivo en casi 170 países, cubriendo aproximadamente 197 millones de hectáreas (FAOSTAT, 2024; Listman & Ordóñez, 2019; Satyanarayana, 2004). Además de su papel como alimento básico, el maíz se utiliza como forraje y como recurso renovable para la producción de biocombustibles. Su valor nutricional, particularmente su contenido de micronutrientes, también contribuye a mejorar la salud pública en regiones en desarrollo (Ranum & Pe, 2014).

A pesar de su importancia agrícola y económica, la producción de maíz enfrenta numerosos estreses bióticos, incluidos plagas de insectos y enfermedades fúngicas. Entre ellas se encuentra el complejo de mancha de asfalto (TSC, por sus siglas en inglés), una enfermedad capaz de reducir los rendimientos hasta en un 50%, lo que representa una seria amenaza para la seguridad alimentaria mundial. TSC es un complejo polimicrobiano asociado principalmente con tres especies de hongos: *Phyllachora maydis*, *Monographella maydis* y *Coniothyrium phyllachorae*, esta última considerada un hiperparásito de *P. maydis* (Ríos Herrera et al., 2007). Sin

embargo, los mecanismos de infección, las interacciones patogénicas y la base genómica de este complejo aún se comprenden de manera limitada.

Entre estos hongos, *Monographella maydis* ha sido particularmente poco estudiado a nivel molecular. Aunque su aislamiento fue reportado por primera vez por Maublanc en 1904, actualmente no existen marcadores moleculares para su identificación rápida, ni recursos genómicos que permitan elucidar su biología. Para cerrar esta brecha de conocimiento, la presente investigación se enfoca en el aislamiento y la caracterización genómica de *Monographella maydis*.

Este estudio abarcó múltiples etapas, incluyendo el aislamiento *in vitro* de *M. maydis*, la extracción de ADN, la secuenciación de lecturas largas mediante la plataforma PacBio Sequel II y el ensamblaje *de novo* del genoma con Canu. Posteriormente, se llevaron a cabo análisis bioinformáticos para evaluar la integridad del genoma (con QUAST y BUSCO), predecir y anotar genes (con Funannotate y EggNOG-mapper), e identificar clústeres de genes biosintéticos y posibles factores de virulencia (mediante antiSMASH y KEGG Mapper). También se realizaron análisis filogenéticos y pangenómicos para explorar las relaciones evolutivas de *M. maydis* dentro de taxones fúngicos relacionados.

En conjunto, este trabajo proporciona la primera anotación genómica de alta calidad e información comparativa sobre *Monographella maydis*. Los datos generados servirán como base para futuros estudios moleculares orientados a comprender su potencial patogénico, sus interacciones con el hospedero y su papel dentro del complejo de mancha de asfalto, contribuyendo finalmente al desarrollo de herramientas diagnósticas y estrategias de manejo de la enfermedad.

## Justificación

El maíz (*Zea mays* L.) es uno de los cereales más importantes del mundo, ya que funciona como alimento básico, fuente de forraje para ganado y materia prima industrial. Sin embargo, su producción se ve amenazada por diversas enfermedades fúngicas, entre las cuales destaca el complejo de mancha de asfalto debido a su creciente incidencia y severidad en las Américas. Esta enfermedad puede causar pérdidas severas en el rendimiento y se ha convertido en una de las principales limitantes para la producción de maíz, especialmente bajo condiciones húmedas favorables.

Aunque *Phyllachora maydis* ha sido identificada como el principal patógeno del complejo, el papel de *Monographella maydis* permanece poco entendido. Su biología, mecanismos de infección y determinantes genómicos no han sido elucidados, lo que limita nuestra capacidad para desarrollar herramientas de diagnóstico precisas y estrategias de manejo efectivas. El conocimiento actual sobre *M. maydis* se restringe a observaciones morfológicas y reportes históricos, sin contar con ninguna caracterización molecular o genómica.

Dada la importancia global del maíz y la naturaleza emergente de la enfermedad de mancha de asfalto, es esencial obtener un entendimiento más profundo de *M. maydis* a nivel genómico. El aislamiento, secuenciación y anotación de su genoma no solo proporcionarán información fundamental sobre su potencial metabólico y virulencia, sino que además permitirán realizar estudios comparativos con otros hongos relacionados. Estos análisis pueden ayudar a identificar genes o clústeres génicos específicos involucrados en la patogenicidad y la interacción con el hospedero, apoyando el diseño futuro de marcadores moleculares para su detección rápida.

Por lo tanto, esta investigación representa un esfuerzo pionero en la caracterización genómica de *Monographella maydis*. Contribuye a llenar un vacío crítico en la genómica fúngica y en la fitopatología, proporcionando información

valiosa que puede guiar estudios futuros sobre epidemiología de la enfermedad, diagnóstico molecular y manejo integrado de enfermedades en maíz.

## Antecedentes

El Complejo Mancha de Asfalto (TSC, por sus siglas en inglés) del maíz ha sido documentado por más de un siglo, siendo reportado por primera vez en 1904 (Maublanc, 1904). Inicialmente, dos hongos—*Phyllachora maydis* y *Coniothyrium phyllachorae*—fueron identificados en México como los agentes causales de la enfermedad. Sin embargo, no fue sino hasta 1984 que *Monographella maydis* fue descrita como un miembro adicional del complejo (Müller & Samuels, 1984). Esta especie se descubrió dentro de las áreas necróticas de hojas de maíz infectadas y se asoció con las características lesiones en “ojo de pescado” que distinguen al TSC de otras enfermedades foliares. El mismo estudio estableció la conexión entre el anamorfo de *M. maydis* (anamorfo—*Microdochium maydis*) y estos síntomas necróticos.

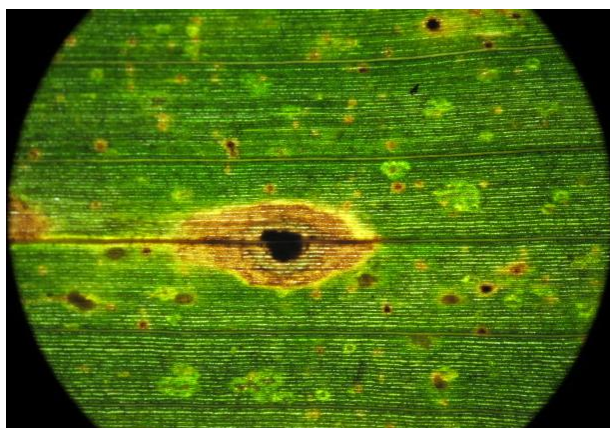


Figura 1. Síntoma de ojo de pescado, visto a través de un estereomicroscopio. Imagen de la autora. Octubre, 2023.

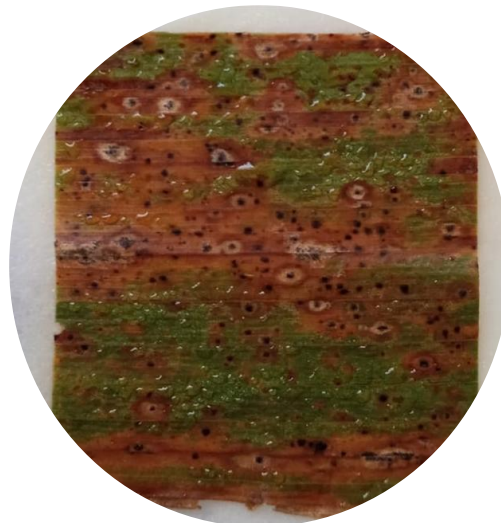
Aunque *M. maydis* ha sido descrita morfológicamente y clasificada taxonómicamente, algunos estudios recientes han cuestionado su papel dentro del TSC. Por ejemplo, un estudio sugirió que *M. maydis* podría no estar directamente involucrada, basándose en la ausencia de aislamientos o detecciones moleculares en ciertas regiones de Norteamérica (Valle-Torres et al., 2020). Una explicación plausible es que *M. maydis* aún no se ha dispersado hacia las zonas más septentrionales del continente, donde predominan únicamente infecciones causadas por *P. maydis*. Estos reportes contrastantes enfatizan la necesidad de un



entendimiento molecular más profundo de *M. maydis*, particularmente mediante estudios genómicos que permitan clarificar su papel dentro del complejo.

### Complejo Mancha de Asfalto

El TSC surge de la interacción sinérgica entre *Phyllachora maydis* y *Monographella maydis* (Hock et al., 1992). Típicamente, *P. maydis* inicia la infección formando los estromas negros y brillosos en la superficie de la hoja. La colonización subsecuente por *M. maydis* incrementa dramáticamente la severidad de la enfermedad, llevando a la formación de las lesiones en “ojo de pescado”: halos necróticos que rodean el punto central de la mancha de asfalto. Inoculaciones experimentales han demostrado que, aunque *P. maydis* puede infectar al maíz por sí sola, la coinfección con *M. maydis* intensifica el desarrollo de los síntomas. Las lesiones en ojo de pescado aparecen aproximadamente entre 2 y 7 días después de la colonización por *P. maydis*, seguidas de una degradación extensa del tejido mediada por *M. maydis* (CIMMYT, 2018). Así, aunque *P. maydis* inicia la infección, *M. maydis* es responsable de la expansión de la lesión, la necrosis y las severas pérdidas de rendimiento asociadas al TSC.



*Figura 2. Hoja con ojos de pescado hidratados.*  
Imagen de la autora. Octubre, 2023.

## Características morfológicas de *Monographella maydis*

*M. maydis* pertenece al filo Ascomycota, caracterizado por la producción de hifas septadas y estructuras sexuales conocidas como peritecios (Tronsmo et al., n.d.). El teleomorfo (*Monographella maydis*) forma peritecios en forma de frasco con ostiolos que liberan ascas con ocho ascosporas hialinas y triseptadas. Su anamorfo (*Microdochium maydis*) produce conidios hialinos, triseptados, que se agrupan en esporodochios. Estas estructuras son inicialmente claras, pero se oscurecen conforme maduran. La descripción de Müller y Samuels (1984) continúa siendo la referencia morfológica más completa de esta especie y ha servido como guía para su identificación en investigaciones subsecuentes.

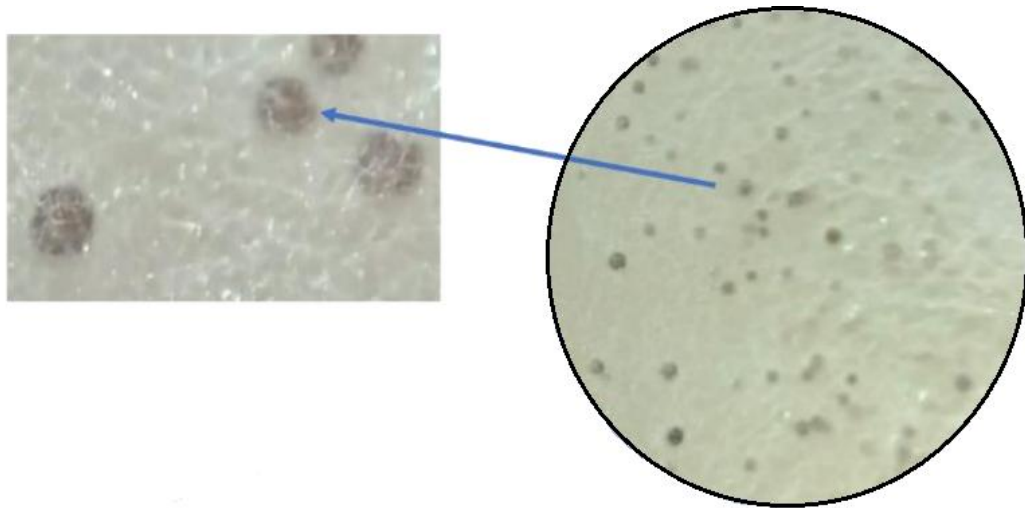


Figura 3. Peritecios vistos a través del estereomicroscopio.

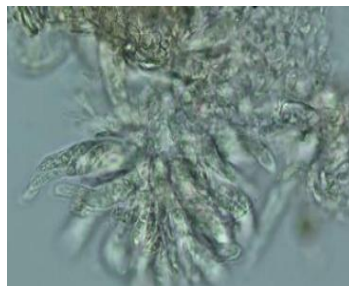


Figura 4. Peritecio abierto; se muestra las ascas que contienen las esporas, visto a través del microscopio compuesto. Imagen de la autora. Octubre, 2023.

En su descripción original (Müller & Samuels, 1984), este hongo fue colocado en el género *Monographella* (Petrak, 1924), describiéndolo de la siguiente manera:

“Un hongo [...] pirenomiceto<sup>1</sup> fue encontrado dentro de las áreas necróticas. Sus ascomas eran inicialmente de color claro pero se tornaban café con la edad; las ascas eran unitunicadas y presentaban un anillo apical amiloide. Cada asca contenía ocho ascosporas incoloras y fragmosporas. Los ascomas estaban acompañados por esporodoquios claros de los cuales surgían células conidiogénicas anelídicas que producían conidios estrechamente clavados a cilíndricos, hialinos y fragmosporos. Esta especie fue fácilmente colocada en *Monographella* (Petrak, 1924) y proponemos *M. maydis* sp. nov. El anamorfo asociado era una especie de *Microdochium* H. Sydow, similar al anamorfo de *Monographella nivalis*...” (Müller & Samuels, 1984).

De aquellos aislamientos mencionados (1984), no existieron preparaciones permanentes, lo que ha llevado a que *Monographella maydis* sea considerada una especie putativa (Mccoy et al., 2019), ya que no existía evidencia adicional de que hubiese sido realmente aislada y subcultivada.

En el presente proyecto, se realizaron preparaciones permanentes de estas estructuras para referencia futura. En los anexos pueden observarse con detalle imágenes del anamorfo y el teleomorfo de *M. maydis*.

En este estudio, estas características morfológicas fueron confirmadas mediante observación microscópica y preparaciones permanentes en portaobjetos, lo que corroboró la identidad de *M. maydis* aislada de hojas de maíz infectadas que mostraban lesiones de ojo de pescado. La identificación se respaldó adicionalmente mediante análisis molecular utilizando marcadores ITS, permitiendo una confirmación fiable del hongo antes de proceder con los análisis genómicos.

---

<sup>1</sup> Palabra en desuso que hacía referencia a una familia de hongos dentro del orden *Xylariales*.



*Figura 5.* Asco aislado visto a través del microscopio compuesto; las ocho ascosporas están perfectamente visibles dentro del asco. En la misma imagen, otro asco (casi vacío) está visible; éstos son unitunicados.

### **Marcadores moleculares**

La identificación precisa de especies fúngicas puede ser un reto debido a sus ciclos de vida complejos y variabilidad morfológica. Los marcadores moleculares, particularmente las regiones ribosomales ITS, se han convertido en herramientas esenciales para la identificación a nivel de especie en hongos (Martin & Rygioewicz, 2005). Sin embargo, en el caso de *M. maydis*, los recursos moleculares son escasos, ya que no existen genomas publicados ni marcadores específicos disponibles para diagnóstico. Esta limitación subraya la importancia de generar un ensamblaje genómico *de novo* que sirva como referencia para estudios futuros. Tal información genómica permitirá análisis comparativos con especies relacionadas, la identificación de posibles genes de virulencia y el diseño de primers específicos para su detección rápida.

### **Impacto de la Mancha de Asfalto**

Aunque *P. maydis* por sí sola causa daños limitados, la infección combinada con *M. maydis* resulta en reducciones severas del rendimiento que pueden alcanzar hasta un 50% o más. El momento de la infección es crucial: cuando el TSC se desarrolla antes de la floración, las pérdidas pueden variar entre 30% y 100%,

especialmente bajo condiciones de humedad. Infecciones después de la polinización aún provocan reducción en el llenado del grano y pérdidas cercanas al 50% (INTAGRI, 2016). Estos efectos hacen del TSC una de las enfermedades del maíz más importantes económicamente en América Latina y una preocupación emergente en Norteamérica.

El TSC puede aparecer en cualquier etapa del ciclo de vida de la planta, y esto determina el grado de impacto. Cuando surge antes de la floración, las pérdidas pueden alcanzar entre el 30 y el 100%; generalmente, esto ocurre cuando existe abundancia de lluvia. La aparición de la mancha durante la floración ocasiona una formación inadecuada de la inflorescencia, lo que puede resultar en la pérdida total del cultivo. Si la inoculación ocurre después de la floración, las pérdidas tienden a ser alrededor del 50%, dado que el grano no completa su desarrollo y permanece chupado (Figura 6) (INTAGRI, 2016).



*Figura 6. Efecto de TSC en hoja y mazorca. (Imagen: CIMMYT).*

## **Distribución geográfica**

El complejo mancha de asfalto es de gran importancia en diversos países del continente americano, como México, Guatemala, El Salvador, Colombia, Nicaragua, Honduras, Bolivia, Costa Rica, República Dominicana, Panamá, Perú, Puerto Rico y Venezuela (Hock et al., 1992; Sol Hernández, 2018), y más recientemente en el Medio Oeste de los Estados Unidos, donde se reportó por primera vez en 2016 (Valle-Torres et al., 2020).



Figura 7. Presencia de TSC de acuerdo a la literatura existente.

Los Estados Unidos son el mayor productor de maíz del mundo, con 384 millones de toneladas en 2021 (Orús, 2023). En 2018, hasta 30 millones de acres pudieron haber sido afectados, lo que representa aproximadamente una tercera parte de la producción nacional. Incluso si las pérdidas fueran mínimas, una reducción del 1% en la producción significaría una pérdida económica de \$231,600,000 USD (CIMMYT, 2018).

Dado que muchos países en desarrollo dependen de las importaciones de maíz provenientes de los Estados Unidos, la expansión del TSC plantea preocupaciones significativas para la seguridad alimentaria (Satyanarayana, 2004).

## Factores ambientales y agronómicos

La diseminación del complejo, aparentemente controlada por *M. maydis*, depende de las condiciones que permiten la proliferación de este hongo. Además de requerir casi obligatoriamente la presencia de *P. maydis* (que actualmente se encuentra en pocas localidades, limitando el nicho de *M. maydis*), necesita rangos de temperatura entre 17 y 22 °C, una humedad relativa superior al 75% y suelos

con un nivel freático alto (INTAGRI, 2016). Por lo tanto, podría pensarse que no existen razones para considerar al TSC como una amenaza importante.

Sin embargo, el cambio climático podría expandir la enfermedad hacia nuevas áreas productoras de maíz. Otro riesgo es la globalización y el intercambio comercial de cultivos. Este riesgo podría disminuirse importando y exportando maíz sin hojas, dado que estas constituyen el principal medio de propagación del hongo; además, el riesgo de infección desde el grano contaminado hacia la planta es mucho menor.

