

UACM

Universidad Autónoma
de la Ciudad de México

Nada humano me es ajeno

UNIVERSIDAD AUTÓNOMA DE LA CIUDAD DE MÉXICO

COLEGIO DE CIENCIA Y TECNOLOGÍA

POSGRADO EN CIENCIAS GENÓMICAS

Interactoma del factor de transcripción EhMyb10: Evidencias de su
papel en la regulación cotranscripcional en el parásito *Entamoeba
histolytica*

TESIS

QUE PARA OPTAR POR EL TÍTULO DE
LICENCIADA EN CIENCIAS GENÓMICAS

PRESENTA:

DANNA PAOLA IXYALIT AGUIRRE CASIMIRO

DIRECTORA: Dra. Elisa Irene Azuara Liceaga

CODIRECTORA: Dra. Patricia Cuellar Silva

POSGRADO
EN CIENCIAS
GENÓMICAS

Ciudad de México, septiembre de 2023.

SISTEMA BIBLIOTECARIO DE INFORMACIÓN Y DOCUMENTACIÓN



UNIVERSIDAD AUTÓNOMA DE LA CIUDAD DE MÉXICO COORDINACIÓN ACADÉMICA

RESTRICCIONES DE USO PARA LAS TESIS DIGITALES

DERECHOS RESERVADOS[©]

La presente obra y cada uno de sus elementos está protegido por la Ley Federal del Derecho de Autor; por la Ley de la Universidad Autónoma de la Ciudad de México, así como lo dispuesto por el Estatuto General Orgánico de la Universidad Autónoma de la Ciudad de México; del mismo modo por lo establecido en el Acuerdo por el cual se aprueba la Norma mediante la que se Modifican, Adicionan y Derogan Diversas Disposiciones del Estatuto Orgánico de la Universidad de la Ciudad de México, aprobado por el Consejo de Gobierno el 29 de enero de 2002, con el objeto de definir las atribuciones de las diferentes unidades que forman la estructura de la Universidad Autónoma de la Ciudad de México como organismo público autónomo y lo establecido en el Reglamento de Titulación de la Universidad Autónoma de la Ciudad de México.

Por lo que el uso de su contenido, así como cada una de las partes que lo integran y que están bajo la tutela de la Ley Federal de Derecho de Autor, obliga a quien haga uso de la presente obra a considerar que solo lo realizará si es para fines educativos, académicos, de investigación o informativos y se compromete a citar esta fuente, así como a su autor ó autores. Por lo tanto, queda prohibida su reproducción total o parcial y cualquier uso diferente a los ya mencionados, los cuales serán reclamados por el titular de los derechos y sancionados conforme a la legislación aplicable.



INTEGRACIÓN DEL JURADO

Presidente: Dra. Ángeles Concepción Tecalco Cruz

Secretario: Dra. Elisa Irene Azuara Liceaga

Vocal: Dra. Patricia Cuellar Silva

Plantel de Adscripción:

PLANTEL DEL VALLE, UACM

Directora

Dra. Elisa Irene Azuara Liceaga

Universidad Autónoma de la Ciudad de México

Codirectora

Dra. Patricia Cuellar Silva

Universidad Autónoma de la Ciudad de México



LECTORES DE TESIS

Dra. Ángeles Concepción Tecalco Cruz

Licenciatura en Ciencias Genómicas
Universidad Autónoma de la Ciudad de México

Dra. Helena Solleiro Villavicencio

Licenciatura en Ciencias Genómicas
Universidad Autónoma de la Ciudad de México

Dra. Lilia López Cánovas

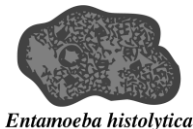
Licenciatura en Ciencias Genómicas
Universidad Autónoma de la Ciudad de México



Entamoeba histolytica

TESIS

Este trabajo fue realizado en el laboratorio 2 del posgrado en Ciencias Genómicas en el plantel del Valle de la UACM, bajo la dirección de la Dra. Elisa Irene Azuara Liceaga, en colaboración con la Dra. Patricia Cuellar Silva. Este trabajo forma parte del proyecto “*Control co-transcripcional de genes relacionados a la virulencia y el enquistamiento de Entamoeba*” (CF-2019-194163), aprobado en la convocatoria Ciencia de Frontera 2019.



AGRADECIMIENTOS

Agradezco a la Universidad Autónoma de la Ciudad de México por las becas otorgadas para realizar estudios de licenciatura semestre 2019-II y para la elaboración de trabajo recepcional o tesis 2023. Así como al apoyo otorgado mediante la convocatoria para participar en foros académicos o culturales nacionales e internacionales 2023 y para la Obtención de Grado 2023-I.

Agradezco a la Sociedad Mexicana de Bioquímica (SMB) por permitirme participar en su XXXIII Congreso Nacional de Bioquímica, llevado a cabo en la ciudad de Mérida, Yucatán, del 16 al 21 de octubre de 2022.

Agradezco a la Dra. Elisa por permitirme formar parte de su equipo de laboratorio, por su apoyo, atención, y paciencia que me brindó a lo largo de mi proyecto de investigación. Su guía experta y compromiso, han sido fundamentales para mi formación académica. Asimismo, le agradezco por ser parte de mi trayectoria académica y por ayudarme a alcanzar mis metas.

Agradezco a la Dra. Patricia por su apoyo incondicional durante mi estancia en el laboratorio 2 del posgrado de Ciencias Genómicas. Su disposición para compartir sus vastos conocimientos, así como su atención y asesoría, que han sido invaluable para mi desarrollo académico. También quiero reconocer el tiempo que dedicó para tener conversaciones enriquecedoras y momentos de risas compartidas.

Agradezco a Alfredo Padilla Barberi por su apoyo en el diseño gráfico del póster presentado en el XXXIII Congreso Nacional de Bioquímica.

Agradezco a mi grupo de trabajo del laboratorio 2 del Posgrado en Ciencias Genómicas, a la Dra. Elizabeth J. Castañeda Ortiz, a Claudia Castillo Mercado, a Landy Marlene Rojas Áviles y a mis compañeros César, Alan y Aldahir. ¡Gracias por hacer de mi tiempo en el laboratorio una experiencia excepcional!



DEDICATORIA

A mis queridos padres, Bertha Casimiro Filio y Francisco Aguirre Hernández; y a mis increíbles hermanos, Junny Annette Aguirre Casimiro, y Julio César Aguirre Casimiro:

“Cada paso que he dado en mi carrera ha sido moldeado por su apoyo constante. Su presencia ha sido mi mayor motivación y fortaleza. Agradezco de todo corazón sus sacrificios y amor, que han sido el cimiento de mis logros”.

A mi Madrina María de Jesús Juárez Rojas:

“En cada paso de nuestro camino, has estado ahí, brindándonos tu apoyo inquebrantable y tu cariño. Tu presencia ha hecho que los desafíos sean más llevaderos y las alegrías sean más significativas. Eres un ejemplo de fortaleza, bondad y amor, y estamos eternamente agradecidos por tenerte en nuestras vidas”.

ÍNDICE

1. RESUMEN	10
2. INTRODUCCIÓN	11
2.1 Transcripción	11
2.2 Regulación de la expresión génica	13
2.3 Factores de transcripción (FT)	16
2.4 Dominios de unión a DNA (DBD)	18
2.5 Factores de transcripción Myb	18
2.5.1 c-Myb	18
2.5.2 Estructura del DBD-MYB	19
2.5.3 Otras funciones de las proteínas Myb	21
2.6 <i>Entamoeba histolytica</i>	23
2.6.1 Taxonomía y antecedentes históricos	23
2.6.2 Epidemiología	25
2.6.3 Ciclo de Vida	27
2.6.4 Genoma de <i>E. histolytica</i>	31
3. ANTECEDENTES PARTICULARES	32
3.1 Factores de transcripción Myb en <i>E. histolytica</i>	32
3.2 EhMyb10	37
4. HIPÓTESIS	40
5. JUSTIFICACIÓN	40
6. OBJETIVOS	40
6.1 Objetivo general	40
6.2 Objetivos particulares	40
7. ESTRATEGIA EXPERIMENTAL	41

8. MATERIALES Y MÉTODOS	42
8.1 Identificación de PTMs, regiones desordenadas y predicción de estructura tridimensional de la proteína EhMyb10	42
8.1.1 Determinación del interactoma <i>in silico</i> de EhMyb10	43
8.1.2 Análisis ontológico (GO) de las proteínas que interactúan con EhMyb10	43
8.1.3 Análisis de expresión de los genes que codifican para proteínas que interactúan con EhMyb10	43
8.2 Preparación de células competentes de <i>Escherichia coli</i> DH5 α y BL21	44
8.3 Transformación de células competentes	44
8.4 Obtención de DNA plasmídico	44
8.5 Análisis de los plásmidos pGEX-6P-1 y pGEX-EhMyb10 mediante restricciones y PCR	45
8.6 Inducción y purificación de las proteínas recombinantes: GST y EhMyb10 (EhMyb10r)	47
8.7 Fraccionamiento subcelular de proteínas de <i>E. histolytica</i>	48
8.8 Ensayo de Pull-Down con la proteína recombinante EhMyb10	48
8.9 SDS-PAGE y Western Blot	49
8.10 Identificación de proteínas por espectrometría de masas (IMS-DIA-MS)	50
8.11 Análisis <i>in silico</i> de las proteínas identificadas en los ensayos de Pull-Down	50
9. RESULTADOS	51
9.1 Análisis <i>in silico</i> de la secuencia de aminoácidos de la proteína EhMyb10	51
9.2 Análisis de PTMs de EhMyb10	51
9.3 Análisis de las regiones desordenadas de EhMyb10	55
9.4 Análisis del interactoma <i>in silico</i> de EhMyb10	57

9.5	Análisis de GO de las proteínas que interactúan con EhMyb10	60
9.6	Análisis de expresión de EhMyb10, y de su interactoma en condiciones de Colonización/Invasión intestinal y Virulencia	62
9.7	Expresión de EhMyb10r	65
9.7.1	Restricción y PCR del gen <i>ehmyb10</i>	65
9.7.2	Inducción de la proteína EhMyb10r en células de <i>E. coli</i> BL21	66
9.7.3	Purificación por cromatografía de afinidad de EhMyb10r	68
9.8	Fraccionamiento subcelular de proteínas de <i>E. histolytica</i>	68
9.9	Ensayo de Pull-Down con la proteína EhMyb10r	70
9.10	Identificación de las proteínas que interactúan con EhMyb10r por espectrometría de masas (IMS-DIA-MS)	70
9.10.1	Análisis de GO de las proteínas identificadas por espectrometría de masas (IMS-DIA-MS)	77
9.10.2	Interactoma <i>in silico</i> de las proteínas que interaccionan con EhMyb10 identificadas por espectrometría de masas (IMS-DIA-MS)	81
9.10.3	Detección de la proteína fibrilarina en el eluido del ensayo de Pull-Down	81
10.	DISCUSIÓN	84
11.	CONCLUSIONES	92
12.	REFERENCIAS	94
13.	ANEXOS	104

ÍNDICE DE TABLAS

Tabla 1. Clasificación de los FT de acuerdo con su modo de acción	17
Tabla 2. Clasificación taxonómica de <i>E. histolytica</i>	23
Tabla 3. Características del genoma de <i>E. histolytica</i>	31
Tabla 4. Condiciones de restricción de los plásmidos pGEX-6P-1 y pGEX-EhMyb10	45
Tabla 5. Condiciones de amplificación del gen <i>ehmyb10</i>	46
Tabla 6. Características del gen y la proteína EhMyb10	51
Tabla 7. Aminoácidos susceptibles a fosforilación en EhMyb10	52
Tabla 8. Sitios predichos de ubiquitinación en EhMyb10	53
Tabla 9. Sitios predichos de SUMOilación en EhMyb10	54
Tabla 10. Anotación y función molecular de las proteínas que se predice que interactúan con EhMyb10.	58
Tabla 11. Listado de proteínas que interaccionan con EhMyb10r identificadas por espectrometría de masas (IMS-DIA-MS)	72

ÍNDICE DE FIGURAS

Figura 1. Proceso de la expresión génica en células eucariotas	14
Figura 2. Niveles de regulación de la expresión génica en organismos eucariotas	15
Figura 3. Esquema de un factor de transcripción prototipo	16
Figura 4. DBD-MYB en organismos eucariotas	20
Figura 5. Diversificación de las funciones de las proteínas Myb	22
Figura 6. Distribución geográfica de casos de amebiasis intestinal y abscesos hepáticos en la República Mexicana	26
Figura 7. Trofozoitos de <i>E. histolytica</i>	27
Figura 8. Quistes de <i>E. histolytica</i>	29
Figura 9. Ciclo de vida de <i>E. histolytica</i>	30
Figura 10. Proteínas EhMyb de <i>E. histolytica</i> .	34
Figura 11. Alineamiento de las secuencia de aminoácidos de los repetidos R2 y R3 de las proteínas EhMybR2R3 y c-Myb	35
Figura 12. Moléculas relacionadas con la virulencia de <i>E. histolytica</i> reveladas a partir de estudios transcriptómicos	36
Figura 13. Estructura tridimensional de las proteínas c-Myb y EhMyb10	37
Figura 14. EhMyb10 se une al elemento de reconocimiento a MYB (ERM)	38
Figura 15. Plásmido pGEX-EhMyb10	46
Figura 16. Residuos potenciales de fosforilación en la secuencia de aminoácidos de EhMyb10	53
Figura 17. Residuos susceptibles a SUMOilación	54
Figura 18. Predicción de regiones intrínsecamente desordenadas en EhMyb10	55
Figura 19. Proteína EhMyb10: estructura tridimensional y modificaciones postraduccionales	56
Figura 20. Interactoma <i>in silico</i> de EhMyb10	57
Figura 21. Análisis de GO por función molecular del interactoma de EhMyb10	60

Figura 22. Análisis de GO por procesos biológicos del interactoma de EhMyb10	61
Figura 23. Cambios en la expresión de los genes que codifican para las proteínas que interactúan con EhMyb10 en un modelo murino	63
Figura 24. Cambios en la expresión de los genes que codifican para las proteínas que interactúan con EhMyb10 en diferentes condiciones de cultivo	64
Figura 25. Verificación del plásmido pGEX-EhMyb10	65
Figura 26. Expresión de la proteína EhMyb10r	66
Figura 27. Análisis de solubilidad de EhMyb10r obtenida de lisados bacterianos	67
Figura 28. Purificación de EhMyb10r y GST	68
Figura 29. Fraccionamiento subcelular de proteínas de <i>E. histolytica</i>	69
Figura 30. Ensayo de Pull-Down	71
Figura 31. Diagrama de Venn de proteínas identificadas mediante espectrometría de masas (IMS-DIA-MS) de EhMyb10 y GST	72
Figura 32. Análisis de GO por función molecular de las proteínas identificadas que interaccionan con EhMyb10r	78
Figura 33. Análisis de GO por procesos biológicos de las proteínas identificadas que interaccionan con EhMyb10r	79
Figura 34. Análisis de GO por compartimento celular de las proteínas identificadas que interaccionan con EhMyb10r	80
Figura 35. Interactoma de las proteínas identificadas que interaccionan con EhMyb10r	82
Figura 36. Detección de la proteína fibrilarina en el eluido del ensayo de Pull-Down	83
Figura 37. Modelo molecular de las interacciones de EhMyb10	91

Abreviaturas

ATP	Adenosín trifosfato
BSA	Albúmina de suero bovino, del inglés bovine serum albumin
CAP	Nucleótido modificado de guanina: 7-metilguanosa
c-Myb	Protooncogen MYB, del inglés Myb proto-oncogene protein
CTD	Dominio carboxilo terminal, del inglés Carboxyl-terminal domain
°C	Grados Celsius
CRD	Dominio regulador C-terminal, del inglés C-terminal regulatory domain
CDC5L	Proteína del ciclo de división celular 5-like, del inglés Cell Division Cycle 5-Like Protein
CaBP	Proteínas de unión a calcio, del inglés calcium-binding proteins
DNA	Ácido desoxirribonucleico, del inglés Deoxyribonucleic acid
dsDNA	DNA bicatenario o de doble cadena, del inglés double strand DNA
DBD	Dominio de unión al DNA, del inglés DNA binding domain
DBD-MYB	Dominio de unión al DNA MYB
ERM	Elemento de reconocimiento a Myb
EhMybr	EhMyb10 recombinante
EhTRF-like	Proteína de <i>E. histolytica</i> similar al TRF humano
EDTA	Ácido etilendiamino tetra acético, del inglés Ethylenediaminetetraacetic acid
EGTA	Ácido etilenglicol tetraacético, del inglés ethylene glycol tetraacetic acid
FS	Factores de <i>splicing</i>
FT	Factor de transcripción
GO	Ontología de genes, del inglés Gene Ontology
GST	Glutación S-transferasa
HRP	Peroxidasa de rábano, del inglés horseradish peroxidase

HAT	Histona acetiltransferasa
HLH	Proteínas hélice-giro-hélice, del inglés Helix-Loop-Helix proteins
HMT	Histona metiltransferasa
HK	Cinasas de histonas, del inglés histone kinases
HDAC	Desacetilasa de histonas, del inglés Histone Deacetylase
IDPs	Regiones proteicas intrínsecamente desordenadas, del inglés Intrinsically Disordered Protein Regions
kDa	Kilodalton
LB	Luria Bertani
MLR3	Repetición similar al repetido R3 Myb
Mb	Megabases
mL	Militros
mRNA	RNA mensajero
MYB	Factor de transcripción relacionado con v-Myb derivado del virus de la Mieloblastosis
mM	Milimolar
M	Molar
μL	Microlitros
μg	Microgramos
μM	Micromolar
nm	Nanómetros
OD	Densidad óptica
OMS	Organización Mundial de la Salud
PRP19	Factor de procesamiento pre-mRNA 19, del inglés pre-mRNA processing factor
PCR	Reacción en cadena de la polimerasa, del inglés Polymerase Chain Reaction

PTMs	Modificaciones postraduccionales, del inglés post-translational modifications
PBS	Solución salina amortiguadora de fosfatos, del inglés phosphate buffered saline
pb	Pares de bases
pH	Potencial de hidrógeno
RNA	Ácido ribonucleico, del inglés Ribonucleic acid
RNA pol	RNA polimerasa
rpm	Revoluciones por minuto
Rpb1	RNA polimerasa II subunidad b1
SUMO	Pequeño modificador similar a la ubiquitina, del inglés Small Ubiquitin-like Modifier
SMH	Proteína de la familia Histona Myb Única, del inglés <i>Single- Myb- Histone</i>
SDS-PAGE	Geles desnaturalizantes de poliacrilamida, del inglés sodium dodecyl sulfate polyacrylamide gel electrophoresis
snRNPs	Ribonucleoproteínas nucleares pequeñas, del inglés Small nuclear ribonucleoproteins
TMK	Cinasas transmembranales, del inglés transmembrane kinases
TRF	Factor de unión a repeticiones teloméricas, del inglés telomeric repeat binding factors
tAD	Dominio de transactivación central, del inglés transactivation domain
WB	Western Blot

1. RESUMEN

E. histolytica es un protozooario entérico que posee 8,201 genes y casi el 30% contiene al menos un intrón, de modo que utiliza la actividad de los factores de transcripción y de *splicing* para regular su expresión génica. Los factores de transcripción (FT) Myb son proteínas recientemente implicadas en la regulación de ambos procesos, así como en el mantenimiento del genoma. En *E. histolytica* las proteínas EhMyb son los FT más abundantes, por lo que podrían regular algunos factores de virulencia necesarios para la invasión del huésped. La proteína EhMyb10 es un FT que pertenece a la familia Myb y que presenta dos repetidos: R2 y R3. Para determinar las funciones en las que participa EhMyb10 en *E. histolytica*, en este trabajo, realizamos un interactoma *in silico* y un ensayo de Pull-Down empleando la proteína recombinante para determinar las proteínas que se unen a este FT. Por otro lado, determinamos *in silico* que EhMyb10 es susceptible a modificaciones postraduccionales que podrían afectar su función, entre ellas: ubiquitinaciones, fosforilaciones y SUMOilaciones. El interactoma *in silico* mostró que EhMyb10 se une a proteínas relacionadas con la transcripción y el *splicing*. Una de las proteínas a la que se une a EhMyb10 es EhCDC5-like. Esta proteína también pertenece a la familia Myb y tiene dos repetidos R2R3; en *Schizosaccharomyces pombe* su función se ha relacionado con la activación del *spliceosoma* más que con la transcripción. También observamos que EhMyb10 interactúa con proteínas relacionadas con ubiquitinación y SUMOilación, así como con proteínas involucradas en la remodelación de la cromatina. Por lo tanto, este FT podría tener un papel importante a nivel cotranscripcional. En la validación experimental de este análisis *in silico* a través de los ensayos de Pull-Down se identificaron por Espectrometría de Masas algunas proteínas involucradas en los procesos de transcripción, *splicing*, remodelación de la cromatina, señalización celular, en la ubiquitinación de proteínas, así como en la biogénesis de ribosomas, traducción, segregación cromosómica, regulación biológica, metabolismo de proteínas y transporte citosólico. Estos resultados indican no sólo la participación de EhMyb10 en eventos cotranscripcionales sino también refleja su participación en una variedad de funciones y procesos celulares claves para este organismo.

2. INTRODUCCIÓN

2.1 Transcripción

La transcripción es un proceso fundamental que ocurre en el núcleo celular. Este proceso involucra a enzimas denominadas RNA polimerasas que sintetizan una cadena de RNA complementaria a una cadena molde de DNA, denominada RNA mensajero (mRNA), con la finalidad de generar una proteína que desempeñe una función en la célula. El producto de RNA no permanece apareado a la cadena molde de DNA; la enzima RNA pol II desplaza la cadena creciente solo unos nucleótidos detrás del que está siendo añadido lo que permite que múltiples RNA polimerasas lo puedan transcribir al mismo tiempo, obteniendo así, un gran número de transcritos de un gen en corto tiempo. La transcripción copia selectivamente solo ciertas partes del genoma y puede producir desde una hasta cientos y miles de copias de una sección en un proceso altamente regulado (Watson *et al.*, 2016).

La transcripción ocurre en una serie de pasos:

1) Reconocimiento de la región promotora. La RNA polimerasa junto con factores generales de transcripción se unen a una secuencia de DNA conocida como promotor. La enzima RNA pol II está conformada por una subunidad mayor (Rpb1) que contiene en su región carboxilo terminal un dominio denominado CTD que consiste en repeticiones de siete residuos aminoacídicos o heptapéptidos (YSPTSPS) que son susceptibles a fosforilación para la regulación de la transcripción a través de la unión con proteínas reguladoras del proceso de biosíntesis de mRNA ([co]activadores, [co]represores) y proteínas que modifican la cromatina, entre otras (Cabrejos *et al.*, 2001). La unión entre ellos es regulada fundamentalmente por el grado de fosforilación de algunos residuos presentes en cada heptapéptido en el CTD. Una vez formado el complejo promotor-polimerasa ocurren cambios estructurales necesarios para iniciar el proceso de transcripción.

2) Elongación. La transcripción, como la replicación, va en sentido de 5'-3', por lo que, los ribonucleótidos son añadidos en el extremo 3' de la cadena creciente. La selección del promotor determina qué fragmento de DNA será transcrito y es el paso donde más comúnmente se impone la regulación. Una vez que la RNA pol II ha sintetizado un fragmento corto de RNA, entra en la fase de elongación. Durante esta fase la enzima realiza la catálisis de la síntesis de la cadena de RNA, desenrolla la doble cadena de DNA por delante y la enrolla por detrás, disocia

a la cadena creciente de RNA molde conforme se va moviendo y realiza la labor de *proofreading* o corrección.

3) Terminación. Una vez que la RNA polimerasa ha transcrito la longitud total del gen (o genes), se detiene y libera el producto de RNA conocido como pre-RNA mensajero (pre-mRNA) y se disocia del DNA, este paso es llamado terminación (Watson *et al.*, 2016).

4) Procesamiento. Posteriormente se lleva a cabo la maduración del mRNA a través de *splicing*, capping (CAP) y poliadenilación, de este modo, el pre-mRNA se procesa a un mRNA maduro y se exporta fuera del núcleo (Watson *et al.*, 2016).

Por otro lado, la transcripción *in vivo* requiere de proteínas reguladoras, un complejo mediador y enzimas modificadoras de los nucleosomas ya que el DNA en los organismos eucariontes está empaquetado formando la cromatina. Esta situación limita el acceso de la RNA pol II y de sus factores asociados al promotor. Es por ello, que se necesitan proteínas reguladoras de la transcripción o FT, que activan o inhiban este proceso.

La regulación génica es un proceso de encendido y apagado de genes que asegura que los genes apropiados se expresen en los momentos adecuados para ayudar a un organismo a responder a su entorno. Este proceso es mediado a través de interacciones entre: los FT unidos al DNA, co-reguladores, los factores remodeladores de la cromatina y proteínas relacionadas con la maquinaria de transcripción (Watson *et al.*, 2016). Una vez sintetizado, el mRNA inmaduro debe ser protegido a través del capping para su exportación al citoplasma (Orphanides *et al.*, 2002). Durante el capping se adiciona una base de guanina metilada en el extremo 5' del RNA a través de un enlace inusual 5' - 5', involucrando tres fosfatos (Watson *et al.*, 2016). Después del capping se disocia la maquinaria y se recluta a la maquinaria necesaria para el *splicing* del pre-mRNA. El *splicing* o empalme es un proceso postranscripcional que conduce a la eliminación de intrones y a la unión de exones en un pre-mRNA para formar un mRNA maduro (House *et al.*, 2008), y es realizado por un gran complejo, que comprende alrededor de 150 proteínas y 5 RNAs (U1, U2, U4, U5 y U6) llamados colectivamente RNAs nucleares pequeños (snRNAs). Estos snRNAs se encuentran en complejos con proteínas (snRNPs). El *spliceosoma* es un gran complejo compuesto de estas snRNPs. Las snRNPs reconocen el extremo del sitio de empalme 5' y del sitio de la ramificación; juntan dichos sitios y catalizan las reacciones de corte y sellado del RNA (Watson *et al.*, 2016).

El evento final del procesamiento es la poliadenilación del extremo 3' del mRNA, lo que está íntimamente ligado a la terminación. En este evento, el CTD de la RNA pol II recluta a las enzimas necesarias para el *splicing* y la poliadenilación. Una vez que la RNA pol II alcanza el final del gen, encuentra una secuencia específica que después de ser transcrita, detona el proceso de poliadenilación a través de cuatro eventos: corte del mensajero, adición de residuos de A al extremo 3', degradación del RNA remanente asociado a la RNA pol II por una ribonucleasa 5' a 3' y terminación de la transcripción. Una vez que el mRNA ha sido procesado es transportado al citoplasma a través de poros en la membrana nuclear. La exportación de mRNA está mediada por factores que se unen al mRNA en el núcleo para su traducción en grandes complejos ribonucleoproteicos llamados ribosomas (Watson *et al.*, 2016).

El nivel de expresión de la mayoría de los genes está regulado por FT que se unen a secuencias reguladoras del DNA situadas río arriba del sitio promotor y a su vez estas proteínas se encuentran reguladas por una amplia gama de vías de señalización. Los FT son comúnmente activados por ligandos, por ejemplo, los receptores de hormonas nucleares. Una vez activados, los FT deben reclutar a proteínas co-reguladoras para la descompactación local de la cromatina y permitir el acceso de la RNA pol II y la maquinaria de transcripción general a la región promotora (Orphanides *et al.*, 2002). La **Figura 1** resume el proceso de expresión génica, desde la activación de los reguladores transcripcionales hasta la síntesis de una proteína funcional (Orphanides *et al.*, 2002).

2.2 Regulación de la expresión génica

Para el control de la transcripción también hay co-reguladores que modulan la estructura de la cromatina y se pueden dividir en dos clases: complejos de remodelación de nucleosomas y co-reguladores. Los primeros son proteínas dependientes de ATP que facilitan el acceso al sitio blanco en el DNA mediante el re-posicionamiento de nucleosomas en el promotor o la inducción de cambios conformacionales en los nucleosomas (Orphanides *et al.*, 2002). Por otro lado, los co-reguladores catalizan la modificación postraducciona de las histonas. Asimismo, la modificación de histonas implicada en la regulación transcripcional puede llevarse a cabo por acetiltransferasas de histonas (HATs), deacetilasas de histonas (HDACs), metiltransferasas de histonas (HMTs) y cinasas de histonas (HKs).

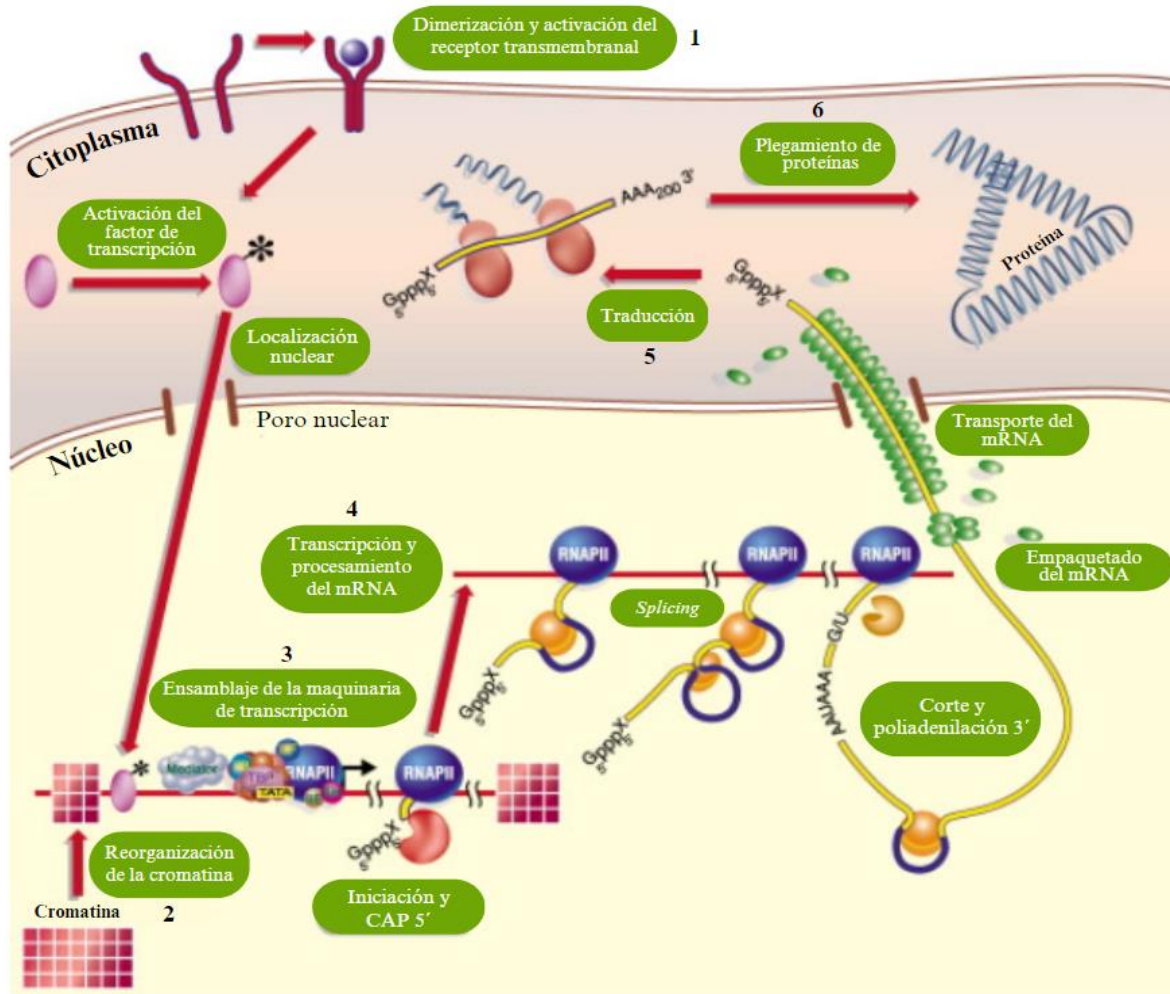


Figura 1. Proceso de la expresión génica en células eucariotas. Se muestran los diferentes eventos para que se lleve a cabo la expresión génica. **1)** Una señal del exterior de la célula inicia una cascada de transducción de señales que tiene efecto en la translocación de un factor de transcripción al núcleo de la célula; **2)** acceso al genoma, lo cual involucra la reorganización de la cromatina; **3)** ensamblaje de la maquinaria de transcripción; **4)** síntesis y procesamiento del RNA: transcripción, adición de CAP, *splicing*, poliadenilación, edición del RNA; **5)** síntesis de proteínas: traducción y **6)** procesamiento de proteínas. Imagen adaptada de Orphanides *et al.*, 2002.