Se cree que el inóculo inicial proviene de estromas que pasan el invierno en hojas en descomposición de ciclos agrícolas previos o en residuos dejados en los campos (Groves et al., 2020).

### **Aproximaciones genómicas y aplicaciones**

Los avances en la secuenciación de ADN han transformado la fitopatología y la sistemática de hongos al permitir una identificación precisa, reconstrucciones filogenéticas confiables y la anotación funcional de genomas. La secuenciación del genoma de *M. maydis* permite explorar sus posibles factores de virulencia, clústeres de biosíntesis de metabolitos secundarios y genes involucrados en la patogenicidad. La genómica comparativa con otros ascomicetos relacionados puede revelar adaptaciones únicas o mecanismos que facilitan su asociación con *P. maydis* dentro del complejo de mancha de asfalto.

Asimismo, contar con registros de genomas de referencia de diversas especies puede ser útil para identificar con precisión un organismo, ya que la identificación exclusivamente morfológica no siempre garantiza certeza, dado que muchas estructuras pueden ser similares entre especies completamente diferentes (evolución convergente). La extracción de ADN de un organismo no identificado, su secuenciación y la comparación con un genoma de referencia constituye un método

mucho más preciso de identificación, aunque más costoso; pero si la precisión se prioriza, esta sería la vía más adecuada (read mapping).

## **Observaciones finales**

A pesar de que ha pasado más de un siglo desde el primer reporte del complejo mancha de asfalto, muchos aspectos de la biología de *Monographella maydis* y su contribución al desarrollo de la enfermedad permanecen sin resolver. Aunque los estudios morfológicos proporcionaron las bases para su identificación, los datos genómicos son esenciales para comprender sus mecanismos moleculares de patogenicidad y su interacción con *Phyllachora maydis*. La generación y el análisis del genoma de *M. maydis* representan, por lo tanto, un paso crucial hacia la comprensión de sus relaciones evolutivas, su potencial de virulencia y su relevancia biotecnológica. Las siguientes secciones de este trabajo describen las estrategias experimentales, los análisis bioinformáticos y los enfoques comparativos utilizados para alcanzar estos objetivos.



## **Hipótesis**

*Monographella maydis* posee características genómicas únicas y conjuntos de genes específicos relacionados con la virulencia y la adaptación metabólica que lo distinguen de otros hongos asociados al complejo de la mancha de asfalto del maíz. La secuencia y anotación de su genoma ofrecerán nuevos conocimientos moleculares sobre su potencial patogénico y sus relaciones evolutivas.

## **Objetivos**

### **Objetivo general**

Aislar, secuenciar y realizar un análisis genómico comparativo e integral de *Monographella maydis*, organismo asociado al complejo mancha de asfalto del maíz, con el fin de dilucidar su composición genética, sus posibles factores de virulencia y sus relaciones evolutivas con especies fúngicas relacionadas.

### **Objetivos específicos**

1. Aislar *M. maydis* de hojas de maíz que presenten síntomas del TSC y confirmar su identidad
2. Estandarizar el cultivo in vitro de *M. maydis*
3. Ensamblar y anotar el genoma de *M. maydis*

## Materiales y métodos

### Aislamiento de *Monographella maydis*

Hojas de maíz que presentaban tanto síntomas de mancha de asfalto como de “ojos de pescado” fueron colectadas en la Estación Experimental del CIMMYT en Poza Rica, Veracruz, México, durante el segundo ciclo agrícola de 2022. El sitio fue seleccionado por sus condiciones climáticas favorables para el desarrollo de la enfermedad y por la persistencia del hongo en el suelo en forma de estromas (INTAGRI, 2016).

Las muestras foliares recolectadas se colocaron en cámaras húmedas para favorecer la rehidratación de los estromas y la esporulación del hongo. Las lesiones necróticas fueron examinadas diariamente con un estereomicroscopio y se extrajeron los estromas de forma aséptica. Tras aproximadamente dos semanas, se observaron ascosporas liberadas desde peritecios formados sobre la superficie foliar; estas fueron aisladas y transferidas a placas de Petri con medio estéril utilizando una aguja fina y estéril.



Figure 8. Zona de muestreo para los aislamientos.

Se realizaron observaciones microscópicas para la identificación de los aislamientos. Las características morfológicas confirmaron la presencia de esporodoquios atribuibles a *M. maydis*, claramente diferenciables de los de *Fusarium spp.*, que ocasionalmente estaban presentes en los mismos tejidos. Aunque *Microdochium maydis* (anamorfo de *Monographella maydis*) es morfológicamente similar a *Fusarium spp.*, rasgos distintivos en la morfología de los

esporodocios permitieron una diferenciación precisa mediante microscopía compuesta (Figuras 9 y 10).

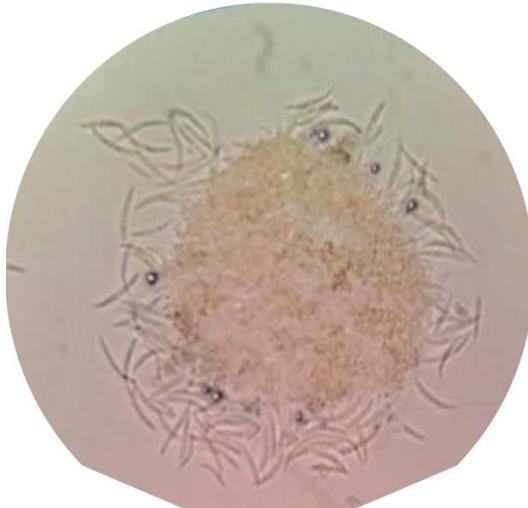


Figure 9. Esporodocio de *Fusarium* spp. und in encontrado en la hoja donde *M. maydis* fue también aislado.



Figure 10. Esporodocio de *Microdochium* maydis.

Tras la identificación, se realizaron aislamientos monospóricos para garantizar la pureza clonal, siguiendo el protocolo estandarizado descrito en la sección siguiente. Los cultivos monospóricos se inocularon en diversos medios —PDA, PDAn, CLA y agar V8— para evaluar condiciones óptimas de crecimiento. El mejor desarrollo micelial y esporulación se observó en medio V8, por lo que este fue seleccionado para el cultivo y subcultivo continuos del organismo.

Tres aislamientos (placas #128, #133 y #137), todos derivados de la muestra foliar #86, mostraron crecimiento consistente y produjeron peritecios blancos *in vitro*. Estos cultivos se utilizaron posteriormente para la extracción de ADN. Los peritecios de estas placas se preservaron sobre papel filtro estéril a 4 °C para uso futuro. Para conservar las características observadas durante la identificación morfológica y disponer de una referencia diagnóstica a largo plazo, se prepararon preparados permanentes de las estructuras de *M. maydis* siguiendo el procedimiento descrito a continuación.

### **Preparaciones permanentes de estructuras de *Monographella maydis***

Para preservar y documentar las estructuras microscópicas observadas durante la identificación morfológica, se prepararon portaobjetos permanentes según el siguiente procedimiento:

1. Una vez confirmada la estructura fungosa deseada (por ejemplo, conidios, esporodoquios o peritecios) al microscopio, se preparó una suspensión homogénea en agua destilada estéril, o la estructura se colocó directamente en una gota de agua estéril sobre un portaobjetos limpio.
2. Se examinó la muestra con un estereomicroscopio para asegurar su correcta colocación.
3. Los portaobjetos se dejaron secar al aire por completo y, posteriormente, se fijaron al calor pasando brevemente sobre la llama de un mechero Bunsen para asegurar el material sobre el vidrio.
4. Las muestras se tiñeron con una solución de azul de metileno:agua destilada (1:4), aplicando una gota directamente sobre el material fijado.
5. Tras 2 minutos de tinción, los portaobjetos se enjuagaron suavemente con un flujo mínimo de agua destilada estéril para eliminar el exceso de colorante y se dejaron secar al aire.
6. Se aplicó un anillo de parafina alrededor de la muestra para crear un sello.
7. Se colocó un cubreobjetos con cuidado encima y se calentó ligeramente para asentar el montaje y obtener una preparación estable.
8. Los portaobjetos fueron rotulados con el número de cepa y la fecha, y almacenados en cajas portaobjetos a temperatura ambiente como referencia a largo plazo y para documentación.

### **Preparación de cámaras húmedas**

Para promover la rehidratación y esporulación de estromas provenientes de tejidos de maíz infectados, se prepararon cámaras húmedas de la siguiente manera:

1. Cortar cuadrados de tejido de 1 × 1 pulgada que contengan lesiones con halos característicos.
2. Colocar cada sección de tejido sobre papel filtro estéril y humedecido dentro de placas Petri.
3. Situar las placas (sin sellar con parafilm) dentro de una caja de plástico cuyo fondo esté forrado con toallas de papel humedecidas para mantener alta humedad sin generar condensación excesiva.
4. Revisar las placas diariamente; los estromas llenos de peritecios se hincharán gradualmente por la hidratación.
5. Bajo estereomicroscopio, extraer un estroma con una aguja fina y montarlo en una gota de agua destilada sobre un portaobjetos.
6. Observar en microscopio compuesto para verificar la presencia de estructuras morfológicas compatibles con *M. maydis* descritas en la literatura.
7. Una vez confirmadas, proceder al aislamiento monospórico para obtener cultivos puros.

### **Aislamiento monospórico**

Para garantizar cultivos fúngicos puros, se realizaron aislamientos monospóricos según el siguiente procedimiento:

1. Con una aguja de insulina rígida y bajo estereomicroscopio, recoger un único esporodoquio de la cultura fúngica.
2. Transferirlo a un tubo microcentrífuga de 1.5 mL que contenga 1 mL de agua destilada estéril.
3. Poner en vortex brevemente para que el esporodoquio libere los conidios.
4. Cuando la suspensión se vea homogénea, pipetear 20 µL y extender uniformemente sobre una placa de Petri con medio de cultivo estéril (en esta etapa puede usarse cualquier medio disponible).
5. Bajo el microscopio, localizar un conidio bien aislado.
6. Cortar el trozo de agar que contiene el conidio con un bisturí estéril.

7. Transferir este fragmento a una nueva placa de Petri con el medio específico (agar V8) e incubar hasta el crecimiento.
8. Monitorear diariamente las placas para detectar contaminación; descartar inmediatamente cualquier placa contaminada para mantener la integridad del cultivo.

### **Confirmación de la identidad de *Monographella maydis***

La región ITS (Internal Transcribed Spacer) se emplea como “código de barras” molecular para la identificación y clasificación de hongos por su alta variabilidad dentro del cluster del ADN ribosómico (Martin & Rygioewicz, 2005). Esta variabilidad hace que las secuencias ITS sean un excelente marcador molecular para diferenciar especies y por lo tanto se utilizó la amplificación y secuenciación del ITS para reconfirmar la identidad de los aislamientos de *M. maydis*.

Se empleó ADN genómico previamente extraído del micelio de los aislamientos seleccionados como templado para PCR. Se utilizaron los primers ITS1 (5'-TCC GTA GGT GAA CCT GCG G-3') e ITS4 (5'-TCC TCC GCT TAT TGA TAT GC-3') con una temperatura de anillamiento ( $T_m$ ) de 64 °C. Se evaluaron dos kits comerciales de polimerasa (Phusion y GoTaq) para comparar rendimiento; la reacción utilizada para las muestras que se secuenciaron se realizó con el kit Phusion.

Los productos de PCR se visualizaron mediante electroforesis en gel de agarosa al 1% y las bandas objetivo se cortaron y purificaron con el kit QIAquick Gel Extraction Kit. Los productos purificados, con concentraciones entre 70–82 ng/μL y valores de 260/280 entre 1.82 y 1.91, fueron enviados para secuenciación a Macrogen Inc. (Seúl, Corea).

Las secuencias obtenidas fueron analizadas mediante BLASTn (NCBI), comparando con organismos de *Monographella spp.* El amplicón ITS verificado se empleó posteriormente como control positivo en ensayos de PCR comparativos,

usando como control negativo *Aspergillus flavus*. Los amplicones de los aislamientos 128, 133 y 137 presentaron la misma longitud que el control positivo, mientras que el control negativo mostró un producto ligeramente mayor, lo cual respalda la confirmación molecular de los aislamientos de *M. maydis* empleados en este estudio.

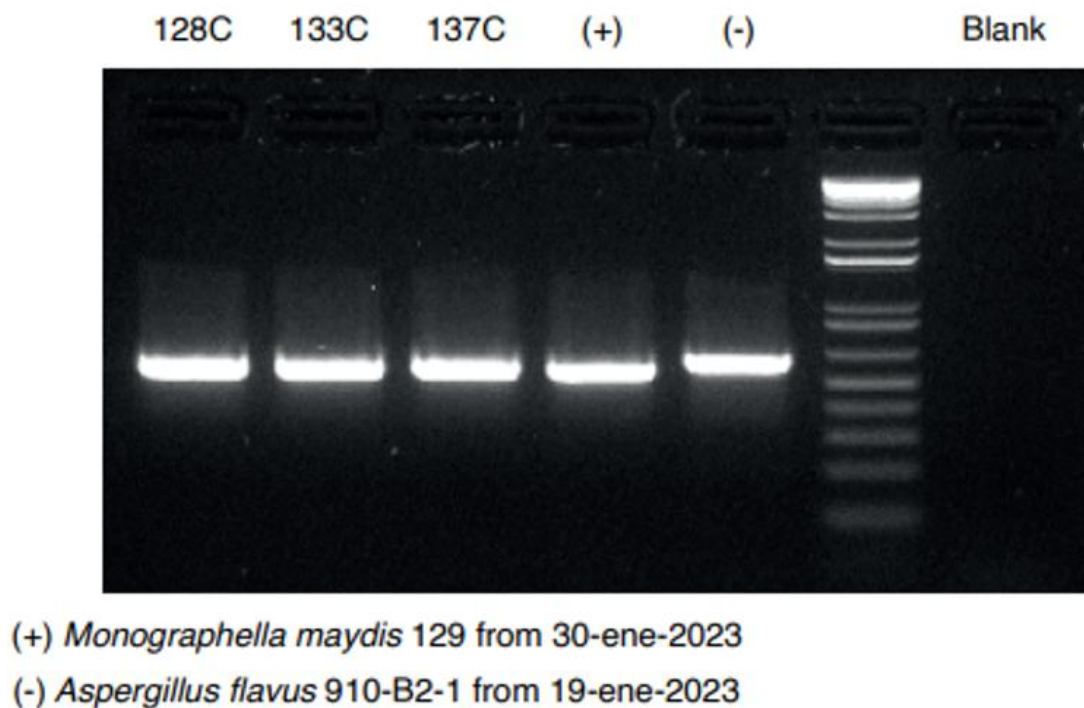


Figura 11. Amplicones de la PCR con primers ITS1/ITS4. Los primeros 3 pozos son de las cepas seleccionadas para el experimento. Se muestran también los amplicones de los controles positivo y negativo.

## Cultivo y subcultivo de *Monographella maydis*

Las tres cepas (128, 133 y 137) preservadas sobre papel filtro estéril fueron reactivadas por subcultivo en medio V8. Se realizaron subcultivos periódicos para mantener la viabilidad de las cepas y generar micelio suficiente para la extracción de ADN. Las placas se incubaron en condiciones controladas (fotoperíodo 12 h luz / 12 h oscuridad a 25 °C) hasta observar crecimiento micelial abundante y formación de peritecios.



## **Extracción de ADN de *Monographella maydis***

Para el subcultivo se transfirieron peritecios asépticamente con aguja de insulina a nuevas placas de V8. Las ascas se recogieron cuidadosamente, se suspendieron en agua destilada estéril y se maceraron con pistilos de microtubo para liberar esporas. Una vez macerada la muestra, 100  $\mu$ L de la suspensión se colocaron sobre placas con medio estéril y se dejaron crecer aproximadamente 5–7 días en incubadores a 25 °C con fotoperíodo 12/12 h.

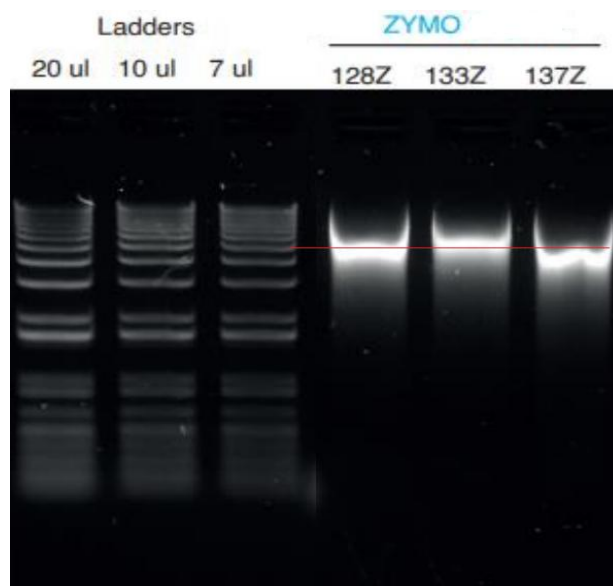
Se implementaron medidas de esterilidad estrictas en todas las transferencias para evitar contaminación; cualquier placa con signos de contaminación bacteriana o fúngica fue descartada de inmediato para garantizar que las extracciones de ADN provinieran exclusivamente de cultivos puros de *M. maydis*.

Se extrajo ADN genómico de micelio fresco obtenido de subcultivos de los aislamientos 128, 133 y 137. El proceso de extracción resultó complejo debido a la intrincada composición de la pared celular fúngica, formada principalmente por quitina, glucanos y glicoproteínas (Lord & Vyas, 2018). La reticulación covalente e iónica de la quitina con otras macromoléculas, como polisacáridos, pigmentos y proteínas, le confiere rigidez, lo que reduce su susceptibilidad a la lisis enzimática (Dyakov & Ozeretskovskaya, 2007) y dificulta la extracción de ácidos nucleicos.

Las primeras extracciones con el método CTAB, comúnmente utilizado en CIMMYT, no alcanzaron la calidad ni el tamaño de fragmentos requeridos para secuenciación de lecturas largas. Se probó posteriormente el kit comercial Quick-DNA™ Fungal/Bacterial Miniprep Kit, pero produjo ADN altamente fragmentado (~5 kb).

Finalmente se optimizó un protocolo alternativo basado en extracción con Fenol:Cloroformo:Alcohol isoamílico (25:24:1) (Servicio Nacional de Sanidad Inocuidad y Calidad Agroalimentaria, 2017), obteniendo mejor rendimiento y fragmentos de mayor tamaño (>15 kb) adecuados para secuenciación de tercera

generación (BiteSize Bio, 2021). A pesar de una ligera degradación en algunas muestras, las concentraciones y la pureza fueron suficientes para la preparación de librerías. Además, este método también demostró ser más eficiente y rentable. El protocolo modificado se encuentra en la siguiente sección.



*Figure 12.* Extracciones hechas con el kit Zymo; se muestra fragmentación de las bandas. El ladder usado como referencia fue el 1 Kb DNA Plus Ladder.

### **Protocolo para extracción de ADN de *Monographella maydis***

1. Recolectar micelio fresco y transferirlo a un mortero limpio y estéril. Congelar con nitrógeno líquido ( $N_2$ ) y moler con pistilo hasta obtener un polvo fino y homogéneo, de color entre rosado claro y blanquecino.
2. Precalentar el buffer de extracción (CTAB 2%) en baño María a 65 °C.
3. Transferir aproximadamente 0.5 mL del material molido a un tubo de microcentrifuga de 2 mL y añadir 600  $\mu$ L del buffer de extracción precalentado.
4. Mezclar suavemente por inversión e incubar a 65 °C durante 30–60 min, invirtiendo cada 10 min.
5. Añadir un volumen igual (600  $\mu$ L) de fenol:cloroformo:alcohol isoamílico (25:24:1) frío; mezclar a mano y luego rotar 5 min a temperatura ambiente.
6. Centrifugar a 14,000 rpm durante 20 min a temperatura ambiente.

7. Transferir cuidadosamente la fase acuosa superior (~750  $\mu$ L) a un tubo limpio de 2 mL sin perturbar la interfase.
8. Añadir 800  $\mu$ L de cloroformo, mezclar en rotor 3 min y centrifugar 10 min a 14,000 rpm a temperatura ambiente.
9. Transferir la fase acuosa a un tubo de 1.5 mL usando una punta recortada para evitar ruptura del ADN.
10. Añadir 3  $\mu$ L de RNasa (10 mg/mL) e incubar a 37 °C durante 20–25 min.
11. Precipitar el ADN añadiendo un volumen equivalente (~700  $\mu$ L) de isopropanol frío. Mezclar suavemente por inversión hasta observar hebras de ADN y dejar en hielo 5–10 min.
12. Centrifugar a 13,000 rpm durante 6 min; descartar el sobrenadante con cuidado, manteniendo el pellet.
13. Lavar el pellet con 500  $\mu$ L de etanol 70% (v/v) y centrifugar 5 min a 13,000 rpm.
14. (Opcional) Repetir el lavado con etanol para eliminar sales y fenol residuales.
15. Eliminar el etanol residual con una pipeta.
16. Secar al aire el pellet a temperatura ambiente durante 20 min y disolver en 100  $\mu$ L de buffer TE 1 $\times$  (10 mM Tris-HCl, 1 mM EDTA, pH 8.0).
17. Incubar la muestra durante la noche a 4 °C para permitir la completa disolución del ADN.
18. Verificar integridad y tamaño en gel de agarosa al 1% (1 h).



*Figura 13.* Micelio obtenido del subcultivo de las cepas originales (cepa 133). La biomasa se usó para las extracciones de ADN.

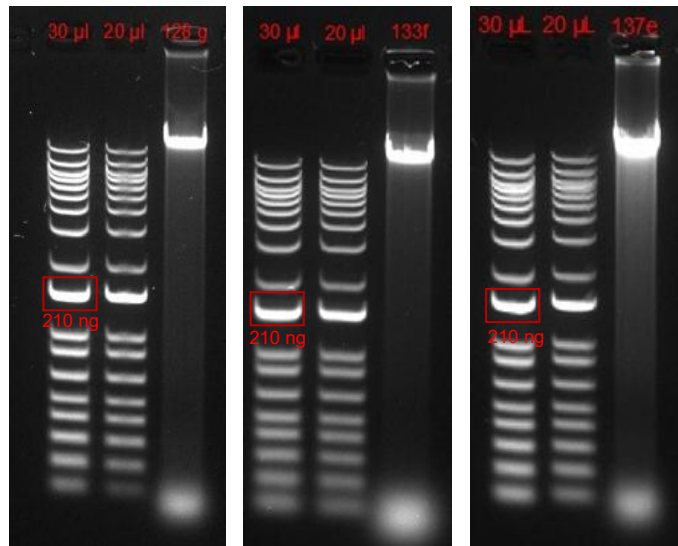
### **Estimación de concentración de ADN**

Debido a la posible interferencia de residuos de CTAB y fenol en las lecturas espectrofotométricas, la concentración de ADN se estimó por electroforesis en gel de agarosa en lugar de NanoDrop.

Para la secuenciación PacBio, la recomendación de muestra es aproximadamente 3  $\mu\text{g}$  por Gb de tamaño genómico (PacBio, 2022). Considerando un tamaño genómico esperado para Ascomycota de  $\sim 37$  Mb (Mohanta & Bae, 2015), se consideró adecuado una concentración superior a 200 ng/ $\mu\text{L}$  y volúmenes superiores a 100  $\mu\text{L}$ .

Para estimar la cantidad de ADN en gel, se cargó 1  $\mu\text{L}$  de muestra en un gel de 1% junto a un marcador 1 Kb Plus DNA Ladder (ThermoFisher, n.d.). La intensidad relativa de la banda de la muestra se comparó con la banda de referencia del marcador ( $\sim 200$  ng por el volumen cargado) para obtener una estimación

semicuantitativa. Si la banda de la muestra resultaba más intensa que la referencia, se infería una concentración superior.



*Figura 14.* Extracción de ADN de las tres cepas que fueron seleccionadas para secuenciación. Referencia de la cantidad de DNA del marcador 1 Kb DNA Plus Ladder en contraste con la intensidad de las bandas.

## Preservación de cepas

Para garantizar la preservación a largo plazo y evitar pérdida de vigor o deriva génica, se realizó criopreservación de los aislamientos. Se observó que el subcultivo prolongado disminuía la capacidad de producir peritecios; por ello se escogió la criopreservación frente a la conservación en aceite mineral, debido a su mejor mantenimiento de la integridad genética (Crahay et al., 2013).

Peritecios y micelio de cada cepa (128, 133, 137) fueron suspendidos en agua destilada estéril, se homogenizó y fueron mezclados en proporción 1:1 con una de las dos soluciones de preservación: V8 medium + glicerol o medio LB + glicerol. Alícuotas de 1.5 mL se almacenaron en criotubos a  $-80^{\circ}\text{C}$ . Se prepararon tubos de respaldo conteniendo tanto micelio como peritecios para proteger contra posible degradación a largo plazo.

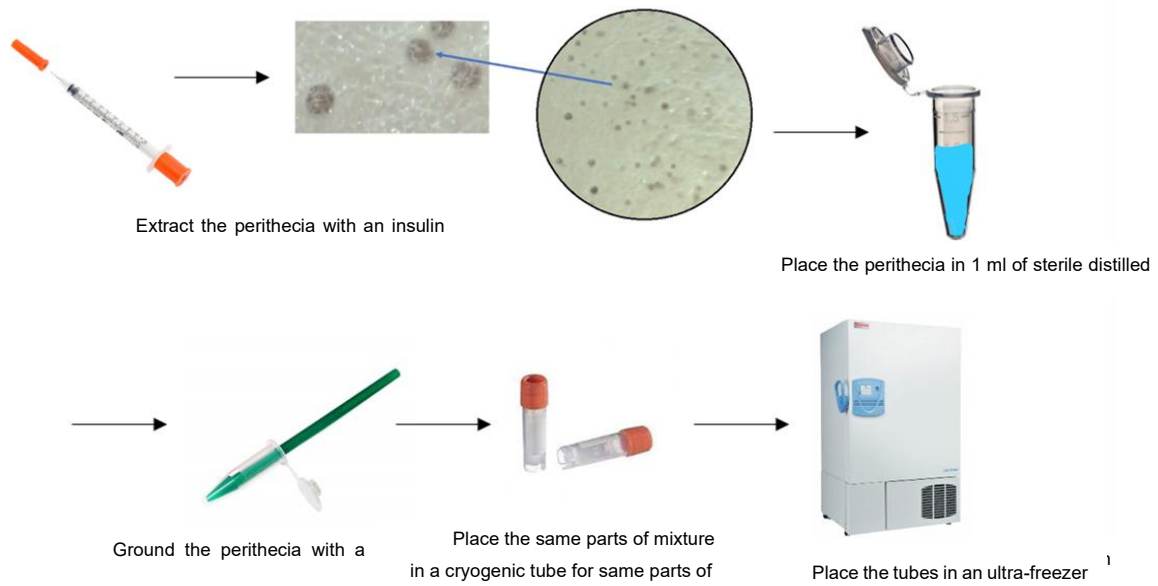


Figura 15. Diagrama de flujo de la conservación de las estructuras fúngicas de *M. maydis*

## Resultados y Discusión

### Secuenciación del genoma completo y ensamblaje *de novo*

Dado que esta fue la primera vez que *M. maydis* se aisló, preservó y cultivó eficientemente *in vitro*, no existían genomas de referencia previos. Para permitir los posteriores análisis de anotación genómica, reconstrucción filogenética y genómica comparativa, fue necesario un ensamblaje *de novo* con el fin de generar un genoma de referencia de alta calidad representativo de la especie (CD Genomics, 2025).

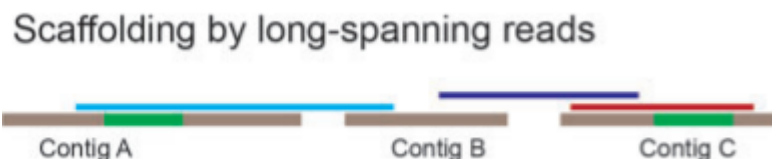
### Estrategia de secuenciación

El ADN genómico de alto peso molecular extraído de los tres aislamientos (MNG128S, MNG133S y MNG137S) se envió a Novogene (USA) para secuenciación del genoma completo usando la plataforma PacBio Sequel II DNA CLR. Este método fue elegido porque las tecnologías de lectura larga superan los retos comunes del ensamblaje *de novo*, como los huecos de cobertura y las regiones repetitivas que suelen permanecer sin resolver con plataformas de lectura corta (Jang-il & Jin-Wu, 2018).

En los ensamblajes *de novo*, las regiones repetitivas como LTRs o LINEs representan un obstáculo importante para la contigüidad (Jang-il & Jin-Wu, 2018). El uso de lecturas largas de secuenciación de molécula única (SMS) de PacBio mitiga estas limitaciones proporcionando cobertura uniforme y uniendo elementos repetitivos extensos (Novogene, 2021). En particular, las lecturas HiFi de PacBio ofrecen fragmentos largos (15–20 kb en promedio) y alta precisión (~99.95%), lo que permite generar ensamblajes continuos y de alta fidelidad (PacBio, 2025).

La longitud de lectura esperada supera el llamado “umbral dorado” de 7 kb, considerado suficiente para resolver complejidad estructural y mejorar la contigüidad en ensamblajes *de novo* (Koren & Phillippy, 2015). Gracias a estas lecturas largas, las variantes estructurales y elementos repetitivos del genoma de

*M. maydis* pudieron resolverse con mayor confianza, sentando una base sólida para los análisis posteriores.



*Figura 16.* Visualización de cómo los problemas de huecos se resuelven cuando un ensamblaje se realiza con lecturas largas. Imagen: Jang-il & Jin-Wu, 2018.

## Control de calidad del ADN

Antes de la secuenciación, la integridad y concentración del ADN genómico extraído se verificaron mediante electroforesis en gel de agarosa, confirmando la obtención de ADN de alto peso molecular. Los fragmentos observados (>15,000 bp) y las concentraciones (>210 ng/μL) estuvieron dentro de los rangos recomendados para PacBio (>200 ng/μL, >3 μg por muestra), lo cual aseguró la idoneidad de los materiales para secuenciación de alta calidad.

## Ensamblaje del genoma y evaluación de calidad

Los genomas de los aislamientos MNG128S, MNG133S y MNG137S se analizaron con GenomeScope (<http://genomescope.org>) para estimar el tamaño genómico, heterocigosidad y contenido repetitivo (Vurture et al., 2017). La longitud haploide estimada osciló entre 45–52 Mb, con una cobertura k-mer de 9–15%, heterocigosidad de aproximadamente 2% y contenido único superior al 80%. Los resultados detallados se encuentran en el Anexo E.

Las lecturas con calidad de mapeo <10 se filtraron antes del ensamblaje. El ensamblaje se realizó usando Canu (Koren et al., 2017), optimizado para lecturas largas PacBio Sequel II. La calidad y completitud del ensamblaje se evaluaron con BUSCO v5.5.0 (Manni et al., 2021) y QUAST (Gurevich et al., 2013). Los ensamblajes resultantes mostraron >97% de completitud y un valor N50 de 1.5 Mb.



## BUSCO

Se generó un ensamblaje consenso, alcanzando 97.9% de completitud y un N50 de 3 Mb, tal como se detalla en la siguiente tabla.

*Tabla 1. Métricas de completitud de BUSCO para el ensamblaje de Monographella maydis. (Sordariomycetes\_odb10, n=1706).*

Resultados de BUSCO		
C:97.9%[S:97.1%,D:0.8%],F:0.5%,M:1.6%,n:1706		
Categoría	Conteo	Porcentaje
<b>Completos (C)</b>	1670	97.9%
Copia única (single-copy) (S)	1656	97.1%
Duplicados (D)	14	0.8%
<b>Fragmentados (F)</b>	9	0.5%
<b>Faltantes (missing) (M)</b>	27	1.6%
<b>Total de grupos buscados (n)</b>	1706	-

BUSCO buscó 1,706 ortólogos universales de copia única, de los cuales 1,670 (97.9%) estuvieron completos (97.1% de copia única y 0.8% duplicados), 9 fragmentados (0.5%) y 27 ausentes (1.6%).

*Tabla 2. Estadísticas de ensamblaje para el genoma de M. maydis.*

Métrica	Valor
Número de scaffolds	90
Número de contigs	90
Longitud total del ensamblaje	39 801 790 bp
Porcentaje de huecos	0.000%
Scaffold N50	3 Mb
Contig N50	3 Mb

## QUAST

QUAST generó métricas estándar de contigüidad y completitud. El análisis generó estadísticas para distribución de tamaño de contig, contenido GC, longitud del ensamblaje y métricas basadas en N (N50, N90, L50, L90), así como el número de bases ambiguas por cada 100 kb.

*Tabla 3. Estadísticas de QUAST del ensamble de M. maydis.*

<b>Métrica</b>	<b>Valor</b>
Número de contigs ( $\geq 0$ bp)	90
Número de contigs ( $\geq 1,000$ bp)	90
Número de contigs ( $\geq 5,000$ bp)	61
Número de contigs ( $\geq 10,000$ bp)	49
Número de contigs ( $\geq 25,000$ bp)	23
Número de contigs ( $\geq 50,000$ bp)	21
Longitud total ( $\geq 0$ bp)	39 801 790
Longitud total ( $\geq 1000$ bp)	39 801 790
Longitud total ( $\geq 5000$ bp)	39 710 908
Longitud total ( $\geq 10000$ bp)	39 630 494
Longitud total ( $\geq 25000$ bp)	39 215 555
Longitud total ( $\geq 50000$ bp)	39 150 514
Contig más largo	5 468 643 bp
Contenido de GC	54.40%
N50	3 014 066 bp
N90	1 058 628 bp
auN	2 849 831.2 bp
L50	6
L90	15
Número de N's por 100 kbp	0

Las estadísticas de QUAST basadas en los umbrales de tamaño indican que el ensamblaje es altamente contiguo y contiene muy pocos contigs pequeños. La

longitud total del ensamblaje se mantiene igual al considerar contigs  $\geq 0$  bp y  $\geq 1,000$  bp, lo que demuestra que no hay contigs muy cortos. Al aumentar progresivamente los umbrales de tamaño ( $\geq 5$  kb,  $\geq 10$  kb,  $\geq 25$  kb y  $\geq 50$  kb), la longitud total del ensamblaje disminuye solo ligeramente en cada paso, lo que demuestra que la mayor parte de la secuencia genómica está contenida en contigs grandes. De manera notable, incluso cuando solo se consideran los contigs  $\geq 50$  kb, la longitud retenida permanece cercana al tamaño total del ensamblaje, lo que confirma que la mayor parte del genoma está ensamblada en secuencias largas y de alta calidad. En conjunto, estos patrones reflejan un genoma robusto y bien ensamblado, con gran continuidad y mínima fragmentación.

El ensamblaje consistió en 90 contigs con un total de 39.8 Mb, y no se realizó ningún scaffolding, resultando en el mismo número de contigs y scaffolds. La ausencia de bases N indica que el ensamblaje está completamente libre de huecos, lo que refleja la alta continuidad de los datos de lectura larga PacBio CLR.

## Aplicaciones downstream

El ensamblaje de alta calidad sirvió como base para la anotación funcional, análisis filogenómicos y pangenómicos. Esta referencia representa la primera caracterización genómica de *M. maydis*, constituyendo un recurso valioso para estudiar su potencial patogénico y sus relaciones evolutivas dentro del complejo de la mancha de asfalto.

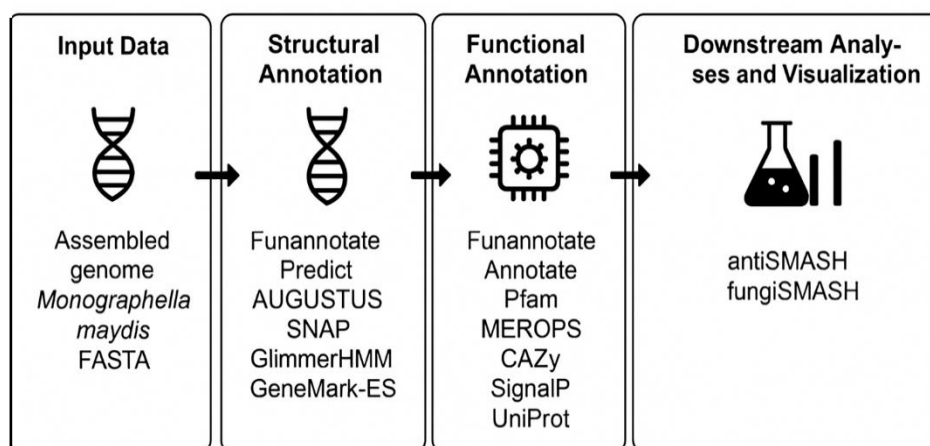


Figura 17. Flujo de trabajo para las lecturas del genoma de *M. maydis*.

Basándose en el ensamblaje obtenido, la siguiente fase del estudio se enfocó en la anotación estructural y funcional del genoma de *M. maydis*, integrando múltiples herramientas bioinformáticas para predecir modelos génicos, asignar funciones funcionales y explorar la base molecular de su interacción con el hospedero y con otros hongos involucrados en el complejo de la mancha de asfalto.