El reclutamiento de HATs y HMTs resulta en la acetilación y metilación, respectivamente, de residuos ubicados en las colas N-terminales de las histonas y está asociado con la activación de la expresión de muchas clases de genes. Por el contrario, el reclutamiento de HDACs conduce a la desacetilación de las colas de histonas lo que generalmente reprime la expresión génica.

La expresión génica está sujeta a una compleja regulación que va desde la síntesis del RNA hasta la degradación de las proteínas. Las células tienen la capacidad de controlar la síntesis de proteínas mediante: **1)** la regulación del momento y la frecuencia de transcripción de genes específicos (control transcripcional), **2)** el control del procesamiento y maduración de los transcritos primarios de RNA (control del procesamiento del RNA), **3)** la selección de los mRNA maduros que serán exportados al citoplasma (control del transporte del RNA), **4)** la selección de los mRNA citoplasmáticos que serán traducidos por los ribosomas (control traduccional), **5)** la desestabilización selectiva de ciertas moléculas de mRNA citoplasmáticos (control de la degradación de los mRNA), y **6)** la activación, inactivación o localización selectiva de las proteínas ya sintetizadas (control de la actividad proteica) (Alberts *et al.* 2010) (**Fig. 2**).

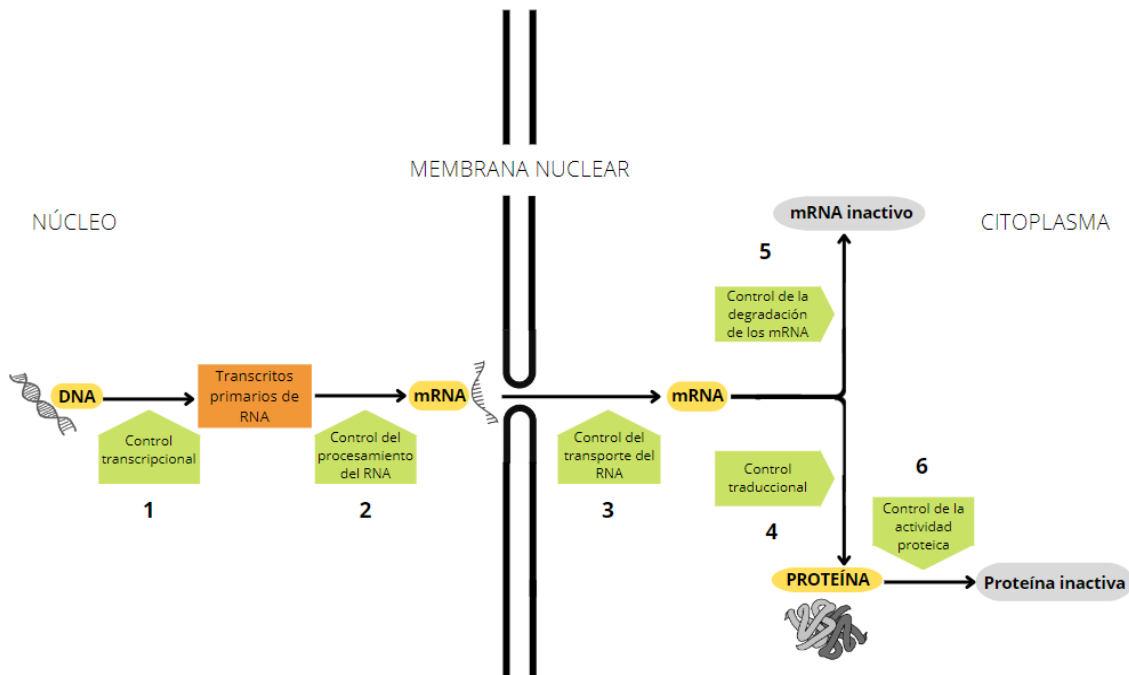


Figura 2. Niveles de regulación de la expresión génica en organismos eucariotas. La expresión génica puede regularse en diversas etapas del proceso que va desde el DNA al RNA y las proteínas mediante: **1)** control transcripcional, **2)** control del procesamiento del RNA, **3)** control del transporte del RNA, **4)** control traduccional, **5)** control de la degradación de los mRNA, y **6)** control de la actividad proteica. Imagen adaptada de Alberts *et al.*, 2010.

2.3 Factores de transcripción (FT)

El inicio de la transcripción es el punto de control más importante en la cascada de la expresión génica y está regulada por dos clases de proteínas que interactúan con el DNA: **1) factores generales de transcripción**, como la RNA pol II, proteínas accesorias y enzimas que intervienen en las modificaciones de las histonas y la metilación del DNA; y **2) factores específicos de los genes**, también llamados "factores de transcripción" (FT) (Suter, 2020). Se denomina FT a cualquier proteína que esté involucrada en la transcripción y que tenga la capacidad de alterar los niveles de expresión génica. Estas proteínas cumplen dos criterios fundamentales: **1)** son capaces de unirse al DNA de manera específica según su secuencia, y **2)** regulan la transcripción de los genes (Lambert *et al.*, 2018) (**Fig. 3**). Esto se consigue mediante dominios específicos que impactan en la transcripción en diferentes niveles, como la accesibilidad del DNA o el reclutamiento de la maquinaria de transcripción (Suter, 2020).

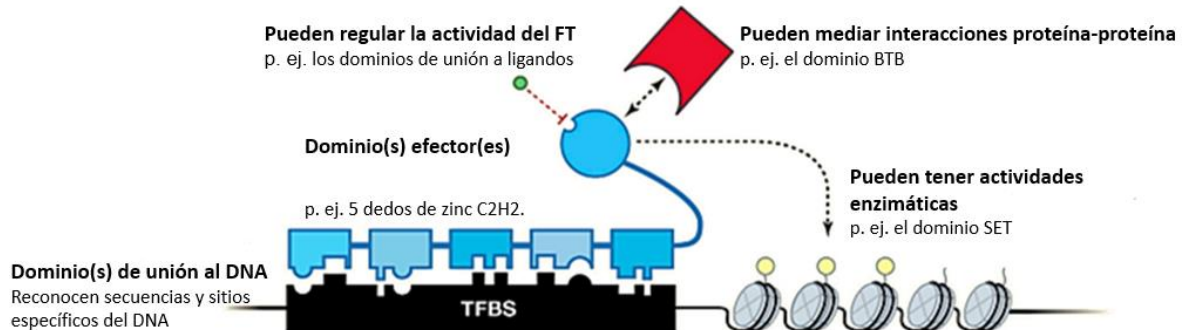


Figura 3. Esquema de un factor de transcripción prototipo. El esquema muestra los dominios efectores que posee un FT canónico. En azul se muestra al dominio de unión al DNA y el dominio efector. Los dominios efectores regulan la actividad del FT al interactuar con sus ligandos, facilitar interacciones proteína-proteína o tener actividad enzimática, como es el caso de los dominios SET en proteínas que modulan la organización de la cromatina y, por ende, controlan de forma transcripcional los estados activos o reprimidos. Imagen adaptada de Lambert *et al.*, 2018.

Los FT reconocen secuencias específicas en el DNA mediante la formación de interacciones no covalentes entre residuos de aminoácidos y los grupos funcionales de las bases nitrogenadas del DNA, lo cual contribuye a la especificidad de unión FT-DNA. Los FT también se pueden clasificar de acuerdo a su modo de acción como: pioneros, colonizadores o migrantes (Slattery *et al.*, 2014) (**Tabla 1**).

Los FT pioneros o *pioneers* son los primeros FT en unirse al DNA, ya que son capaces de unirse a regiones que normalmente estarían inaccesibles debido a su estructura compacta (heterocromatina), y su función es la de establecer un sitio de unión para otros FT. Al unirse, modifican la estructura de la cromatina y permiten el acceso a otros FT. Los pioneros son importantes en la diferenciación celular y en el desarrollo de órganos y tejidos (Slattery *et al.*, 2014).

Por otro lado, los FT colonizadores o *settlers*, se unen al DNA después de los pioneros, y su función principal es la de estabilizar y mantener la estructura abierta de la cromatina para permitir la unión de otros FT. Los colonizadores pueden unirse a nuevos sitios en el DNA siendo importantes en la regulación de genes específicos durante el desarrollo embrionario y la diferenciación (Slattery *et al.*, 2014).

Por último, los FT migrantes o *migrants*, se unen al DNA después de los pioneros y colonizadores, y son importantes en la regulación de la expresión gen-específica en respuesta a señales ambientales, como cambios en la concentración de hormonas o la exposición a estímulos externos (Slattery *et al.*, 2014). Por lo tanto, los FT pioneros, colonizadores y migrantes juegan papeles distintos pero interrelacionados en la regulación de la expresión génica.

Tabla 1. Clasificación de los FT de acuerdo con su modo de acción.

Factor de transcripción	Características	Ejemplo
Pioneros	Se pueden unir a sitios inaccesibles al DNA	c-Myb, FoxA, GATA
Colonizadores	Se unen al mismo tipo de secuencias, pero deben estar accesibles.	HLH, PAS, TBOX, bZIP, MAX
Migrantes	Requieren de la presencia de cofactores para la unión a la secuencia blanco	TEAD4, ALX4, POU2F1

Elaborada a partir de Ehsani *et al.*, 2016.

2.4 Dominios de unión a DNA (DBD)

Un dominio de unión al DNA es una región de una proteína que corresponde a una secuencia de aminoácidos que se une a una región específica del DNA. Los dominios de unión al DNA adoptan diferentes estructuras, como por ejemplo: los dedos de zinc, la hélice-giro-hélice (HTH) y la cremallera de leucina (entre muchos otros) que facilitan la unión al ácido nucleico. También hay ejemplos de activadores de la transcripción menos comunes, como son los efectores y la interacción del DBD con el DNA puede establecerse a través de hélices alfa, hojas beta o regiones desordenadas. Los DBD se denominan según sus características estructurales y la mayoría de los organismos contienen varios FT que comparten el mismo tipo de dominio de unión al DNA. En consecuencia, los FT se clasifican en familias que suelen recibir el nombre del respectivo DBD (Charoensawan *et al.*, 2010).

Los FT de la misma familia tienen especificidades de unión al DNA más similares que los que pertenecen a familias diferentes. Sin embargo, a menudo se observan variaciones en la especificidad de unión al DNA dentro de la misma familia y éstas se deben principalmente a cambios en residuos específicos del dominio de unión al DNA (Berger *et al.*, 2008).

2.5 Factores de transcripción Myb

2.5.1 c-Myb

Las proteínas Myb (acrónimo virus de la mieloblastosis aviar) son FT que se identificaron por primera vez en el virus de la mieloblastosis aviar y se han encontrado en todas las células eucariotas (Millard *et al.*, 2019). El FT c-Myb de *Homo sapiens* es el factor Myb canónico, controla la expresión génica de varios tipos celulares y participa en procesos que van desde el control del ciclo celular hasta la diferenciación. En las células hematopoyéticas, c-Myb es codificado por un proto-oncogen que se encuentra altamente expresado y tiene un papel importante en la proliferación y diferenciación de progenitores mieloides y linfoides. La proteína c-Myb consta de tres dominios: el dominio de unión al DNA (DBD) en la región N-terminal, un dominio de transactivación central (tAD) y un dominio regulador C-terminal (CRD).

El factor c-Myb es un FT pionero debido a su capacidad de reconocer su secuencia consenso en la cromatina cerrada e influir en la conformación del DNA, ya que se ha demostrado que puede reclutar a acetiltransferasas de histonas (HATs) como las proteínas p300 y CBP e interactuar con su dominio KIX a través del dominio tAD. CBP y p300 están involucradas en la acetilación de residuos conservados de lisina de las histonas; lo cual neutraliza las cargas positivas de las histonas y ocasiona un cambio en la interacción con el DNA, permitiendo una mayor accesibilidad a los FT (Lemma *et al.*, 2021).

2.5.2 Estructura del DBD-MYB

En particular, los FT pertenecientes a la familia hélice-giro-hélice (HTH) donde encontramos a los factores Myb, se encuentran implicados en procesos fundamentales, como la regulación del metabolismo, señalización celular en respuesta a diferentes estímulos, la diferenciación celular y el control de la proliferación (Arratia *et al.*, 2013).

El DBD-MYB, consiste en una secuencia conservada de 50 a 53 aminoácidos, que puede estar de 1 a 4 veces repetida en la misma proteína. Cada uno de estos repetidos contienen tres residuos de triptófano que están separados por 18-19 aminoácidos, los cuales le confieren estabilidad a la unión del DNA y un ambiente hidrofóbico el cual estabiliza su plegamiento. Estos repetidos forman tres α -hélices (H1, H2, H3) en donde la segunda y tercera α -hélice están involucradas en la unión al DNA (**Fig. 4A**).

La clasificación de las proteínas Myb se basa en la similitud y el número de repetidos: 1R-MYB, 2R-MYB (R2R3-Myb), 3R-MYB y 4R-MYB (Arratia *et al.*, 2013) (**Fig. 4A**). El DBD-MYB reconoce una secuencia consenso ($^C/_T$ A AACTG), denominada elemento de reconocimiento a MYB o ERM. Por otro lado, se ha observado que el DBD-MYB se encuentra altamente conservado en diferentes organismos eucariotas como plantas, animales, protozoarios y hongos, típicamente cerca de la región N-terminal (Arratia *et al.*, 2013) (**Fig. 4B**).

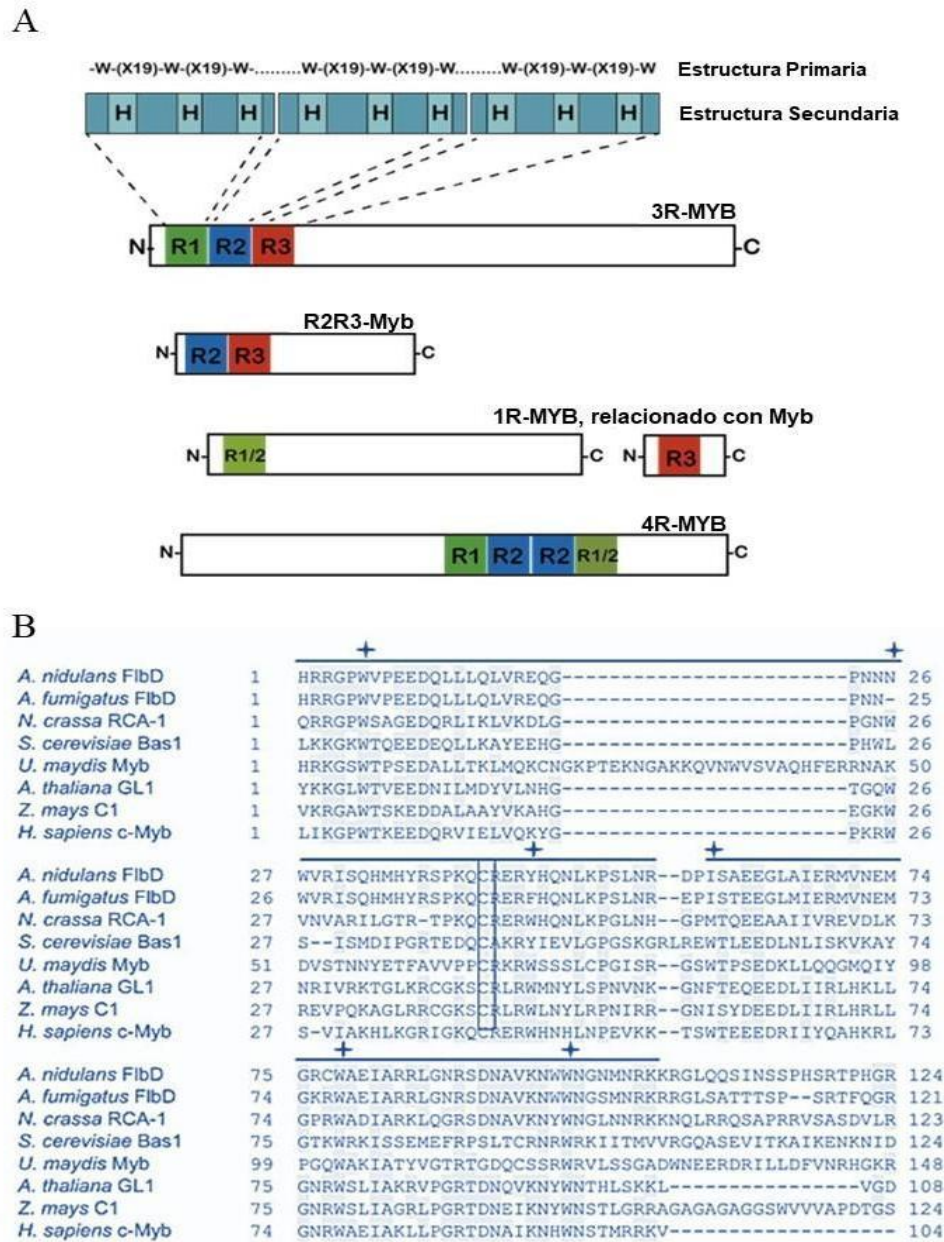


Figura 4. DBD-MYB en organismos eucariotas. A) Esquema de la estructura del DBD-MYB de las proteínas clasificadas según el número de secuencias repetidas adyacentes. Las secuencias repetidas son señaladas por R1, R2, R3. Se indica la estructura secundaria (hélice) mediante las barras H1, H2, H3. B) Los triptófanos conservados en los repetidos R2R3 en distintos grupos filogenéticos. Se indica la longitud de la secuencia repetida R2 y R3 con líneas. Los asteriscos representan los residuos de triptófano espaciados característicos de los dominios Myb. Se indica en un rectángulo vertical los residuos de cisteína conservados dentro del dominio Myb. Imagen adaptada de Arratia *et al.*, 2013.

2.5.3 Otras funciones de las proteínas Myb

Además de participar como FT, se ha demostrado que las proteínas Myb participan en eventos cotranscripcionales como el proceso de *splicing*. La familia de proteínas CDC5 representa una subfamilia conservada de proteínas relacionadas con Myb necesarias para el empalme de pre-mRNA. Tienen un DBD-MYB similar al c-Myb humano que consta de dos repetidos Myb (R1 y R2) seguidos de un repetido similar a Myb (MLR3, del inglés Myb Like Repeat). Asimismo, estas proteínas tienen un papel importante en el control del ciclo celular a través de la regulación transcripcional. Por otro lado, se ha identificado que CDC5 se colocaliza con factores de empalme de pre-mRNA en los núcleos de células de mamífero, co-immunoprecipita con componentes centrales de la maquinaria de *splicing* de extractos nucleares e interactúa con el *spliceosoma* *in vitro* (Burns *et al.*, 1999). En *E. histolytica* se identificó una proteína con homología a la familia CDC5 a la cual se le denominó EhCDC5-like, la cual podría desempeñar una función esencial en el control del ciclo celular y en el *splicing* (Meneses *et al.*, 2010).

Las proteínas Myb también actúan como factores de unión a repetidos teloméricos (TRF). Los TRF contienen un dominio similar a Myb en su C terminal que les permite unirse al DNA telomérico de doble cadena (Rendón *et al.*, 2018). En humanos, estas proteínas nucleares se denominan TRF1 y TRF2, las cuales son clave para la regulación de la longitud de los telómeros, para proteger los extremos de los cromosomas y detectar daños en el DNA, entre otras funciones (Chong *et al.*, 1995; Broccoli *et al.*, 1997; de Lange, 2005; Palm y de Lange, 2008). Asimismo, forman parte del complejo *Shelterin* el cual tiene la función de proteger, regular, prevenir la degradación y fusión de los telómeros, manteniendo así la integridad del genoma (Palm y de Lange, 2008). En *E. histolytica*, se han identificado tres proteínas homólogas a las proteínas TRF1 y TRF2 humanas, que se unen al DNA y se cree que están involucradas en el mantenimiento de los telómeros formando parte del complejo de *Shelterin* (Rendón *et al.*, 2018). Estas proteínas, llamadas EhTRF-like I, II y III, están presentes tanto en el núcleo como en el citoplasma de la célula y se ubican en la periferia nuclear donde interactúan con la proteína laminina B1 y H4K20 trimetilado. Además, se ha encontrado que EhTRF-like III se localiza en regiones de cromatina más condensada, lo que sugiere una posible función en la organización de la estructura de la cromatina.

Además, EhTRF-like III puede formar complejos con secuencias relacionadas con telómeros, lo que sugiere una posible función en la protección de los extremos cromosómicos (Rendón *et al.*, 2018), sin embargo, se necesitan más estudios para entender completamente su papel en el mantenimiento de la longitud de los cromosomas de este parásito.

Por otro lado, también se han descrito a otras proteínas con un motivo telebox en su dominio Myb con funciones análogas a las del complejo de “*Shelterin*” humano en los telómeros de la planta *Arabidopsis thaliana*, las cuales se denominan proteínas SMH (*Single-Myb-Histona*). Las proteínas SMH se unen preferencialmente al dsDNA telomérico y su DBD-MYB se localiza en el N-terminal. También poseen en la parte central de la proteína un dominio H1/5 similar a una histona y un dominio coiled-coil en el extremo carboxilo. En *A. thaliana* la familia de proteínas SMH codifica para cinco proteínas denominadas AtTRB1-5, de las cuales solo AtTRB2 y AtTRB3 se han caracterizado previamente y son capaces de unirse preferentemente al dsDNA telomérico de la planta. Por otro lado, también se ha determinado la interacción preferencial de AtTRB1 con el DNA telomérico de doble cadena la cual está mediada por el dominio Myb, mientras que el dominio H1/5 está involucrado en la interacción DNA-proteína no específica y en la multimerización de AtTRB1 (Mozgová *et al.*, 2008) (**Fig. 5**).

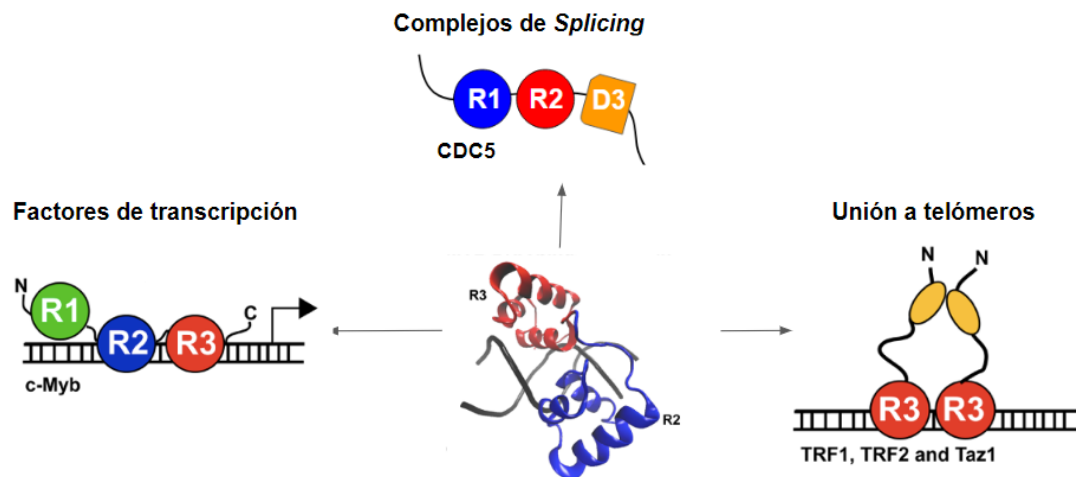


Figura 5. Diversificación de las funciones de las proteínas Myb. El DBD-MYB puede estar presente en proteínas que realizan diversas funciones como factores de transcripción (c-Myb), proteínas involucradas en *splicing* (CDC5) y proteínas de unión a telómeros (TRF y Taz). Imagen elaborada por la Dra. Elisa Irene Azuara Liceaga de la Licenciatura en Ciencias Genómicas en la Universidad Autónoma de la Ciudad de México (2023).

2.6 *Entamoeba histolytica*

2.6.1 Taxonomía y antecedentes históricos

Este organismo pertenece al reino de los protistas y se encuentra en el filo *Sarcomastigophora*, que engloba a organismos unicelulares móviles que poseen estructuras flageladas o ameboides. Dentro del subfilo *Sarcodina*, que incluye a las amebas, *E. histolytica* se clasifica en la clase *Lobosea*. A nivel de orden, se encuentra en *Amoebida*, y en la familia *Entamoebidae*. Finalmente, su nombre científico, que incluye el género y la especie, es *Entamoeba histolytica* (Romero *et al.*, 2008) (**Tabla 2**).

Tabla 2. Clasificación taxonómica de *E. histolytica*

Clasificación taxonómica	
Phyllum	<i>Sarcomastigophora</i>
Subphyllum	<i>Sarcodina</i>
Clase	<i>Lobosea</i>
Familia	<i>Entamoebidae</i>
Género y especie	<i>Entamoeba histolytica</i>

Tabla adaptada de Romero *et al.*, 2008.

E. histolytica fue identificada en 1875 por Friedrich Lösch en las heces de un paciente con disentería y en las úlceras del colon durante una autopsia. Aunque no relacionó las lesiones del colon con las amebas, su trabajo fue fundamental para el reconocimiento de la ameba como el agente causante de la disentería amebiana y otros trastornos intestinales en humanos. Los trabajos de Kartulis (1886), Hava (1887) y Consilman (1891) aportaron las evidencias necesarias para demostrar que el parásito era el agente etiológico de la disentería y del absceso hepático (Romero *et al.*, 2008). En 1913, Walker y Sellards llevaron a cabo un estudio en Filipinas en el que demostraron que *E. histolytica* es la causa de la colitis amebiana y que *Entamoeba coli* es un comensal intestinal inofensivo.

El estudio involucró la ingesta de quistes de amebas en tres grupos de “voluntarios” de una prisión en Manila y concluyó que los portadores asintomáticos de amebas pueden transmitir la enfermedad y desarrollarla en cualquier momento.

En 1925, el parasitólogo francés Emile Brumpt propuso la existencia de *E. histolytica* como un complejo de especies. Brumpt denominó *E. dysenteriae* a la especie que causa infección sintomática y *E. dispar* a la que se encuentra en personas asintomáticas. Sin embargo, este planteamiento fue rechazado inicialmente por la comunidad científica internacional en esa época. Durante los años siguientes, se realizaron varias investigaciones sobre el estudio de *E. histolytica*. Esto se debió a la posibilidad de cultivo que se logró gracias a Beck, Dbrhlav, Balamuth y Diamond en 1961. Estos investigadores obtuvieron diferentes tipos de cultivo hasta contar con los medios axénicos para el cultivo de esta ameba. Esto permitió su estudio experimental de múltiples formas. Posteriormente, en 1968, Diamond introdujo el medio monofásico TPS-1 que comenzó a ser ampliamente utilizado, pero que luego fue reemplazado por el medio líquido TYI-S-33. Para el aislamiento de cepas de *E. histolytica* en medios de cultivo mixtos, se utilizó el medio bifásico descrito por Robinson (Pinilla *et al.*, 2008).

Es importante señalar que el desarrollo de pruebas inmunológicas para el diagnóstico de amebiasis fue posible gracias al desarrollo del cultivo axénico de *E. histolytica*. Entre esas pruebas inmunológicas se han descrito técnicas que incluyen hemoaglutinación indirecta, inmunodifusión, fijación de complemento, contrainmunolectroforesis (CIEF), inmunofluorescencia indirecta, ELISA y WB. Por otro lado, en 1968 Bischoff y Reeves pudieron diferenciar las especies patógenas y no patógenas de *Entamoeba* empleando perfiles isoenzimáticos. Sin embargo, con esta técnica no pudieron distinguir entre cepas patógenas y no patógenas de *E. histolytica*. En los comienzos de los años setenta, se empezaron a recopilar datos que apoyaban la hipótesis de Brumpt, la cual afirmaba que existían dos especies diferentes de *E. histolytica*. Fue solo en 1978, luego de varios años de investigación, que Sargeant y Williams diferenciaron por primera vez cepas de *E. histolytica* aisladas de pacientes con manifestaciones clínicas de amebiasis y portadores asintomáticos mediante estudios electroforéticos de isoenzimas. Esto confirmó que *E. histolytica* está compuesta por cepas patógenas y no patógenas (Pinilla *et al.*, 2008).

Posteriormente, en 1993, Diamond y Clark redescubrieron la hipótesis de Brumpt de 1925 a la luz de estudios bioquímicos, inmunológicos y genéticos. Llegaron a la conclusión de que existen dos especies morfológicamente idénticas: una patógena y otra no patógena que corresponden a *E. histolytica* y *E. dispar*, respectivamente. Fue solo en 1997 que la OMS aceptó esta hipótesis por recomendación de un comité de expertos reunidos en Ciudad de México, que reglamentaron que se trata de dos especies morfológicamente idénticas sólo diferenciables mediante patrones isoenzimáticos y por determinación de una adhesina en materia fecal o técnicas moleculares (Pinilla *et al.*, 2008).

El descubrimiento de *E. histolytica* fue el resultado de una investigación exhaustiva y compleja que involucró a varios científicos durante siglos. Este proceso permitió una mejor comprensión del parásito *E. histolytica* y de la patogénesis de la disentería amebiana, lo que resultó fundamental para el desarrollo de tratamientos eficaces contra esta enfermedad.

2.6.2 Epidemiología

El parásito *E. histolytica* es el agente causal del amebiasis, una enfermedad que afecta a millones de personas, lo que la convierte en la causa más frecuente de diarrea y se calcula que provoca más de 55,000 muertes al año (Shirley *et al.*, 2018). Esta enfermedad tiene una mayor prevalencia en México, América Central, América del Sur, Asia y África. A nivel mundial, la amebiasis es la tercera causa de muerte entre las enfermedades parasitarias, especialmente en países en desarrollo debido a las malas condiciones sanitarias y la falta de acceso a agua potable. Sin embargo, su distribución es cosmopolita, ya que en países desarrollados también se presentan casos de amebiasis con frecuencia, principalmente debido a la migración de poblaciones procedentes de zonas altamente endémicas o a viajeros que regresan de esas zonas (Romero *et al.*, 2008). A nivel mundial se reportan anualmente alrededor de 500 millones de personas infectadas (Fundación IO, 2020). Ambos sexos se ven afectados por igual por esta enfermedad, pero la variante hepática se da entre 7 y 10 veces más frecuentemente en el sexo masculino. Los niños son el grupo de mayor riesgo de desarrollar formas fulminantes, lo que conlleva una elevada tasa de mortalidad infantil.

La prevalencia de amebiasis aguda intestinal y hepática en los menores de edad refleja las condiciones socioeconómicas y la falta de educación sanitaria, ya que esta infección suele ser el resultado de factores interrelacionados, incluidos aspectos sociales, económicos, culturales, históricos y políticos (Romero *et al.*, 2008).

Según el Boletín Epidemiológico del Sistema Nacional de Vigilancia Epidemiológica (SINAVE), entre los años 2021 y 2022 se registraron en la República Mexicana 224,634 casos de amebiasis intestinal y 2,717 casos de abscesos hepáticos como consecuencia de amebiasis extraintestinal. Hasta el momento, en lo que va del año 2023, se han sumado 48,756 casos de amebiasis y 228 casos de abscesos hepáticos (SINAVE, 2023). Los datos correspondientes a la semana epidemiológica 22 del año 2023 revelaron que los estados céntricos como la Ciudad de México, Hidalgo, Estado de México y Puebla, así como los estados costeros como Oaxaca, Veracruz, Guerrero, Tabasco, Jalisco, Nayarit, Sinaloa, Yucatán y Chiapas, presentaron un mayor número de casos de amebiasis intestinal. Entre ellos, Chiapas registró la mayor incidencia con un total de 8,654 casos. En cuanto a los abscesos hepáticos, los estados con mayor incidencia durante este mismo período fueron Sinaloa, Chihuahua y Jalisco, siendo Jalisco el estado con la mayor cantidad de casos reportados hasta el momento, con un total de 32 casos en el año 2023 (SINAVE, 2023) (**Fig. 6**).

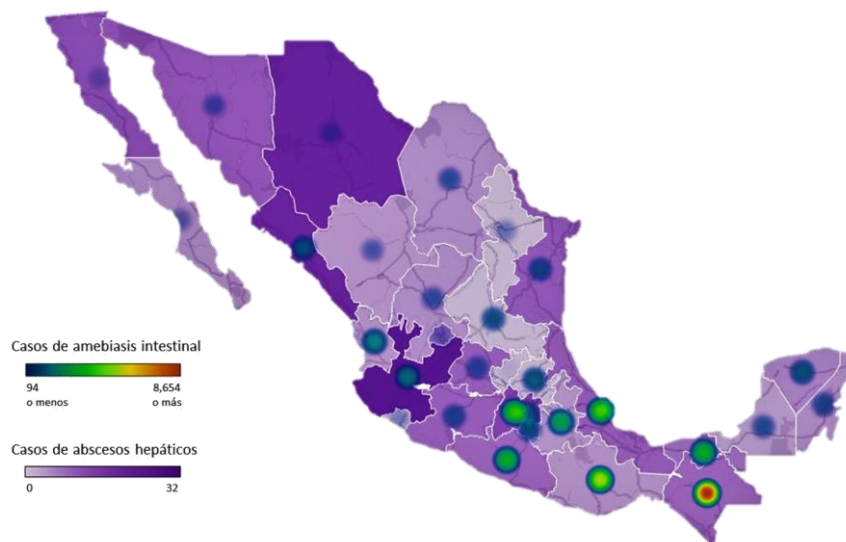


Figura 6. Distribución geográfica de casos de amebiasis intestinal y abscesos hepáticos en la República Mexicana. El mapa muestra la distribución de los casos de amebiasis intestinal y abscesos hepáticos registrados en todo México durante el periodo de enero a junio de 2023. Imagen de elaboración propia a partir de datos obtenidos de la semana epidemiológica 22 del año 2023 del Boletín Epidemiológico del SINAVE.

2.6.3 Ciclo de Vida

Este parásito tiene dos fases de desarrollo con morfologías distintas. El primero es el trofozoito que es la forma infectiva. El tamaño del trofozoito oscila entre 20 y 40 μm y se caracteriza por tener un núcleo esférico de unos 4 a 7 μm de diámetro. El RNA se encuentra en la periferia del núcleo, mientras que el DNA se localiza en su parte central. Dentro del núcleo, se pueden encontrar pequeñas esferas nucleares que miden alrededor de 0.2-0.5 μm de diámetro. Cabe mencionar que la función de estas esferas nucleares aún es desconocida.

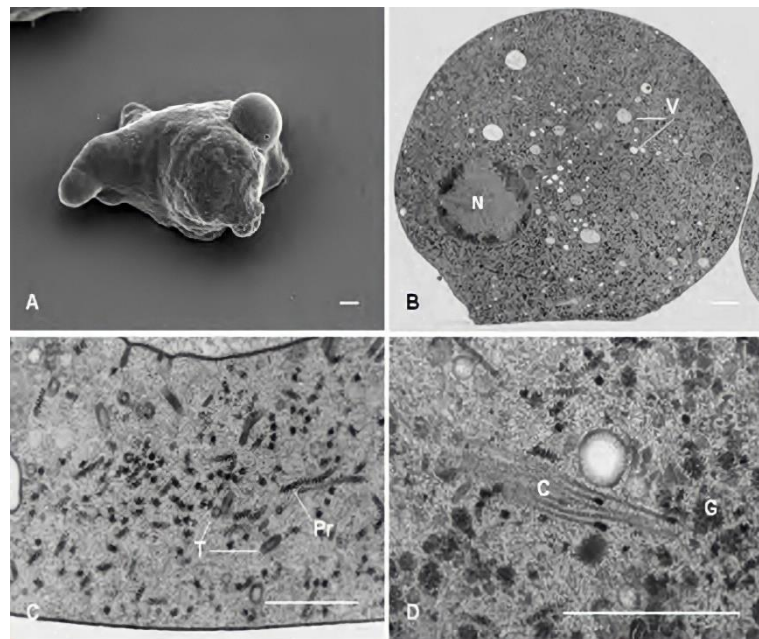


Figura 7. Trofozoitos de *E. histolytica*. A) Trofozoito observado a través del microscopio electrónico de barrido. B) y D): Microscopía electrónica de transmisión. B) Corte fino de un trofozoito. Se aprecia el núcleo (N) y vacuolas (V). C) Polirribosomas (Pr) y perfiles de delgados túbulos citoplasmáticos (T). D) Gránulos de glucógeno (G) y un grupo de cisternas aplanadas (C) similares al Golgi. Imagen obtenida de Chávez *et al.*, 2013.

El citoplasma de este organismo tiene varias vacuolas, las cuales están limitadas por una membrana. Las vacuolas tienen un papel importante en el transporte de moléculas que son internalizadas mediante los procesos de pinocitosis y fagocitosis. En el citoplasma también hay gránulos de glucógeno y agregados de ribosomas, los cuales llevan a cabo la síntesis de proteínas.

No obstante, hasta el momento no se han podido identificar organelos como mitocondrias o aparato de Golgi, aunque se han hallado proteínas asociadas con las funciones de estos organelos. Por otro lado, la membrana plasmática del trofozoito tiene un grosor de 10 nanómetros y está compuesta por lípidos específicos que permiten que la amiba se mueva y se dirija a las células blanco o a otros sitios específicos dentro del hospedero (Chávez *et al.*, 2013) (**Fig. 7**). En la membrana ocurre la formación del casquete membranar, que le permite a la amiba evadir la respuesta inmunitaria del hospedero y desprenderse de moléculas dañinas. La plasticidad de la membrana, junto con la reorganización del citoesqueleto y la fluidez del citoplasma, le permite cambiar su forma constantemente, así como desplazarse. La amiba forma estructuras llamadas pseudópodos, mediante las cuales captura los alimentos líquidos por pinocitosis y sólidos por fagocitosis. Finalmente, en los trofozoitos pueden observarse componentes similares a partículas virales, sin embargo, esta función aún es desconocida (Olivos *et al.*, 2011).

El segundo estadio de vida es el quiste, el cual es la forma infectiva y resistente del parásito, ya que permite que la amiba sobreviva fuera del hospedero, así como resistir a las distintas condiciones ambientales mientras halla a su próximo hospedero. Los quistes tienen una forma esférica y miden aproximadamente de 10 a 12 μm . La pared del quiste tiene entre 120 a 150 nm de espesor y está formada principalmente de quitina, el cual es un polímero de carbohidratos formado por unidades de N-acetilcolina-D glucosamina. En el citoplasma del quiste se encuentran de 1 a 4 núcleos que dependen de la madurez del quiste. El quiste también contiene vacuolas, pero en menor cantidad y agregados de polirribosomas, denominados cuerpos cromatoides (Olivos *et al.*, 2011) (**Fig. 8**).

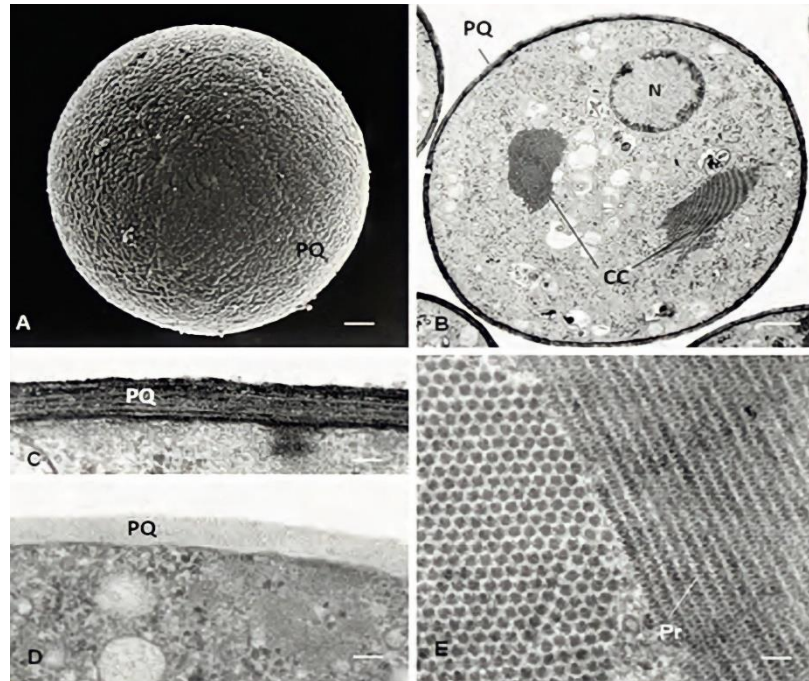


Figura 8. Quistes de *E. histolytica*. **A)** Superficie rugosa de la pared del quiste por microscopía de barrido. **B)** Citoplasma del quiste donde se observan vacuolas, el núcleo (N), dos cuerpos cromatoides y la pared del quiste (PQ). **C)** Cuerpos cromatoides a mayor amplificación, que muestran polirribosomas cortados transversal y longitudinalmente. **D)** y **E)** Pared del quiste (PQ) con rojo de rutenio **D)** y sin rojo de rutenio **E)**. Imagen obtenida de Chávez *et al.*, 2013.

El ciclo de vida de este parásito comienza con la ingesta de alimentos contaminados con quistes, que al ser consumidos reiniciará la infección; una vez en el tracto digestivo los jugos gástricos desintegran la pared del quiste liberando la forma vegetativa del parásito (trofozoito), el cual se va adherir a la mucosa del intestino y comenzará a duplicarse por fisión binaria (**Fig. 9**). El trofozoito viajará hacia el colon, donde comenzarán su proceso de enquistamiento para finalmente ser excretado con las heces y reiniciar su ciclo de vida en otro hospedero (Chávez *et al.*, 2013).

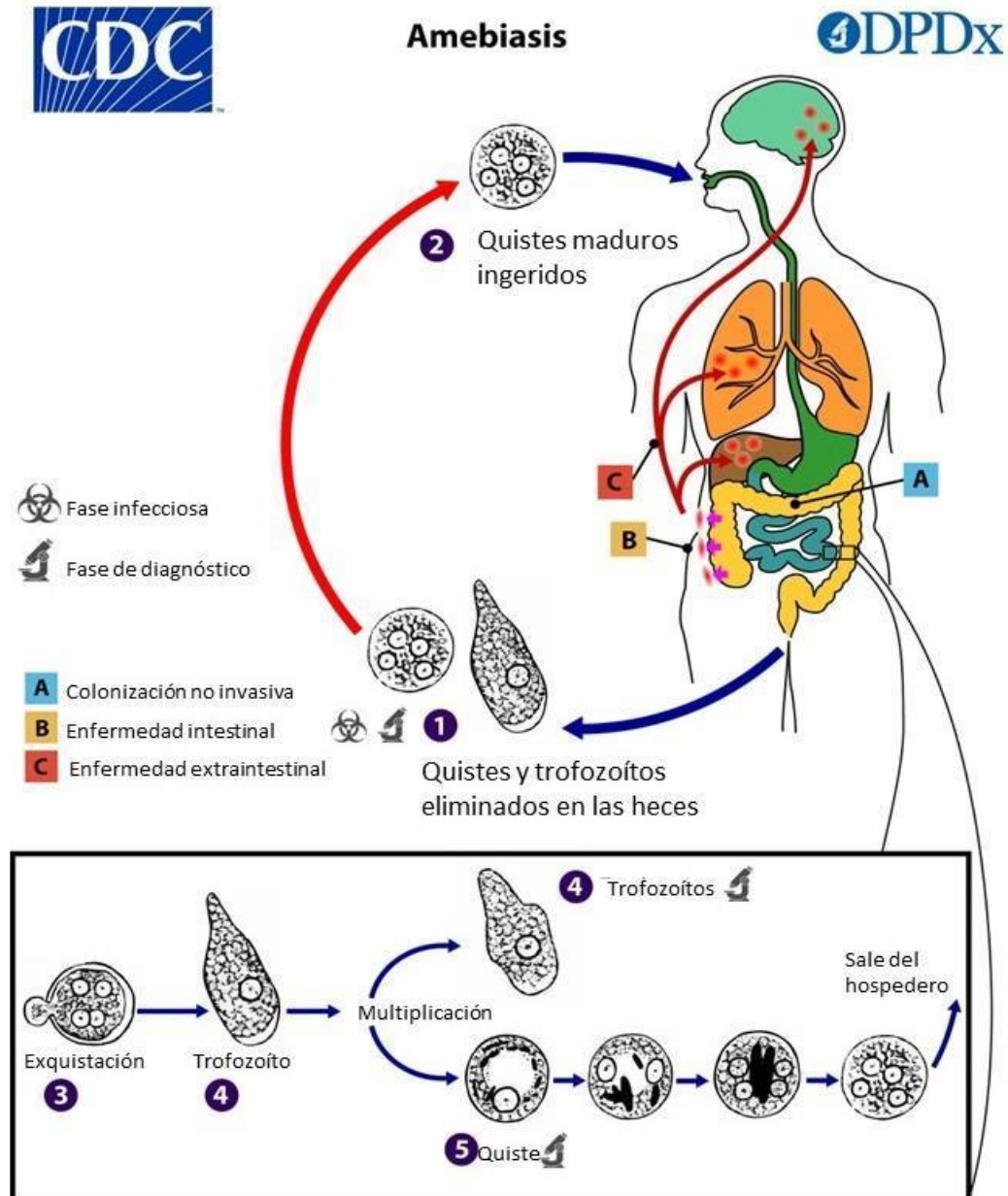


Figura 9. Ciclo de vida de *E. histolytica*. El ciclo incluye dos estadios del parásito: el quiste (forma de resistencia) que ingresa al hospedero a través de alimentos o agua contaminados; y el trofozoito (forma invasiva). La colonización no invasiva (A) ocurre en el colon, donde en algunos casos la ameba provoca amebiasis intestinal sintomática (B). Puede además diseminarse a otros órganos provocando infecciones extraintestinales (C), como en el hígado, los pulmones y el cerebro, donde pueden formarse abscesos. Bajo ciertas condiciones el trofozoito se enquista y los quistes son liberados en las heces. Imagen adaptada de CDC *Public Health Image Library*, 2015.

2.6.4 Genoma de *E. histolytica*

En el año 2005 Loftus y colaboradores, completaron la primera secuenciación del genoma de *E. histolytica*. Posteriormente, en 2010, se realizó un nuevo análisis y re- anotación del genoma, revelando que el tamaño del genoma de *E. histolytica* era notablemente más pequeño en comparación con el análisis inicial, abarcando un tamaño de 20 Mb distribuido en 1,486 scaffolds. Además, se estimó que este organismo albergaba 8,201 genes que codifican para proteínas, teniendo el gen más grande un tamaño de 15,210 pb y el más pequeño de 147 pb. Asimismo, se estima que alrededor de 2,000 genes contienen intrones y el promedio de intrones/gen es de 1.28. De manera interesante, el 60% de estos genes se mantuvieron idénticos a la anotación previa, mientras que el 40% presentó cambios estructurales notables. También se observó que el 60% de los genes experimentaron cambios en su clasificación funcional al compararlos con secuencias más recientes y nuevos datos experimentales. Además, solo se asignaron funciones putativas a 3,788 proteínas, lo que equivale al 46% del proteoma predicho de *E. histolytica*. Dentro de estas proteínas, se identificaron 897 familias compuestas por cinco o más miembros, de las cuales 58 familias no mostraron homología con proteínas conocidas. Esto sugiere la posible especificidad de estas familias para la especie *Entamoeba*. Sin embargo, para ambas anotaciones no se logró una cobertura completa debido a que este parásito presenta aneuploidía y una gran cantidad de regiones repetidas (Lorenzi *et al.*, 2010) (**Tabla 3**).

Tabla 3. Características del genoma de *E. histolytica*

Genoma	
Tamaño pb	20,799,072
Contenido GC (%)	24.2
Número de genes	8,201
Longitud promedio de genes (pb)	1,260.9
Número de genes/10 kbp	3.9
Porcentaje codificante (%)	49.7
Porcentaje de genes con intrones (%)	24.4

Exones	
Número	10,754
Promedio de número por gen	1.3
Contenido de GC (%)	19.3
Longitud promedio (pb)	962
Longitud total (pb)	10,340,284
Intrones	
Número	2,553
Contenido de GC (%)	19.3
Longitud promedio (pb)	74.1
Longitud total (pb)	189,260
Regiones intergénicas	
Contenido de GC (%)	20.5
Longitud promedio (pb)	708.7

Tabla adaptada de Lorenzi *et al.* 2010.

3. ANTECEDENTES PARTICULARES

3.1 Factores de transcripción Myb en *E. histolytica*

Las proteínas con DBD-MYB se han estudiado principalmente en organismos superiores y plantas. Sin embargo, recientemente se han investigado en organismos unicelulares como *Dictyostelium discoideum* (Stober *et al.*, 1992), *Euplotes aediculatus* (Yang *et al.*, 2013 y Yang *et al.*, 2003), *Sterkiella histriomuscorum* (Yang *et al.*, 2003), *Giardia lamblia* (Sun *et al.*, 2002), *Trichomonas vaginalis* (Ong *et al.*, 2006), *Trypanosoma brucei* (Li *et al.*, 2005) y *Plasmodium falciparum* (Boschet *et al.*, 2004).

Estas investigaciones han revelado que las proteínas con dominio MYB están involucradas en la regulación de genes que controlan el crecimiento celular, la diferenciación, la adhesión y la protección de las regiones teloméricas cromosómicas en estos parásitos.

La secuenciación del genoma de *E. histolytica* ha revelado la presencia de genes que codifican para FT canónicos que contienen un dominio de unión al DNA, entre ellos MYB, HMG1, MADS, GATA, HSF y STAT. Destaca especialmente la abundancia del dominio MYB en el genoma de *E. histolytica*, ya que se ha observado que presenta expansiones específicas de linaje en este parásito (Meneses *et al.*, 2010). Estas expansiones pueden ser relevantes en términos de la función y regulación de los genes asociados a este dominio, lo que podría tener implicaciones en los procesos biológicos o adaptaciones específicas del parásito. Previamente, se identificaron en el genoma de *E. histolytica* mediante análisis *in silico* 34 marcos de lectura abiertos (ORF) que codifican para 32 proteínas con DBD-MYB homólogos al factor c-Myb humano (Meneses *et al.*, 2010). Estas proteínas se agrupan en tres grupos monofiléticos pertenecientes a tres familias: La familia I denominada EhMybR2R3, la familia II denominada de unión a telómeros y la familia III o EhMybSHAQKYF (Meneses *et al.*, 2010) (**Fig. 10**).

La familia I EhMybR2R3 está conformada por 14 proteínas que están numeradas secuencialmente (EhMyb1 a EhMyb14). Estas proteínas tienen una longitud que varía de 97 a 231 aminoácidos, con una longitud promedio de aproximadamente 160 aminoácidos. Estas proteínas conservan solo los repetidos R2 y R3 en su DBD-MYB, la arquitectura de su dominio R2R3 se asemeja al de las proteínas Myb en plantas (**Fig. 11**). Cabe destacar que todas las proteínas que contienen el DBD-MYB R2R3 carecen de los dominios relevantes en el extremo C-terminal presentes en c-Myb (Meneses *et al.*, 2010). Por otro lado, las proteínas de unión a telómeros constan de 5 proteínas que tienen un solo repetido R3 y se relacionan con la protección de los extremos terminales de los cromosomas (Meneses *et al.*, 2010). La familia EhMybSHAQKYF contiene 9 proteínas similares a las proteínas SHAQKYF Myb que están presentes principalmente en las plantas, algas y *D. discoideum*. Estas proteínas tienen una sola repetición R3 al igual que la familia II con la diferencia de que en la región correspondiente al triptófano conservado de la tercera hélice alfa es reemplazada por un motivo $S_{/T}HAQK^Y_{/F}$ (Meneses *et al.*, 2010 y Rendón *et al.*, 2018). Por último, se identificaron 3 proteínas que no pertenecían a ninguna familia, una de estas proteínas presentó un dominio R2R3 a la cual se le denominó EhCDC5-like ya que comparte similitud con proteínas de la familia **CELL DIVISION CYCLE 5-LIKE PROTEIN** y las otras dos proteínas presentan sólo un repetido R3 (Meneses *et al.*, 2010).

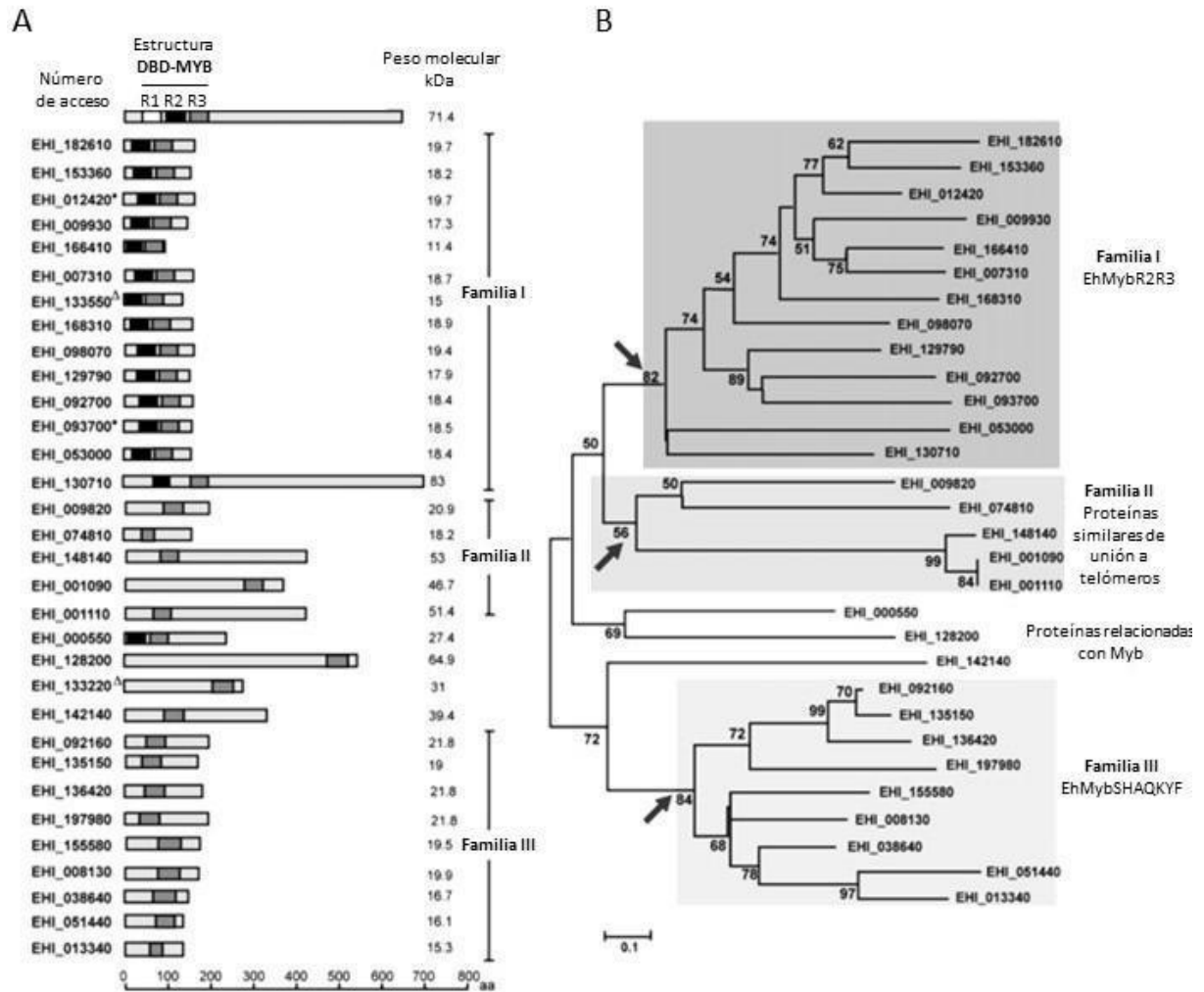


Figura 10. Proteínas EhMyb de *E. histolytica*. **A)** Diagrama de la proteína c-Myb y las proteínas que contienen el DBD-MYB de *E. histolytica*. Organización del dominio DBD-MYB: R1, blanco; R2, negro sólido; y R3, gris. *Indica proteínas codificadas por genes duplicados, y Δ indica proteínas codificadas por genes truncos. **B)** Árbol filogenético de las proteínas que contienen el dominio DBD-MYB de *E. histolytica* (no incluye las proteínas de genes duplicados o truncos). Los cuadros grises claros muestran las familias específicas de *E. histolytica*. Imagen adaptada de Meneses *et al.*, 2010.

Debido a la homología que las proteínas EhMyb presentan con el FT c-Myb, ellas podrían reconocer al ERM. Por lo que los autores analizaron la presencia del ERM en los promotores de genes que podrían activarse por las proteínas EhMyb y encontraron genes relacionados con transducción de señales, transporte vesicular, respuesta al choque térmico y virulencia. Esto indica que los FT Myb en *E. histolytica* podrían estar relacionados con la regulación transcripcional de estos genes (Meneses *et al.*, 2010).

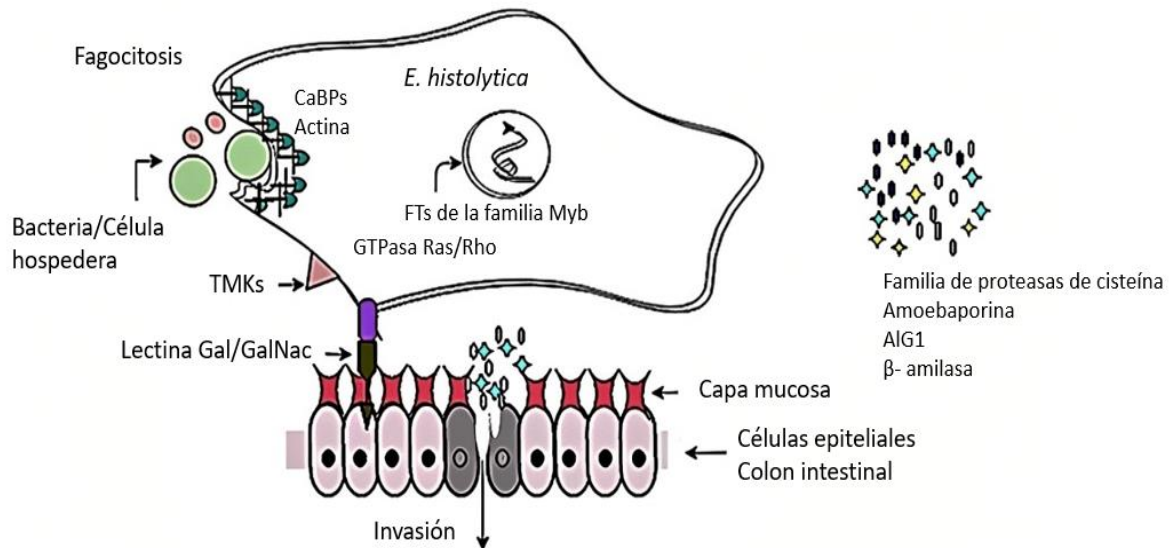


Figura 12. Moléculas relacionadas con la virulencia de *E. histolytica* reveladas a partir de estudios transcriptómicos. Las moléculas que contribuyen a la virulencia y patogenicidad de *E. histolytica* están representadas por: Cisteína proteasas, contribuyen a la invasión de tejidos; β -amilasa, ayuda en la eliminación y utilización de la mucosidad; AIG1, funciona en defensa contra bacterias intestinales; Amebaporos, median la lisis celular; Lectina Gal/GalNac, facilita el contacto de los ligandos con la célula huésped; Cinasas transmembranales (TMK) y las GTPasas Ras/Rho, participan en la transducción de señales; Factores de transcripción de la familia Myb, regulan la expresión de múltiples vías; Proteínas de unión a calcio (CaBP), median la reorganización de la actina durante la fagocitosis de bacterias/células huésped. Imagen adaptada de Naiyer *et al.*, 2019.

Otros genes que modificaron su expresión fueron los FT de la familia Myb, lo cual los hace factores de transcripción relevantes para el control de la expresión génica en este parásito. Por ejemplo, el gen *ehmyb10* mostró una expresión diferencial (aumento de 2.87 veces) en trofozoitos obtenidos en el día uno de colonización intestinal en un modelo murino (Gilchrist *et al.*, 2006).

Adicionalmente, los FT Myb desempeñan un papel crucial durante el enquistamiento de *E. histolytica*. En un estudio por Ehrenkauffer (2009), se identificó a EhMyb-dr como el primer FT que probablemente esté involucrado en este proceso. EhMyb-dr pertenece a la familia EhMybSHAQKYF. La sobreexpresión de EhMyb-dr indujo la expresión de sólo un subconjunto de genes específicos de quistes, lo que sugiere que es probable que se requieran otros reguladores transcripcionales, junto con factores de remodelación de la cromatina como las HATs, para que ocurra la conversión de trofozoito a quiste. La proteína EhMyb-dr muestra una expresión baja en la etapa de trofozoito y aumenta su expresión en el enquistamiento.

EhMyb-dr reconoce al motivo CCCCCC presente en los promotores de los genes relacionados con enquistamiento (Ehrenkaufner *et al.*, 2009).

Por tanto, la comprensión de los FT Myb es esencial para entender la patogenicidad, la virulencia y el enquistamiento de *E. histolytica* y para desarrollar estrategias terapéuticas eficaces contra esta infección parasitaria.

3.2 EhMyb10

El factor de transcripción EhMyb10 pertenece a la familia EhMybR2R3. El gen *ehmyb10* codifica una proteína de 153 aminoácidos con un PM de ~17.9 kDa. La comparación entre la estructura terciaria del DBD de c-Myb y EhMyb10, muestran una notable similitud. El DBD de EhMyb10 conserva la misma conformación hélice-giro-hélice en los repetidos R2 y R3, al igual que c-Myb (Meneses *et al.*, 2010) (**Fig. 13**).

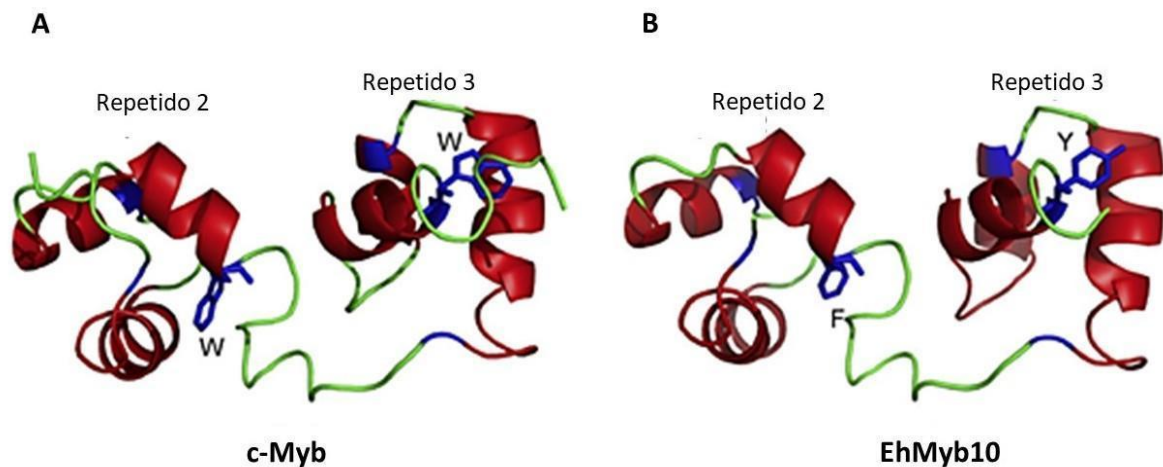


Figura 13. Estructura tridimensional de las proteínas c-Myb y EhMyb10. A) Estructura tridimensional de c-Myb. B) Estructura tridimensional de EhMyb10. “W” representa a los residuos de triptófano y “F” y “Y”, residuos de los fenilalanina y tirosina respectivamente. Imagen adaptada de Meneses *et al.*, 2010.

Por medio de ensayos de disminución de la movilidad electroforética (EMSA) ha sido demostrado que EhMyb10 se une al ERM del promotor del gen *mim-1* y forma complejos de DNA-proteína de forma específica (**Fig. 14A**), ya que al interactuar con una secuencia de *mim-1* mutada (*mim-1M*) o con secuencias no relacionadas (*Adhp*) no forma complejos de DNA-proteína (Meneses *et al.*, 2010) (**Fig. 14B**).

La especificidad de esta interacción se confirmó mediante ensayos de competencia con sondas no marcadas donde se observó que compitió específicamente la interacción DNA-proteína utilizando un exceso molar de 100 veces de la sonda no radiactiva *mim-1* (Fig. 14C, carril 3), mientras que no se observó competencia con el mismo exceso molar de las sondas *mim-1M* o *Adhp* utilizadas como competidores (Fig. 14C, carriles 4 y 5) (Meneses *et al.*, 2010).