### **Predicción génica y anotación funcional**

La predicción génica se realizó sobre el genoma ensamblado utilizando FunAnnotate (v1.8.15) (Palmer & Stajich, 2020). Las secuencias repetitivas fueron enmascaradas con la herramienta de enmascaramiento de FunAnnotate, y la predicción génica se llevó a cabo mediante el algoritmo GeneMark-ES. La completitud del genoma se evaluó utilizando la base de datos fúngica de BUSCO (Ascomycota\_odb10) (Manni et al., 2021). Las secuencias codificantes predichas fueron posteriormente sometidas a búsquedas con BLAST+ (v2.14.1) para homología y anotación funcional. Las secuencias proteicas extraídas del archivo GFF3 que contenía las predicciones de GeneMark fueron consultadas contra la base de datos de proteínas UniProt para identificar proteínas homólogas y asignar funciones putativas (Ter-Hovhannisyan et al., 2008; The UniProt Consortium, 2025). A partir del conjunto de datos de UniProt se generó una base de datos BLAST local, y se ejecutó BLASTp para comparar contra ella las secuencias proteicas predichas.

### **Anotación estructural y funcional**

La anotación estructural del genoma de *M. maydis* se realizó utilizando Funannotate (v1.8.15), un conjunto integral de herramientas optimizado para la anotación de genomas fúngicos. Los análisis se llevaron a cabo localmente en el clúster de cómputo de alto rendimiento MITLA del CCG-UNAM.

Este proceso empleó herramientas de predicción génica como AUGUSTUS, SNAP, GlimmerHMM y parcialmente GeneMark-ES, generando 10,933 genes

predichos. La salida incluyó archivos de anotación estándar (.gff3, .gbk, .faa, .ffn y .fna), así como registros de errores y parámetros (Bromberg & Rost, 2007; Majoros et al., 2004; Stanke et al., 2004; Ter-Hovhannisyan et al., 2008).

Este paso incorporó múltiples bases de datos y recursos, incluyendo Pfam, MEROPS, CAZy, SignalP y UniProt (a través de DIAMOND) (Lombard et al., n.d.; Mistry et al., n.d.; Rawlings et al., 2014; Teufel et al., 2022; The UniProt Consortium, 2025). Debido a limitaciones de configuración local, no se ejecutaron InterProScan ni eggNOG-mapper.

*Tabla 4. Resultados generales de la anotación estructural y funcional del genoma de M. maydis.*

<b>Categoría</b>	<b>Cantidad</b>
Genes anotados	<b>10,933</b>
Dominios Pfam	<b>11,971</b>
Coincidencias con UniProt	<b>773</b>
Entradas MEROPS	<b>368</b>
Familias CAZy	<b>794</b>
Proteínas secretadas predichas (SignalP)	<b>1016</b>

## **Análisis de metabolitos secundarios**

La predicción de clústeres de genes biosintéticos (BGCs) y de rutas de metabolitos secundarios se realizó utilizando antiSMASH (v4.1.0) (Blin et al., 2019). Aunque se intentó ejecutar funannotate remote -m antismash, el análisis se completó finalmente mediante la carga manual del archivo GenBank anotado en el servidor fungiSMASH.

El análisis identificó exitosamente múltiples clústeres biosintéticos, potencialmente involucrados en la producción de metabolitos secundarios. Los resultados se resumen en la siguiente imagen.

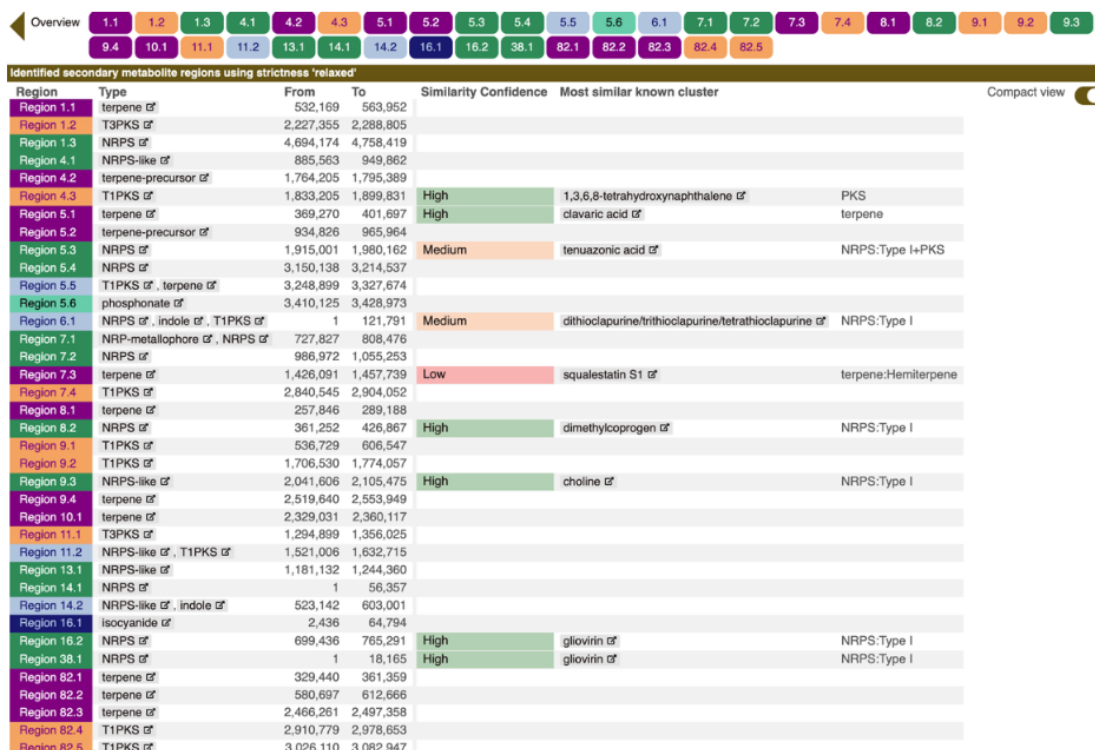


Figura 18. Análisis de antiSMASH v5.0 del genoma de *M. maydis*. Visualización de la predicción de clusters de genes biosintéticos (BGCs).

Cada bloque de color representa un tipo distinto de clúster génico (p. ej., NRPS-like, PKS, terpenos). El análisis reveló múltiples clústeres potencialmente implicados en la síntesis de metabolitos secundarios asociados con la patogenicidad y la adaptación fúngica.

## Análisis de ortología

La anotación funcional y la identificación de grupos ortólogos se realizaron utilizando EggNOG-mapper (v2.1.9), que asigna proteínas predichas a grupos ortólogos e identificadores de Ortología KEGG (KO) basados en la base de datos EggNOG (Cantalapiedra et al., n.d.; Kanehisa & Sato, 2019). Las proteínas predichas de *M. maydis* fueron formateadas en FASTA y analizadas utilizando Diamond para un alineamiento rápido contra la base de datos EggNOG.

El análisis generó anotaciones KO de KEGG, categorías funcionales e información sobre rutas metabólicas. KEGG Mapper (<https://www.genome.jp/kegg/mapper/reconstruct.html>) se utilizó para visualizar interactivamente las rutas anotadas. Todos los KOs asignados y sus funciones correspondientes se encuentran en el Anexo I.

Entre las funciones más representadas se encontraron las siguientes, entre otras, lo que refleja la diversidad metabólica y los posibles roles funcionales dentro del genoma de *M. maydis*.

*Tabla 5. Los 10 KOs más frecuentes.*

KO ID	Conteo	Descripción
K00505	30	Glucokinasa [EC:2.7.1.2]
K01426	24	Adenosina desaminasa [EC:3.5.4.4]
K16261	21	Oxidoreductasa dependiente de FAD
K01183	17	Alfa-glucosidasa [EC:3.2.1.20]
K05349	15	Factor de transcripción tipo zinc-finger fúngico
K00799	15	Glutamina sintetasa [EC:6.3.1.2]
K08286	14	Proteína F-box
K08141	14	Proteína quinasa
K00480	11	Subunidad I de citocromo c oxidasa [EC:7.1.1.9]
K01230	10	Beta-galactosidasa [EC:3.2.1.23]

## KEGG Mapper Reconstruction Result

Pathway (430)

Brite (46)

Brite Table (7)

Module (91)

Show matched objects

### Metabolism

Global and overview maps

- 01100 Metabolic pathways (830)
- 01110 Biosynthesis of secondary metabolites (328)
- 01120 Microbial metabolism in diverse environments (173)
- 01200 Carbon metabolism (84)
- 01210 2-Oxocarboxylic acid metabolism (37)
- 01212 Fatty acid metabolism (23)
- 01230 Biosynthesis of amino acids (98)
- 01232 Nucleotide metabolism (36)
- 01250 Biosynthesis of nucleotide sugars (19)
- 01240 Biosynthesis of cofactors (118)
- 01220 Degradation of aromatic compounds (12)
- 01310 Nitrogen cycle (2)
- 01320 Sulfur cycle (7)

Carbohydrate metabolism

- 00010 Glycolysis / Gluconeogenesis (26)
- 00020 Citrate cycle (TCA cycle) (21)
- 00030 Pentose phosphate pathway (19)
- 00040 Pentose and glucuronate interconversions (22)
- 00051 Fructose and mannose metabolism (24)
- 00052 Galactose metabolism (15)
- 00053 Ascorbate and aldarate metabolism (12)
- 00500 Starch and sucrose metabolism (27)
- 00620 Pyruvate metabolism (31)
- 00630 Glyoxylate and dicarboxylate metabolism (23)
- 00640 Propanoate metabolism (20)
- 00650 Butanoate metabolism (17)
- 00660 C5-Branched dibasic acid metabolism (3)
- 00562 Inositol phosphate metabolism (24)

Figura 19. Visión general de la reconstrucción de rutas metabólicas KEGG basada en las anotaciones de EggNOG-mapper. Las rutas metabólicas se reconstruyeron utilizando la herramienta en línea KEGG Mapper, empleando los identificadores KO generados por EggNOG.

## Análisis Pangenómico y Filogenómico

Para inferir las relaciones evolutivas e identificar familias génicas conservadas, se realizó un análisis genómico comparativo que incluyó *Microdochium bolleyi* (GCA\_001566295), *Microdochium trichocladiopsis* (GCA\_020744255),



*Microdochium nivale* (GCA\_026745475), *Phyllachora maydis* (GCA\_029339225) y *Monographella maydis* (este estudio). Los datos genómicos en formato GenBank (.gbff) se analizaron utilizando Get\_Homologues (v14112024), empleando los algoritmos BBH, COG y OrthoMCL para identificar clústeres ortólogos (Contreras-Moreira & Vinuesa, 2013).

Un total de 2,034 grupos ortólogos centrales, cada uno con una proteína por especie, se conservaron y se procesaron posteriormente con Get\_Phylomarkers (v2.8.1.4), obteniendo 208 grupos de proteínas homólogas adecuados para la inferencia filogenética (Vinuesa et al., 2018).

Se realizaron alineamientos múltiples de secuencias, y el árbol filogenético se reconstruyó utilizando RAxML (v8.2.12) bajo el modelo de sustitución GTR (Kozlov et al., 2019). El árbol final se visualizó usando IQ-TREE (v2.1.2) e iTOL, utilizando el método de Máxima Verosimilitud (MLE), ilustrando las relaciones evolutivas entre las cinco especies fúngicas en un árbol en formato Newick (Letunic & Bork, 2021; Minh et al., 2020).

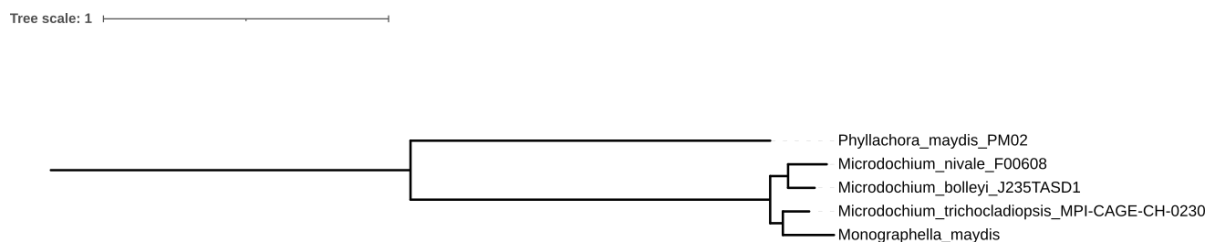


Figura 20. Relaciones filogenéticas entre *Monographella maydis* y especies de hongos relacionadas.

El análisis filogenético muestra a *Phyllachora maydis*\_PM02 ramificándose por separado de las especies de *Microdochium* y *Monographella*, indicando que *P. maydis* es el taxón más divergente dentro del conjunto de datos. Las especies restantes forman un clado bien soportado en el que *Monographella maydis* se agrupa estrechamente con *Microdochium trichocladiopsis*, sugiriendo un ancestro común relativamente reciente. Ramas ligeramente más profundas separan a *M. bolleyi* y *M. nivale*, que se agrupan como taxones hermanos.

En conjunto, la topología refleja una clara separación filogenética entre *Phyllachora* y el grupo *Microdochium/Monographella*; la identidad de *Monographella maydis* también se confirma con este árbol filogenético, mostrando los resultados esperados dentro de su género.

## Conclusiones

Este estudio representa la primera caracterización genómica integral de *Monographella maydis*, un hongo asociado con el complejo de mancha de asfalto del maíz (TSC). A través de su aislamiento exitoso, cultivo *in vitro* y secuenciación genómica de alta calidad utilizando tecnología PacBio de lecturas largas, se obtuvo un ensamblaje genómico completo y casi contiguo. La anotación de este genoma proporcionó un catálogo robusto de genes predichos, dominios funcionales y posibles factores de virulencia, ofreciendo información valiosa sobre los mecanismos moleculares subyacentes a su biología. Los análisis comparativos utilizando enfoques filogenómicos y pangenómicos revelaron que *M. maydis* comparte relaciones evolutivas con otros miembros del género *Microdochium/Monographella*, mientras que también exhibe características genómicas únicas que podrían explicar su especialización ecológica y su posible papel en el TSC.

La integración de herramientas bioinformáticas como BUSCO, QUAST, FunAnnotate y EggNOG-mapper permitió una anotación estructural y funcional detallada del genoma, lo que condujo a la identificación de genes potencialmente vinculados con la patogenicidad, adaptación a estreses y de metabolismo secundario. Estos hallazgos no solo proporcionan un recurso fundamental para la especie, sino que también contribuyen a una comprensión más amplia de la diversidad fúngica asociada con las enfermedades del maíz.

Sin embargo, aunque los datos genómicos y comparativos avanzan nuestra comprensión de *M. maydis*, la interacción patogénica entre *Phyllachora maydis* y *M. maydis* permanece sin resolverse. Las aproximaciones experimentales futuras deberían orientarse a elucidar el orden cronológico y la dependencia de la infección entre estos dos hongos —si *M. maydis* requiere una inoculación previa por *P. maydis*, o si ambos pueden actuar de manera independiente bajo condiciones ambientales específicas. Los ensayos de doble inoculación, el perfilamiento transcriptómico y los estudios de infección basados en microscopía podrían aclarar

estas dinámicas, revelando así la base molecular de su patogenicidad sinérgica dentro del complejo de mancha de asfalto.

En conjunto, este trabajo establece un marco genómico y analítico para *Monographella maydis*, abriendo nuevas vías para comprender la evolución, el potencial patógeno y la adaptación ecológica de los hongos involucrados en las enfermedades del maíz.

## Referencias

- BiteSize Bio. (2021). *The Basics: How Phenol Extraction of DNA Works*.
- Blin, K., Shaw, S., Villebro, K. S. R., Ziemert, N., Lee, S. Y., Medema, M. H., & Weber, T. (2019). antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Research*. doi: 10.1093/nar/gkz310.
- Bromberg, Y., & Rost, B. (2007). SNAP: predict effect of non-synonymous polymorphisms on function. *Nucleic Acids Research*, 35(11), 3823–3835. <https://doi.org/10.1093/nar/gkm238>
- Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P., & Huerta-Cepas, J. (n.d.). eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Molecular Biology and Evolution*, 38(12), 5825–5829. <https://doi.org/10.1093/molbev/msab293>
- CD Genomics. (2025). *A Guide to De Novo Genome Assembly*. <https://www.cd-genomics.com/a-guide-to-de-novo-genome-assembly.html>
- CIMMYT. (2018). *El complejo de la mancha de asfalto podría causar una gran mancha negra en la economía de maíz de EUA*.
- CONABIO. (2022). *Razas de maíz de México*. Biodiversidad Mexicana.
- Contreras-Moreira, B., & Vinuesa, P. (2013). GET\_HOMOLOGUES, a versatile software package for scalable and robust microbial pan-genome analysis. *Applied and Environmental Microbiology*, 79(24), 7696–7701.
- Crahay, C., Declerck, S., Colpaert, J., Pigeon, M., & Munaut, F. (2013). Viability of ectomycorrhizal fungi following cryopreservation. *Fungal Biology*.
- Dyakov, Y., & Ozeretskovskaya, O. (2007). Vertical pathosystem: avirulence genes and their products. In *Comprehensive and Molecular Phytopathology* (pp. 181–215).
- FAOSTAT. (2024). Agricultural Production Statistics 2010-2023. *FAOSTAT Analytical Brief* 96.