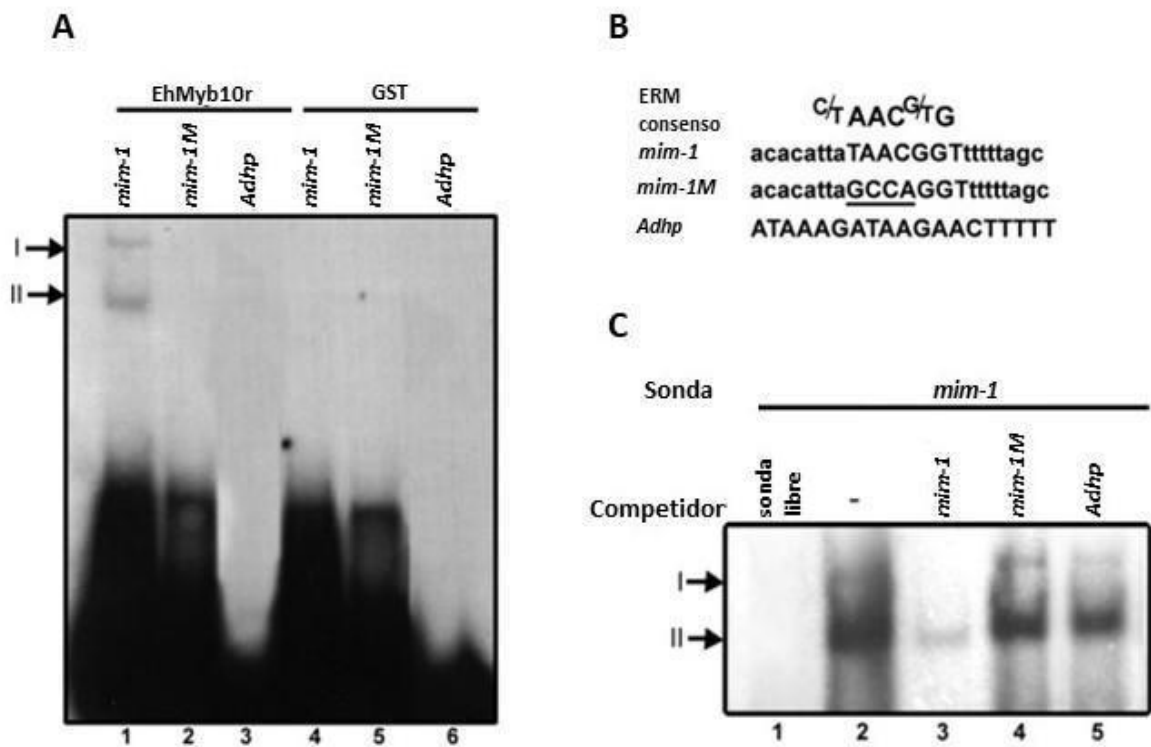


Figura 14. EhMyb10 se une al elemento de reconocimiento a MYB (ERM). A) Ensayo EMSA de la proteína EhMyb10r. B) Secuencias de los oligonucleótidos utilizados en los ensayos EMSA. C) Ensayos de competencia. Los complejos de DNA-proteína I y II se indican mediante flechas. Imagen adaptada de Meneses *et al.*, 2010.

Asimismo, en extractos nucleares de trofozoitos de *E. histolytica* Meneses y cols. detectaron la unión de proteínas a la secuencia ERM consenso. Por otro lado, mediante análisis *in silico* se observó que el ERM se encuentra en los promotores de 246 genes de *E. histolytica* relacionados con transducción de señales, transporte vesicular, respuesta a choque térmico y virulencia. Por lo tanto, el factor de transcripción EhMyb10 de *E. histolytica* podría participar en la regulación transcripcional de genes relacionados con estas funciones (Meneses *et al.*, 2010).



Es por ello, que en este trabajo nos propusimos determinar las interacciones proteicas del FT EhMyb10 para tener una comprensión de la función de EhMyb10 en la regulación de la expresión génica y en la virulencia de la amiba. Esto podría revelar proteínas clave que participan en vías reguladoras críticas relacionadas con la virulencia. La comprensión de las interacciones de EhMyb10 puede ofrecer oportunidades para desarrollar terapias dirigidas a interrumpir o modular su función. Esto podría conducir al desarrollo de tratamientos más efectivos contra las infecciones amebianas. Por otro lado, el estudio de EhMyb10 no solo tiene implicaciones clínicas, sino que también contribuirá a nuestra comprensión fundamental de la biología de la amiba y de los factores de transcripción en general.

4. HIPÓTESIS

La proteína EhMyb10 interactúa con proteínas nucleares que están relacionadas con el control de la transcripción en *E. histolytica*.

5. JUSTIFICACIÓN

La familia de proteínas con DBD-MYB en *E. histolytica* es una de las familias más abundantes relacionadas con la regulación de la transcripción. Los genes que codifican para estas proteínas modifican su expresión en respuestas a diferentes eventos como lo son la virulencia, el enquistamiento, la respuesta al estrés y durante el crecimiento. De estas proteínas, EhMyb10 es de las que presenta mayor expresión y además se sobreexpresa en condiciones de colonización intestinal. Es por ello, que es importante determinar las interacciones proteína-proteína de este factor de transcripción para dilucidar su papel en la regulación de la expresión génica.

6. OBJETIVOS

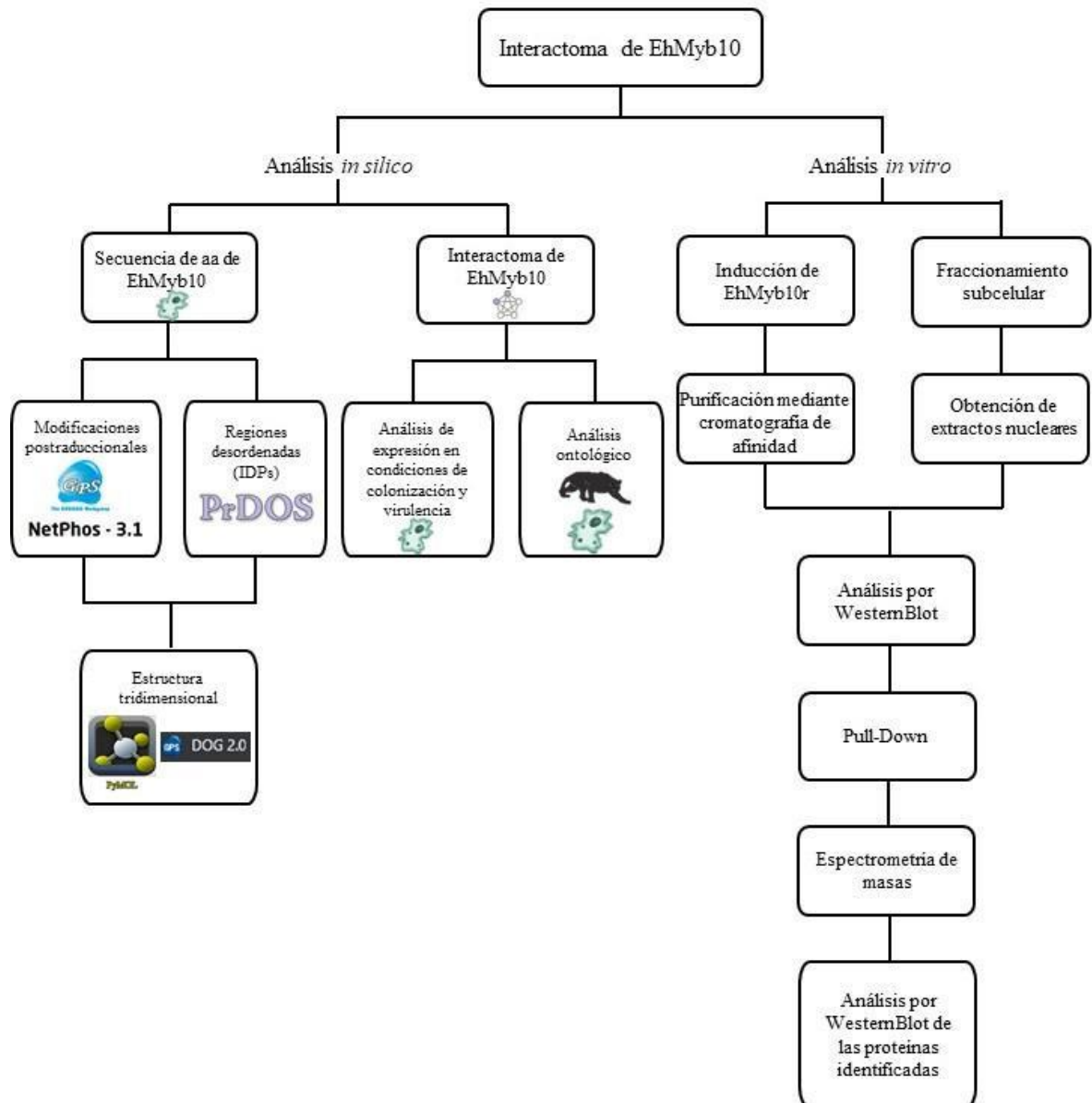
6.1 Objetivo general

Caracterizar mediante análisis *in silico* e *in vitro* el interactoma del factor de transcripción EhMyb10.

6.2 Objetivos particulares

1. Analizar *in silico* la secuencia de aminoácidos de EhMyb10.
2. Determinar *in silico* el interactoma de EhMyb10.
3. Analizar la expresión *in silico* de los genes que codifican para proteínas que interaccionan con EhMyb10 en condiciones de estrés y durante la colonización intestinal.
4. Identificar *in vitro* las proteínas que se unen a EhMyb10.

7. ESTRATEGIA EXPERIMENTAL



8. MATERIALES Y MÉTODOS

8.1 Identificación de PTMs, regiones desordenadas y predicción de estructura tridimensional de la proteína EhMyb10.

Para obtener los datos del gen y la proteína EhMyb10 se utilizó la base de datos de AmoebaDB (<https://amoebadb.org/amoeba/app/>) utilizando su número de acceso (EHI_129790). Asimismo, se obtuvo de la misma base de datos la secuencia en formato FASTA de la proteína para obtener los datos del peso molecular (PM) y el punto isoeléctrico (pI) mediante el servidor de ExPASy (<https://www.expasy.org/>). Para la determinación de modificaciones postraduccionales de la proteína EhMyb10, se realizaron diversos análisis *in silico* con diferentes predictores. Para determinar los residuos susceptibles a fosforilación, se utilizaron dos servidores: NetPhos-3.1-services (<https://services.healthtech.dtu.dk/service.php?NetPhos-3.1>) y GPS 6.0 (<http://gps.biocuckoo.cn/index.php>). Asimismo, se emplearon iUbiq-Lys (<http://www.jci-bioinfo.cn/iUbiq-Lys>) y BDM-PUB: Prediction of Ubiquitination sites with Bayesian Discriminant Method (<http://bdmpub.biocuckoo.org/index.php>), para identificar los sitios susceptibles a ubiquitinación. Por último, se evaluó la posible SUMOilación de la proteína, empleando los servidores SUMOgo (<http://predictor.nchu.edu.tw/SUMOgo/>), JASSA V4 (<http://www.jassa.fr>) y SUMOplot™ Analysis Program (<https://www.abcepta.com/sumoplot>). Para determinar las regiones intrínsecamente desordenadas presentes en la proteína, se utilizaron los servidores PrDOS (<https://prdos.hgc.jp/cgi-bin/top.cgi>) y IUPred2A (<https://iupred2a.elte.hu/plot>). A continuación, se obtuvo la estructura tridimensional de EhMyb10 por medio de la plataforma AlphaFold (<https://alphafold.ebi.ac.uk/>) en formato PDB. Para incluir las modificaciones postraduccionales y las regiones desordenadas en la estructura tridimensional se utilizó el programa de PyMOL 2.5 (<https://pymol.org/2/>). La estructura del dominio DBD-MYB que incluye los sitios susceptibles a modificaciones postraduccionales y las regiones intrínsecamente desordenadas se modeló con el programa DOG 2.0- Visualization of Protein Domain Structures (<http://dog.biocuckoo.org/>).

8.1.1 Determinación del interactoma *in silico* de EhMyb10

Para determinar la interacción proteína-proteína se utilizó la herramienta de STRING (versión 10.5, 11 y 12) (<https://string-db.org>), seleccionando un puntaje de interacción de confianza media (0.400) como valor de corte, y un número máximo de no más de 10 interactuantes.

8.1.2 Análisis ontológico (GO) de las proteínas que interactúan con EhMyb10

Para los análisis ontológicos en función molecular, proceso biológico y compartimento celular de los genes, que codifican para proteínas que interactúan con EhMyb10, se utilizó los servidores PantherDB (<http://www.pantherdb.org>), AmoebaDB (<https://amoebadb.org/amoeba/app/>) y UniProt (<https://www.uniprot.org/>).

8.1.3 Análisis de expresión de los genes que codifican para proteínas que interactúan con EhMyb10

Por otro lado, se llevó a cabo el análisis de los niveles de expresión del interactoma de EhMyb10 en dos contextos diferentes: colonización/invasión intestinal y virulencia. Para ello, se utilizó el transcriptoma de *E. histolytica* durante la Colonización /Invasión intestinal evidenciado por Microarrays por los autores Gilchrist *et al.*, 2006; y el transcriptoma de *E. histolytica* axénica (normal, atenuada, virulenta) y cultivada en colon evidenciado por RNA-seq por los autores Weber *et al.*, 2016. Los datos obtenidos se visualizaron en un mapa de calor obtenido mediante el servidor Heatmapper (<http://www.heatmapper.ca/>) utilizando el coeficiente de correlación de Pearson como método de medición de distancia y el método de agrupamiento de enlace promedio (Average Linkage).

8.2 Preparación de células competentes de *Escherichia coli* DH5α y BL21

Para preparar células competentes, se seleccionó una colonia de la cepa DH5α o BL21 de *Escherichia coli*, la cual fue utilizada para inocular 5 mL de medio Luria Bertani (LB) líquido sin ampicilina. Este cultivo se incubó a 37 °C y 180 rpm durante toda la noche. Al día siguiente, se realizó una segunda inoculación con el volumen deseado (entre 5 y 50 mL), y se incubó a 37 °C y 200 rpm hasta que alcanzó una densidad óptica (OD) de 0.4-0.6 a 600 nm. Luego, se procedió a centrifugar el cultivo a 1000 rpm (con un límite máximo de 2000 rpm para evitar estrés en las células) durante 5 min, y se resuspendió la pastilla resultante en CaCl₂ (50 mM) frío. Posteriormente, se incubó la suspensión a 4 °C durante 20 min, y se volvió a centrifugar a 1000 rpm durante 5 min. La pastilla obtenida se suspendió nuevamente en CaCl₂ (1 mL por cada 10 mL de cultivo) y las bacterias competentes se almacenaron a 4 °C durante 7 días. Después de este periodo, las bacterias pierden su competencia.

8.3 Transformación de células competentes

En un vial de células competentes *E. coli* DH5α o BL21 (de 50-100 μL), se colocó 50 ng del DNA plasmídico (pGEX-6P-1 y pGEX-EhMyb10). Las bacterias se crecieron en medio LB líquido, seleccionando con ampicilina (100 μg/mL) y se incubaron de 5-30 min en hielo, mezclando ligeramente el DNA una vez que las células están descongeladas. Transcurrido el tiempo, se dio un choque térmico de 45-50 seg a 42 °C. Después los tubos se colocaron en hielo por 2 min, se agregaron 300 μL de medio LB ambientado a 37 °C y se incubó a esa temperatura a 200 rpm durante 1 hora. Al terminar la incubación, se plaquearon 30-50 μL de cada transformación en una caja petri con agar LB y ampicilina (100 μg/mL) para realizar la selección de las transformantes.

8.4 Obtención de DNA plasmídico

El DNA plasmídico fue extraído utilizando el Kit Midi (Qiagen) siguiendo las instrucciones del fabricante.

Las bacterias transformadas se cultivaron y cosecharon para que la pastilla resultante fuera concentrada y se agregaron 300 μL de buffer P1, procediendo a resuspensión. Luego, se añadieron 300 μL de buffer P2, se mezcló por inversión y se incubó durante 5 min a temperatura ambiente (si la muestra estuvo bien resuspendida tomará un color azul). Una vez transcurrido el tiempo, se agregaron 300 μL de buffer P3 frío, se mezcló hasta que el color desapareciera y se incubó sobre hielo durante 15 min. Posteriormente, se centrifugó a 13,000 rpm durante 30 min a 4 °C. Si se observó un exceso de proteína desprendida, se volvió a centrifugar o filtrar para obtener únicamente el sobrenadante. Al sobrenadante se le añadieron 560 μL de isopropanol frío y se centrifugó a 13,000 rpm durante 30 min. Por último, la pastilla resultante se lavó con 600 μL de etanol al 70%, se dejó secar y se resuspendió en 100-200 μL de agua inyectable. El DNA plasmídico se cuantificó utilizando el NanoDrop™ de Thermo Scientific™.

8.5 Análisis de los plásmidos pGEX-6P-1 y pGEX-EhMyb10 mediante restricciones y PCR

Después de cuantificar el DNA plasmídico, se llevó a cabo una reacción de restricción utilizando las enzimas *Bam*HI y *Xho*I (New England BioLabs). La reacción de restricción se realizó siguiendo las condiciones indicadas en la **Tabla 4**.

Tabla 4. Condiciones de restricción de los plásmidos pGEX-6P-1 y pGEX-EhMyb10

Reacción	Plásmido	DNA plasmídico (ng)	Buffer Cut Smart 10x (μL)	<i>Bam</i> HI (μL)	<i>Xho</i> I (μL)	BSA (μL)	H ₂ O (μL)	VF* (μL)
1	pGEX-6P-1	500	2	1	1	0.2	13.8	20
2	pGEX-6P-1 S/R	500	2	-	-	0.2	15.8	20
3	pGEX-EhMyb10	500	2	1	1	0.2	13.8	20
4	pGEX-EhMyb10 S/R	500	2	-	-	0.2	15.8	20

La condición de incubación fue de 37 °C por 24 hrs *VF: volumen final.

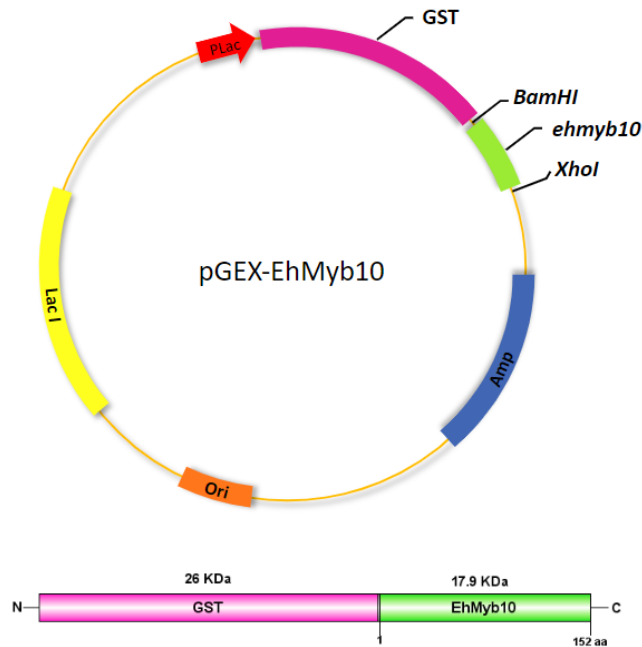


Figura 15. Plásmido pGEX-EhMyb10. Mapa del plásmido utilizado para la obtención de la proteína recombinante EhMyb10 acoplada a una etiqueta de GST (Glutatión S-transferasa). El gen *ehmyb10* tiene un tamaño de 459 pb mientras que la etiqueta de GST tiene un tamaño de 7440 pb. Imagen de elaboración propia a partir de datos obtenidos de Addgene: Vector Database (<https://www.addgene.org/>).

Para la amplificación del gen se realizó una PCR utilizando el termociclador GeneAmp[®] PCR System 9700 (Applied Biosystems) (**Tabla 5**). La reacción se llevó a un volumen de 12.5 μ L: 6.25 μ L de Master Mix (Roche), 0.5 μ M de primer sentido (5'-GAGTCGGGATCCATGAACGTGACTATTGAACAA-3'), y antisentido (5'-GAGTCGCTCGAGCTACCAAGTGGACCCTCCATT-3') y 100 ng de DNA genómico de *E. histolytica* como molde.

Tabla 5. Condiciones de amplificación del gen *ehmyb10*

PCR	Temperatura °C	Duración	Ciclos
Desnaturalización inicial	94 °C	3 min	1x
Desnaturalización	94 °C	30 seg	
Alineamiento	55 °C	45 seg	25x
Elongación	72 °C	1 min	
Elongación final	72 °C	7 min	1x

Las restricciones y la PCR, se analizaron mediante electroforesis en geles de agarosa al 1.5% corridos a 90 V por 30 min y teñidos con Gelred Nucleic Acid Stain 1x (Biotium). Los geles se analizaron en el fotodocumentador Quantity One XR (BioRad).

8.6 Inducción y purificación de las proteínas recombinantes: GST y EhMyb10 (EhMyb10r)

Las células *E. coli* BL21 Rosetta transformadas se incubaron a 37 °C a 180 rpm en medio LB líquido con ampicilina (100 µg/mL) durante 3 horas hasta llegar a una OD entre 0.4-0.6 a 600 nm. Una vez obtenida esta densidad se añadió IPTG (1mM) y se incubó a 16 °C a 180 rpm durante toda la noche. Al día siguiente el cultivo se centrifugó y la pastilla celular se almacenó a -20 °C, para su posterior uso. Para la obtención de la proteína recombinante de la fracción soluble a 50 mL de cultivo empastillado se añadieron 7.5 mL de buffer de lisis [PBS 10x, 0.5% Tritón x-100, 10 mM EDTA] y 0.5 mM de PMSF y se incubó durante 10 min en hielo. Posteriormente, el cultivo se sonicó en ciclos de 30 s al 80% por 3 min y se centrifugó a 7500 rpm por 5 min para separar la fracción soluble de la insoluble. Posteriormente, la fracción soluble se purificó en condiciones nativas utilizando cromatografía de afinidad con Glutathione Sepharose[®] High Performance. Para ello, 500 µL de slurry se lavaron dos veces con 9 mL de PBS pH 7.5 [140 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄] centrifugando a 1000 rpm por 5 min en cada lavado. La resina se mezcló con la fracción soluble del extracto de bacterias adicionado con inhibidores [Complete Mini Tablets de ROCHE 1x, 0.5 mM PMSF y 1 mM de DTT] mediante vórtex y se incubó toda la noche a 4 °C. A continuación, se centrifugó a 8000 rpm por 5 minutos y el material recuperado (proteínas unidas a la resina) se lavó tres veces con 5 mL de PBS frío. Por último, se realizaron tres eluciones para cada muestra de proteínas-resina con 200 µL de Buffer de elución [0.1 % Tritón x-100, 50 mM Tris-HCl pH 7.5, 40 mM Glutatió n reducido, 0.2 M NaCl, 1 mM DTT]. Los eluidos se recuperaron y almacenaron a -20°C. Las muestras fueron analizadas mediante SDS-PAGE al 10% y WB usando anticuerpos α-GST.

8.7 Fraccionamiento subcelular de proteínas de *E. histolytica*

Se resuspendieron diez millones de trofozoitos en 1.33 mL de Buffer A frío [10 mM HEPES pH 7.9, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 1 mM DTT, 0.5 mM PMSF, Complete protease inhibitor cocktail (ROCHE) 1x, y 10 mM E64 (Sigma Aldrich)]. Los trofozoitos se incubaron por 5 min en hielo, se añadieron 83 μ L de NP-40 al 10% y se mezclaron mediante vórtex por 30 seg. El lisado se centrifugó por 10 min a 14000 rpm, para separar el sobrenadante (extractos citoplasmáticos) del pellet (núcleos). El pellet se suspendió en 3 mL de buffer A y se colocó sobre 1 mL de buffer A con 0.34 M de sacarosa. Después se centrifugó por 10 min a 14000 rpm. La pastilla se resuspendió en 300 μ L de buffer C [20 mM HEPES pH 7.9, 400 mM NaCl, 1mM EDTA, 1mM EGTA, 1 mM DTT, 0.5 mM PMSF, Complete 1x, y 10 μ M E-64], y se agitó vigorosamente a 4 °C por 30 min en un vórtex. La muestra se centrifugó a 13,000 rpm por 10 min a 4 °C. El sobrenadante se recuperó como fracción nuclear. Las muestras se cuantificaron mediante un ensayo de Lowry compatible con detergente (BIO-RAD), y se almacenaron a -70°C hasta su uso.

8.8 Ensayo de Pull-Down con la proteína EhMyb10r

Para los ensayos de Pull-Down, las proteínas recombinantes (GST y EhMyb10r) se inmovilizaron a una resina de glutatión (Glutathione Sepharose® High Performance). Por otro lado, 25 μ g de extractos nucleares se clarificaron con 100 μ L de resina de glutatión en presencia de inhibidores [Complete protease inhibitor cocktail, (ROCHE) 1x] durante 30 min a 4 °C en un vórtex. Posteriormente, se tomaron 50 μ L del sobrenadante clarificado y se incubaron con 250 μ L de la resina unida a las proteínas toda la noche a 4 °C en un vortex. La resina se lavó con 250 μ L de PBS, y se eluyó incubando a 4 °C durante 30 min con 200 μ L de buffer de elución [0.1% Tritón x-100, 50 mM Tris-HCl pH 7.5, 40 mM glutatión reducido, 0.2 mM NaCl, 1 mM DTT]. Este paso se repitió 3 veces. Las fracciones obtenidas se resolvieron mediante SDS-PAGE al 12%, y luego se visualizaron mediante tinción con Coomassie® Brilliant blue G 250.

8.9 SDS-PAGE y Western Blot

Para el análisis de las proteínas recombinantes, se utilizó la electroforesis desnaturizante en una dimensión en geles de poliacrilamida al 10% a 120 V, y para el análisis de los extractos nucleares y las muestras del Pull-Down, se utilizaron geles al 12%. Se empleó el marcador Precision Plus Protein™ y Kaleidoscope™ (BIO-RAD) para determinar los pesos moleculares de las proteínas. Los geles se tiñeron con Coomassie® Brilliant blue G 250, y se capturaron con el fotodocumentador Chemi-Doc XRS (BIO-RAD). Para el análisis de WB los geles se sumergieron en buffer de transferencia [48 mM Tris, 39 mM Glicina, 0.04% SDS, y 20% Metanol] al igual que las membranas de NTC y el papel Whatman, y se formó un “sándwich” de la siguiente manera: papel-membrana-gel-papel. El sándwich se colocó en una cámara de transferencia semiseca (BIO-RAD) del lado del cátodo durante 40 min a 15V. Posteriormente, las membranas se tiñeron con Rojo de Ponceau al 0.1% (Sigma), y se lavaron con agua destilada. Las membranas se bloquearon con leche Svelty (Nestlé) al 5% en PBS-Tween 20 al 0.1% a 37 °C durante 1 hora en agitación constante a 4°C durante toda la noche. Posteriormente, las membranas se lavaron 5 veces con PBS-Tween 0.1%, se secaron y se guardaron a 4 °C hasta su uso.

Para el análisis de las proteínas recombinantes, las membranas se incubaron con el anticuerpo α -GST hecho en cabra (GE Healthcare, 1:5000). Las membranas se incubaron toda la noche a 4 °C. Al día siguiente las membranas se lavaron 5 veces con PBS-Tween 20, y se incubaron con un anticuerpo secundario acoplado a HRP. Para el análisis de los extractos nucleares, se utilizó el anticuerpo α -histona H4K20me3 (Abcam, 1:1000) a una concentración de 1:1000, para los extractos citoplasmáticos el anticuerpo α -CP112 (no comercial, 1:2500) y como control de carga el α -actina (Sigma, 1:1000). Para la confirmación de las proteínas identificadas por espectrometría de masas, se utilizó como control el anticuerpo α -histona H4K20me3 (Abcam, 1:1000) y el anticuerpo α -fibrilarina (Santa Cruz Biotechnology Inc., 1:3000), y se capturaron en el fotodocumentador Chemi-Doc XRS (BIO-RAD).

8.10 Identificación de proteínas por espectrometría de masas (IMS-DIA-MS)

Las eluciones obtenidas del ensayo de Pull-Down, se cargaron en un gel SDS-PAGE y se dejaron correr aproximadamente 0.7-1 cm. Después, el gel se tiñó con Coomassie® Brilliant blue G 250 compatible con espectrometría de masas en un recipiente de vidrio previamente lavado con solución crómica. Todos los reactivos utilizados para preparar los geles de poliacrilamida, la solución de Coomassie y la solución desteñidora, se filtraron y esterilizaron previamente para evitar la contaminación de las muestras. Posteriormente, el gel se envió en la Unidad de Genómica, Proteómica y Metabolómica LaNSE-Cinvestav para su análisis por espectrometría de masas con ionización por movilidad iónica (IMS) y adquisición de datos independiente (DIA), también conocida como IMS-DIA-MS (Ion Mobility Spectrometry-Data Independent Acquisition-Mass Spectrometry). Esta técnica separa proteínas en fase gaseosa antes de su análisis en el espectrómetro de masas, lo que aumenta la sensibilidad y la resolución. Además, la técnica de adquisición independiente de datos (DIA), proporciona una identificación más completa de proteínas en la muestra en comparación con otras técnicas de adquisición de datos (Krasny *et al.*, 2021).

8. 11 Análisis *in silico* de las proteínas identificadas en los ensayos de Pull-Down

Una vez identificadas las proteínas de interés, se procedió a realizar un análisis ontológico exhaustivo para clasificar las proteínas con base en su función molecular, proceso biológico, y localización celular. Este análisis se llevó a cabo utilizando las bases de datos AmoebaDB (<https://amoebadb.org/amoeba/app/>), PantherDB (<http://www.pantherdb.org/>) y Uniprot (<https://www.uniprot.org/>). Adicionalmente, se llevó a cabo un análisis de interacciones utilizando la herramienta STRING versión 12.0 (<https://version-12-0.string-db.org/>), para determinar las posibles asociaciones y relaciones entre las proteínas identificadas con un valor de confianza de 0.150. Esta estrategia nos permitió obtener una visión más completa y detallada de cómo estas proteínas interactúan entre sí.

9. RESULTADOS

9.1 Análisis *in silico* de la secuencia de aminoácidos de la proteína EhMyb10

El transcrito del gen *ehmyb10* tiene una longitud de 459 pb, lo que coincide con la longitud de su secuencia genómica. Esta correspondencia indica que el gen no contiene intrones. A partir de su secuencia de mRNA, se sintetiza una proteína compuesta por 152 aminoácidos, con un PM estimado de 17.9 kDa y un punto isoeléctrico (pI) de 9.86. Esto significa que en un entorno con un pH cercano a 9.86, la carga neta de la proteína es igual a cero, lo cual es importante determinar, ya que puede tener implicaciones en su solubilidad, interacciones con otras moléculas, procesos de purificación y su comportamiento en diferentes entornos biológicos (Tabla 6).

Tabla 6. Características del gen y la proteína EhMyb10

Gen <i>ehmyb10</i>	
Tamaño de la secuencia genómica	459 pb
Tamaño del transcrito (mRNA)	459 pb
Número de intrones	0
Proteína	
Tamaño de la proteína (aminoácidos)	152 aa
Peso molecular (PM)	17.9 kDa
Punto isoeléctrico (pI)	9.86

Tabla elaborada a partir de datos obtenidos de las bases de datos AmoebaDB y Expasy.

9.2 Análisis de PTMs de EhMyb10

Las PTMs desempeñan un papel crucial en la regulación de la actividad biológica de las proteínas. Por lo tanto, con el objetivo de comprender mejor las características de EhMyb10, se realizó un análisis para determinar las PTMs que presenta esta proteína, como fosforilaciones, ubiquitinaciones y SUMOilaciones como ocurre en la proteína c-Myb.

Los análisis bioinformáticos revelaron que 14 residuos de aminoácidos son susceptibles a fosforilación. De estos, 10 corresponden a residuos de serina, 2 a treonina, y 2 a tirosina (**Tabla 7** y **Fig. 16**). La fosforilación de estos residuos es común en las proteínas Myb y desempeña un papel importante en la regulación de su actividad. En el caso de EhMyb10, se observó que los residuos de serina son los más propensos a ser fosforilados. Esto puede tener un impacto en la interacción de EhMyb10 con otros componentes de la maquinaria de transcripción, su capacidad de unirse al DNA y su estabilidad.

Tabla 7. Aminoácidos susceptibles a fosforilación en EhMyb10

Residuo	Posición	Cinasa	Puntaje	Umbral
NNDTYNYSI	12	-	0.73	0.5
NYSITYKSE	17	PKC	0.88	0.5
YSITYKSEK	18	-	0.51	0.5
ITYKSEKVK	20	PKC	0.90	0.5
FGEKSWVEV	52	CKII	0.58	0.5
KFVGTRSRK	62	PKC	0.86	0.5
VGTRSRKQC	64	PKC	0.70	0.5
KRPWSSEED	86	CKII	0.62	0.5
		RSK	0.50	
RPWSSEEDE	87	CKII	0.61	0.5
GSKWSEISK	105	PKC	0.72	0.5
		cdc2	0.50	
WSEISKKLS	108	PKC	0.64	0.5
SKKLSGRTA	112	PKC	0.59	0.5
YKALSIRNK	127	PKC	0.86	0.5
RNKLSCFIE	133	PKA	0.54	0.5

Tabla elaborada con datos obtenidos del servidor NetPhos 3.1 (2023).

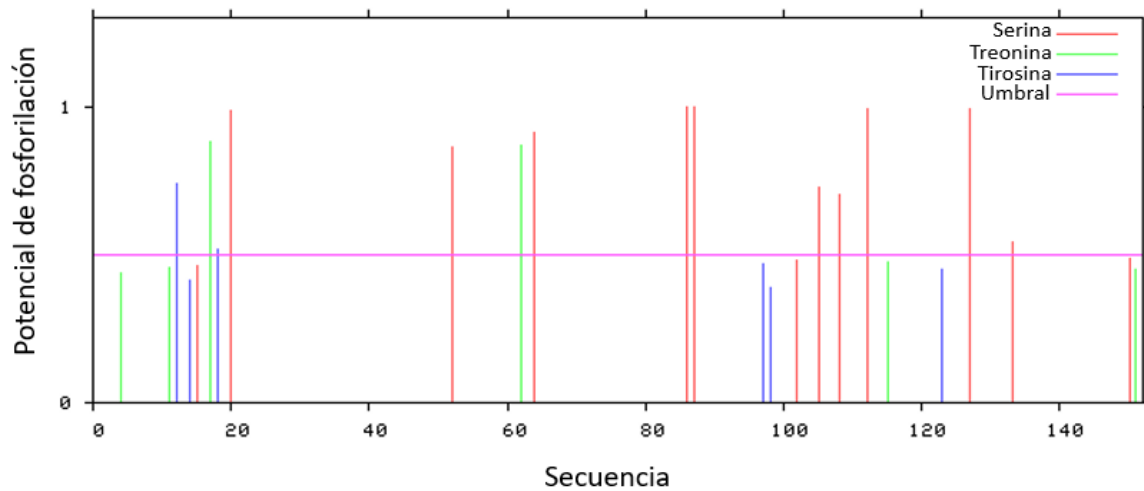


Figura 16. Residuos potenciales de fosforilación en la secuencia de aminoácidos de EhMyb10. Se utilizó el predictor NetPhos 3.1 Server - DTU para identificar fosforilaciones en EhMyb10. Punto de corte 0.5.

Por otro lado, los resultados del análisis utilizando el servidor BDM-PUB, revelaron que EhMyb10 posee 7 sitios potencialmente susceptibles a ubiquitinación, los cuales corresponden a residuos de lisina (**Tabla 8**). Estos hallazgos sugirieron que la ubiquitinación podría ser un mecanismo de regulación para EhMyb10, al igual que ocurre con su homólogo c-Myb. Por tanto, el proceso de ubiquitinación puede desempeñar un papel importante en la regulación de la estabilidad y degradación de EhMyb10 por el sistema de ubiquitina-proteasoma, asegurando así su nivel adecuado y actividad funcional en la célula.

Tabla 8. Sitios predichos de ubiquitinación en EhMyb10

Motivo	Posición	Puntaje	Umbral
15-SITYKSEK K VKKQRKA-29	22	2.66	0.3
17-TYKSEK V KKQRKAGI-31	24	1.90	0.3
18-YKSEK V KKQRKAGIA-32	25	1.00	0.3
21-EKVKKQR K AGIAWRK-35	28	1.83	0.3
59-FVGTRSR K QCRERFI-73	66	1.43	0.3
96-GYYHHGS K WSEISKK-110	103	0.47	0.3
103-KWSEISK K LSGRTAR-117	110	1.42	0.3

Tabla elaborada con datos obtenidos de BDM-PUB (2023). La lisina (K) se muestra en color rojo ya que es el residuo que se ubiquitina.

Por último, se determinó si la secuencia de aminoácidos de EhMyb10 posee sitios susceptibles a SUMOilación (Small Ubiquitin-like Modifier). En el caso de EhMyb10, se identificaron cinco sitios que podrían ser susceptibles a SUMOilación, pero con una baja probabilidad (**Fig. 17, Tabla 9**). Esto indica que estos sitios tienen características estructurales o secuenciales que los hacen potencialmente reconocibles por el sistema de SUMOilación.

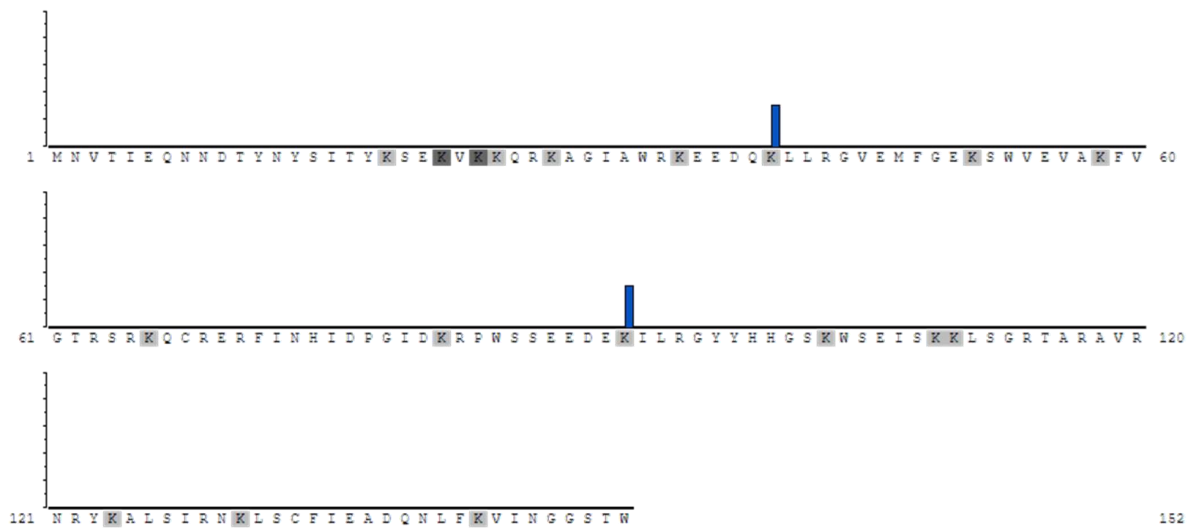


Figura 17. Residuos susceptibles a SUMOilación. Se utilizaron los predictores JASSA V4 para identificar SUMOilaciones en EhMyb10. Los rectángulos verticales indican los sitios susceptibles a SUMO en la secuencia de aminoácidos de EhMyb10. El color Azul indica una baja probabilidad y el Rojo indica una alta probabilidad de los sitios con posible susceptibilidad a SUMO utilizando un valor de corte de $PS_{max} = 38.183$ | Corte=0.269.

Tabla 9. Sitios predichos de SUMOilación en EhMyb10

Motivo	Posición	Score
29-AGIAWR K EEDQKLL-42	35	0.44
77-IDPGID K RPWSSEE-90	83	0.39
22-KV K KQR K AGIAWRK-35	28	0.27

Tabla elaborada con datos obtenidos de SUMOplot™ Analysis Program (2023). La lisina (K) se muestra en color en rojo ya que es el residuo que se SUMOíla.

9.3 Análisis de las regiones desordenadas de EhMyb10

Las regiones desordenadas (IDPs, del inglés Intrinsically Disordered Protein Regions) son segmentos en la secuencia de la proteína que no tienen una estructura tridimensional definida. Estas regiones carecen de una conformación estable y adoptan múltiples conformaciones en función de su entorno o de las interacciones que establecen con otras moléculas. Debido al papel que desempeña EhMyb10 como FT, se evaluó la presencia de regiones desordenadas a través de los predictores PrDOS y IUPred2A. EhMyb10 presenta regiones desordenadas en los extremos amino (1-32 aa) y carboxilo terminal (142-152 aa.) (**Fig. 18**). Estos resultados sugirieron que EhMyb10 tiene la capacidad de interactuar con otras proteínas y, por lo tanto, desempeñar diferentes funciones celulares.

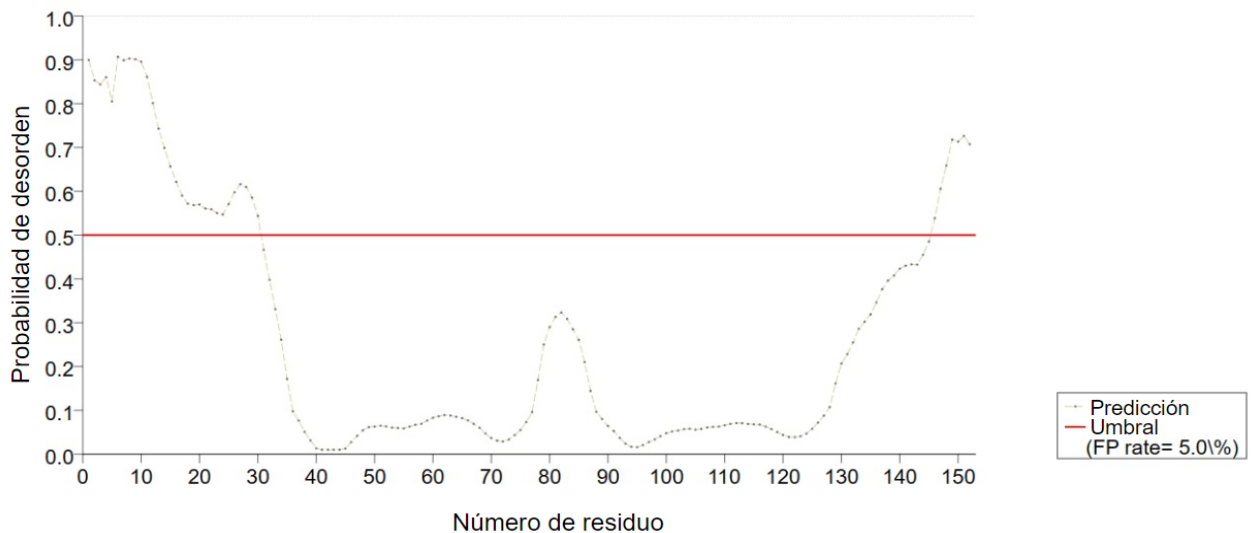


Figura 18. Predicción de regiones intrínsecamente desordenadas en EhMyb10. Predicción de las regiones desordenadas a través del Predictor PrDOS. La línea punteada representa la secuencia de EhMyb10, y la roja representa el umbral de predicción (mayor a 0.5).

Por último, utilizando los datos obtenidos y basándonos en la estructura tridimensional de la proteína, se realizó una representación esquemática de la ubicación potencial de las PTMs y las IDPs en el factor de transcripción EhMyb10 a través de las herramientas PyMOL y DOG 2.0 (**Fig. 19**).

9.4 Análisis del interactoma *in silico* de EhMyb10

Las interacciones de proteínas con el FT EhMyb10 se determinaron mediante un análisis de interactoma en el servidor de STRING. Se observó que EhMyb10 interactúa principalmente con proteínas relacionadas con la transcripción, lo cual concuerda con su función como FT (**Fig. 20**). De manera interesante, EhMyb10 también interactúa con proteínas involucradas en eventos de *splicing*, organización de la cromatina y ubiquitinación. Esta amplia gama de interacciones sugiere un posible papel de EhMyb10 en eventos cotranscripcionales, y además fortalece los hallazgos de residuos susceptibles a ubiquitinación y SUMOilación (**Tabla 8 y 9, Fig. 17**).

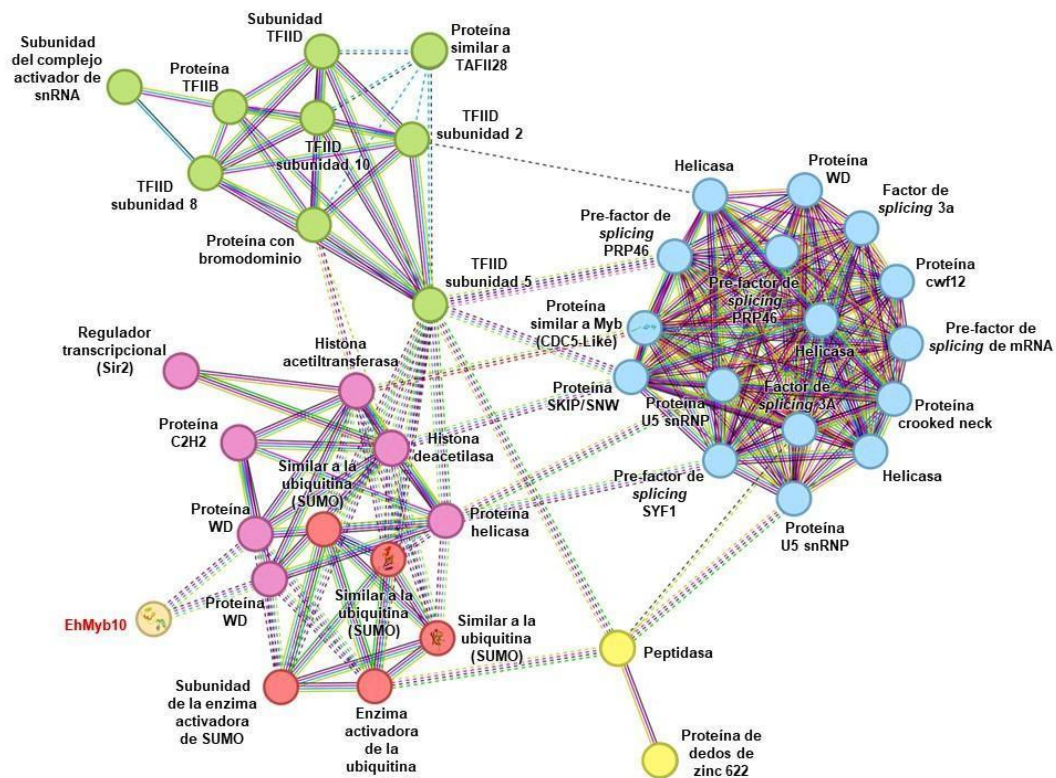


Figura 20. Interactoma *in silico* de EhMyb10. Interactoma obtenido mediante la herramienta STRING empleando un puntaje de interacción de confianza media (0.400) como valor de corte y un número máximo de no más de 10 interactuantes. Las esferas representan a las proteínas que interactúan con EhMyb10 (naranja) y que participan en diversos procesos: Transcripción (verde), *splicing* (azul), organización de la cromatina (morado), modificación postraduccional (rojo), y otras (amarillo). Número de nodos: 41, número de aristas: 265, valor p de enriquecimiento PPI $10e-16$.

Entre las proteínas que interactúan con EhMyb10 (**Tabla 10**), se encontró a la proteína de 2 repetidos EhCDC5-like. En otros organismos esta proteína funciona tanto en regulación transcripcional como en el proceso de *splicing*. Asimismo, EhMyb10 interactúa con otras proteínas que participan en el *splicing* por ejemplo: el pre-factor de *splicing* del mRNA SYF1, proteína que contiene dominio WD, helicasas, proteína que contiene el dominio helicasa DNA2/NAM7, proteína de control del ciclo celular cwf12, proteína U5 snRNP específica, factor de *splicing* 3a subunidad 2, proteína de dominio SKIP/SNW, factor de *splicing* 3A subunidad 3, pre-factor de *splicing* del mRNA PRP46, proteína *crooked neck*, y el pre-factor de *splicing* del mRNA, que están involucradas en el procesamiento y el ensamblaje del *spliceosoma* y en el reconocimiento del sitio de *splicing* en el mRNA.

Tabla 10. Anotación y función molecular de las proteínas que se predice que interactúan con EhMyb10

Gen ID	Anotación
<i>Splicing</i>	
EHI_073300	Pre-factor de <i>splicing</i> del mRNA SYF1
EHI_126260	Proteína que contiene dominio WD
EHI_122790	Helicasa, putativa
EHI_200000	Proteína que contiene el dominio helicasa DNA2/NAM7
EHI_187260	Proteína de control del ciclo celular cwf12, putativo
EHI_021380	Proteína U5 snRNP específica
EHI_045170	Proteína U5 snRNP específica, putativa
EHI_148930	Helicasa, putativa
EHI_117940	Factor de <i>splicing</i> 3a subunidad 2, putativo
EHI_000550	Proteína que contiene un dominio de unión al DNA similar a Myb
EHI_199600	Proteína de dominio SKIP/SNW
EHI_038600	Factor de <i>splicing</i> 3A subunidad 3, putativo
EHI_027560	Pre-factor de <i>splicing</i> del mRNA PRP46, putativo
EHI_175230	Proteína <i>crooked neck</i>
EHI_013330	Pre-factor de <i>splicing</i> del mRNA, putativo
EHI_153530	Pre-factor de <i>splicing</i> del mRNA PRP46, putativo

Transcripción

EHI_050980	Proteína que contiene un dominio tipo TFIIB
EHI_108470	Factor de iniciación de la transcripción TFIID subunidad 2
EHI_184680	Factor de iniciación de la transcripción TFIID subunidad 8
EHI_023040	Proteína que contiene el dominio de proteína similar a TAFII28
EHI_093770	Proteína de bromodominio, putativo
EHI_105280	Factor de iniciación de la transcripción TFIID subunidad 10
EHI_137090	Subunidad TFIID, putativo
EHI_198960	Factor de iniciación de la transcripción TFIID subunidad 5, putativo
EHI_197110	Subunidad del complejo proteico activador de snRNA, putativo

Remodelación de la cromatina

EHI_067880	Proteína que contiene un dominio de tipo C2H2
EHI_151300	Regulador transcripcional de la familia Sir2, putativo
EHI_142000	Histona acetiltransferasa
EHI_020290	Proteína que contiene un dominio WD
EHI_192950	Proteína que contiene un dominio WD
EHI_119320	Histona desacetilasa
EHI_031370	Proteína que contiene el dominio helicasa C-terminal

Modificación postraducciona

EHI_170060	Proteína similar a la ubiquitina
EHI_147470	Proteína de la familia de enzimas conjugadoras de ubiquitina
EHI_004850	Enzima activadora de la ubiquitina, putativa
EHI_035540	Subunidad de la enzima activadora de SUMO
EHI_178500	Proteína de la familia de enzimas conjugadoras de ubiquitina

Biogénesis de la subunidad grande ribosómica

EHI_154230	Proteína con dedos de zinc 622, putativo
------------	--

Proteólisis

EHI_175020	Peptidasa, putativo (proteína 2G4 asociada a la proliferación, putativa [^])
------------	--

[^] De acuerdo a los ortólogos de la misma familia en la base de datos AmoebaDB. Tabla elaborada con datos obtenidos del servidor STRING, AmoebaDB y PantherDB (2023).

Por otro lado, EhMyb10 interactúa con proteínas de la remodelación de la cromatina, tales como: la proteína con dominio C2H2, histona acetiltransferasa, Sir2, histona desacetilasa, proteínas con dominio WD y proteína con dominio helicasa C-terminal.

Estos resultados sugirieron la participación de EhMyb10 en la regulación cotranscripcional y en la remodelación de la cromatina. Asimismo, que EhMyb10 puede estar sujeto a regulación a través de procesos de ubiquitinación y SUMOilación.

9.5 Análisis de GO de las proteínas que interactúan con EhMyb10

En el análisis de GO del interactoma de EhMyb10, se identificó que el 15% de las proteínas están involucradas en la remodelación de la cromatina, lo que implica su contribución en la modificación de la estructura y accesibilidad del DNA. El 42.5% participa en *splicing* (GO:1990935), mientras que el 25% está relacionado con la transcripción (GO:0016563), lo que sugiere la participación de EhMyb10 en eventos cotranscripcionales. Por último, el 12.5% de las proteínas tienen una función de unión a proteínas (GO:0005515), lo que sugiere su interacción con otras moléculas para desempeñar roles específicos en procesos biológicos (**Fig. 21**)

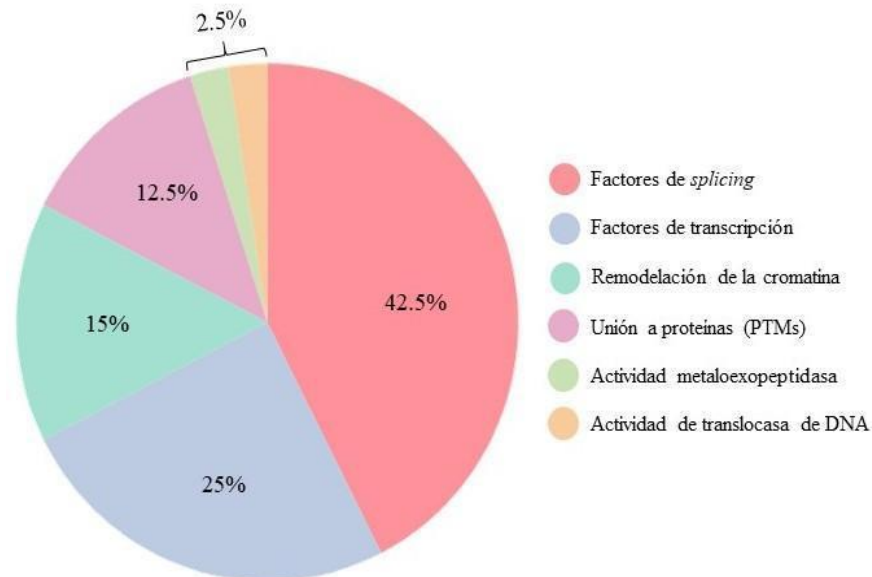


Figura 21. Análisis de GO por función molecular del interactoma de EhMyb10. Gráfica de pastel con la distribución de las proteínas asociadas a EhMyb10 clasificadas por función molecular a través del programa Panther DB.

Cuando las proteínas se analizan por proceso biológico, un 34.2% de estas proteínas desempeñan un papel principal en el procesamiento del mRNA (GO:0006397); un 10.5% está implicado en la iniciación de la transcripción (GO:0006352), mientras que un 7.9% se relaciona con la desacetilación de histonas (GO:0016575). Además, se encontró que un 5.3% de las proteínas está involucrado en la generación del *spliceosoma* (GO:0000387), el ensamblaje del complejo de preiniciación, la regulación de la transcripción, el ciclo de la ubiquitina (GO:0006512) y la organización de la cromatina (GO:0006325). Por lo tanto, el GO por proceso biológico del interactoma de EhMyb10, reveló una amplia diversidad de funciones moleculares desempeñadas (**Fig. 22**).

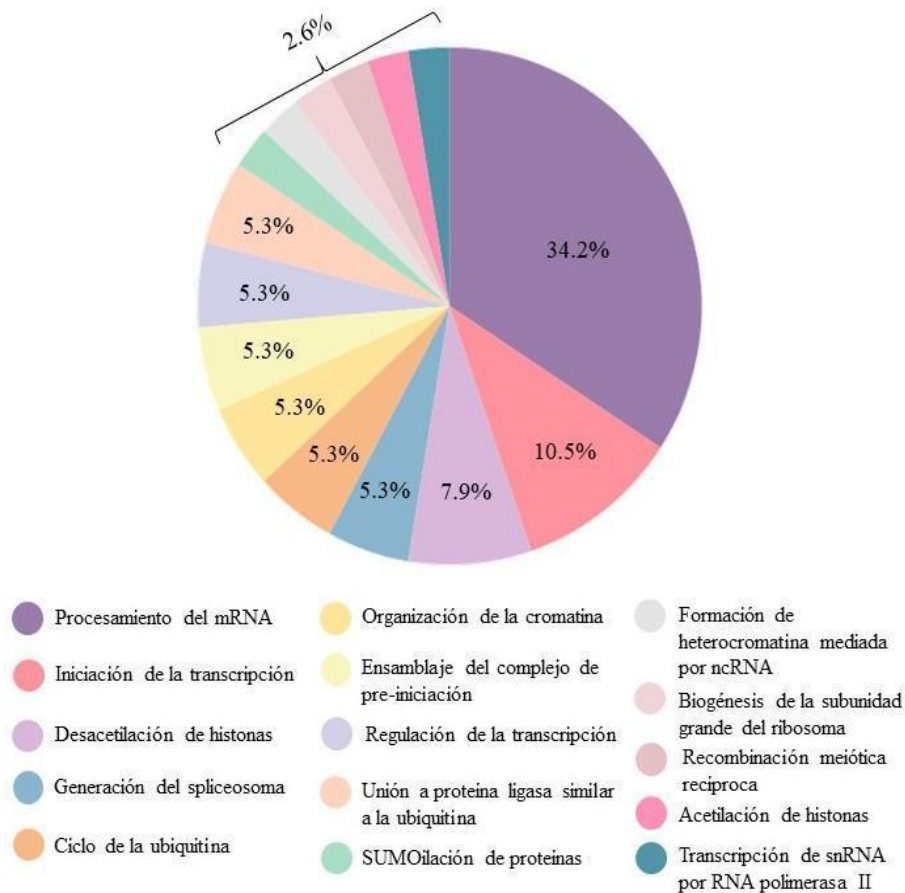


Figura 22. Análisis de GO por procesos biológicos del interactoma de EhMyb10. Gráfica de pastel con la distribución de las proteínas asociadas a EhMyb10 clasificadas por procesos biológicos a través del programa PantherDB.

9.6 Análisis de expresión de EhMyb10 y su interactoma en condiciones de Colonización/Invasión intestinal y Virulencia.

Para analizar los cambios de expresión de los genes que codifican para las proteínas que se asocian con EhMyb10, se utilizaron los transcriptomas reportados en la base de datos AmoebaDB por Gilchrist y cols 2006. Durante la colonización/invasión intestinal, se observaron genes que aumentan su expresión en comparación con la condición basal (EHI_038600, EHI_050980, EHI_073300, EHI_091460, EHI_093770, EHI_151300, EHI_184680, EHI_199600). Estos genes presentan un patrón de expresión similar a *ehmyb10* y sus productos proteicos tienen funciones relacionadas a la transcripción, *splicing*, remodelación de la cromatina, y ubiquitinación. En la **Tabla 8** se desglosa la anotación y posible función de estos genes. Por otro lado, se observó un descenso de la expresión de *ehmyb10* en el día 29 de la colonización intestinal (**Fig. 23**). Las proteínas con las que interactúa EhMyb10 están involucradas en eventos cotranscripcionales, por lo que es plausible inferir que EhMyb10 podría estar también involucrada en estos procesos, influyendo en la formación de complejos multiproteicos en el sitio de inicio de la transcripción, afectar la unión del mRNA a las proteínas del *splicing*, o estar implicada en la remodelación de la cromatina.

Asimismo, cuando se analizan los niveles de expresión en diferentes condiciones de virulencia, se observó que *ehmyb10* presenta una expresión alta en cepas virulentas en explante de colon por Weber y cols 2016. Al igual que en la colonización intestinal, se observaron genes con patrones de expresión similares a EhMyb10 (EHI_199600, EHI_020290, EHI_153530, EHI_021380, EHI_184680, EHI_013330, EHI_148930, EHI_093770, EHI_147470, EHI_091460, EHI_137090, EHI_067880, EHI_119320, EHI_000550 y EHI_122790). En la **Tabla 8** se desglosa la anotación y posible función de estos genes. Las proteínas que son codificadas por estos genes están relacionadas a la transcripción, *splicing*, remodelación de la cromatina, y ubiquitinación; por tanto, se puede inferir que participan en la regulación de genes específicos relacionados con la virulencia y la colonización intestinal. Esta regulación coordinada podría ser esencial para la adaptación y virulencia de la amiba en el ambiente intestinal (**Fig. 24**).

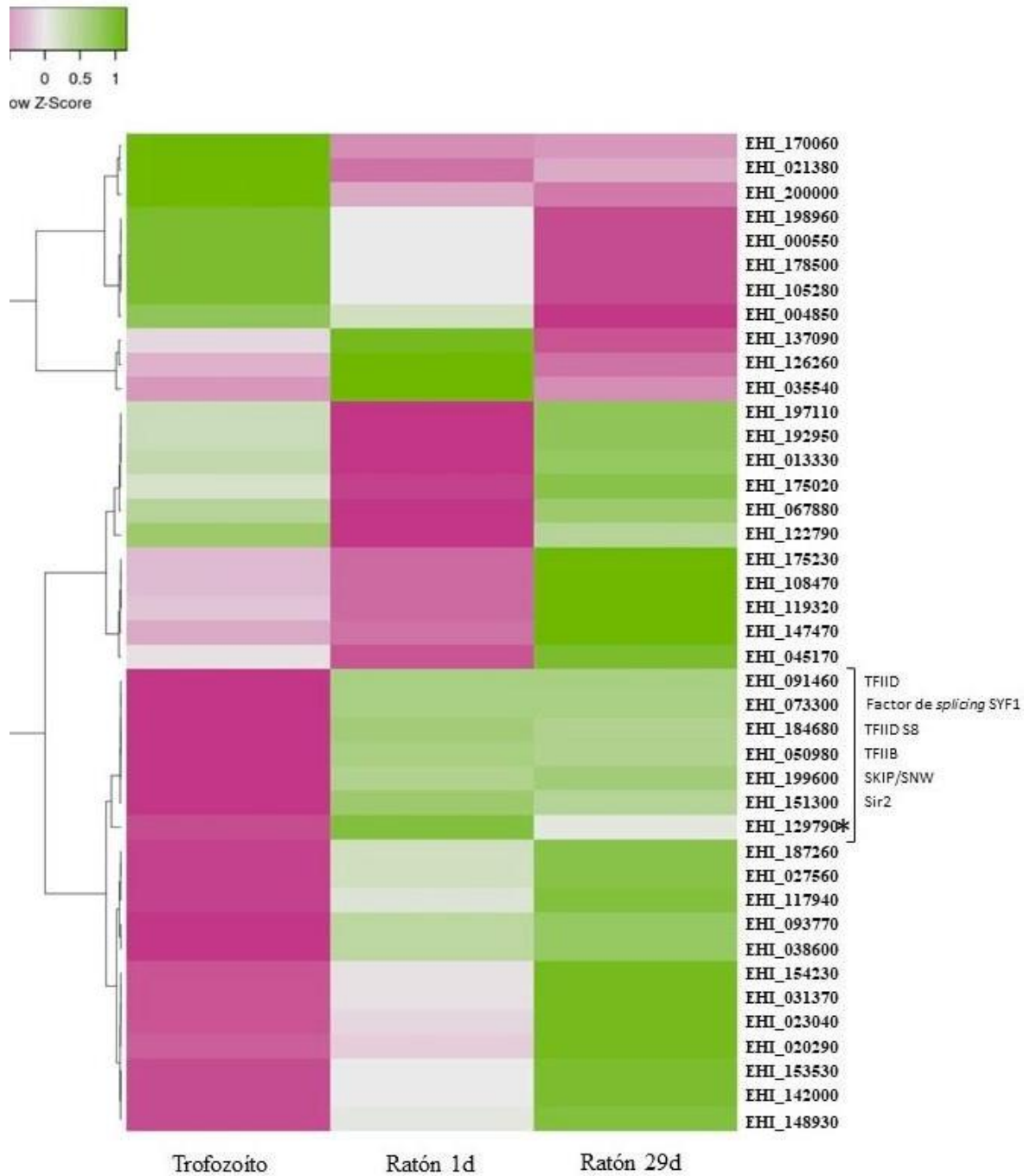


Figura 23. Cambios en la expresión de los genes que codifican para las proteínas que interactúan con EhMyb10 en un modelo murino. Mapa de calor que representa los niveles de expresión obtenidos de los análisis de transcripción reportados en AmoebaDB durante la interacción de trofozoitos con un modelo murino (Gilchrist, *et al.*, 2006), y construido mediante la herramienta Heatmapper (<http://www.heatmapper.ca/expression/>). EhMyb10 es señalado con un asterisco, mientras que los corchetes indican las proteínas que presentan un patrón de expresión similar a EhMyb10.