<https://openknowledge.fao.org/server/api/core/bitstreams/df90e6cf-4178-4361-97d4-5154a9213877/content>

- Groves, C. L., Kleczewski, N. M., Telenko, D. E. P., Chilvers, M. I., & Smith, D. L. (2020). *Phyllachora maydis* ascospore release and germination from overwintered corn residue. *Plant Health Prog*, 21, 26–30.
- Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*. <https://doi.org/10.1093/bioinformatics/btt086>
- Hock, J., Dittrich, U., Renfro, B., & Kranz, J. (1992). Sequential development of pathogens in the maize tarspot disease complex. *Mycopathologia*, 157–161.
- INTAGRI. (2016). *El Complejo de la Mancha de Asfalto en el Cultivo de Maíz*.
- Jang-il, S., & Jin-Wu, N. (2018). The present and future of de novo whole-genome assembly. *Briefings in Bioinformatics*, Volume 19(1), 23–40. <https://doi.org/10.1093/bib/bbw096>
- Kanehisa, M., & Sato, Y. (2019). KEGG Mapper for inferring cellular functions from protein sequences. *Protein Sci*, 29(1), 28–35. <https://doi.org/10.1002/pro.3711>
- Koren, S., & Phillippy, A. M. (2015). One chromosome, one contig: complete microbial genomes from long-read sequencing and assembly. *Current Opinion in Microbiology*, 23, 110–120. <https://doi.org/https://doi.org/10.1016/j.mib.2014.11.014>
- Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res*, 27(5), 722–736. <https://doi.org/10.1101/gr.215087.116>
- Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAXML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, 35(21), 4453–4455. <https://doi.org/10.1093/bioinformatics/btz305>

- Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, 49(W1), W293–W296. <https://doi.org/10.1093/nar/gkab301>
- Listman, M., & Ordóñez, R. (2019). *Ten things you should know about maize and wheat*.
- Lombard, V., Ramulu, H. G., Drula, E., Coutinho, P. M., & Henrissat, B. (n.d.). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research*, 42(D1), D490–D495. <https://doi.org/10.1093/nar/gkt1178>
- Lord, A., & Vyas, J. (2018). Host Defenses to Fungal Pathogens. In *Clinical Immunology*.
- Majoros, W. H., Pertea, M., & Salzberg, S. L. (2004). TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. *Bioinformatics*, 20(16), 2878–2879. <https://doi.org/10.1093/bioinformatics/bth315>
- Manni, M., Berkeley, M., Seppey, M., Simao, F., & Zdobnov, E. (2021). *BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes*. <https://doi.org/arXiv:2106.11799> [q-bio] [Internet]
- Martin, K. J., & Rygioewicz, P. (2005). Fungal-specific PCR primers developed for analysis of the ITS region of environmental DNA extracts. *BMC Microbiology*. <https://doi.org/10.1186/1471-2180-5-28>
- Maublanc, A. (1904). Espèces nouvelles de champignons inférieurs. *Bulletin de La Societe Phytopathologique Française*, 72.
- Mccoy, A. G., Roth, M. G., Shay, R., Noel, Z. A., Jayawardana, M. A., Longley, R. W., & Bonito, G. (2019). *Identification of Fungal Communities Within the Tar Spot Complex of Corn in Michigan via Next-Generation Sequencing*. 3(3), 235–243. <https://doi.org/10.1094/PBIOMES-03-19-0017-R>
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New models and efficient

- methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, 37(5), 1530–1534.
- Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D., & Bateman, A. (n.d.). Pfam: The protein families database in 2021. *Nucleic Acids Research*, 49(D1), D412–D419. <https://doi.org/10.1093/nar/gkaa913>
- Mohanta, T. K., & Bae, H. (2015). The diversity of fungal genome. *Biological Procedures Online*.
- Müller, E., & Samuels, J. G. (1984). *Monographiella maydis* sp. nov. and its connection to the tar-spot disease of *Zea mays*. *Nova Hedwigia*, 40.
- Novogene. (2021). *WGS vs WES: Which Genetic Sequencing Method is Right for You?*
- Orús, A. (2023). *Ranking de los principales productores de maíz a nivel mundial en 2021*. Statista.
- PacBio. (2022). *Considerations for using the low and ultra-low DNA input workflows for whole genome sequencing*.
- PacBio. (2025). *Whole genome sequencing*. <https://www.pacb.com/products-and-services/applications/whole-genome-sequencing/>
- Palmer, J. M., & Stajich, J. E. (2020). Funannotate v1.8.1: Eukaryotic genome annotation. *Zenodo*. <https://doi.org/10.5281/zenodo.4054262>
- Petrak, F. (1924). Mykologische Notizen. VII. *Annales Mycologici*, 22.
- Ranum, P., & Pe, J. P. (2014). *Global maize production , utilization , and consumption*. 1312, 105–112. <https://doi.org/10.1111/nyas.12396>
- Rawlings, N. D., Waller, M., Barrett, A. J., & Bateman, A. (2014). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research*, 42(D1), D503–D509. <https://doi.org/10.1093/nar/gkt953>
- Ríos Herrera, E. N., Ochoa Fuentes, Y. M., Cerna Chávez, E., Landeros Flores, J.,



- Cepeda Siller, M., & Rodríguez Guerra, R. (2007). Hongos asociados a la mancha de asfalto en el cultivo de maíz en México. *SciELO*, 8.
- Satyanarayana, E. (2004). *Importance of Maize and Utilisation*. CornIndia.
- Secretaría de Agricultura y Desarrollo Rural. (2023). *Maíces nativos, patrimonio biológico, agrícola, cultural y económico*.
- Servicio Nacional de Sanidad Inocuidad y Calidad Agroalimentaria. (2017). *PROTOCOLO DE EXTRACCIÓN DE DNA Y RNA, A PARTIR DE SEMILLAS Y HOJAS*.
- Sol Hernández, G. (2018). *Diversidad genética de Phyllachora maydis y Monographella maydis y respuesta del maíz a la inoculación artificial*. UNIVERSIDAD AUTÓNOMA DE CHIAPAS.
- Stanke, M., Steinkamp, R., Waack, S., & Morgenstern, B. (2004). AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Research*, 32, W309–W312.
- Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y., & Borodovsky, M. (2008). Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Research*, 18(12), 1979–1990. <https://doi.org/10.1101/gr.081612.108>
- Teufel, F., Armenteros, J. J. A., Johansen, A. R., Gíslason, M. H., Pihl, S. I., Konstantinos D. Tsirigos, Winther, O., Søren Brunak, Heijne, G. von, & Nielsen, H. (2022). SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nature Biotechnology*, 40, 1023–1025. <https://doi.org/10.1038/s41587-021-01156-3>
- The UniProt Consortium. (2025). UniProt: the Universal Protein Knowledgebase in 2025. *Nucleic Acids Research*, 53(D1), D609–D617. <https://doi.org/10.1093/nar/gkae1010>
- ThermoFisher. (n.d.). *How to Use Phenol/Chloroform for DNA Purification*.

- Tronsmo, A. M., Collinge, D. B., Djurle, A., Munk, L., Yuen, J., & Tronsmo, A. (n.d.). Plant pathology and plant diseases. In *2020* (pp. 48–59).
- Valle-Torres, J., Ross, T. J., Plewa, D., Avellaneda, M. C., Check, J., Chilvers, M. I., Cruz, A. P., Dalla Lana, F., Groves, C., Gongora-Canul, C., Henriquez-Dole, L., Jamann, T., Kleczewski, N., Lipps, S., Malvick, D., McCoy, A. G., Mueller, D. S., Paul, P. A., Puerto, C., ... Cruz, C. D. (2020). Tar Spot: An Understudied Disease Threatening Corn Production in the Americas. *APS Publications*.
- Vinuesa, P., Ochoa-Sanchez, L.-E., & Contreras-Moreira, B. (2018). GET\_PHYLOMARKERS, a software package to select optimal orthologous clusters for phylogenomics and inferring pan-genome phylogenies, used for a critical geno-taxonomic revision of the genus *Stenotrophomonas*. *Frontiers in Microbiology*, 9. <https://doi.org/doi:10.3389/fmicb.2018.00771>
- Vurture, G. W., Sedlazeck, F. J., Nattestad, M., Underwood, C. J., Fang, H., Gurtowski, J., & Schatz, M. C. (2017). GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics*, 33(14), 2202–2204. <https://doi.org/10.1093/bioinformatics/btx153>

## Anexos

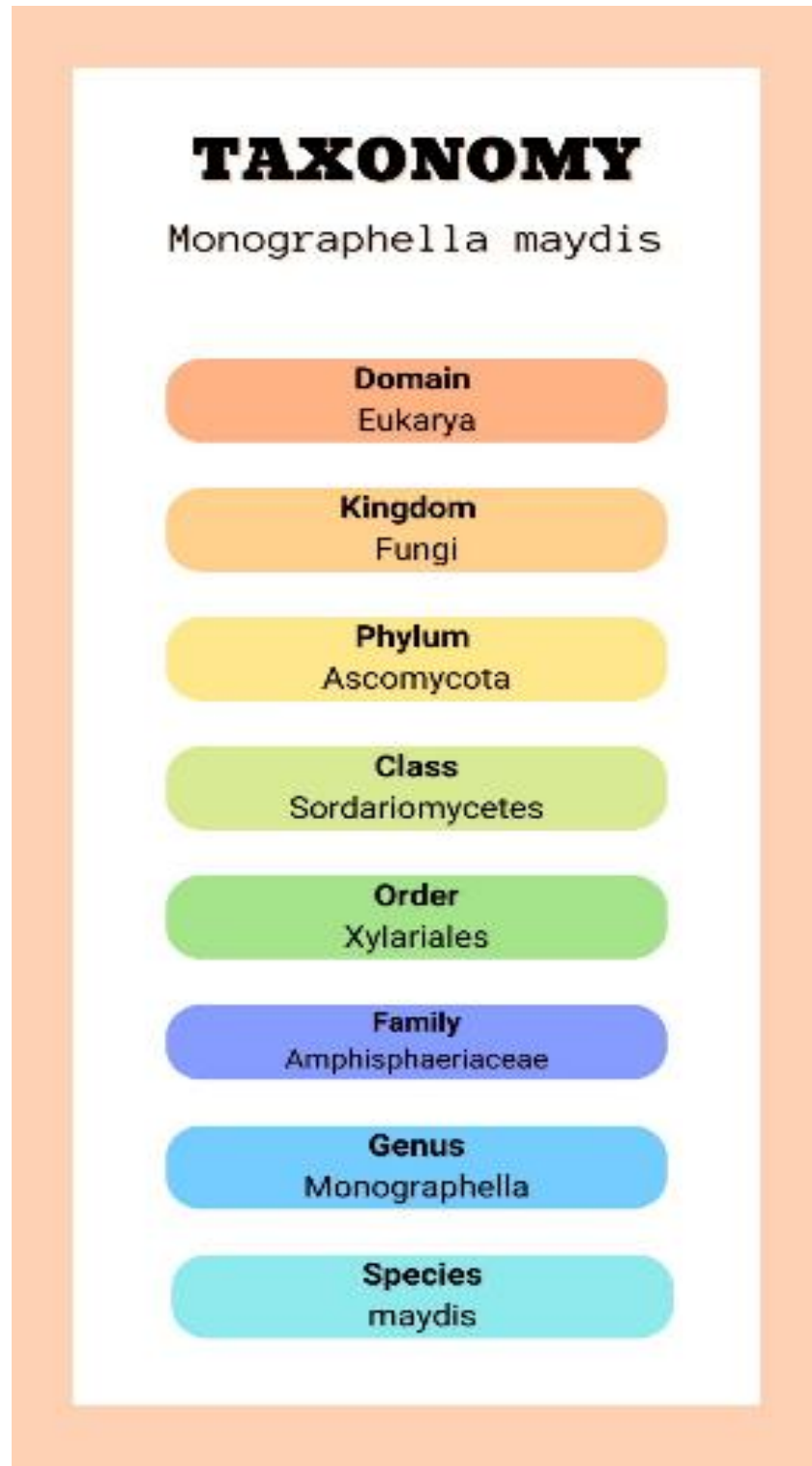
### Anexo A. Glosario

1. Ensamblaje genómico *de novo*. Proceso mediante el cual se ensambla el genoma completo de un organismo a partir de lecturas de secuenciación—largas o cortas—sin utilizar un genoma de referencia.
2. Genoma de referencia. Ensamblaje genómico continuo, preciso y anotado de una especie, utilizado como base para comparaciones y análisis.
3. Anotación del genoma. Proceso de identificación de genes, regiones codificantes y características regulatorias o estructurales dentro de un genoma.
4. Hiperparásito. Parásito cuyo hospedero es a su vez otro parásito.
5. Cebadores (primers). Secuencias cortas de ADN monocatenario que se unen a regiones específicas y permiten la amplificación de ADN en PCR.
6. Estroma. Tejido fúngico compacto que contiene estructuras reproductivas como ascocarpos o peritecios.
7. Ascós. Sacos cilíndricos o alargados característicos de los Ascomycetes, donde se desarrollan y almacenan las ascosporas durante la reproducción sexual.
8. Triseptado. Que posee tres septos internos; se refiere comúnmente a conidios o hifas con tres divisiones transversales.
9. Conidios. Esporas asexuales producidas por hongos, típicamente dispersadas para la reproducción y diseminación.
10. Himenio. Capa fértil de un ascocarpo compuesta por ascos y paráfisis.
11. Ascocarpo. Estructura reproductiva sexual de hongos Ascomycetes donde se producen los ascos y las ascosporas. Puede presentarse en diversas formas (peritecios, apotecios, cleistotecios) y es responsable del desarrollo de esporas sexuales.
12. Peritecios. Ascocarpos con forma de matraz que presentan una abertura pequeña (ostíolo) por la cual se liberan las ascosporas maduras. Suelen estar incrustados o unidos a un estroma en muchos hongos Ascomycetes.

13. Parásito obligado. Organismo que solo puede completar su ciclo de vida dentro de un hospedero vivo.
14. Mapeo de lecturas (read mapping). Alineamiento de lecturas de secuenciación contra un genoma de referencia conocido para comparaciones o identificación de variantes.
15. Internal Transcribed Spacer (ITS) (Espaciador Interno Transcrito). Regiones no codificantes altamente variables del ADN ribosomal usadas como marcadores moleculares para la identificación de hongos.
16. Melting temperature,  $T_m$  (Temperatura de fusión,  $T_m$ ). Temperatura a la cual el 50% de las hebras de ADN están desnaturalizadas (sencillas) y el 50% permanecen doblemente enlazadas.
17. Single Nucleotide Variant, SNV (Variante de un solo nucleótido). Sustitución de una base nucleotídica en el genoma, utilizada frecuentemente en genómica comparativa.
18. *In vitro*. Proceso realizado fuera de un organismo vivo bajo condiciones controladas de laboratorio.
19. LINEs (Elementos Nucleares Dispersos Largos). Elementos móviles repetitivos de ADN capaces de copiarse e insertarse dentro del genoma.
20. LTRs (Repeticiones Terminales Largas). Secuencias repetidas de ADN que flanquean retrotransposones; esenciales para su movilidad.
21. Contenido GC. Porcentaje de bases guanina y citosina en una molécula de ADN; indicador de estabilidad genómica.
22. Haploide. Que contiene un solo conjunto de cromosomas.
23. K-mer. Subcadena de longitud  $k$  extraída de una secuencia de ADN; utilizada para análisis de ensamblaje, estimación de tamaño de genoma y corrección de errores.
24. N50. Longitud del contig más corto que representa el 50% del tamaño total del genoma ensamblado; métrica de calidad.
25. N90. Longitud de los contigs que representan el 90% del ensamblaje; indicador de calidad más estricto.

26. auN. Métrica de continuidad del ensamblaje genómico que integra todas las estadísticas  $N_x$  ( $N_{10}$ – $N_{90}$ ) en un solo valor, ofreciendo una medida integral de la distribución del tamaño de contigs y la contigüidad general del ensamblaje.
27. L50. Métrica que indica el número mínimo de contigs o andamios que juntos contienen el 50% de la longitud total del ensamblaje. Un L50 más pequeño refleja un ensamblaje más contiguo.
28. L90. Métrica que indica el número mínimo de contigs o andamios que juntos contienen el 90% de la longitud total del ensamblaje. Un L90 menor sugiere un ensamblaje más completo y menos fragmentado.
29. Anamorfos. Etapa asexual de un hongo.
30. Teleomorfos. Etapa sexual de un hongo.
31. Cobertura del genoma. Número promedio de veces que cada base del genoma es secuenciada.
32. Scaffold (Andamio). Secuencia compuesta por contigs conectados cuyo orden y orientación se conocen.
33. BUSCO score. Métrica que indica el grado de completitud del genoma basada en ortólogos conservados de copia única.
34. Coverage depth (Profundidad de cobertura). Medida del número de veces que un nucleótido dado es secuenciado.
35. Análisis downstream. Pasos analíticos realizados después del ensamblaje y anotación del genoma, como análisis filogenéticos u ontologías génicas.
36. Árbol filogenético. Diagrama que representa relaciones evolutivas entre especies o cepas.
37. Contig. Secuencia continua de ADN ensamblada a partir de lecturas superpuestas.
38. Modelo de sustitución GTR. Modelo evolutivo “General Time Reversible” utilizado para inferencia filogenética, que permite tasas diferentes de sustitución entre todas las posibles parejas de nucleótidos.
39. Formato Newick. Formato de texto estándar para representar árboles filogenéticos mediante paréntesis anidados.

## Anexo B. Taxonomía de *M. maydis*



## Anexo C. Imágenes de estructuras de *M. maydis*



Figura 21. Conidio aislado obtenido de esporodoquio de *M. maydis*



Figura 22. Himenio de *M. maydis*.



Figura 23. Hifas de *Monographella maydis*, unidas a un peritecio abierto.