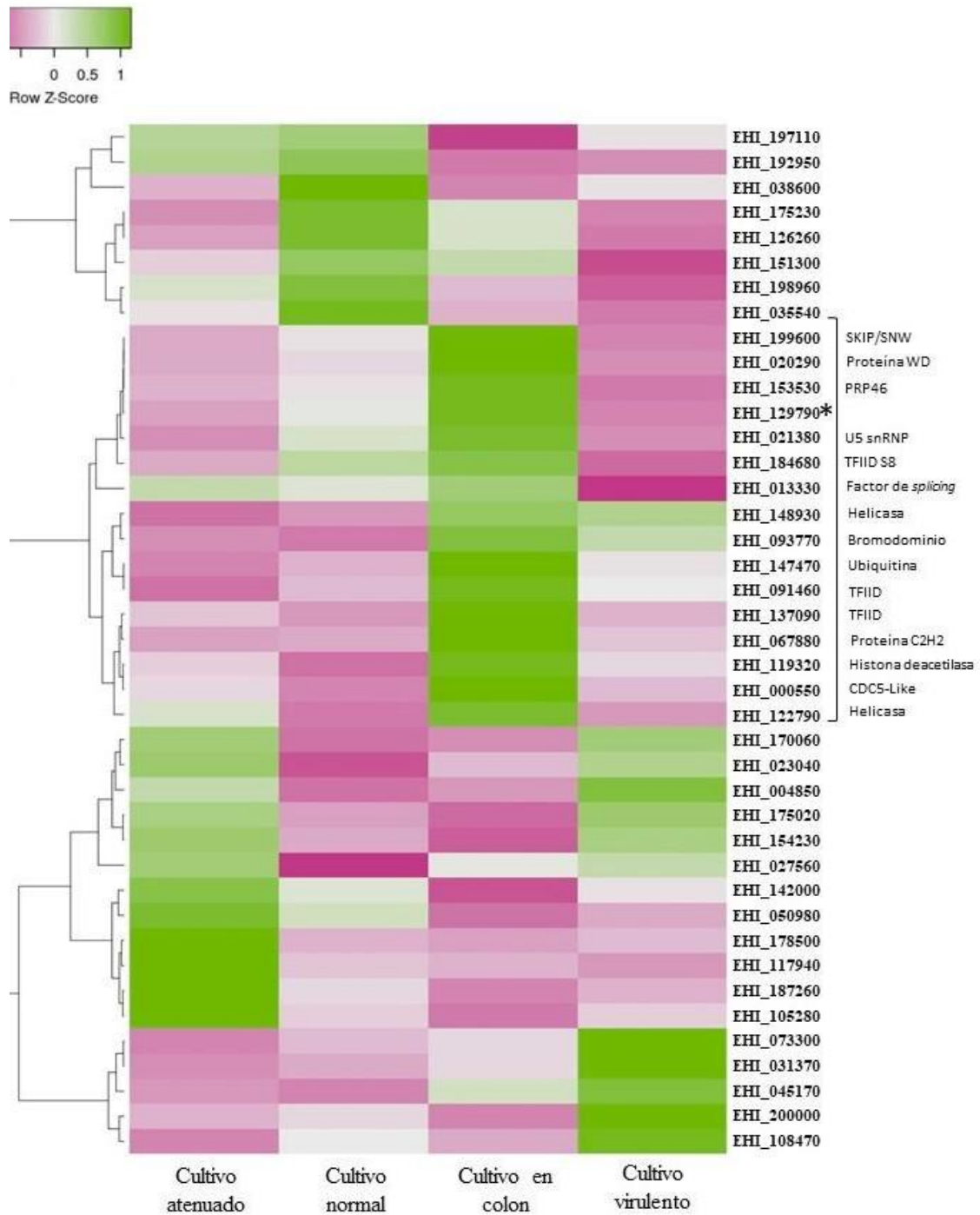


Figura 24. Cambios en la expresión de los genes que codifican para las proteínas que interactúan con EhMyb10 en diferentes condiciones de cultivo. Mapa de calor que representa los niveles de expresión obtenidos de los análisis de transcripción reportados en AmoebaDB en diferentes condiciones relacionadas con virulencia (Weber *et al.*, 2016), y construido mediante la herramienta Heatmapper (<http://www.heatmapper.ca/expression/>). EhMyb10 es señalado con un asterisco, mientras que los corchetes indican las proteínas que presentan un patrón de expresión similar a EhMyb10.

9.7 Expresión de EhMyb10r

9.7.1 Restricción y PCR del gen *ehmyb10*

En el laboratorio de la Dra. Elisa Azuara, se contaba con un plásmido denominado pGEX-EhMyb10, el cual tiene clonada la secuencia codificante del gen *ehmyb10*, y permite expresar de manera recombinante a la proteína EhMyb10, fusionada a un tag de GST. Para ello, primeramente, se verificó que el plásmido contuviera al gen *ehmyb10*, por lo que se analizó empleando enzimas de restricción y un ensayo de PCR. Las enzimas de restricción *Bam*HI y *Xho*I, permitieron la liberación de un inserto de 459 pb, correspondiente al tamaño del gen completo, indicando que el plásmido contaba con el gen *ehmyb10* (**Fig. 25A**). Posteriormente, se corroboró la amplificación del gen completo mediante oligonucleótidos específicos y se obtuvo un amplicón de 459 pb. En este ensayo se incluyó, como control, positivo al plásmido pKT3M-EhMyb10, que contiene al mismo gen, pero en otro vector. En conjunto, estos resultados respaldan la idoneidad del plásmido para la expresión de la proteína EhMyb10r en los siguientes experimentos (**Fig. 25B**).

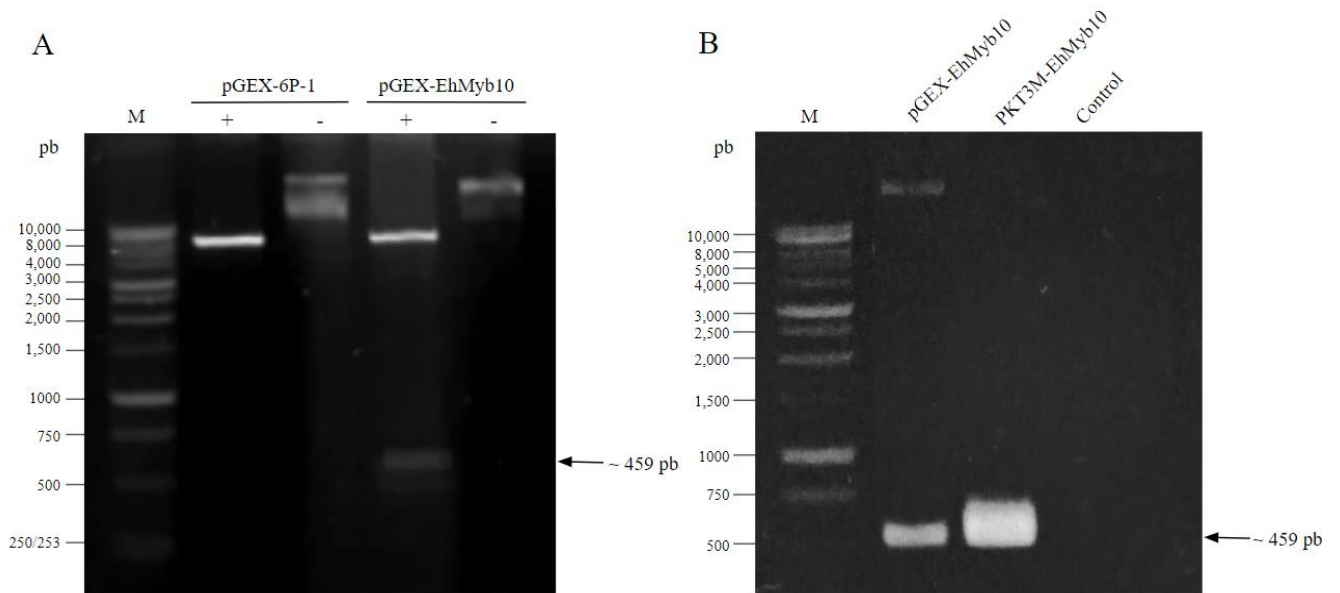


Figura 25. Verificación del plásmido pGEX-EhMyb10. **A)** Digestión de los plásmidos pGEX-6P-1 y pGEX-EhMyb10 con las enzimas de restricción *Bam*HI y *Xho*I. “+”, restricción; “-”, sin restricción. **B)** Ensayo de PCR de los plásmidos pGEX-EhMyb10 y pKT3M-EhMyb10. Carril 1, pGEX-EhMyb10; Carril 2, pKT3M-EhMyb10 (control +); Carril 3, control sin DNA plasmídico. M; marcador de 1kb. Electroforesis en gel de agarosa al 1%. Las flechas indican el tamaño en pb del gen *ehmyb10*.

9.7.2 Inducción de la proteína EhMyb10r en células de *E. coli* BL21

Después de verificar la integridad del plásmido pGEX-EhMyb10, se procedió a inducir la expresión de la proteína EhMyb10r en células *E. coli* BL21 por medio de IPTG. Al analizar el patrón de proteínas en las bacterias transformadas con el plásmido pGEX-EhMyb10, se observó la presencia de una banda de 43 kDa en la muestra inducida, que coincide con el peso teórico esperado para la proteína fusionada con GST (18 + 25 kDa) (**Fig. 26A**). La identidad de la proteína, se confirmó mediante un ensayo de WB, utilizando un anticuerpo α -GST específico, el cual reconoció también a la banda de 43 kDa (**Fig. 26B**). En las bacterias transformadas con el vector vacío (pGEX-6P-1), se observó la presencia de una banda de 25 kDa, correspondiente a la etiqueta de GST, la cual también fue reconocida con el anticuerpo α -GST, y se utilizó como control en los diferentes ensayos. En conjunto, estos resultados muestran la estandarización de las condiciones para la expresión de las proteínas GST y EhMyb10r.

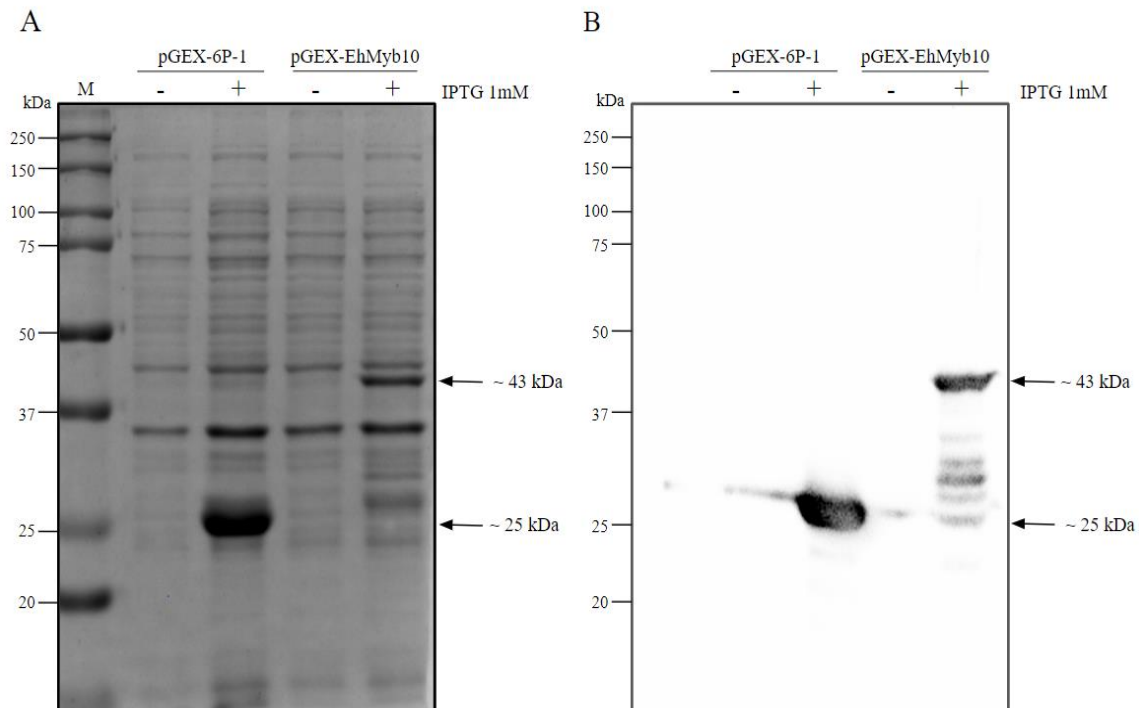


Figura 26. Expresión de la proteína EhMyb10r. A) Electroforesis en gel de poliacrilamida-SDS al 10% teñido con Coomassie® Brilliant blue G 250 de los extractos de *E. coli* de la cepa BL21, transformada con los plásmidos e inducida. El signo (-) representa los cultivos bacterianos sin inducir; el signo (+) representan los cultivos después de 24h de inducción con IPTG 1 mM. B) WB con el anticuerpo α -GST. M, representa el marcador molecular. Se utilizó como control al vector vacío que solo expresa la etiqueta de GST de ~25 kDa. Las fechas indican el PM de las proteínas inducidas.

Una parte esencial para llevar a cabo los ensayos de Pull-Down, es contar con la proteína EhMyb10r en forma soluble. Por tanto, se evaluó la solubilidad de la proteína obtenida durante el proceso de inducción y purificación. El análisis de las fracciones solubles e insolubles, reveló que la proteína se encontraba mayoritariamente en la fracción insoluble (**Fig. 27A**). No obstante, al utilizar un tampón que contenía tritón al 0.5%, se solubilizó una parte de la proteína, lo cual se evidenció mediante WB al emplear el anticuerpo α -GST como bandas de 25 y 43 kDa, que contienen en las fracciones solubles las proteínas GST y EhMyb10r, respectivamente (**Fig. 27B**).

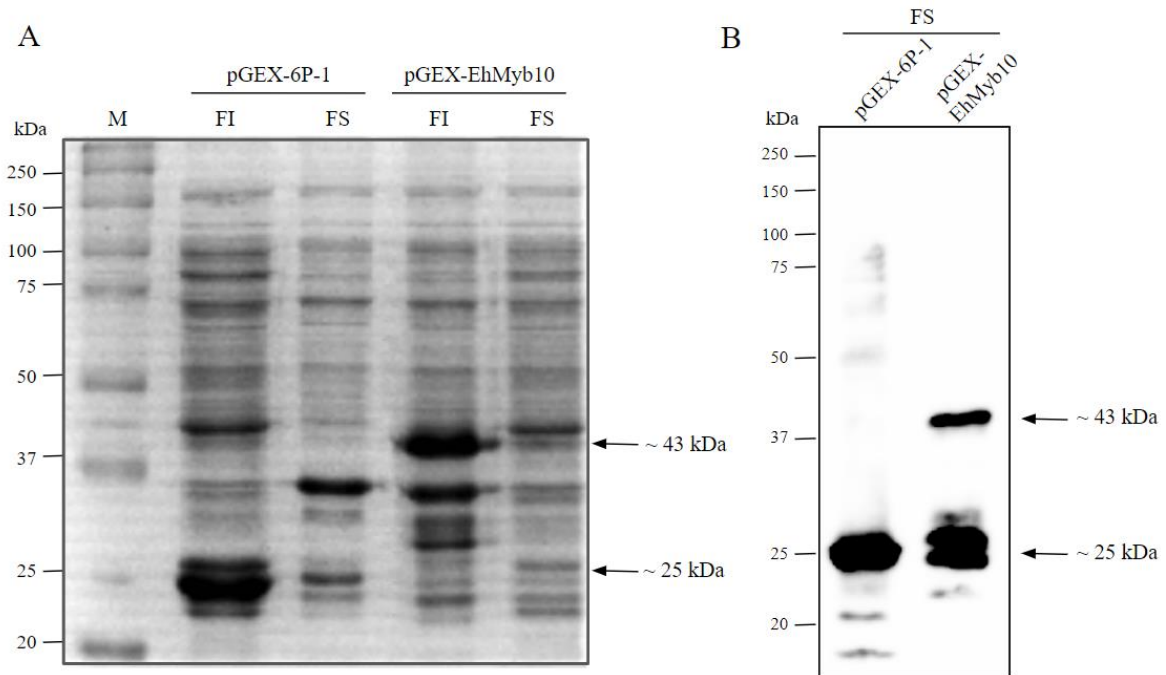


Figura 27. Análisis de la solubilidad de EhMyb10r obtenida de lisados bacterianos. **A)** Electroforesis en geles de poliacrilamida-SDS al 10% teñido con Coomassie® Brilliant blue G 250 de las fracciones solubles e insolubles. “FI”, representa la fracción insoluble, y “FS” la fracción soluble. **B)** WB de la fracción soluble con el anticuerpo α -GST. Se utilizó como control al vector vacío que solo expresa la etiqueta de GST de 25 kDa. Las flechas indican el PM de las proteínas inducidas. M, corresponde al marcador molecular.

9.7.3 Purificación por cromatografía de afinidad de EhMyb10r

Una vez corroborado que EhMyb10 se encontraba soluble, se purificó a través de cromatografía de afinidad utilizando una resina de sepharosa-glutation. La proteína EhMbyb10r se observó en el eluido 1, no así en los lavados de la resina (**Fig. 28A**), mientras que la proteína GST se observó en los tres eluidos (**Fig. 28B**). Por lo tanto, la cromatografía de afinidad permitió obtener a las proteínas recombinantes con suficiente pureza, ya que no se observaron bandas adicionales. Los resultados obtenidos permitirán la caracterización y el análisis del interactoma, utilizando la proteína EhMyb10r y análisis de Pull-Down.

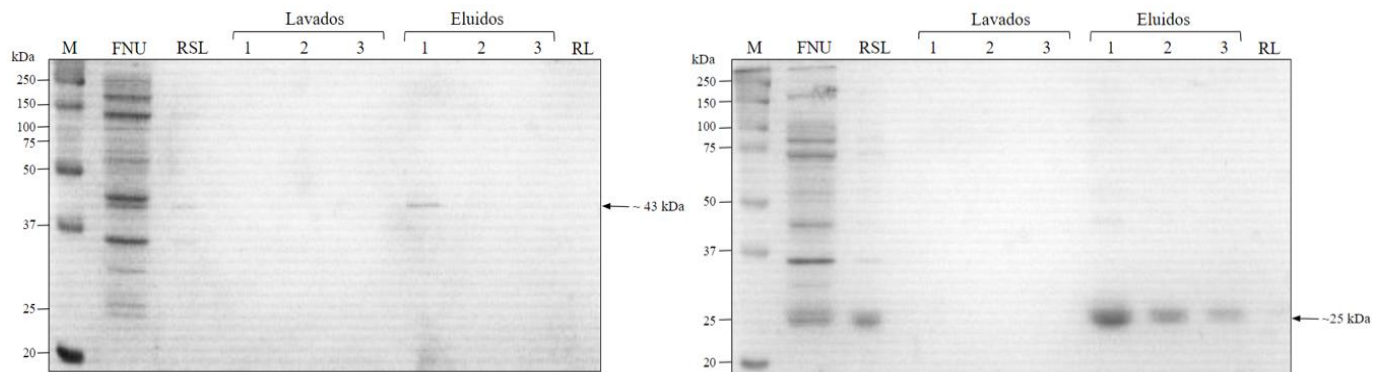


Figura 28. Purificación de EhMyb10r y GST. Electroforesis SDS-PAGE al 10% teñida con Coomassie® Brilliant blue G 250 de las diferentes fracciones del proceso de purificación de **A)** EhMyb10r y **B)** GST. Carril FNU, corresponde a la fracción no unida a la resina; Carril RSL, resina antes de ser lavada; Lavados 1-3, corresponde a una muestra de cada lavado; Eluidos 1-3, corresponde a una muestra de cada eluido; Carril RL, resina después de la elución; M, corresponde al marcador molecular. Las flechas indican los pesos moleculares de las proteínas purificadas.

9.8 Fraccionamiento subcelular de proteínas de *E. histolytica*

Para poder realizar la interacción de EhMyb10r con las proteínas nucleares de *E. histolytica*, se realizó un fraccionamiento subcelular de acuerdo a Schreiber et al. (1989). En el fraccionamiento subcelular, se obtuvieron 3 fracciones distintas: fracción citoplasmática, fracción nuclear soluble, y nuclear insoluble. Una parte del lisado se guardó como extracto total. Cada fracción representa un subconjunto de proteínas enriquecidas. Es importante mencionar que, aunque la técnica de fraccionamiento subcelular enriquece los diferentes extractos obtenidos, algunas proteínas se localizan en más de una fracción.

Las fracciones obtenidas se analizaron mediante un gel SDS-PAGE, en el cual se observaron bandas diferenciales para las tres fracciones subcelulares (**Fig. 29A**). Asimismo, se realizó un análisis de WB para confirmar el enriquecimiento de cada fracción. Para ello, se utilizaron anticuerpos específicos para detectar proteínas de cada una de las fracciones como se describió previamente.

El anticuerpo α -histona H4K20me3, reconoció a la histona H4 en la fracción nuclear soluble con un PM alrededor de los 48 kDa, así como otra banda de ~ 27 kDa. También se detectó una banda menos enriquecida en la fracción nuclear insoluble de ~ 35 kDa (**Fig. 29B**). Se ha reportado, en la base de datos (AmoebaDB), que la histona H4 de *E. histolytica* tiene un PM de 12.9 kDa, sin embargo, las histonas experimentan diversas PTMs y forman complejos de multímeros, o agregados lo que puede modificar su PM detectado en un WB.

Por otro lado, utilizamos un anticuerpo α -CP112 que reconoce a una proteasa citoplasmática y otra de superficie. Se observó inmunodetección solo en la fracción total, al detectar dos bandas de ~ 38 y 35 kDa y en la fracción citoplasmática una banda de ~ 25 kDa (**Fig. 29C**). Estas diferencias de pesos moleculares pueden deberse a que la CP112 se sintetiza como zimógeno en forma inactiva, y requiere de la escisión proteolítica para activarse. Por último, utilizamos el anticuerpo α -actina como control de carga, ya que es una proteína presente en una variedad de compartimentos subcelulares. Observamos diferentes bandas en las tres fracciones celulares, con PM aproximados de 125, 55 y 38 kDa (**Fig. 29D**).

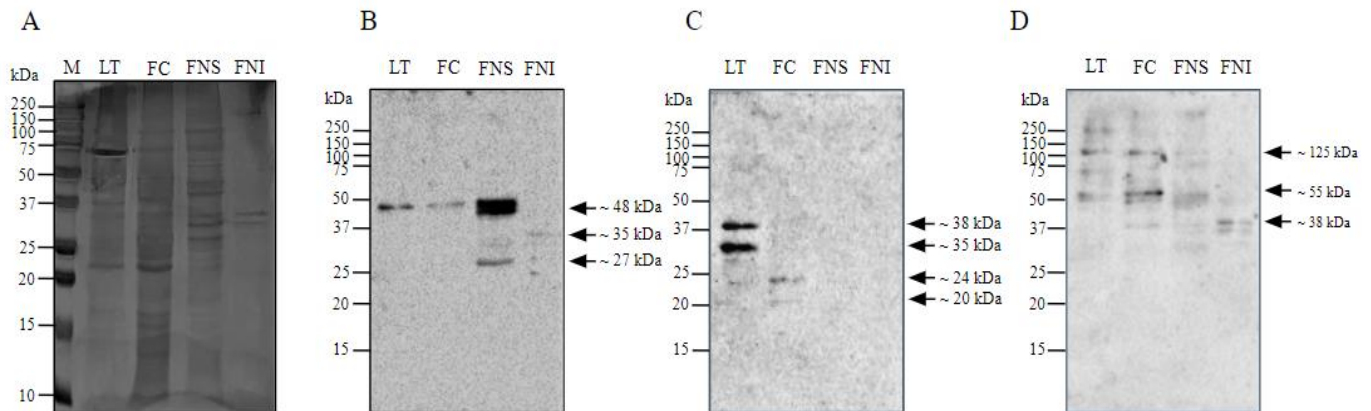


Figura 29. Fraccionamiento subcelular de proteínas de *E. histolytica*. A) SDS-PAGE en gel al 12% teñido con Coomassie® Brilliant blue G 250. B) Inmunodetección empleando los anticuerpos α -histona H4K20me3 (1:100), C) α -CP112 (1:2500). D) α -actina (1:1000). LT, lisado total; FC, fracción citoplasmática; FNS, fracción nuclear soluble; FNI, fracción nuclear insoluble. M corresponde al marcador molecular.

Estos reconocimientos se deben a la presencia de diferentes isoformas producto de múltiples genes codificantes para actina en este parásito. Debido al enriquecimiento de proteínas nucleares solubles, empleamos esta fracción celular en los ensayos de Pull-Down.

9.9 Ensayo de Pull-Down con la proteína EhMyb10r

Para identificar a las proteínas que interactúan con la proteína EhMyb10, se realizó un ensayo de Pull-Down, utilizando a EhMyb10r y extractos nucleares solubles de *E. histolytica*. Para este ensayo, las proteínas purificadas se acoplaron a la resina para interaccionar los extractos nucleares. Durante el proceso, se recolectaron varias fracciones, las cuales fueron analizadas por SDS-PAGE, y teñidas con Coomassie® Brilliant blue G 250 (**Fig. 30A y B**). En las muestras eluidas de la interacción de EhMyb10r, se observaron bandas diferenciales, cuyos pesos moleculares oscilaban entre 16-60 kDa (**Fig. 30B**).

Para aumentar la sensibilidad y visualizar proteínas en baja cantidad en el gel, se realizó una tinción con plata que evidenció bandas de 18, 23 y 38 kDa en la condición de EhMyb10 (**Fig. 30C**). Este experimento, se realizó por duplicado mostrando patrones reproducibles, tanto para EhMyb10 como para GST. Por lo tanto, los resultados obtenidos sugirieron una interacción específica entre EhMyb10 y proteínas nucleares de *E. histolytica*.

9.10 Identificación de las proteínas que interactúan con EhMyb10r por espectrometría de masas (IMS-DIA-MS)

Las diluciones obtenidas de los ensayos de Pull-Down, se concentraron mediante electroforesis SDS-PAGE en gel al 12%, y se mandaron analizar mediante IMS-DIA-MS en la Unidad de Genómica, Proteómica y Metabolómica LaNSE-Cinvestav. En estos análisis, se identificaron 122 proteínas en las muestras del ensayo de Pull-Down del control de GST, y 192 proteínas en las muestras de EhMyb10. Interesantemente, 93 proteínas fueron específicas de EhMyb10 (**Fig. 31**).

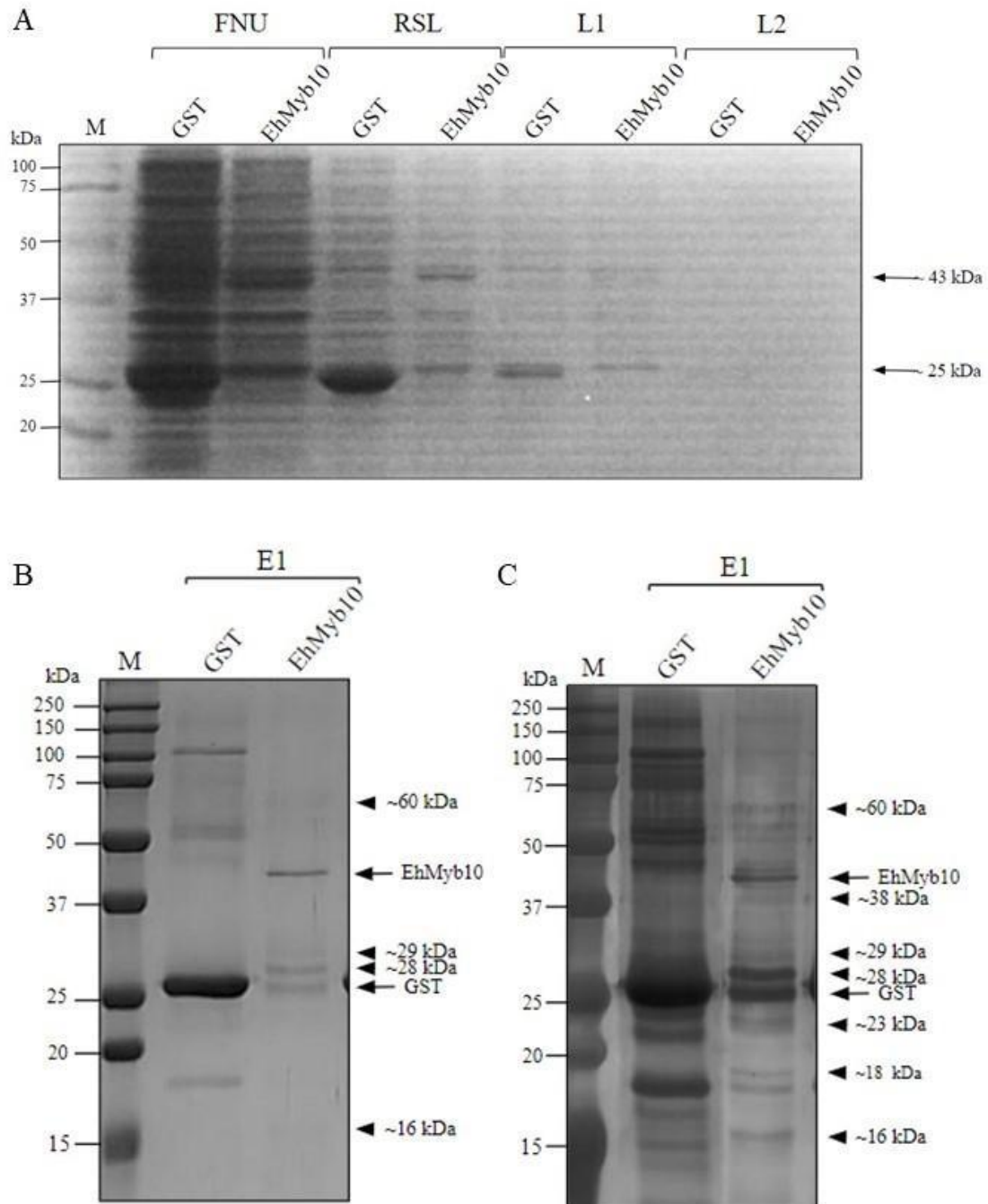


Figura 30. Ensayo de Pull-Down. Corrimiento electroforético en SDS-PAGE en gel al 12% de **A)** Fracciones obtenidas del ensayo de Pull-Down de EhMyb10 y GST. FNU, fracción no unida a la resina; RSL, resina sin lavar; L1 y L2, lavados. **B)** Primera elución. Los gels se tiñeron con Coomassie® Brilliant blue G 250. **C)** Primera elución teñida con plata. Las bandas diferenciales se señalan con cabezas de flecha con el PM aproximado. M corresponde al marcador de PM.

Cabe destacar que, durante el análisis, se identificaron isoformas de proteínas codificadas por diferentes genes (ejemplo: EHI_182990 y EHI_188730 que codifican para histona H2B) y que comparten el mismo número de acceso en UniProt. Por otro lado, con relación a la cobertura de las secuencias de las proteínas identificadas, se obtuvo un rango de 2.13-42.71. Asimismo, el rango de péptidos identificados fue de 2-34, y el de péptidos digeridos de 7-123 (**Tabla 11**).

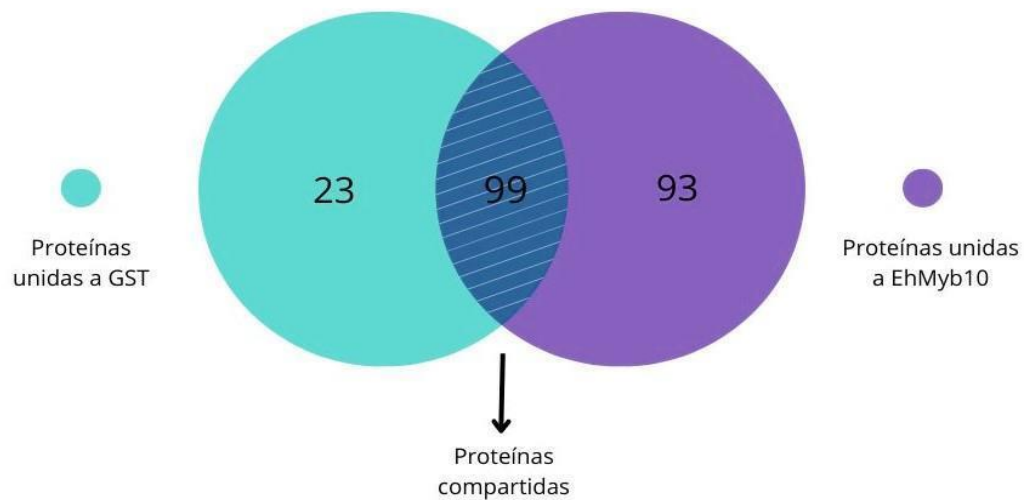


Figura 31. Diagrama de Venn de proteínas identificadas mediante espectrometría de masas (IMS-DIA-MS) de EhMyb10 y GST. El Diagrama de Venn representa las proteínas que coinciden en ambas muestras, y las proteínas específicas de GST y EhMyb10.