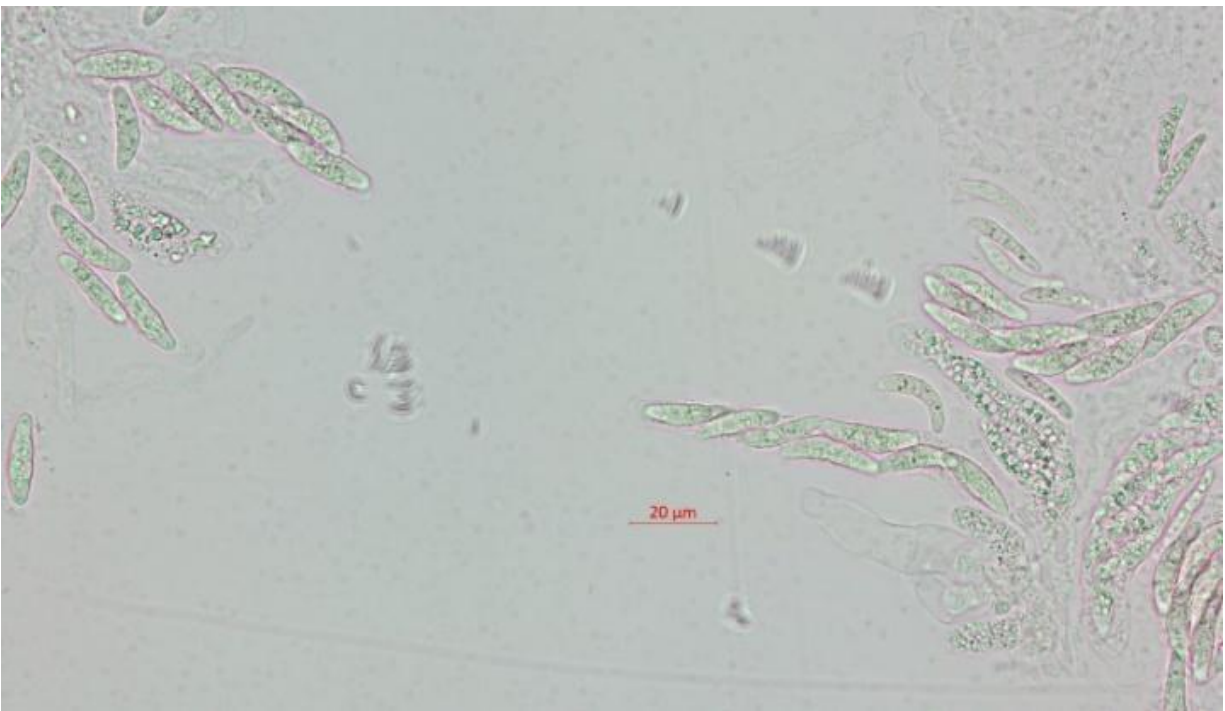
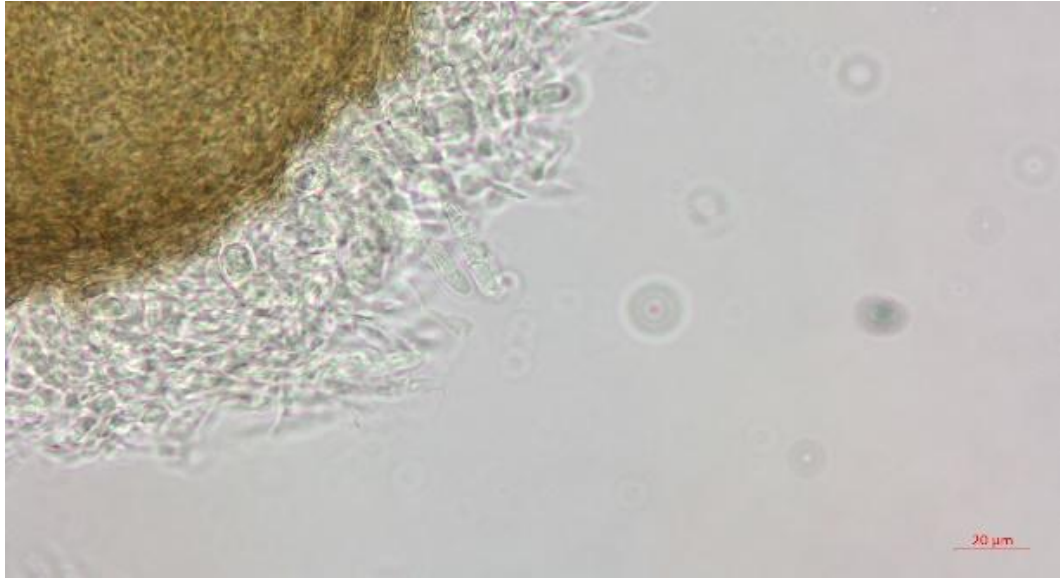


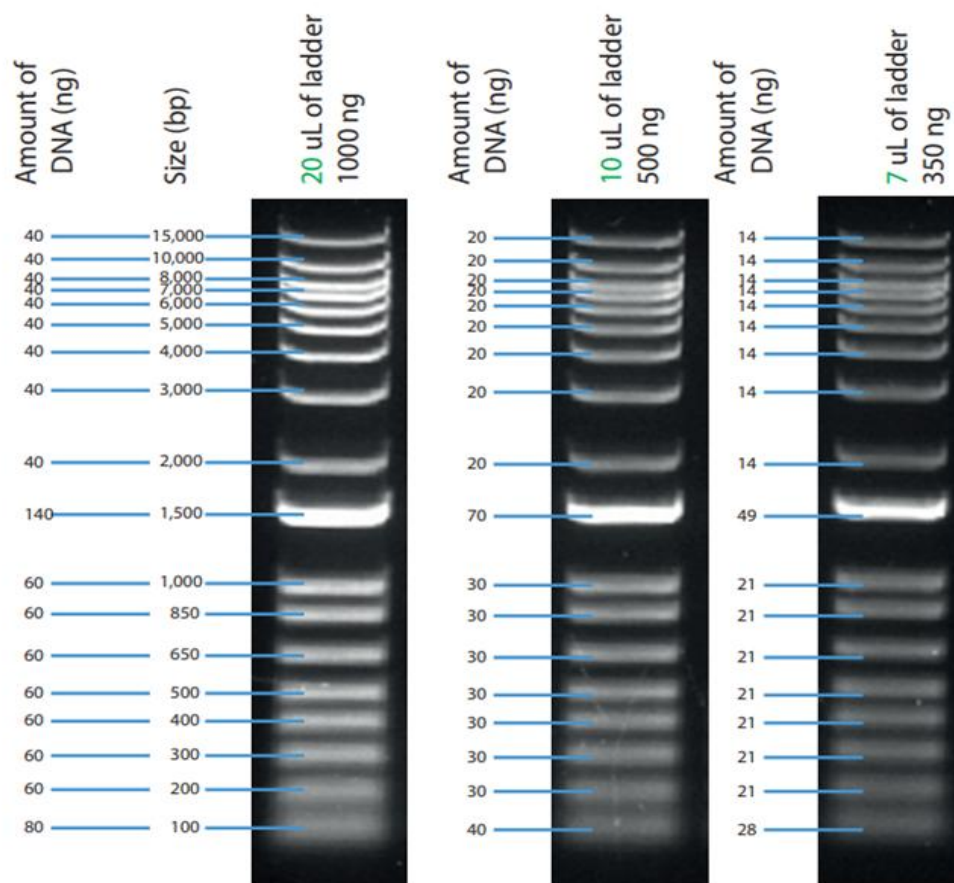
Figura 24. Ascas diferenciadas. En cada saco hay o ascosporas bien definidas.





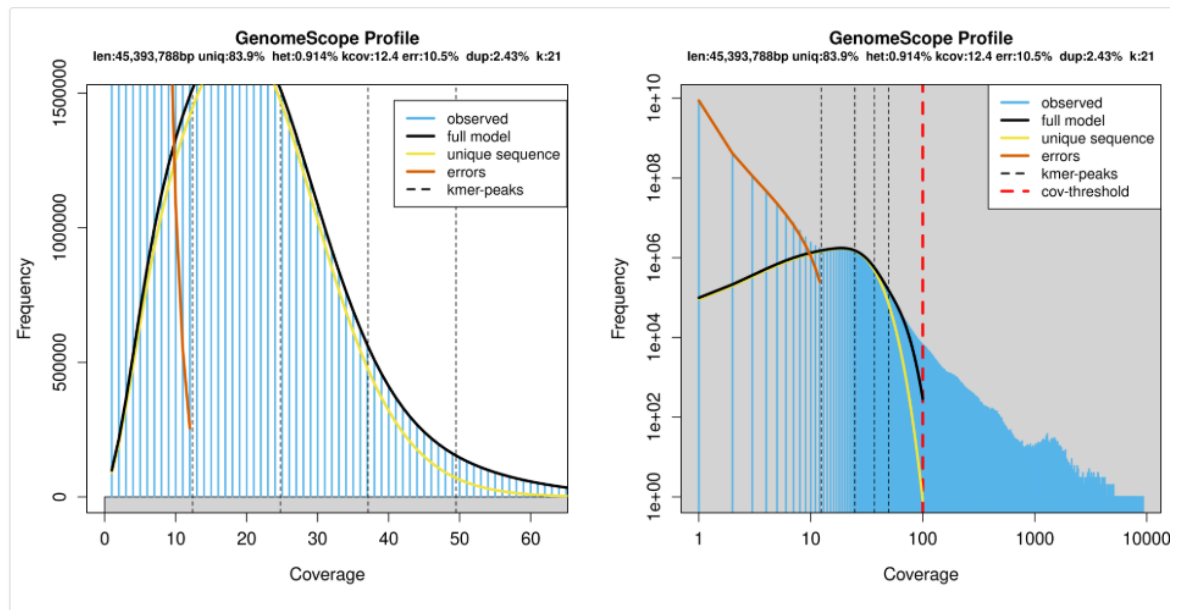
*Figura 25.* Peritecio con ascas saliendo de él. Las 8 ascosporas triseptadas puede verse en cada célula.

## Anexo D. Referencia de ladder usados para los geles de agarosa



## Anexo E. Análisis de K-meros usando GenomeScope

### MNG128 S



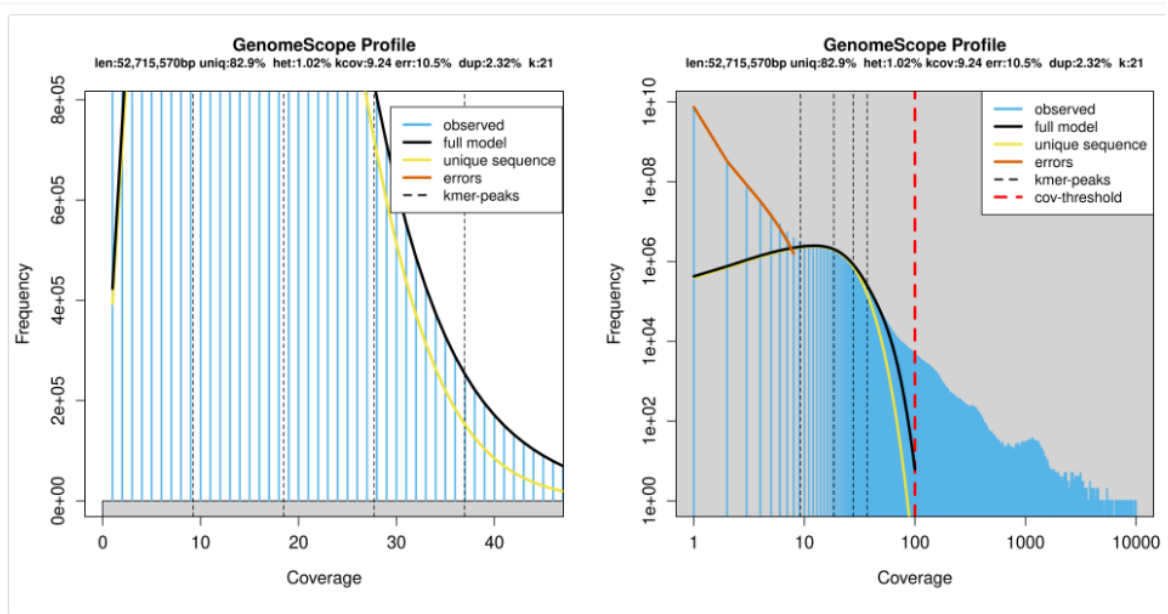
#### Results

GenomeScope version 1.0  
k = 21

property	min	max
Heterozygosity	0.85377%	0.973336%
Genome Haploid Length	44,837,867 bp	45,393,788 bp
Genome Repeat Length	7,285,005 bp	7,254,337 bp
Genome Unique Length	37,632,861 bp	38,099,451 bp
Model Fit	97.9456%	98.544%
Read Error Rate	10.5198%	10.5198%

<http://genomescope.org/analysis.php?code=XRAjw6uDy6pi9ZQJMedE>

## MNG133S



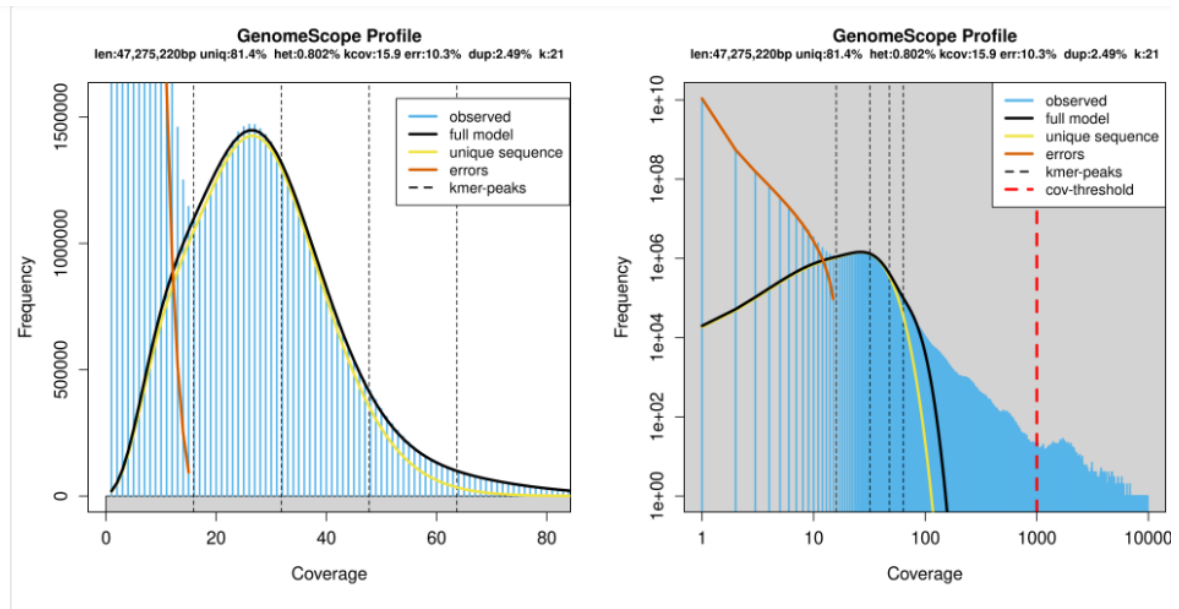
### Results

GenomeScope version 1.0  
k = 21

property	min	max
Heterozygosity	0.581476%	1.45799%
Genome Haploid Length	48,706,361 bp	52,715,570 bp
Genome Repeat Length	8,342,861 bp	9,029,594 bp
Genome Unique Length	40,363,501 bp	43,685,976 bp
Model Fit	93.2142%	93.8625%
Read Error Rate	10.4644%	10.4644%

<http://genomescope.org/analysis.php?code=md1yYETImSdTma6Qx0HD>

## MNG137S



### Results

GenomeScope version 1.0

k = 21

property	min	max
Heterozygosity	0.787027%	0.816983%
Genome Haploid Length	47,091,185 bp	47,275,220 bp
Genome Repeat Length	8,746,938 bp	8,781,121 bp
Genome Unique Length	38,344,248 bp	38,494,099 bp
Model Fit	96.7874%	98.1271%
Read Error Rate	10.3009%	10.3009%

<http://genomescope.org/analysis.php?code=leVpzH5oRnkp4kNAg8z8>

## Anexo F. Resumen de los resultados de la anotación de *M. maydis*

Category	Description	Count / Value	Tool / Database
<b>Genome assembly size</b>	Total assembled length	39.8 Mb	QUAST
<b>N50 / L50</b>	Assembly contiguity metrics	3.0 Mb / 6	QUAST
<b>BUSCO completeness</b>	Complete (C) / Single (S) / Duplicated (D)	97.9% (97.1% S, 0.8% D)	BUSCO v5.5.0
<b>Predicted genes</b>	Total number of gene models	10,933	Funannotate
<b>Predicted proteins</b>	Number of translated CDS	10,933	Funannotate
<b>Pfam domains</b>	Functional protein domains detected	11,971	Pfam v35.0
<b>CAZy families</b>	Carbohydrate-active enzymes	794	dbCAN / CAZy
<b>MEROPS entries</b>	Proteolytic enzymes	368	MEROPS
<b>SignalP hits</b>	Predicted secreted proteins	1,016	SignalP
<b>UniProt matches</b>	Homologous proteins identified	773	UniProt (DIAMOND)
<b>Biosynthetic gene clusters (BGCs)</b>	Predicted secondary metabolite clusters	X (to fill in once you check antiSMASH results)	antiSMASH v4.1.0
<b>Major BGC types</b>	Example: PKS, NRPS, Terpene, Hybrid	X	antiSMASH v4.1.0

## **Anexo G. Workflow del análisis bioinformático**

### **Proposed Workflow Structure**

#### **1. Input Data**

- Assembled genome (*Monographella maydis* FASTA)
- Sequencing platform: PacBio Sequel II (long reads)

#### **2. Assembly Quality Assessment**

- Tools: **QUAST**, **BUSCO**
- Metrics: completeness, contiguity (N50), GC content

#### **3. Structural Annotation**

- Tool: **Funannotate Predict**
- Gene predictors: **AUGUSTUS**, **SNAP**, **GlimmerHMM**, **GeneMark-ES**
- Output: GFF3, GBK, CDS, mRNA, protein FASTA files

#### **4. Functional Annotation**

- Tool: **Funannotate Annotate**
- Databases: **Pfam**, **MEROPS**, **CAZy**, **SignalP**, **UniProt**
- Output: Annotated GBK, GFF3, protein and CDS files

#### **5. Secondary Metabolite Analysis**

- Tool: **antiSMASH / fungiSMASH**
- Input: annotated GBK file
- Output: identified biosynthetic gene clusters (BGCs)

## 6. **Downstream Analyses and Visualization**

- Data interpretation (Pfam, CAZy, SignalP categories)
- Summary statistics and figure generation



## Anexo H. Herramientas bioinformáticas y parámetros

*Tabla 6. Versiones de software usadas para el análisis de BUSCO*

<b>Tool</b>	<b>Version</b>
hmmsearch	3.3
BBTools	39.01
MetaEuk	GITDIR-NOTFOUND (version not detected)
BUSCO	5.5.0

## Anexo I. Salidas de KEGG Mapper mostrando vías metabólicas reconstruidas a partir de anotaciones de EggNOG

### KEGG Mapper Reconstruction Result

Pathway (430)	Brite (46)	Brite Table (7)	Module (91)
---------------	------------	-----------------	-------------

[Show matched objects](#)

#### Metabolism

Global and overview maps

- [01100 Metabolic pathways \(830\)](#)
- [01110 Biosynthesis of secondary metabolites \(328\)](#)
- [01120 Microbial metabolism in diverse environments \(173\)](#)
- [01200 Carbon metabolism \(84\)](#)
- [01210 2-Oxocarboxylic acid metabolism \(37\)](#)
- [01212 Fatty acid metabolism \(23\)](#)
- [01230 Biosynthesis of amino acids \(98\)](#)
- [01232 Nucleotide metabolism \(36\)](#)
- [01250 Biosynthesis of nucleotide sugars \(19\)](#)
- [01240 Biosynthesis of cofactors \(118\)](#)
- [01220 Degradation of aromatic compounds \(12\)](#)
- [01310 Nitrogen cycle \(2\)](#)
- [01320 Sulfur cycle \(7\)](#)

Carbohydrate metabolism

- [00010 Glycolysis / Gluconeogenesis \(26\)](#)
- [00020 Citrate cycle \(TCA cycle\) \(21\)](#)
- [00030 Pentose phosphate pathway \(19\)](#)
- [00040 Pentose and glucuronate interconversions \(22\)](#)
- [00051 Fructose and mannose metabolism \(24\)](#)
- [00052 Galactose metabolism \(15\)](#)
- [00053 Ascorbate and aldarate metabolism \(12\)](#)
- [00500 Starch and sucrose metabolism \(27\)](#)
- [00620 Pyruvate metabolism \(31\)](#)
- [00630 Glyoxylate and dicarboxylate metabolism \(23\)](#)
- [00640 Propanoate metabolism \(20\)](#)
- [00650 Butanoate metabolism \(17\)](#)
- [00660 C5-Branched dibasic acid metabolism \(3\)](#)
- [00562 Inositol phosphate metabolism \(24\)](#)