Tabla 11. Listado de proteínas que interaccionan con EhMyb10r identificadas por espectrometría de masas (IMS-DIA-MS)

Gene ID	Anotación	PM (kDa)	Péptidos identificados /digeridos	Cobertura de la secuencia (%)	Uniprot ID
EHI_062510	Proteína con dominio SUN [^]	40.8	10/29	8.86	C4M1N7
EHI_092620	Espermidina sintasa ^{^^}	47.1	11/41	6.02	C4LV22
EHI_118840	Fibrilarina, putativa	32.7	9/28	15.08	C4LYN2
EHI_122840	Subunidad Mu, isoforma b ⁺	40.5	10/27	2.29	A0A8U0WQ59
EHI_137090	Subunidad TFIID, putativa	20	3/21	10.34	C4M5Y7
EHI_083430	Subunidad del complejo 3 de proteína relacionada con el adaptador (AP-3), putativa	98.5	18/110	6.2	B1N424

EHI_004910	Subunidad 2 del complejo condensina [^]	53.9	10/30	3.28	C4M086
EHI_174240	Proteína de ribonucleoproteína nucleolar pequeña U3 IMP4, putativo ⁺	13.7	3/7	12.71	C4M271
EHI_167310	Grainin 2	24.2	8/23	21.6	A0A8U0WPU5
EHI_111720	Grainin 2	14.4	5/13	23.81	B1N3T3
EHI_012510	Proteína con dominio PH [^]	94.4	14/67	7.4	C4LTN8
EHI_125880	Proteína con dominio metalo-beta-lactamasa	29	10/17	8.56	C4LVX4
EHI_069080	Proteína con dominio SET y MYND ⁺	52.6	12/39	07.08	C4LWJ8
EHI_152970	Proteína que contiene el dominio PB1 [^]	25.3	4/18	24.09	C4LS90
EHI_195100	Proteína con dominio Brix ⁺	33.9	3/25	05.05	B1N3M3
EHI_038960	Proteína con dominio LisH [^]	61.8	10/46	15.81	C4LZR8
EHI_069560	Proteína con dominio MRH ^{^^}	23.2	8/18	11.59	C4LWP4
EHI_101240	Proteína con dominio nucleasa ⁺	96.3	26/99	18.86	C4M8B5
EHI_029090	Proteína DEK ⁺	28.9	8/21	15.51	C4M4W6
EHI_005657	Chaperona Clpb ⁺	97.8	23/100	8.2	B1N4V0
EHI_174650	Proteína que contiene dominio de dinamina ⁺	91.8	17/79	3.67	C4LZL0
EHI_112830	Proteína que contiene un dominio similar a la nucleoplasmina [^]	23.3	2/23	7.92	C4M9M2
EHI_129960	Metiltransferasa, proteína de la familia YaeB/AF_0241 ⁺	38	7/31	12.46	C4M5R2
EHI_005150	Proteína que contiene dominio Tudor ⁺	99.7	34/123	24.85	C4M0A8
EHI_153260	Proteína YIF1A ⁺	30.6	6/15	21.89	C4LSB8
EHI_191090	Proteína de diferenciación celular rcd1 ^{^^}	28.8	6/20	19.12	B1N534
EHI_194540	Péptido formador de poros amebaporo B, EH-APP B	10.8	4/10	42.71	Q24824
EHI_147740	Serina/treonina quinasa, putativa	13.7	4/14	10.53	B1N4D0*
EHI_147850	Serina/treonina quinasa, putativa	13.7	4/14	10.53	B1N4D0*

EHI_024450	Proteína tirosina fosfatasa, putativa	40.2	9/26	17.25	C4M0I7
EHI_148900	Proteína quinasa, putativa	55.5	13/42	10.44	C4LT57
EHI_015260	Proteína activadora de GTPasa Rab, putativa	35.2	3/21	19.87	C4M0W4
EHI_022860	Proteína que contiene dominio RhoGAP	70.4	10/58	3.54	C4LUJ9
EHI_174570	Reparación de ruptura de doble cadena de DNA Rad50 ATPasa ^^	28.4	5/53	6.44	C4LZK5
EHI_168400	tRNA (Guanina-N(1)-)metiltransferasa TRM10	32.4	5/23	15.87	C4M4Q0
EHI_152460	Histona-arginina N-metiltransferasa, putativa +	38.1	7/30	2.13	C4LST3
EHI_182990	Histona H2B	14.5	8/15	36.84	C4LV75*
EHI_188730	Histona H2B	14.5	8/15	36.84	C4LV75*
EHI_188960	Histona H2A	17.6	4/21	6.29	C4M2W3
EHI_083270	Ubiquitina putativa	87	2/9	20.78	C4M760*
EHI_083410	Ubiquitina putativa	87	2/9	20.78	C4M760*
EHI_156660	Ubiquitina putativa	87	2/9	20.78	C4M760*
EHI_166800	Ubiquitina putativa	11.1	2/10	16.49	C4LY85
EHI_178340	Ubiquitina putativa	87	2/9	20.78	C4M760*
EHI_170060	Proteína similar a la ubiquitina	12.7	6/19	26.32	C4M1C8
EHI_017350	Proteína de choque térmico, putativo	97.8	24/97	8.2	B1N523
EHI_022620	Proteína de choque térmico, putativo	97.8	27/97	8.2	B1N2L8
EHI_028920	Proteína de choque térmico, putativo	97.8	24/97	8.2	B1N3R1
EHI_034710	Proteína de choque térmico, putativo	97.8	24/97	8.2	C4MB78
EHI_042860	Proteína de choque térmico, putativa	97.8	24/97	8.2	B1N3X2*
EHI_072140	Proteína de choque térmico, putativa	97.8	24/97	8.2	B1N3X2*
EHI_156560	Proteína de choque térmico, putativa	97.8	24/97	8.2	B1N3X2*
EHI_178230	Proteína de choque térmico 101, putativa	28.2	9/25	15.87	B1N5D4

EHI_147220	Proteína trucada de la familia hsp70	13.1	5/12	13.79	C4MBG4
EHI_148320	Proteína ribosómica 40S S3, putativa	27.2	8/30	29.51	C4LT02*
EHI_146340	Proteína ribosómica 40S S3, putativa	27.2	8/30	29.51	C4LT02*
EHI_003940	Proteína ribosómica 40S S4, putativa	36.9	6/35	19.33	C4LYF4*
EHI_118170	Proteína ribosómica 40S S4, putativa	35.9	6/36	19.81	B1N336
EHI_098840	Proteína ribosómica 40S S4, putativa	36.9	6/35	19.33	C4LYF4*
EHI_008210	Proteína ribosómica 40S S4, putativa	36.9	6/35	19.33	C4LYF4
EHI_006860	Proteína ribosómica 60S L5, putativa	32.6	6/30	24.48	C4LU56*
EHI_068660	Proteína ribosómica 60S L5, putativa	32.6	6/30	24.48	C4LU56*
EHI_045300	Proteína ribosómica 60S L6, putativa	23.0	8/19	22.55	C4LTE3
EHI_140550	Proteína ribosómica 60S L6, putativa	23.0	8/19	22.55	O15595
EHI_093580	Proteína ribosómica 60S L6, putativa	23.0	7/19	22.55	C4LYF9
EHI_067520	Proteína ribosómica 60S L6, putativa	23.0	8/19	22.55	C4M9F8
EHI_029530	Proteína ribosómica 60S L7a	32.3	8/24	7.69	B1N3R9*
EHI_169360	Proteína ribosómica 60S L7a	32.3	8/24	7.69	B1N3R9*
EHI_103310	Proteína ribosómica 60S L7a	32.4	8/24	7.69	C4LVH6
EHI_035090	Proteína ribosómica 60S L7a	32.3	8/24	7.69	B1N3R9*
EHI_150470	Proteína ribosómica 60S L2/L8, putativa	28.1	3/20	12.45	C4LWZ3*
EHI_181180	Proteína ribosómica 60S L2/L8, putativa	28.1	3/20	12.45	C4LWZ3*
EHI_127200	Proteína ribosómica 60S L2/L8, putativa	28.1	3/20	12.45	O15574
EHI_126140	Proteína ribosómica 60S L9, putativa	22.4	12/23	18.78	C4LVZ9
EHI_193080	Proteína ribosómica 60S L9, putativa	22.4	11/23	18.78	C4M9B0

EHI_038580	Proteína ribosómica 60S L11	19.7	7/18	32.76	C4LZP2*
EHI_199970	Proteína ribosómica 60S L11	19.8	7/18	32.76	C4LZG9*
EHI_040660	Proteína ribosómica 60S L11, putativa	19.7	7/18	32.76	C4LZP2*
EHI_124300	Proteína ribosómica 60S L11, putativa	19.8	7/18	32.76	C4LZG9*
EHI_012360	Proteína ribosómica 40S S14, putativa	15.8	6/15	20.55	B1N2I3*
EHI_074090	Proteína ribosómica 40S S14, putativa	15.8	6/15	20.55	B1N2I3*
EHI_050130	Proteína ribosómica 60S L14, putativa	13.5	7/12	11.57	C4LTQ0
EHI_184470	Proteína ribosómica 60S L14, putativa	15.2	6/13	10.37	C4M6S2
EHI_185580	Proteína ribosómica 60S L14, putativa	15.2	6/13	10.37	C4M8C9
EHI_141940	Proteína ribosómica 60S L22	16	4/11	18.25	C4LW40
EHI_146370	Proteína ribosómica 60S L27	15.8	4/15	21.58	C4M9C4
EHI_183480	Proteína ribosómica 60S L27	15.8	5/14	21.58	C4M727
EHI_068200	Proteína ribosómica 60S L31, putativa	17	3/15	9.33	C4LU13*
EHI_058080	Proteína ribosómica 60S L31, putativa	17	2/14	9.33	C4M8A9
EHI_140370	Proteína ribosómica 60S L31, putativa	17	3/15	9.33	C4LU13*
EHI_148850	Proteína ribosómica ácida 60S P0	34.5	10/29	25.95	B1N2Z3*
EHI_175460	Proteína ribosómica ácida 60S P0	34.5	10/29	25.95	B1N2Z3*
EHI_177630	Proteína ribosómica ácida 60S P0	34.5	10/29	25.95	B1N2Z3*
EHI_090400	Proteína ribosómica ácida 60S P0	34.5	10/29	25.95	B1N2Z3*

* indica las proteínas que comparten el mismo número de acceso de Uniprot.

+ De acuerdo a los ortólogos de la misma familia en la base de datos AmoebaDB

^ Indica la anotación obtenida por Uniprot

^^ De acuerdo a los ortólogos de la misma familia en la base de datos Uniprot

9.10.1 Análisis de GO de las proteínas identificadas por espectrometría de masas (IMS-DIA-MS)

Para determinar las funciones en las que participan las proteínas que interaccionan con EhMyb10, se realizó un análisis de GO basado en procesos biológicos, funciones moleculares y compartimento celular, empleando las bases de datos Panther, Uniprot y AmoebaDB.

Los análisis por función molecular mostraron que, de las 92 proteínas identificadas, un 35.8% desempeñan una función principal de unión a RNA. Estas proteínas, comprenden principalmente aquellas que se enlazan al RNA ribosómico (rRNA), así como en menor medida al RNA mensajero (mRNA) en la región 5'-UTR y al snoRNA, lo que sugiere su implicación en la regulación de la expresión génica a nivel de transcripción, *splicing* y traducción. Además, un 10.8% de las proteínas identificadas, forman parte de los complejos de pre-ensamblaje de los ribosomas. Otros roles funcionales destacables, son la actividad de unión a DNA (GO:0043566), a cromatina (GO:0003682) y como factor de transcripción (GO:0001144); relacionadas con PTMs como actividad de metiltransferasas (GO:0016766) y unión a ubiquitina (GO:0004842). Finalmente, el 9.7% de las proteínas, no tienen funciones moleculares asignadas, debido a que no están anotadas y no existen ortólogos en otros organismos eucariotas. Estos hallazgos, subrayan la complejidad y la versatilidad de las proteínas interactuantes de EhMyb10 en los trofozoitos, apuntando hacia un papel diverso, abarcando desde la regulación de la expresión génica hasta su implicación en el mantenimiento de la integridad nuclear, entre otras funciones (**Fig. 32**).

Cuando se realizó el análisis GO por proceso biológico, el 40.8% de estas proteínas participa activamente en la biogénesis de ribosomas (GO:004254), lo que implica su involucramiento tanto en la síntesis de proteínas como en la formación de los complejos de pre-ensamblaje de los ribosomas.

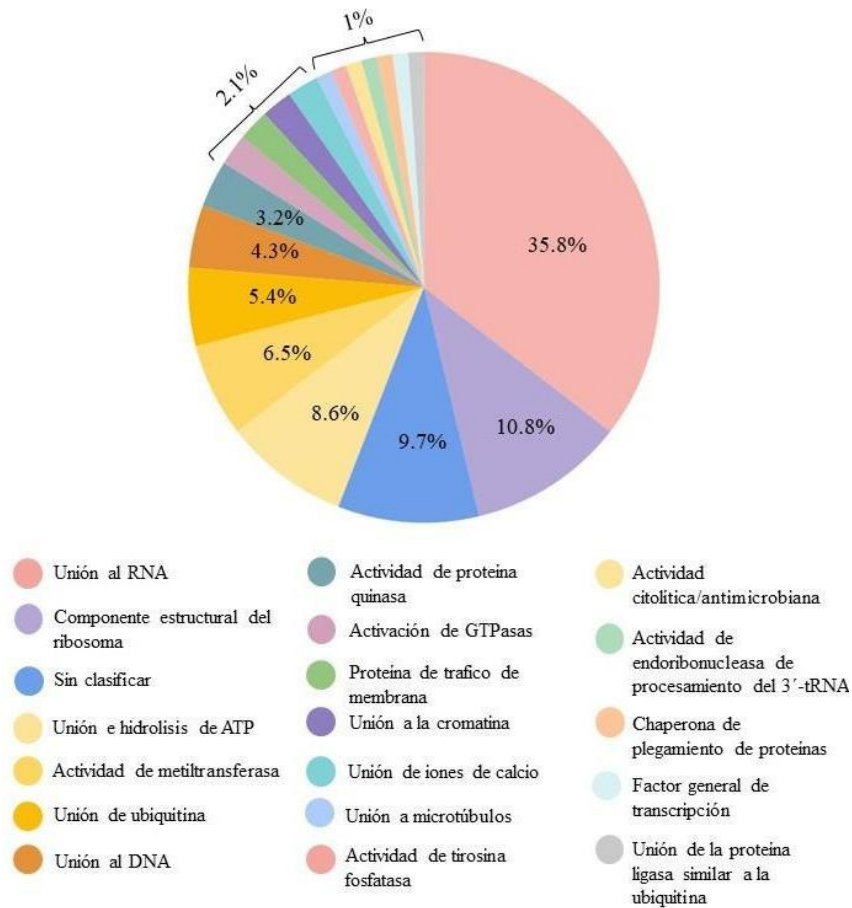


Figura 32. Análisis de GO por función molecular de las proteínas identificadas que interaccionan con EhMyb10r. Gráfica de pastel con la distribución de las proteínas asociadas a EhMyb10, clasificadas por función molecular a través del programa PantherDB.

Por otro lado, se ha evidenciado que un 9.6% de las proteínas, están directamente relacionadas con procesos asociados a la respuesta al estrés (GO:0006950) y a la modificación de proteínas (GO:0036211). Este último aspecto, sugirió una participación clave en la regulación de la función y actividad proteica, incluyendo modificaciones tales como la ubiquitinación, fosforilación, desfosforilación y SUMOilación, las cuales afectan la función de las proteínas, incluyendo a EhMyb10. Otro proceso biológico identificado, corresponde a la metilación de histonas (GO:0016571) como una modificación proteica relevante, relacionada con la organización de la cromatina. Este hallazgo, coincide con la función molecular donde EhMyb10 participa en la regulación de la expresión génica a nivel epigenético.

En el mismo contexto, se identificaron proteínas que participan en otros procesos biológicos de importancia. Por ejemplo, algunas proteínas contribuyen en un 1% a una variedad de eventos biológicos, abarcando desde el transporte lisosomal y mediado por vesículas hasta la transducción de señales impulsada por GTPasas y la respuesta defensiva ante bacterias (**Fig. 33**).

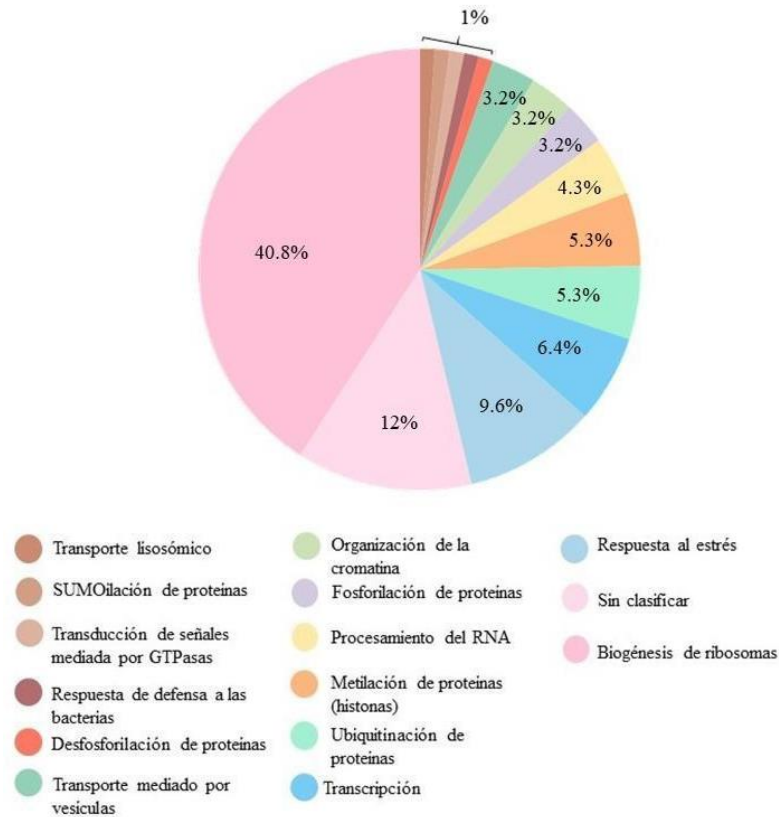


Figura 33. Análisis de GO por procesos biológicos de las proteínas identificadas que interaccionan con EhMyb10r. Gráfica de pastel con la distribución de las proteínas asociadas a EhMyb10, clasificadas por procesos biológicos a través del programa PantherDB.

Finalmente, se llevó a cabo un análisis para determinar en qué compartimento celular se localizan estas proteínas. El análisis de GO, reveló que el 14% de las proteínas se localizan exclusivamente en el núcleo (GO:0005634), y un 9.5% se podrían localizar en núcleo y nucléolo (GO:0005730), sugiriendo que participan en funciones nucleares como la transcripción y la regulación génica. Interesantemente, también se identificaron proteínas que se encuentran en la envoltura nuclear. El 38.7% de las proteínas, se encuentran en complejos ribonucleoproteicos relacionados con los ribosomas (GO:0005840).

El 18.3%, no se pudieron ubicar en un compartimento celular al no estar anotados. En el citoplasma, se encuentran un 12.9% (GO:0005737) por la función metabólica que desempeñan. Por último, también encontramos proteínas de la membrana del retículo endoplasmático (GO:0005791) y de microtúbulos (GO:1905720), lo cual indica la participación de las mismas en el transporte intracelular, el mantenimiento de la estructura celular, y la comunicación entre compartimentos subcelulares (**Fig. 34**). Por lo tanto, el análisis de localización celular, reveló que estas proteínas tienen una distribución amplia y diversa en los diferentes compartimentos celulares, reflejando su participación en una variedad de funciones y procesos celulares clave.

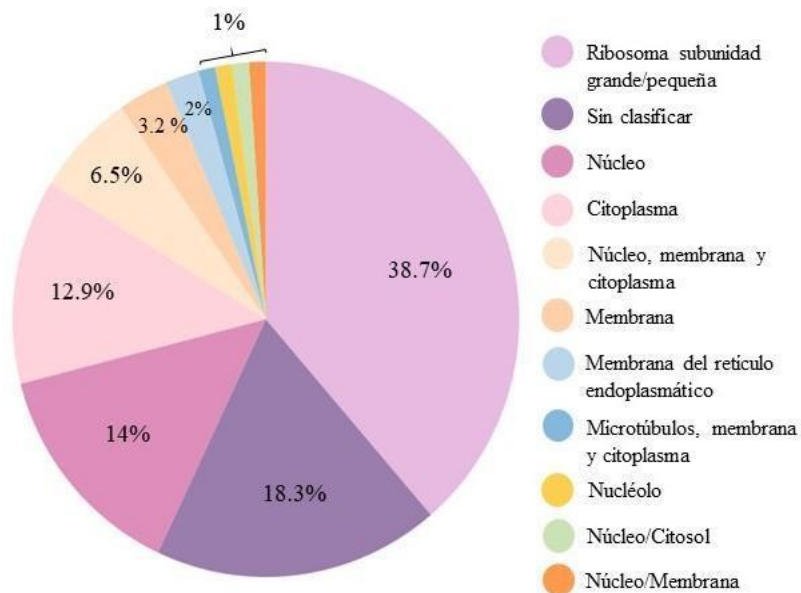


Figura 34. Análisis de GO por compartimento celular de las proteínas identificadas que interaccionan con EhMyb10r. Gráfica de pastel con la distribución de las proteínas asociadas a EhMyb10, clasificadas por compartimento celular a través del programa PantherDB.

9.10.2 Interactoma *in silico* de las proteínas que interactúan con EhMyb10 identificadas por espectrometría de masas (IMS-DIA-MS)

Para una mejor comprensión de las redes de interacción entre las proteínas identificadas, se llevó a cabo un análisis de interacción, utilizando la herramienta STRING. Este análisis mostró una estrecha interacción de EhMyb10 con proteínas ribosomales de las subunidades 60s y 40s, nucleoplasmina y fibrilarina, así como con proteínas de unión a snoRNA, como la Proteína Brix, que podrían estar relacionadas con la biogénesis de ribosomas ubicándola en el nucléolo. Observamos, además, la interacción de EhMyb10 con proteínas asociadas a la modificación de la cromatina y a la formación de dominios topológicos, por ejemplo, el complejo condensina, histonas H2B y H2A, proteínas modificadoras como ubiquitinas, cinasas y metiltransferasas, factores generales de transcripción, nucleasas, DEK y rcd1. Es relevante mencionar que se observó una interacción entre las proteínas SUN y espermidina sintasa que en otros organismos desempeña una función de metiltransferasa (**Fig. 35**).

Finalmente, se observó la interacción de EhMyb10 con proteínas involucradas en respuestas de estrés (*Heat shock*), proteasas, proteínas de activación de GTPasas (RhoGAP), metalo-beta-lactamasa, dinamina, proteínas de tráfico de membrana (AP-3), MRH, PH, amebaporo B, YIF1, LisH, Rad50 ATPasa, PB1 y Rab GTPasa, que no manifestaron interacciones con EhMyb10 en el interactoma *in silico*, aunque fueron identificadas por masas a partir del ensayo de Pull-Down.

9. 10. 3 Detección de la proteína fibrilarina en el eluido del ensayo de Pull-Down

Para confirmar las interacciones de las proteínas identificadas por espectrometría de masas con EhMyb10, llevamos a cabo un WB, utilizando un anticuerpo α -fibrilarina, la cual es una de las proteínas reportadas en este estudio. Para ello, analizamos las fracciones nucleares solubles e insolubles, y las eluciones del ensayo de Pull-Down con EhMyb10 y GST (**Fig. 36A**).

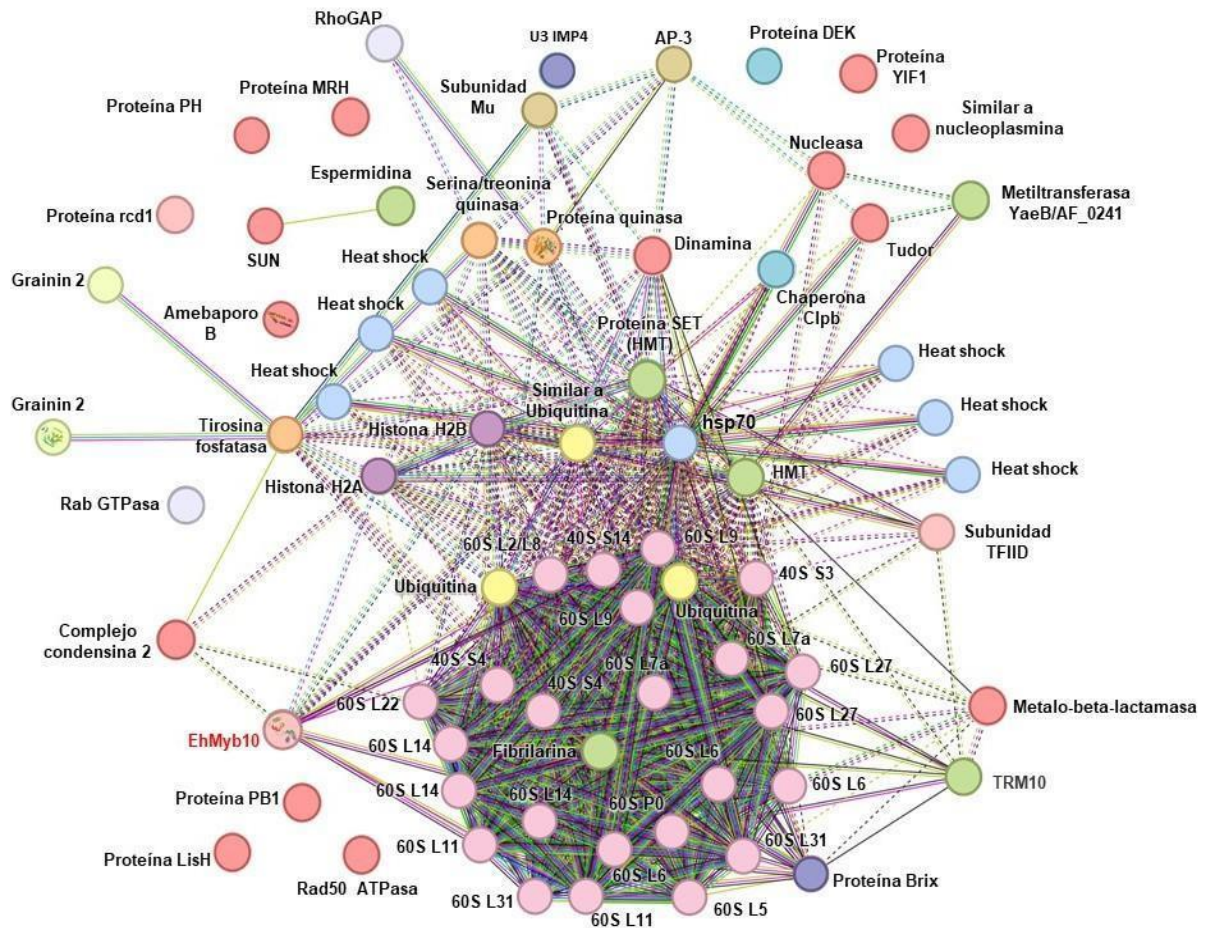


Figura 35. Interactoma de las proteínas identificadas que interactúan con EhMyb10r. Interactoma realizado con la herramienta STRING, empleando un puntaje de interacción de confianza media (0.400) como valor de corte y un número máximo de no más de 10 interactuantes. Cada color indica los cluster de proteínas que comparten la misma función molecular: Componente estructural del ribosoma (lila), unión a snoRNA (azul marino), factores de transcripción (rosa), actividad de metiltransferasa (verde), ubiquitinación y SUMOilación (amarillo), unión a iones calcio (amarillo claro), proteínas de choque térmico (azul claro), actividad de chaperona (azul fuerte), actividad de cinasa y fosfatasa (naranja), histonas (morado), proteínas de tráfico de membrana (café), proteínas con actividad de GTPasa (gris), y otras funciones (rojo). Número de nodos: 72, número de aristas: 645, número esperado de aristas: 570, valor p de enriquecimiento PPI < 0.00107.

El anticuerpo α -fibrilarina, detectó una banda 35 kDa en los extractos nucleares solubles, en concordancia con lo reportado por Lozano-Amado (2016). Sin embargo, en la elución del ensayo de Pull-Down cuando se utilizó la proteína EhMyb10 con la fracción nuclear insoluble, se reconoce una banda de aproximadamente 10 kDa. Esta diferencia, sugiere la posibilidad de una proteólisis de la proteína en la fracción insoluble, lo cual confirmó la posible interacción específica de EhMyb10 y fibrilarina (**Fig. 36B**). Como control de las fracciones, se empleó el anticuerpo α -H4K20me3 que reconoció tres señales con distintos pesos moleculares (53, 38 y 35 kDa), correspondientes a la proteína histona en los extractos nucleares solubles como se esperaba (**Fig. 36C**).

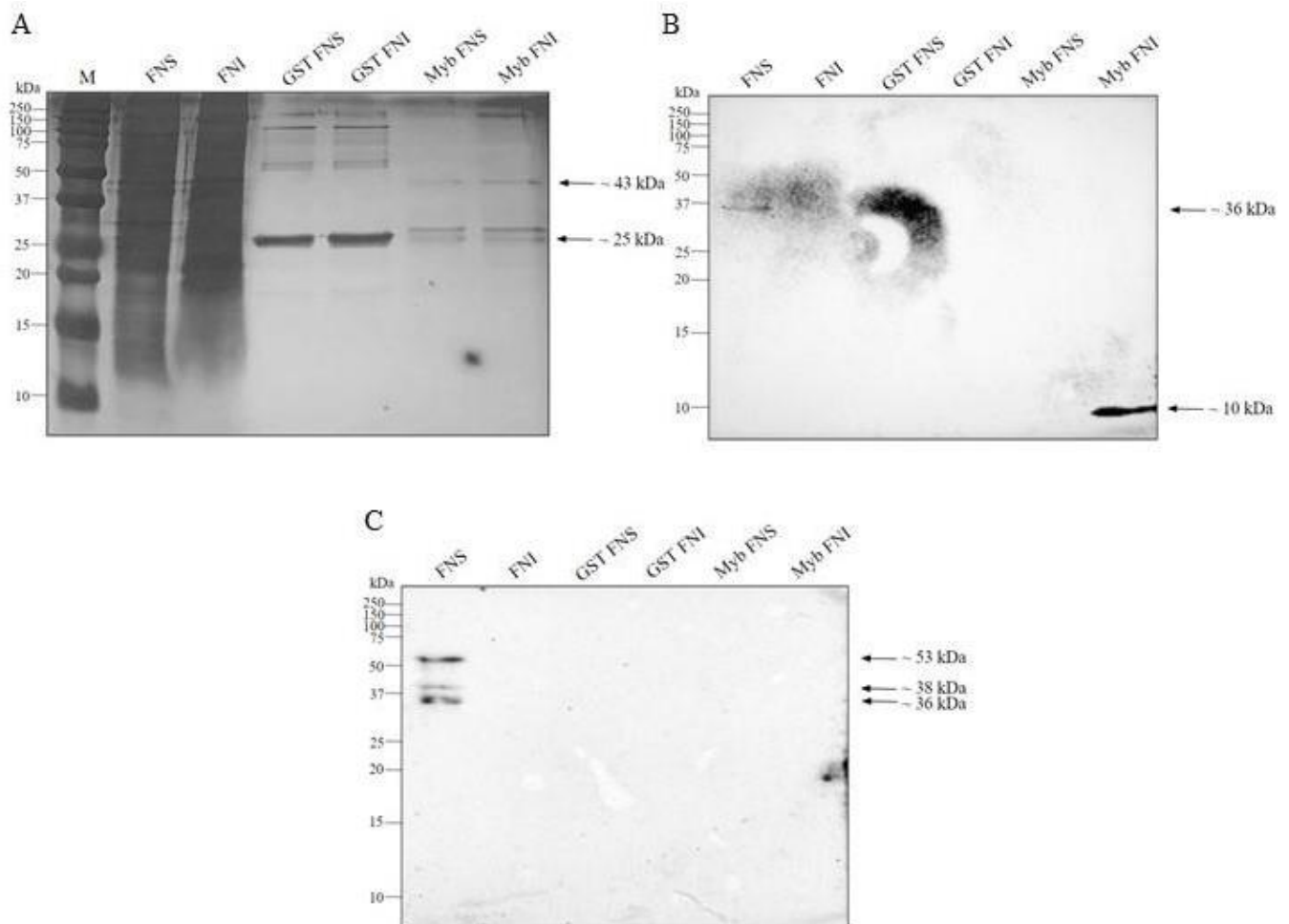


Figura 36. Detección de la proteína fibrilarina en el eluido del ensayo de Pull-Down. **A)** SDS-PAGE en gel al 15% de las fracciones nucleares solubles (FNS) e insolubles (FNI), y la elución del ensayo de Pull-Down de GST o Myb10 con FNS o FNI. Las bandas en el gel se visualizaron mediante tinción con plata. **B)** WB de fibrilarina. **C)** WB de histona H4K20me3. Los pesos moleculares se indican con flechas.