Energy metabolism

- [00190 Oxidative phosphorylation \(70\)](#)
- [00710 Carbon fixation by Calvin cycle \(17\)](#)
- [00720 Other carbon fixation pathways \(9\)](#)

- 00680 Methane metabolism (19)
- 00910 Nitrogen metabolism (13)
- 00920 Sulfur metabolism (16)
- Lipid metabolism
  - 00061 Fatty acid biosynthesis (8)
  - 00062 Fatty acid elongation (10)
  - 00071 Fatty acid degradation (13)
  - 00073 Cutin, suberine and wax biosynthesis (2)
  - 00074 Mycolic acid biosynthesis (1)
  - 00100 Steroid biosynthesis (17)
  - 00120 Primary bile acid biosynthesis (4)
  - 00140 Steroid hormone biosynthesis (5)
  - 00561 Glycerolipid metabolism (23)
  - 00564 Glycerophospholipid metabolism (37)
  - 00565 Ether lipid metabolism (9)
  - 00600 Sphingolipid metabolism (19)
  - 00590 Arachidonic acid metabolism (8)
  - 00591 Linoleic acid metabolism (3)
  - 00592 alpha-Linolenic acid metabolism (5)
  - 01040 Biosynthesis of unsaturated fatty acids (11)
- Nucleotide metabolism
  - 00230 Purine metabolism (51)
  - 00240 Pyrimidine metabolism (32)
- Amino acid metabolism
  - 00250 Alanine, aspartate and glutamate metabolism (28)
  - 00260 Glycine, serine and threonine metabolism (32)
  - 00270 Cysteine and methionine metabolism (38)
  - 00280 Valine, leucine and isoleucine degradation (23)
  - 00290 Valine, leucine and isoleucine biosynthesis (10)
  - 00300 Lysine biosynthesis (13)
  - 00310 Lysine degradation (23)
  - 00220 Arginine biosynthesis (22)
  - 00330 Arginine and proline metabolism (20)
  - 00340 Histidine metabolism (13)
  - 00350 Tyrosine metabolism (19)
  - 00360 Phenylalanine metabolism (13)
  - 00380 Tryptophan metabolism (22)
  - 00400 Phenylalanine, tyrosine and tryptophan biosynthesis (19)
- Metabolism of other amino acids
  - 00410 beta-Alanine metabolism (14)
  - 00430 Taurine and hypotaurine metabolism (7)
  - 00440 Phosphonate and phosphinate metabolism (5)

- 00450 Selenocompound metabolism (8)
- 00460 Cyanoamino acid metabolism (9)
- 00470 D-Amino acid metabolism (3)
- 00480 Glutathione metabolism (20)
- Glycan biosynthesis and metabolism
  - 00520 Amino sugar and nucleotide sugar metabolism (28)
  - 00541 Biosynthesis of various nucleotide sugars (7)
  - 00510 N-Glycan biosynthesis (29)
  - 00513 Various types of N-glycan biosynthesis (22)
  - 00515 Mannose type O-glycan biosynthesis (1)
  - 00514 Other types of O-glycan biosynthesis (4)
  - 00531 Glycosaminoglycan degradation (4)
  - 00563 Glycosylphosphatidylinositol (GPI)-anchor biosynthesis (22)
  - 00603 Glycosphingolipid biosynthesis - globo and isoglobo series (2)
  - 00604 Glycosphingolipid biosynthesis - ganglio series (1)
  - 00511 Other glycan degradation (7)
  - 00540 Lipopolysaccharide biosynthesis (2)
- Metabolism of cofactors and vitamins
  - 00730 Thiamine metabolism (11)
  - 00740 Riboflavin metabolism (11)
  - 00750 Vitamin B6 metabolism (7)
  - 00760 Nicotinate and nicotinamide metabolism (17)
  - 00770 Pantothenate and CoA biosynthesis (17)
  - 00780 Biotin metabolism (4)
  - 00785 Lipoic acid metabolism (15)
  - 00790 Folate biosynthesis (19)
  - 00670 One carbon pool by folate (21)
  - 00830 Retinol metabolism (4)
  - 00860 Porphyrin metabolism (18)
  - 00130 Ubiquinone and other terpenoid-quinone biosynthesis (11)
- Metabolism of terpenoids and polyketides
  - 00900 Terpenoid backbone biosynthesis (18)
  - 00909 Sesquiterpenoid and triterpenoid biosynthesis (2)
  - 00904 Diterpenoid biosynthesis; Including: Gibberellin biosynthesis (1)
  - 00906 Carotenoid biosynthesis (4)
  - 00981 Insect hormone biosynthesis (1)
  - 00908 Zeatin biosynthesis (1)
  - 00903 Limonene degradation (2)
  - 00907 Pinene, camphor and geraniol degradation (2)
  - 01051 Biosynthesis of ansamycins (1)
  - 01057 Biosynthesis of type II polyketide products (1)
  - 00523 Polyketide sugar unit biosynthesis (2)

01055 Biosynthesis of vancomycin group antibiotics (1)

Biosynthesis of other secondary metabolites

00940 Phenylpropanoid biosynthesis (4)

00944 Flavone and flavonol biosynthesis (1)

00946 Degradation of flavonoids (2)

00901 Indole alkaloid biosynthesis (3)

00950 Isoquinoline alkaloid biosynthesis (6)

00960 Tropane, piperidine and pyridine alkaloid biosynthesis (5)

00232 Caffeine metabolism (2)

00965 Betalain biosynthesis (2)

00966 Glucosinolate biosynthesis (1)

00311 Penicillin and cephalosporin biosynthesis (2)

00332 Carbapenem biosynthesis (2)

00261 Monobactam biosynthesis (4)

00521 Streptomycin biosynthesis (6)

00524 Neomycin, kanamycin and gentamicin biosynthesis (1)

00525 Acarbose and validamycin biosynthesis (1)

00401 Novobiocin biosynthesis (2)

00404 Staurosporine biosynthesis (2)

00405 Phenazine biosynthesis (1)

00333 Prodigiosin biosynthesis (2)

00254 Aflatoxin biosynthesis (3)

00998 Biosynthesis of various antibiotics; Including: Kanosamine biosynthesis, Aurachin biosynthesis, Bacilysin biosynthesis, Puromycin biosynthesis, Dapdiamides biosynthesis, Fosfomycin biosynthesis, Cremeomycin biosynthesis, Fumagillin biosynthesis, Pentalenolactone biosynthesis, Terpentecin biosynthesis, Roseoflavin biosynthesis, Cycloserine biosynthesis (2)

00999 Biosynthesis of various plant secondary metabolites; Including: Crocin biosynthesis, Ginsenoside biosynthesis, Saponin adjuvant biosynthesis, Cannabidiol biosynthesis, Mugineic acid biosynthesis, Pentagalloylglucose biosynthesis, Benzoxazinoid biosynthesis, Gramine biosynthesis, Coumarin biosynthesis, Furanocoumarin biosynthesis, Hordatine biosynthesis, Podophyllotoxin biosynthesis (4)

Xenobiotics biodegradation and metabolism

00362 Benzoate degradation (7)

00627 Aminobenzoate degradation (8)

00364 Fluorobenzoate degradation (2)

00625 Chloroalkane and chloroalkene degradation (5)

00361 Chlorocyclohexane and chlorobenzene degradation (4)

00623 Toluene degradation (3)

00643 Styrene degradation (6)

00791 Atrazine degradation (3)

00930 Caprolactam degradation (6)

00621 Dioxin degradation (1)

00626 Naphthalene degradation (3)

00624 Polycyclic aromatic hydrocarbon degradation (1)

00984 Steroid degradation (1)

00980 Metabolism of xenobiotics by cytochrome P450 (8)  
00982 Drug metabolism - cytochrome P450 (7)  
00983 Drug metabolism - other enzymes (16)

## **Genetic Information Processing**

### Transcription

03020 RNA polymerase (27)  
03022 Basal transcription factors (30)  
03040 Spliceosome (90)

### Translation

03010 Ribosome (122)  
00970 Aminoacyl-tRNA biosynthesis (25)  
03013 Nucleocytoplasmic transport (58)  
03015 mRNA surveillance pathway (46)  
03008 Ribosome biogenesis in eukaryotes (68)

### Folding, sorting and degradation

03060 Protein export (21)  
04141 Protein processing in endoplasmic reticulum (71)  
04130 SNARE interactions in vesicular transport (17)  
04120 Ubiquitin mediated proteolysis (54)  
04122 Sulfur relay system (10)  
03050 Proteasome (36)  
03018 RNA degradation (51)

### Replication and repair

03030 DNA replication (33)  
03410 Base excision repair (30)  
03420 Nucleotide excision repair (48)  
03430 Mismatch repair (20)  
03440 Homologous recombination (24)  
03450 Non-homologous end-joining (10)  
03460 Fanconi anemia pathway (22)

### Chromosome

03082 ATP-dependent chromatin remodeling (35)  
03083 Polycomb repressive complex (11)

### Information processing in viruses

03250 Viral life cycle - HIV-1 (15)  
03264 Virion - Flavivirus and Alphavirus (1)  
03265 Virion - Ebolavirus, Lyssavirus and Morbillivirus (1)  
03273 Virion - Lassa virus and SFTS virus (1)  
03272 Virion - Hepatitis viruses (8)  
03266 Virion - Herpesvirus (1)

## Environmental Information Processing

### Membrane transport

- 02010 ABC transporters (9)
- 03070 Bacterial secretion system (2)

### Signal transduction

- 02020 Two-component system (19)
- 04010 MAPK signaling pathway (23)
- 04013 MAPK signaling pathway - fly (15)
- 04016 MAPK signaling pathway - plant (4)
- 04011 MAPK signaling pathway - yeast (68)
- 04012 ErbB signaling pathway (9)
- 04014 Ras signaling pathway (23)
- 04015 Rap1 signaling pathway (15)
- 04310 Wnt signaling pathway (16)
- 04330 Notch signaling pathway (3)
- 04340 Hedgehog signaling pathway (5)
- 04341 Hedgehog signaling pathway - fly (7)
- 04350 TGF-beta signaling pathway (10)
- 04390 Hippo signaling pathway (10)
- 04391 Hippo signaling pathway - fly (8)
- 04392 Hippo signaling pathway - multiple species (7)
- 04370 VEGF signaling pathway (13)
- 04371 Apelin signaling pathway (16)
- 04630 JAK-STAT signaling pathway (4)
- 04064 NF-kappa B signaling pathway (6)
- 04668 TNF signaling pathway (9)
- 04066 HIF-1 signaling pathway (15)
- 04068 FoxO signaling pathway (18)
- 04020 Calcium signaling pathway (13)
- 04070 Phosphatidylinositol signaling system (21)
- 04072 Phospholipase D signaling pathway (14)
- 04071 Sphingolipid signaling pathway (27)
- 04024 cAMP signaling pathway (19)
- 04022 cGMP-PKG signaling pathway (16)
- 04151 PI3K-Akt signaling pathway (27)
- 04152 AMPK signaling pathway (28)
- 04150 mTOR signaling pathway (41)
- 04075 Plant hormone signal transduction (4)

### Signaling molecules and interaction

- 04080 Neuroactive ligand-receptor interaction (3)
- 04082 Neuroactive ligand signaling (5)

- 04081 Hormone signaling (6)
- 04517 IgSF CAM signaling (25)

## Cellular Processes

### Transport and catabolism

- 04144 Endocytosis (70)
- 04145 Phagosome (33)
- 04142 Lysosome (29)
- 04146 Peroxisome (45)
- 04140 Autophagy - animal (50)
- 04138 Autophagy - yeast (71)
- 04136 Autophagy - other (22)
- 04137 Mitophagy - animal (27)
- 04139 Mitophagy - yeast (32)
- 04148 Efferocytosis (23)

### Cell growth and death

- 04110 Cell cycle (59)
- 04111 Cell cycle - yeast (95)
- 04112 Cell cycle - Caulobacter (3)
- 04113 Meiosis - yeast (68)
- 04114 Oocyte meiosis (33)
- 04210 Apoptosis (13)
- 04214 Apoptosis - fly (14)
- 04215 Apoptosis - multiple species (2)
- 04216 Ferroptosis (9)
- 04217 Necroptosis (21)
- 04115 p53 signaling pathway (7)
- 04218 Cellular senescence (27)

### Cellular community - eukaryotes

- 04510 Focal adhesion (20)
- 04520 Adherens junction (14)
- 04530 Tight junction (28)
- 04540 Gap junction (9)
- 04550 Signaling pathways regulating pluripotency of stem cells (5)

### Cellular community - prokaryotes

- 02024 Quorum sensing (10)
- 02025 Biofilm formation - Pseudomonas aeruginosa (2)
- 02026 Biofilm formation - Escherichia coli (1)

### Cell motility

- 04814 Motor proteins (30)
- 04820 Cytoskeleton in muscle cells (10)
- 04810 Regulation of actin cytoskeleton (28)



## Organismal Systems

### Immune system

- 04611 Platelet activation (13)
- 04613 Neutrophil extracellular trap formation (22)
- 04620 Toll-like receptor signaling pathway (4)
- 04624 Toll and Imd signaling pathway (7)
- 04621 NOD-like receptor signaling pathway (18)
- 04622 RIG-I-like receptor signaling pathway (7)
- 04623 Cytosolic DNA-sensing pathway (16)
- 04625 C-type lectin receptor signaling pathway (10)
- 04650 Natural killer cell mediated cytotoxicity (10)
- 04612 Antigen processing and presentation (8)
- 04660 T cell receptor signaling pathway (16)
- 04658 Th1 and Th2 cell differentiation (4)
- 04659 Th17 cell differentiation (6)
- 04657 IL-17 signaling pathway (9)
- 04662 B cell receptor signaling pathway (8)
- 04664 Fc epsilon RI signaling pathway (9)
- 04666 Fc gamma R-mediated phagocytosis (24)
- 04670 Leukocyte transendothelial migration (12)
- 04062 Chemokine signaling pathway (20)

### Endocrine system

- 04911 Insulin secretion (5)
- 04910 Insulin signaling pathway (25)
- 04922 Glucagon signaling pathway (19)
- 04923 Regulation of lipolysis in adipocytes (6)
- 04920 Adipocytokine signaling pathway (6)
- 03320 PPAR signaling pathway (12)
- 04929 GnRH secretion (3)
- 04912 GnRH signaling pathway (13)
- 04913 Ovarian steroidogenesis (4)
- 04915 Estrogen signaling pathway (12)
- 04914 Progesterone-mediated oocyte maturation (22)
- 04917 Prolactin signaling pathway (6)
- 04921 Oxytocin signaling pathway (20)
- 04926 Relaxin signaling pathway (11)
- 04935 Growth hormone synthesis, secretion and action (11)
- 04918 Thyroid hormone synthesis (9)
- 04919 Thyroid hormone signaling pathway (19)
- 04928 Parathyroid hormone synthesis, secretion and action (11)
- 04916 Melanogenesis (8)

- 04924 Renin secretion (5)
- 04614 Renin-angiotensin system (1)
- 04925 Aldosterone synthesis and secretion (9)
- 04927 Cortisol synthesis and secretion (3)
- Circulatory system
  - 04260 Cardiac muscle contraction (16)
  - 04261 Adrenergic signaling in cardiomyocytes (15)
  - 04270 Vascular smooth muscle contraction (12)
- Digestive system
  - 04970 Salivary secretion (5)
  - 04971 Gastric acid secretion (6)
  - 04972 Pancreatic secretion (11)
  - 04976 Bile secretion (7)
  - 04973 Carbohydrate digestion and absorption (3)
  - 04974 Protein digestion and absorption (6)
  - 04975 Fat digestion and absorption (5)
  - 04979 Cholesterol metabolism (6)
  - 04977 Vitamin digestion and absorption (1)
  - 04981 Folate transport and metabolism (8)
  - 04980 Cobalamin transport and metabolism (1)
  - 04978 Mineral absorption (6)
- Excretory system
  - 04962 Vasopressin-regulated water reabsorption (13)
  - 04960 Aldosterone-regulated sodium reabsorption (5)
  - 04961 Endocrine and other factor-regulated calcium reabsorption (9)
  - 04964 Proximal tubule bicarbonate reclamation (2)
  - 04966 Collecting duct acid secretion (11)
- Nervous system
  - 04724 Glutamatergic synapse (10)
  - 04727 GABAergic synapse (11)
  - 04725 Cholinergic synapse (7)
  - 04728 Dopaminergic synapse (18)
  - 04726 Serotonergic synapse (11)
  - 04720 Long-term potentiation (11)
  - 04730 Long-term depression (8)
  - 04723 Retrograde endocannabinoid signaling (29)
  - 04721 Synaptic vesicle cycle (23)
  - 04722 Neurotrophin signaling pathway (18)
- Sensory system
  - 04744 Phototransduction (2)
  - 04745 Phototransduction - fly (3)
  - 04740 Olfactory transduction (3)

- 04742 Taste transduction (2)
- 04750 Inflammatory mediator regulation of TRP channels (8)

#### Development and regeneration

- 04320 Dorso-ventral axis formation (4)
- 04360 Axon guidance (16)
- 04361 Axon regeneration (12)
- 04380 Osteoclast differentiation (8)
- 04382 Cornified envelope formation (10)

#### Aging

- 04211 Longevity regulating pathway (18)
- 04212 Longevity regulating pathway - worm (19)
- 04213 Longevity regulating pathway - multiple species (19)

#### Environmental adaptation

- 04710 Circadian rhythm (7)
- 04713 Circadian entrainment (6)
- 04711 Circadian rhythm - fly (1)
- 04712 Circadian rhythm - plant (2)
- 04714 Thermogenesis (87)
- 04626 Plant-pathogen interaction (6)

### Human Diseases

#### Cancer: overview

- 05200 Pathways in cancer (39)
- 05202 Transcriptional misregulation in cancer (10)
- 05206 MicroRNAs in cancer (21)
- 05205 Proteoglycans in cancer (21)
- 05204 Chemical carcinogenesis - DNA adducts (5)
- 05207 Chemical carcinogenesis - receptor activation (15)
- 05208 Chemical carcinogenesis - reactive oxygen species (69)
- 05203 Viral carcinogenesis (41)
- 05230 Central carbon metabolism in cancer (11)
- 05231 Choline metabolism in cancer (16)
- 05235 PD-L1 expression and PD-1 checkpoint pathway in cancer (9)

#### Cancer: specific types

- 05210 Colorectal cancer (15)
- 05212 Pancreatic cancer (11)
- 05225 Hepatocellular carcinoma (18)
- 05226 Gastric cancer (10)
- 05214 Glioma (10)
- 05216 Thyroid cancer (5)
- 05221 Acute myeloid leukemia (5)
- 05220 Chronic myeloid leukemia (7)

- 05217 Basal cell carcinoma (2)
- 05218 Melanoma (5)
- 05211 Renal cell carcinoma (12)
- 05219 Bladder cancer (3)
- 05215 Prostate cancer (10)
- 05213 Endometrial cancer (9)
- 05224 Breast cancer (8)
- 05222 Small cell lung cancer (6)
- 05223 Non-small cell lung cancer (9)
- Infectious disease: viral
  - 05166 Human T-cell leukemia virus 1 infection (41)
  - 05170 Human immunodeficiency virus 1 infection (30)
  - 05161 Hepatitis B (13)
  - 05160 Hepatitis C (18)
  - 05171 Coronavirus disease - COVID-19 (75)
  - 05164 Influenza A (20)
  - 05162 Measles (13)
  - 05168 Herpes simplex virus 1 infection (17)
  - 05163 Human cytomegalovirus infection (23)
  - 05167 Kaposi sarcoma-associated herpesvirus infection (16)
  - 05169 Epstein-Barr virus infection (30)
  - 05165 Human papillomavirus infection (41)
- Infectious disease: bacterial
  - 05110 Vibrio cholerae infection (21)
  - 05120 Epithelial cell signaling in Helicobacter pylori infection (17)
  - 05130 Pathogenic Escherichia coli infection (33)
  - 05132 Salmonella infection (53)
  - 05131 Shigellosis (51)
  - 05135 Yersinia infection (23)
  - 05133 Pertussis (6)
  - 05134 Legionellosis (11)
  - 05152 Tuberculosis (22)
  - 05100 Bacterial invasion of epithelial cells (16)
- Infectious disease: parasitic
  - 05146 Amoebiasis (6)
  - 05145 Toxoplasmosis (8)
  - 05140 Leishmaniasis (6)
  - 05142 Chagas disease (6)
  - 05143 African trypanosomiasis (2)
- Immune disease
  - 05322 Systemic lupus erythematosus (8)
  - 05323 Rheumatoid arthritis (13)

- 05340 Primary immunodeficiency (2)
- Neurodegenerative disease
  - 05010 Alzheimer disease (121)
  - 05012 Parkinson disease (107)
  - 05014 Amyotrophic lateral sclerosis (140)
  - 05016 Huntington disease (123)
  - 05017 Spinocerebellar ataxia (58)
  - 05020 Prion disease (107)
  - 05022 Pathways of neurodegeneration - multiple diseases (137)
- Substance dependence
  - 05030 Cocaine addiction (5)
  - 05031 Amphetamine addiction (10)
  - 05032 Morphine addiction (5)
  - 05033 Nicotine addiction (1)
  - 05034 Alcoholism (22)
- Cardiovascular disease
  - 05417 Lipid and atherosclerosis (27)
  - 05418 Fluid shear stress and atherosclerosis (14)
  - 05410 Hypertrophic cardiomyopathy (4)
  - 05412 Arrhythmogenic right ventricular cardiomyopathy (2)
  - 05414 Dilated cardiomyopathy (3)
  - 05415 Diabetic cardiomyopathy (72)
  - 05416 Viral myocarditis (6)
- Endocrine and metabolic disease
  - 04930 Type II diabetes mellitus (6)
  - 04940 Type I diabetes mellitus (2)
  - 04950 Maturity onset diabetes of the young (1)
  - 04936 Alcoholic liver disease (15)
  - 04932 Non-alcoholic fatty liver disease (48)
  - 04931 Insulin resistance (14)
  - 04933 AGE-RAGE signaling pathway in diabetic complications (13)
  - 04934 Cushing syndrome (13)
- Drug resistance: antimicrobial
  - 01501 beta-Lactam resistance (1)
  - 01503 Cationic antimicrobial peptide (CAMP) resistance (1)
- Drug resistance: antineoplastic
  - 01521 EGFR tyrosine kinase inhibitor resistance (11)
  - 01524 Platinum drug resistance (16)
  - 01523 Antifolate resistance (6)
  - 01522 Endocrine resistance (7)

## KEGG Mapper Reconstruction Result

[Pathway \(430\)](#)[Brite \(46\)](#)[Brite Table \(7\)](#)[Module \(91\)](#)

[Show matched objects](#)

### Genes and Proteins

Orthologs and modules

[ko00001](#) KEGG Orthology (KO) ([3427](#))

Protein families: metabolism

[ko01000](#) Enzymes ([1431](#))

[ko01001](#) Protein kinases ([91](#))

[ko01009](#) Protein phosphatases and associated proteins ([82](#))

[ko01002](#) Peptidases and inhibitors ([116](#))

[ko01003](#) Glycosyltransferases ([41](#))

[ko01005](#) Lipopolysaccharide biosynthesis proteins ([2](#))

[ko01011](#) Peptidoglycan biosynthesis and degradation proteins ([1](#))

[ko01004](#) Lipid biosynthesis proteins ([25](#))

[ko01008](#) Polyketide biosynthesis proteins ([5](#))

[ko01006](#) Prenyltransferases ([14](#))

[ko01007](#) Amino acid related enzymes ([39](#))

[ko00199](#) Cytochrome P450 ([10](#))

Protein families: genetic information processing

[ko03000](#) Transcription factors ([73](#))

[ko03021](#) Transcription machinery ([156](#))

[ko03019](#) Messenger RNA biogenesis ([179](#))

[ko03041](#) Spliceosome ([147](#))

[ko03011](#) Ribosome ([121](#))

[ko03009](#) Ribosome biogenesis ([213](#))

[ko03016](#) Transfer RNA biogenesis ([122](#))

[ko03012](#) Translation factors ([65](#))

[ko03110](#) Chaperones and folding catalysts ([78](#))

[ko04131](#) Membrane trafficking ([374](#))

[ko04121](#) Ubiquitin system ([140](#))

[ko03051](#) Proteasome ([48](#))

[ko03032](#) DNA replication proteins ([80](#))

[ko03036](#) Chromosome and associated proteins ([302](#))

[ko03400](#) DNA repair and recombination proteins ([159](#))

[ko03029](#) Mitochondrial biogenesis ([191](#))

Protein families: signaling and cellular processes

- [ko02000](#) Transporters (145)
- [ko02044](#) Secretion system (18)
- [ko02022](#) Two-component system (6)
- [ko03037](#) Cilium and associated proteins (14)
- [ko04812](#) Cytoskeleton proteins (75)
- [ko04147](#) Exosome (175)
- [ko02048](#) Prokaryotic defense system (3)
- [ko04030](#) G protein-coupled receptors (2)
- [ko04040](#) Ion channels (7)
- [ko04031](#) GTP-binding proteins (37)
- [ko04052](#) Cytokines and neuropeptides (2)
- [ko04090](#) CD molecules (3)
- [ko00535](#) Proteoglycans (1)
- [ko00536](#) Glycosaminoglycan binding proteins (4)
- [ko00537](#) Glycosylphosphatidylinositol (GPI)-anchored proteins (7)
- [ko04091](#) Lectins (7)
- [ko04990](#) Domain-containing proteins not elsewhere classified (10)

## KEGG Mapper Reconstruction Result

[Pathway \(430\)](#)[Brite \(46\)](#)[Brite Table \(7\)](#)[Module \(91\)](#)

[Show matched objects](#)

### Annotation guide

[br01610](#) Ribosomal proteins (121)

[br01611](#) RNA polymerases (26)

[br01612](#) DNA polymerases (18)

[br01613](#) Aminoacyl-tRNA synthetases (22)

[br01601](#) Enzymes of 2-oxocarboxylic acid metabolism (23)

[br01620](#) Photosynthetic and chemosynthetic capacities (1)

### Compounds and Reactions

#### Reaction information

[br08203](#) Glycosyltransferase reactions (39)



## KEGG Mapper Reconstruction Result

[Pathway \(430\)](#)[Brite \(46\)](#)[Brite Table \(7\)](#)[Module \(91\)](#)

View ☒ complete only ☐ including 1 block missing ☐ including any incomplete

[Exec](#)[Show matched objects](#)

### Pathway modules

#### Carbohydrate metabolism

Central carbohydrate metabolism

- [M00001](#) Glycolysis (Embden-Meyerhof pathway), glucose => pyruvate ([10](#)) (complete 9/9)
- [M00002](#) Glycolysis, core module involving three-carbon compounds ([6](#)) (complete 5/5)
- [M00003](#) Gluconeogenesis, oxaloacetate => fructose-6P ([8](#)) (complete 7/7)
- [M00307](#) Pyruvate oxidation, pyruvate => acetyl-CoA ([4](#)) (complete 1/1)
- [M00009](#) Citrate cycle (TCA cycle, Krebs cycle) ([15](#)) (complete 8/8)
- [M00010](#) Citrate cycle, first carbon oxidation, oxaloacetate => 2-oxoglutarate ([4](#)) (complete 3/3)
- [M00011](#) Citrate cycle, second carbon oxidation, 2-oxoglutarate => oxaloacetate ([11](#)) (complete 5/5)
- [M00004](#) Pentose phosphate pathway (Pentose phosphate cycle) ([9](#)) (complete 6/6)
- [M00006](#) Pentose phosphate pathway, oxidative phase, glucose 6P => ribulose 5P ([3](#)) (complete 2/2)
- [M00007](#) Pentose phosphate pathway, non-oxidative phase, fructose 6P => ribose 5P ([5](#)) (complete 4/4)
- [M00005](#) PRPP biosynthesis, ribose 5P => PRPP ([1](#)) (complete 1/1)

Other carbohydrate metabolism

- [M00630](#) D-Galacturonate degradation (fungi), D-galacturonate => glycerol ([5](#)) (complete 4/4)
- [M00081](#) Pectin degradation ([3](#)) (complete 3/3)
- [M00632](#) Galactose degradation, Leloir pathway, galactose => alpha-D-glucose-1P ([4](#)) (complete 4/4)
- [M00854](#) Glycogen biosynthesis, glucose-1P => glycogen/starch ([4](#)) (complete 2/2)
- [M00855](#) Glycogen degradation, glycogen => glucose-6P ([3](#)) (complete 3/3)
- [M00012](#) Glyoxylate cycle ([5](#)) (complete 5/5)

#### Energy metabolism

Methane metabolism

- [M00344](#) Formaldehyde assimilation, xylulose monophosphate pathway ([4](#)) (complete 4/4)

Nitrogen metabolism

- [M00531](#) Assimilatory nitrate reduction, nitrate => ammonia ([2](#)) (complete 2/2)

Sulfur metabolism

- [M00176](#) Assimilatory sulfate reduction, sulfate => H<sub>2</sub>S ([5](#)) (complete 3/3)

ATP synthesis

- [M00143](#) NADH dehydrogenase (ubiquinone) Fe-S protein/flavoprotein complex, mitochondria ([10](#)) (complete 1/1)
- [M00148](#) Succinate dehydrogenase (ubiquinone) ([4](#)) (complete 1/1)

## **Lipid metabolism**

### **Fatty acid metabolism**

- M00082 Fatty acid biosynthesis, initiation (4) (complete 2/2)
- M00083 Fatty acid biosynthesis, elongation (4) (complete 1/1)
- M00415 Fatty acid elongation in endoplasmic reticulum (5) (complete 4/4)
- M00086 beta-Oxidation, acyl-CoA synthesis (1) (complete 1/1)
- M00087 beta-Oxidation (5) (complete 3/3)

### **Sterol biosynthesis**

- M00102 Ergocalciferol biosynthesis, FPP => ergosterol/ergocalciferol (13) (complete 13/13)

### **Lipid metabolism**

- M00089 Triacylglycerol biosynthesis (7) (complete 4/4)
- M00098 Acylglycerol degradation (5) (complete 2/2)
- M00090 Phosphatidylcholine (PC) biosynthesis, choline => PC (3) (complete 3/3)
- M00091 Phosphatidylcholine (PC) biosynthesis, PE => PC (2) (complete 1/1)
- M00092 Phosphatidylethanolamine (PE) biosynthesis, ethanolamine => PE (3) (complete 3/3)
- M00093 Phosphatidylethanolamine (PE) biosynthesis, PA => PS => PE (3) (complete 3/3)
- M00094 Ceramide biosynthesis (4) (complete 4/4)
- M00099 Sphingosine biosynthesis (5) (complete 5/5)
- M00100 Sphingosine degradation (2) (complete 2/2)

## **Nucleotide metabolism**

### **Purine metabolism**

- M00048 De novo purine biosynthesis, PRPP + glutamine => IMP (8) (complete 8/8)
- M00049 Adenine ribonucleotide biosynthesis, IMP => ADP,ATP (6) (complete 4/4)
- M00050 Guanine ribonucleotide biosynthesis, IMP => GDP,GTP (4) (complete 4/4)
- M00053 Deoxyribonucleotide biosynthesis, ADP/GDP/CDP/UDP => dATP/dGTP/dCTP/dUTP (3) (complete 2/2)

### **Pyrimidine metabolism**

- M00051 De novo pyrimidine biosynthesis, glutamine (+ PRPP) => UMP (9) (complete 3/3)
- M00052 Pyrimidine ribonucleotide biosynthesis, UMP => UDP/UTP,CDP/CTP (3) (complete 3/3)

## **Amino acid metabolism**

### **Serine and threonine metabolism**

- M00020 Serine biosynthesis, glycerate-3P => serine (3) (complete 3/3)
- M00018 Threonine biosynthesis, aspartate => homoserine => threonine (5) (complete 5/5)
- M00621 Glycine cleavage system (3) (complete 3/3)
- M00555 Betaine biosynthesis, choline => betaine (2) (complete 1/1)

### **Cysteine and methionine metabolism**

- M00338 Cysteine biosynthesis, homocysteine + serine => cysteine (2) (complete 2/2)
- M00017 Methionine biosynthesis, aspartate => homoserine => methionine (7) (complete 7/7)
- M00034 Methionine salvage pathway (9) (complete 8/8)
- M00035 Methionine degradation (4) (complete 4/4)

### **Branched-chain amino acid metabolism**

- M00019 Valine/isoleucine biosynthesis, pyruvate => valine / 2-oxobutanoate => isoleucine (5) (complete 4/4)
- M00570 Isoleucine biosynthesis, threonine => 2-oxobutanoate => isoleucine (7) (complete 5/5)
- M00432 Leucine biosynthesis, 2-oxoisovalerate => 2-oxoisocaproate (3) (complete 3/3)

M00036 Leucine degradation, leucine => acetoacetate + acetyl-CoA (11) (complete 6/6)

Lysine metabolism

M00030 Lysine biosynthesis, AAA pathway, 2-oxoglutarate => 2-aminoadipate => lysine (8) (complete 8/8)

M00433 Lysine biosynthesis, 2-oxoglutarate => 2-oxoadipate (4) (complete 3/3)

Arginine and proline metabolism

M00028 Ornithine biosynthesis, glutamate => ornithine (5) (complete 4/4)

M00844 Arginine biosynthesis, ornithine => arginine (3) (complete 3/3)

M00015 Proline biosynthesis, glutamate => proline (3) (complete 2/2)

M00970 Proline degradation, proline => glutamate (2) (complete 1/1)

M00972 Proline metabolism (3) (complete 3/3)

Polyamine biosynthesis

M00134 Polyamine biosynthesis, arginine => ornithine => putrescine (2) (complete 2/2)

Histidine metabolism

M00026 Histidine biosynthesis, PRPP => histidine (8) (complete 6/6)

Aromatic amino acid metabolism

M00022 Shikimate pathway, phosphoenolpyruvate + erythrose-4P => chorismate (6) (complete 4/4)

M00023 Tryptophan biosynthesis, chorismate => tryptophan (5) (complete 3/3)

M00024 Phenylalanine biosynthesis, chorismate => phenylpyruvate => phenylalanine (3) (complete 2/2)

M00025 Tyrosine biosynthesis, chorismate => HPP => tyrosine (3) (complete 2/2)

Other amino acid metabolism

M00027 GABA (gamma-Aminobutyrate) shunt (4) (complete 3/3)

M00118 Glutathione biosynthesis, glutamate => glutathione (3) (complete 2/2)

**Glycan metabolism**

Nucleotide sugar biosynthesis

M00892 UDP-GlcNAc biosynthesis, eukaryotes, Fru-6P => UDP-GlcNAc (4) (complete 4/4)

M00549 UDP-Glc biosynthesis, Glc => UDP-Glc (3) (complete 3/3)

M00999 UDP-GlcA/GalA biosynthesis, UDP-Glc => UDP-GlcA => UDP-GalA (2) (complete 2/2)

M00554 UDP-Gal biosynthesis, Gal => UDP-Gal (2) (complete 2/2)

M01000 GDP-Man biosynthesis, Fru-6P => GDP-Man (3) (complete 3/3)

M01015 GDP-Man biosynthesis, Man => GDP-Man (3) (complete 2/2)

Glycan biosynthesis

M00055 N-glycan precursor biosynthesis (13) (complete 13/13)

M00072 N-glycosylation by oligosaccharyltransferase (6) (complete 1/1)

**Metabolism of cofactors and vitamins**

Cofactor and vitamin metabolism

M00911 Riboflavin biosynthesis, fungi, GTP => riboflavin/FMN/FAD (8) (complete 8/8)

M00916 Pyridoxal-P biosynthesis, R5P + glyceraldehyde-3P + glutamine => pyridoxal-P (2) (complete 1/1)

M00120 Coenzyme A biosynthesis, pantothenate => CoA (5) (complete 3/3)

M00881 Lipoic acid biosynthesis, plants and bacteria, octanoyl-ACP => dihydrolipoyl-E2/H (2) (complete 2/2)

M00126 Tetrahydrofolate biosynthesis, GTP => THF (7) (complete 5/5)

M00880 Molybdenum cofactor biosynthesis, GTP => molybdenum cofactor (6) (complete 3/3)

M00141 C1-unit interconversion, eukaryotes (2) (complete 2/2)

M00868 Heme biosynthesis, animals and fungi, glycine => heme (8) (complete 8/8)

## Biosynthesis of terpenoids and polyketides

Terpenoid backbone biosynthesis

M00095 C5 isoprenoid biosynthesis, mevalonate pathway (7) (complete 7/7)

M00367 C10-C20 isoprenoid biosynthesis, non-plant eukaryotes (3) (complete 3/3)

Terpenoid biosynthesis

M00097 Lycopene biosynthesis, geranylgeranyl-PP => lycopene (2) (complete 2/2)

M01042 Neurosporaxanthin biosynthesis, fungi, lycopene => neurosporaxanthin (4) (complete 4/4)

## Signature modules

### Module set

Metabolic capacity

M00615 Nitrate assimilation (1) (complete 2/2)