10. DISCUSIÓN

La amiba es un organismo parasitario altamente adaptable, capaz de responder de manera dinámica a las cambiantes condiciones de su entorno en el hospedero. Este microorganismo, conocido por su capacidad para colonizar y sobrevivir en distintos ambientes, se enfrenta a la necesidad constante de ajustar su comportamiento, función y actividad metabólica, para asegurar su adaptación y persistencia en el interior del huésped. Para ello, las células requieren detectar constantemente cambios en las condiciones internas y externas, que pueden transmitirse a través de vías de señalización que finalizan con PTMs reversibles o irreversibles en las proteínas. Las PTMs, pueden encender o apagar la función biológica de una proteína, modificar su localización dentro de la célula, influir en su capacidad para interactuar con otras proteínas, e incluso determinar si la proteína será degradada o no (Beltrao *et al.*, 2013). En nuestro trabajo, observamos que entre las PTMs, que pueden modificar la actividad de EhMyb10, se encuentran las fosforilaciones, ubiquitinaciones y SUMOilaciones. Entre las proteínas cinasas que se identificaron como posibles efectoras de fosforilación de los residuos de serina en EhMyb10, se destacan las proteínas cinasas C (PKC) y CKII (CK2). Se ha demostrado que PKC está involucrada en numerosos procesos de señalización celular que controlan la proliferación, diferenciación, supervivencia, invasión, migración y apoptosis celular (Kawano *et al.*, 2021), mientras que CK2, participa en varias funciones esenciales, incluida la transducción de señales, el crecimiento, la diferenciación celular, la expresión génica, la apoptosis, y en la regulación de las HDAC (Martínez *et al.*, 2020). Se ha observado que la fosforilación de c-Myb en los residuos de serina 11 y 12 por CK2, se encuentra implicada con la regulación de la unión del factor de transcripción al DNA (Cures *et al.*, 2001).

Por lo tanto, se puede inferir que CK2, podría estar regulando la unión de EhMyb10 al DNA, de manera análoga a lo que ocurre con su homólogo c-Myb (Bies *et al.*, 2001). Por otro lado, las cinasas RSK son una familia de proteína cinasas de serina/treonina que fosforilan la proteína S40 de la 6S del ribosoma y son efectoras río abajo de la vía de señalización Ras/ERK (cinasa regulada por señales extracelulares) (Ludwik *et al.*, 2016 y Romeo *et al.*, 2012).

Estas enzimas, están involucradas en la regulación de procesos celulares complejos, incluyendo la señalización celular y la transcripción, ya que pueden fosforilar y activar FT; también están involucradas en la proliferación, la supervivencia celular y la motilidad (Frödin *et al.*, 1999; Youn *et al.*, 2021). Las proteínas cinasas A (PKA) también son posibles efectoras de EhMyb10, se ha observado que la fosforilación por PKA en el residuo de serina 116, ubicado en el DBD-MYB del factor de transcripción c-Myb, conduce a una menor afinidad de unión al ERM en la región promotora, por tanto, la fosforilación de S116, tiene consecuencias bioquímicas y funcionales en la proteína lo que sugiere que al igual que CK2, PKA, regularía la actividad de unión EhMyb10 al DNA y modular negativamente la expresión de sus genes diana (Andersson *et al.*, 2003). Por último, las cinasas cdc2, pertenecen a la familia de proteínas cinasas Ser/Thr que desempeñan un papel clave en la regulación del ciclo celular, ya que promueven la transición de la fase G2/M del ciclo celular. La activación de cdc2, requiere su asociación con la subunidad de ciclina B (cdc2-ciclina B), esta asociación facilita la entrada en la fase mitótica mediante la fosforilación de sus sustratos (Kim *et al.*, 2020). Dado que EhMyb10 podría fosforilarse por estas cinasas, cada una de ellas podría modular su actividad a través de la fosforilación de los residuos de serina en eventos clave en la regulación génica y la señalización celular.

Con respecto a la ubiquitinación, esta PTM podría estar modulando la estabilidad de EhMyb10 como ocurre con c-Myb promoviendo su degradación a través de la vía 26S del proteosoma, la cual podría estar regulando la estabilidad de al controlar su vida media de manera irreversible, favoreciendo un recambio rápido de estas moléculas en la célula (Bies *et al.*, 1997). Por lo tanto, la ubiquitinación puede tener un impacto directo en la disponibilidad de este FT en la regulación génica. En concordancia, con la presencia de residuos susceptibles a ubiquitinación, encontramos en el interactoma *in silico* que EhMyb10 interactúa con proteínas relacionadas con este proceso, además con una proteasa del tipo methionil aminopeptidasa tipo 1 (MetAP 1), la cual podría estar relacionada con la remoción de la metionina del extremo amino terminal, lo cual es necesario para unirse a proteínas ribosomales (Walker, 2013).

Por otro lado, la SUMOilación, es un proceso de conjugación de la proteína SUMO (Small Ubiquitin-like Modifier) a residuos de lisina ubicados en un motivo ψ -K-X-D/E en las proteínas objetivo, lo que puede afectar su función, localización celular, y la interacción con otras proteínas. Los ciclos de SUMOilación y ubiquitinación son similares, pero cumplen funciones distintas. La modificación de la proteína diana con ubiquitina, principalmente marca la proteína para su reconocimiento y degradación por el proteosoma, mientras que la modificación con SUMO, confiere mayor estabilidad a la proteína modificada, y modula las interacciones proteína-proteína para regular la localización y la función de las proteínas diana. Además, la SUMOilación es esencial para mantener la integridad genómica, regular la transcripción y la transducción de señales intracelulares (Han *et al.*, 2018). Las posibles SUMOilaciones presentes en la secuencia de EhMyb10, podrían influir en su actividad de varias maneras: modulando las interacciones proteína-proteína, afectando su estabilidad, y localización dentro de la célula y su capacidad para unirse a elementos reguladores del DNA y regular la transcripción de genes específicos, como se ha demostrado con c-Myb (Bies *et al.*, 2002). Esto puede deberse a que la amiba pudiera presentar motivos de SUMOilación divergentes, ya que recientemente se ha analizado la presencia de la maquinaria de SUMOilación en protozoarios parásitos, encontrándose algunas enzimas en *E. histolytica*, lo que podría sugerir la relevancia de estas modificaciones en el contexto biológico de este parásito (Karpivyevich and Artavanis-Tsakonas, 2020). Por lo que, EhMyb10, es un FT susceptible de PTMs que pueden modificar su actividad biológica, su localización subcelular, o las interacciones proteína-proteína. Sería interesante realizar ensayos experimentales para corroborar la presencia, y determinar las condiciones a las cuales ocurren estas modificaciones en EhMyb10.

Por otro lado, observamos que EhMyb10 presenta, además del DBD-MYB característico conformado por los repetidos R2 y R3, regiones IDPs. Las IDPs, juegan un papel importante en funciones como la regulación de la transcripción y traducción, ya que facilita o median las interacciones proteína-proteína. Estas regiones desordenadas, participan en la regulación de procesos celulares, la señalización celular, la formación de complejos proteicos y la respuesta a estímulos ambientales. La presencia de IDPs en los factores de transcripción, proporcionan plasticidad estructural y funcional, lo que le permite participar en diversas interacciones y desempeñar diferentes funciones en el contexto celular (Dyson *et al.*, 2005).

En las proteínas Myb de *E. histolytica*, se han identificado IDPs en los extremos amino y carboxilo terminal (Cárdenas *et al.*, 2021).

Posteriormente, analizamos *in silico* las proteínas que interactúan con EhMyb10, e identificamos a una proteína con DBD-MYB con homología a CDC5 de *H. sapiens* a la cual se denominó EhCDC5-like. Las proteínas CDC5, son parte integral del complejo PRP19-CDC5L, que forma parte del *spliceosoma* y es necesario para activar el empalme de pre-mRNA. Además, se ha sugerido que el complejo PRP19-CDC5L, podría tener implicaciones en la respuesta al daño del DNA. En este análisis, también se observaron interacciones con otras proteínas relacionadas con *splicing*, como PRP46, el factor de *splicing* 3A y una proteína *crooked neck*. Asimismo, se identificaron proteínas relacionadas con la transcripción, como las subunidades 2, 5, 8 y 10, del factor general de transcripción TFIID, una proteína de bromodominio y proteínas con dominios similares al dominio de la proteína TAFII28 y TFIIB (**Fig. 37A**). El predominio de proteínas involucradas en los procesos de *splicing* y transcripción, resalta la importancia de EhMyb10 en la regulación de la expresión génica a nivel cotranscripcional. Estos hallazgos, subrayan la complejidad de la red de interacciones proteicas en las funciones biológicas en las que participa EhMyb10, lo que enfatiza la necesidad de comprender estas funciones moleculares y su relación con EhMyb10, para desentrañar los mecanismos de regulación y la contribución de EhMyb10 en los procesos biológicos específicos en los que está involucrado. En consecuencia, estas interacciones revelan la posible implicación de EhMyb10, y estas proteínas en procesos cotranscripcionales (Grote *et al.*, 2010). Otra proteína de unión a EhMyb10 es sirtuina 2 o Sir2 que participa en la desacetilación de histonas dependiente de NAD⁺, lo cual promueve la condensación de la cromatina reprimiendo la transcripción. Sir2 desempeña un papel importante en la regulación de la expresión génica y el mantenimiento de la estabilidad del genoma (Buck *et al.*, 2004). Por lo tanto, la interacción de EhMyb10 con Sir2, puede afectar la accesibilidad del DNA reprimiendo la expresión de genes específicos.

Para corroborar los interactomas *in silico*, se realizaron ensayos de Pull-Down con la proteína recombinante EhMyb10, y se identificaron las proteínas que se unieron a EhMyb10 mediante espectrometría de masas. Dentro de las proteínas que interactúan con EhMyb10, se encontraron a proteínas relacionadas con el inicio de la transcripción como a una subunidad del factor general de transcripción TFIID y al cofactor de transcripción Rcd1.

Rcd1 puede afectar el inicio de la transcripción de manera positiva o negativa, ya que interacciona con moléculas del complejo SAGA que presentan actividad de acetiltransferasa, tal como ocurre con la contraparte en *H. sapiens*, c-Myb (Has *et al.*, 2004). Adicionalmente, se identificaron chaperonas de histonas como la proteína DEK que tiene efectos en la arquitectura de la cromatina, lo que podría acercar o alejar regiones en el DNA (Waldmann *et al.*, 2004). Entre otros remodeladores de cromatina, se identificó a una PRMT y a una SMYD (SET and MYND Domain-Containing Proteins, por sus siglas en inglés), las cuales pueden afectar la organización de la cromatina a través de la metilación de histonas, ya sea en residuos de arginina o lisina, respectivamente. Interesantemente, también identificamos proteínas TUDOR que interaccionan con las histonas metiladas. Finalmente, se identificaron proteínas que están relacionadas con dominios topológicamente activos (TADs, por sus siglas en inglés), como es la proteína SUN1 y las condensinas (**Fig. 37A**).

Por otro lado, entre las moléculas identificadas que interaccionan con EhMyb10, encontramos proteínas relacionadas con la biogénesis de los ribosomas. La biogénesis de los ribosomas ocurre en el nucléolo y se encuentra relacionada con el crecimiento y la división celular. El nucléolo es la mayor estructura que se encuentra en el núcleo, es muy dinámico y su naturaleza depende de la cantidad de las moléculas que lo componen (Hernández *et al.*, 2006). La biogénesis de los ribosomas, implica la transcripción del DNAr, el procesamiento del rRNA y el ensamblaje de las proteínas ribosómicas. En los organismos eucariontes, la biogénesis de los ribosomas se lleva a cabo a través de la RNA pol I para transcribir un pre-rRNA 45S, que funciona como un andamio para el ensamblaje de complejos ribonucleoproteicos, en donde participan proteínas y RNAs pequeños nucleolares (snoRNP), para formar una partícula pre-ribosomal 90S en el nucléolo. Durante los pasos de maduración, el pre-rRNA 45S en la partícula pre-ribosomal 90S, se procesa en 18S, 5.8S, y 25S rRNA mediante la eliminación de las regiones espaciadoras (Henras *et al.*, 2015). De manera interesante, en este trabajo identificamos a la fibrilarina, la cual es una proteína considerada como un marcador molecular de la RNA polimerasa I transcripcionalmente activa (**Fig. 37B**). Esta proteína, tiene un peso molecular de 34 a 38 kDa. Es una enzima con actividad de S-adenosilmetionina (SAM) metiltransferasa (MTasas) que participa los primeros pasos del procesamiento preribosómico y necesarios para la estabilidad estructural del ribosoma (Rodríguez *et al.*, 2015).

Ya que en el parásito *E. histolytica*, se ha detectado a la fibrilarina utilizando un anticuerpo heterólogo dirigido a esta proteína (Lozano et al., 2016). En este trabajo, también se utilizó para validar una de las proteínas identificadas en los ensayos de Pull-down, sin embargo, en la fracción eluida de extractos insolubles, se inmunodetectó una banda de un peso molecular menor (~10 kDa) al que se inmunodetectó en los extractos totales (~35 kDa). De manera interesante, está reportada la presencia de actividad de proteasa, asociada con la remodelación del andamiaje nuclear con una selectividad por las láminas A, B y C, durante diferentes eventos celulares (Tokes *et al.*, 1989). Eso podría explicar el bajo peso molecular que identificamos para la proteína fibrilarina, la cual se detectó en la elución en el ensayo de Pull-Down. Cuando analizamos la secuencia de aminoácidos de fibrilarina, encontramos sitios susceptibles a proteólisis mediante la herramienta de PeptideCutter, lo que podría explicar el tamaño de las moléculas en la banda identificada.

También se identificaron proteínas que contienen el dominio Brix (del inglés, biogenesis of ribosomes in *Xenopus* domain), la cual participa en el procesamiento del rRNA pre-18S al unirse a RNA nucleolar U3 (U3 snoRNP), y a su proteína asociada (proteína de unión a U3 snoRNP) (Eisenhaber *et al.*, 2001). En este ensayo, también se identificó a la proteína nucleolina, la cual es la fosfoproteína multifuncional más abundante del nucléolo (**Fig. 37B**). Los niveles de nucleolina se correlacionan con la tasa de actividad funcional del nucléolo en células de crecimiento. La nucleolina es una helicasa que se une al DNA y al RNA. Regula varios aspectos de su metabolismo, la estructura de la cromatina, la transcripción del DNAr, la maduración del rRNA, la citocinesis, la nucleogénesis, la proliferación, el crecimiento celular, el plegamiento, la maduración, y el ensamblaje de ribosomas y el transporte nucleocitoplasmático de pre-RNAs recién sintetizados (Tajrishi *et al.*, 2011). También se identificó a la proteína IMP4, la cual forma un complejo ternario con IMP3 y MPP10 que interactúa con U3 snoRNA, necesario para los primeros pasos del procesamiento del pre-rRNA (**Fig. 37B**) (Granneman *et al.*, 2003). Finalmente, identificamos diferentes proteínas ribosomales, las cuales están relacionadas con el pre-ensamble de los ribosomas, entre las que se encuentran las proteínas ribosómicas 40 y 60S, las cuales actúan como chaperonas que ayudan al correcto plegamiento de los pre-rRNA nacientes, así como el procesamiento y ensamblaje (Henras *et al.*, 2015).

A la fecha no se ha descrito la interacción directa de proteínas Myb con las proteínas relacionadas con la biogénesis de los ribosomas, pero se ha descrito a la proteína Mybbp1a, la cual interactúa con c-Myb. La proteína Mybbp1a, se expresa de manera ubicua y se localiza predominantemente en el nucléolo, modulando la actividad de varios reguladores de la transcripción dependiente de la RNA pol I, y también se encuentra relacionada con la maquinaria de biogénesis de los ribosomas (Hochstatter *et al.*, 2012). De manera interesante en el genoma de *E. histolytica*, existe un gen que codifica para una proteína con alta identidad con la proteína Mybbp1a de *H. sapiens* (EHI_030450). Esto sugiere que EhMyb10, podría también regular la transcripción de los genes de rRNA y participar en el procesamiento temprano de la subunidad pequeña ribosomal, la cual contiene la U3 snoRNA, y que también se identificó que interactúa con EhMyb10 (**Fig. 37B**).

Este trabajo provee las primeras evidencias de las interacciones de la proteína EhMyb10 en *E. histolytica* y su papel en la regulación cotranscripcional, y en la biogénesis de los ribosomas.

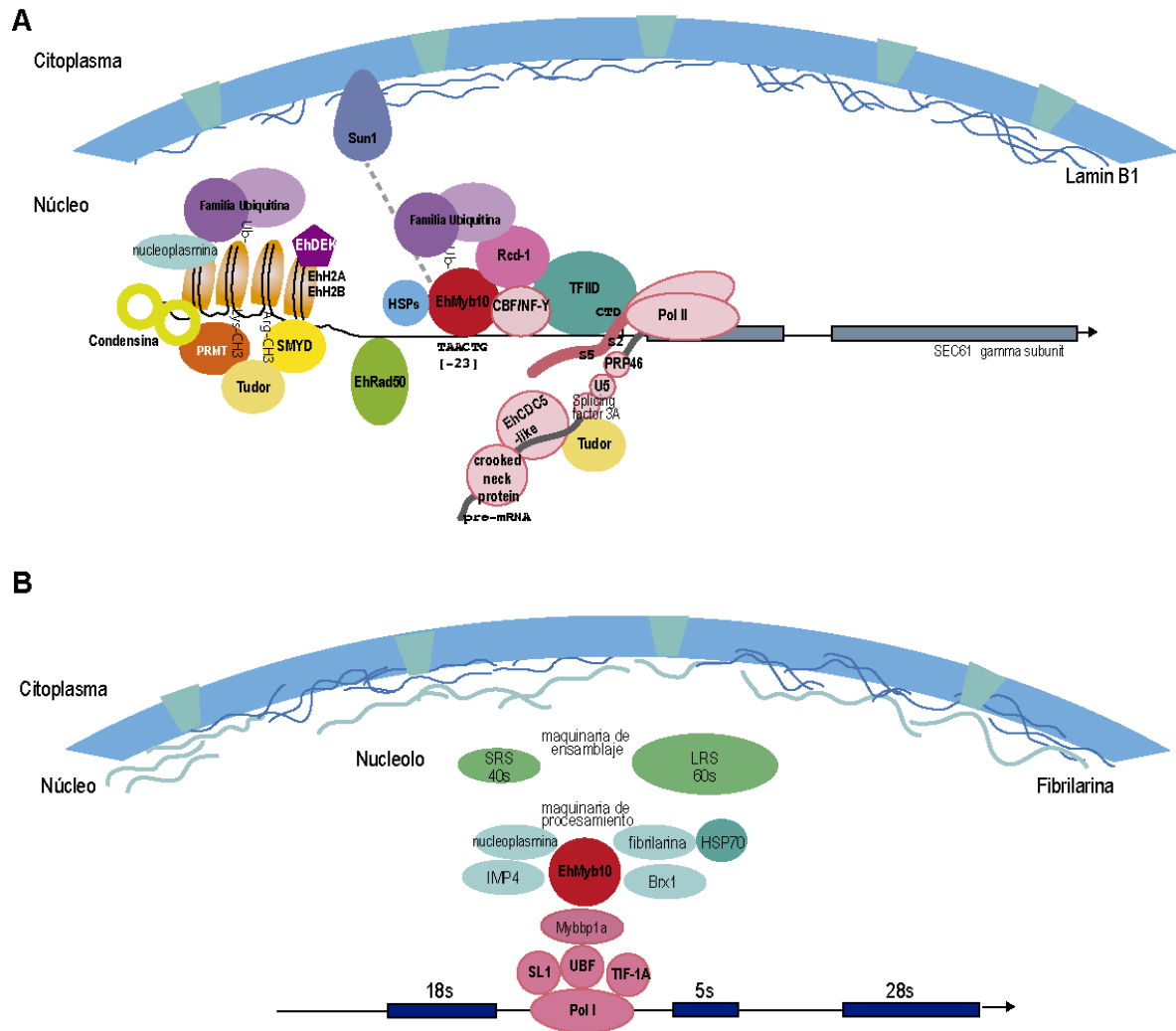


Figura 37. Modelo molecular de las interacciones de EhMyb10. A) EhMyb10 y sus interacciones con proteínas relacionadas con factores de transcripción generales y factores de *splicing*, lo cual permite proponer su papel en mecanismos cotranscripcionales. EhMyb se une a la secuencia ERM (C/TAAGTC) del promotor del gen que codifica para la proteína Sec61gamma. TFIID interactúa con la RNA pol II, con un coactivador Rcd1 y con el factor CBF/NF-Y. Los factores de *splicing* PRP46, U5, 3A, el factor de transcripción EhCDC5-like y la proteína *crooked neck*, participan en el empalme del pre-mRNA. B) EhMyb10 podría interactuar con Mybbp1a y proteínas que participan en el procesamiento y ensamblaje de los ribosomas.

11. CONCLUSIONES

- La proteína EhMyb10 es susceptible a PTMs que pueden desempeñar un papel crucial en la regulación de su actividad y función, influyendo en su unión al DNA, interacción con otras proteínas y estabilidad.
- La presencia de regiones IDPs en EhMyb10, podría proporcionar plasticidad estructural y funcional, permitiéndole interactuar con diversas proteínas y participar en múltiples funciones celulares.
- Por ensayos de Pull-Down e identificación por masas, se ha revelado la posible participación de EhMyb10 en la regulación de la expresión génica, a través de procesos cotranscripcionales mediante la interacción con proteínas relacionadas con la modificación de la cromatina, formación del complejo de iniciación, factores de transcripción y factores de *splicing*.
- EhMyb10 interacciona con proteínas relacionadas con la biogénesis de ribosomas, lo cual sugiere su participación en la regulación de la transcripción de genes que codifican para rRNA y en el procesamiento de las subunidades ribosómicas.

PERSPECTIVAS

Una de las principales limitaciones de este estudio fue en el origen de la proteína recombinante utilizada, la cual proviene de un sistema bacteriano de *E. coli*. A diferencia de las proteínas de origen eucariota, las proteínas bacterianas generalmente carecen de las modificaciones postraduccionales típicas en eucariotas, como la fosforilación, la glicosilación, SUMOilación y la acetilación (Wingfield, 2015). Estas modificaciones, desempeñan un papel crucial en la regulación de las interacciones proteína-proteína y en la actividad de los FT. Por lo tanto, la falta de estas modificaciones postraduccionales en la proteína recombinante, podría limitar la capacidad del FT para interactuar con proteínas nucleares, por lo que las interacciones dependientes de ubiquitinación, fosforilación o SUMOilación no se encontrarían presentes. Esto podría resolverse realizando ensayos de inmunoprecipitación de EhMyb10 con trofozoitos que sobreexpresen a este factor de transcripción, o cambiando el sistema de expresión de una célula procarionte a una eucarionte.

Las interacciones identificadas, ofrecen nuevas evidencias sobre las redes de regulación en las que EhMyb10 puede estar involucrada y subrayan la necesidad de experimentos adicionales para comprender completamente su función en la biología del parásito. Dado que el estudio se basa en un sistema *in vitro* utilizando proteínas recombinantes, es esencial validar las interacciones identificadas en un entorno celular relevante. Esto implica realizar experimentos adicionales, como co-inmunoprecipitaciones empleando extractos totales de trofozoitos. Asimismo, un ensayo de inmunofluorescencia, permitiría confirmar la interacción y colocalización del FT con algunas de las proteínas identificadas. Este estudio proporciona una base sólida para futuras investigaciones que podrían profundizar en los mecanismos mediante los cuales *E. histolytica* regula la expresión génica para desarrollar procesos biológicos clave en su virulencia.

12. REFERENCIAS

1. Arratia, J., & Aguirre, J. (2013). Los factores de transcripción tipo MYB, una familia de reguladores de la diferenciación celular conservada en los organismos eucariontes. *TIP*, *16*(2), 98-108.
2. Andersson, K. B., Kowenz-Leutz, E., Brendeford, E. M., Tygset, A. H., Leutz, A., & Gabrielsen, O. S. (2003). Phosphorylation-dependent down-regulation of C-MYB DNA binding is abrogated by a point mutation in the V-MYB Oncogene. *Journal of Biological Chemistry*, *278*(6), 3816-3824.
3. Burns, C. G., Ohi, R., Krainer, A. R., & Gould, K. L. (1999b). Evidence that MYB-related CDC5 proteins are required for pre-mRNA *splicing*. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(24), 13789-13794.
4. Buck, S. W., Gallo, C. M., & Smith, J. S. (2004). Diversity in the SIR2 family of protein deacetylases. *Journal of Leukocyte Biology*, *75*(6), 939-950.
5. Beltrao, P., Bork, P., Krogan, N. J., & Van Noort, V. (2013). Evolution and functional cross-talk of protein post-translational modifications. *Molecular Systems Biology*, *9*(1).
6. Berger, M. F., Badis, G., Gehrke, A. R., Talukder, S., Philippakis, A., Peña-Castillo, L., Alleyne, T. M., Mnaimneh, S., Botvinnik, O., Chan, E. T., Khalid, F., Zhang, W., Newburger, D. E., Jaeger, S., Morris, Q., Bulyk, M. L., & Hughes, T. R. (2008). Variation in homeodomain DNA binding revealed by High-Resolution Analysis of Sequence preferences. *Cell*, *133*(7), 1266-1276.

7. Boschet, C., Gissot, M., Briquet, S., Hamid, Z., Claudel-Renard, C., & Vaquero, C. (2004). Characterization of PFMYB1 transcription factor during erythrocytic development of 3D7 and F12 *plasmodium falciparum* clones. *Molecular and Biochemical Parasitology*, *138*(1), 159-163.
8. Bies, J., Feiková, S., Markus, J., & Wolff, L. (2001). Phosphorylation-Dependent conformation and proteolytic stability of C-MYB. *Blood Cells Molecules and Diseases*.
9. Bies, J., & Wolff, L. (1997). Oncogenic activation of C-MYB by carboxyl-terminal truncation leads to decreased proteolysis by the ubiquitin-26S proteasome pathway. *Oncogene*, *14*(2), 203-212.
10. Bies, J., Markus, J., & Wolff, L. (2002). Covalent attachment of the SUMO-1 protein to the negative regulatory domain of the C-MYB transcription factor modifies its stability and transactivation capacity. *Journal of Biological Chemistry*, *277*(11), 8999-9009.
11. Broccoli, D., Smogorzewska, A., Chong, L., & de Lange, T. (1997). Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nature genetics*, *17*(2), 231–235.
12. Cárdenas-Hernández, H., Titaux-Delgado, G., Castañeda-Ortiz, E. J., Torres-Larios, A., Briebe, L. G., Del Río-Portilla, F., & Azuara-Liceaga, E. (2021). Genome-wide and structural analysis of the MYB-SHAQKYF family in *Entamoeba histolytica* *Biochimica Et Biophysica Acta - Proteins And Proteomics*, *1869*(4), 140601.
13. Chávez, B., & Gonzales, A. (2013). *Entamoeba histolytica*: La estructura interna de un destructor por naturaleza. *Ciencia*, *64*(2), 42–49.
14. CDC - DPDx - Amebiasis. (2019). Centers for Disease Control and Prevention.

15. Cabrejos, M., Tamayo, C., & Maldonado, M. (2001). Análisis molecular del proceso de transcripción de genes en eucariontes. *Revista chilena de pediatría*, 72(5).
16. Charoensawan, V., Wilson, D. A., & Teichmann, S. A. (2010). Genomic repertoires of DNA-binding transcription factors across the tree of life. *Nucleic Acids Research*, 38(21), 7364-7377.
17. Cures, A., House, C. M., Kanei-Ishii, C., Kemp, B. E., & Ramsay, R. G. (2001). Constitutive C-MYB amino-terminal phosphorylation and DNA binding activity uncoupled during entry and passage through the cell cycle. *Oncogene*, 20(14), 1784-1792.
18. Chong, L., van Steensel, B., Broccoli, D., Erdjument-Bromage, H., Hanish, J., Tempst, P., & de Lange, T. (1995). A human telomeric protein. *Science (New York, N.Y.)*, 270(5242), 1663–1667.
19. Dyson, H. J., & Wright, P. E. (2005). Intrinsically unstructured proteins and their functions. *Nature Reviews Molecular Cell Biology*, 6(3), 197-208.
20. de Lange T. (2005). Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes & development*, 19(18), 2100–2110.
21. Ehrenkauf, G., Hackney, J. A., & Singh, U. (2009). A developmentally regulated MYB domain protein regulates expression of a subset of stage-specific genes in *Entamoeba histolytica*. *Cellular Microbiology*, 11(6), 898-910.
22. Ehsani, R., Bahrami, S., & Drabløs, F. (2016). Feature-based classification of human transcription factors into hypothetical sub-classes related to regulatory function. *BMC Bioinformatics*, 17(1).

23. Eisenhaber, F., Wechselberger, C., & Kreil, G. (2001). The Brix domain protein family – a key to the ribosomal biogenesis pathway? *Trends in Biochemical Sciences*, 26(6), 345-347.
24. Fundacion iO. (2020). *Entamoeba histolytica* - *Fundación IO*. Fundación iO.
25. Frödin, M., & Gammeltoft, S. (1999). Role and regulation of 90 KDA ribosomal S6 kinase (RSK) in signal transduction. *Molecular and Cellular Endocrinology*, 151(1-2), 65-77.
26. Grote, M., Wolf, E., Will, C. L., Lemm, I., Agafonov, D. E., Schomburg, A., Fischle, W., Urlaub, H., & Lührmann, R. (2010). Molecular architecture of the human PRP19/CDC5L complex. *Molecular and Cellular Biology*, 30(9), 2105-2119.
27. Granneman, S. (2003). The human IMP3 and IMP4 proteins form a ternary complex with HMPP10, which only interacts with the U3 SNORNA in 60-80S ribonucleoprotein complexes. *Nucleic Acids Research*, 31(7), 1877-1887.
28. Gilchrist, C. A., Houpt, E., Trapaidze, N., Fei, Z., Crasta, O., Asgharpour, A., Evans, C., Martino-Catt, S., Baba, D. J., Stroup, S., Hamano, S., Ehrenkaufner, G., Okada, M., Singh, U., Nozaki, T., Mann, B. J., & Petri, W. A., Jr (2006). Impact of intestinal colonization and invasion on the *Entamoeba histolytica* transcriptome. *Molecular and biochemical parasitology*, 147(2), 163–176.
29. Haas, M., Siegert, M., Schürmann, A., Sodeik, B., & Wolfes, H. (2004). C-MYB protein interacts with RCD-1, a component of the CCR4 transcription mediator complex. *Biochemistry*, 43(25), 8152-8159.

30. Han, Z., Feng, Y., Gu, B. H., Li, Y. M., & Chen, H. (2018). The Post-translational Modification, SUMOylation, and Cancer (Review). *International Journal of Oncology*.
31. Hochstatter, J., Hölzel, M., Michaela, R., Schermelleh, L., Leonhardt, H., Keough, R. A., Gonda, T. J., Imhof, A., Eick, D., Längst, G., & Németh, A. (2012). MYB-binding protein 1A (MyBBP1A) regulates levels and processing of pre-ribosomal RNA. *Journal of Biological Chemistry*, 287(29), 24365-24377.
32. Hernandez-Verdun, D. (2006). The Nucleolus: a model for the organization of nuclear functions. *Histochemistry and Cell Biology*, 126(2), 135-148.
33. Henras, A. K., Plisson-Chastang, C., O'Donohue, M., Chakraborty, A., & Gleizes, P. (2014). An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdisciplinary Reviews - Rna*, 6(2), 225-242.
34. Kariyevich, M., & Artavanis-Tsakonas, K. (2020). Ubiquitin-Like Modifiers: Emerging Regulators of Protozoan Parasites. *Biomolecules*, 10(10), 1403.
35. Kawano, T., Inokuchi, J., Eto, M., Murata, M., & Kang, J. H. (2021). Activators and inhibitors of protein kinase C (PKC): their applications in clinical trials. *Pharmaceutics*, 13(11), 1748.
36. Kim, H. M., He, L., Lee, S., Park, C., 김동현, Han, H. S., Han, J., Hwang, J., Cha-Molstad, H., Lee, K. H., Ko, S. K., Jang, J. H., Ryoo, I., Blenis, J., Lee, H. G., Ahn, J. S., Kwon, Y. T., Soung, N., & Kim, B. Y. (2020). Inhibition of osteoclasts differentiation by CDC2-induced NFATC1 phosphorylation. *Bone*, 131, 115153.
37. Krasny, L., & Huang, P. H. (2021). Data-Independent Acquisition Mass Spectrometry (DIA-MS) for proteomic applications in oncology. *Molecular omics*, 17(1), 29-42.

38. Lemma, R. B., Ledsaak, M., Fuglerud, B. M., Sandve, G. K., Eskeland, R., & Gabrielsen, O. S. (2021). Chromatin occupancy and target genes of the haematopoietic Master Transcription Factor MYB. *Scientific Reports*, *11*(1).
39. Lorenzi, H., Puiu, D., Miller, J. R., Brinkac, L., Amedeo, P., Hall, N., & Caler, E. (2010). New assembly, reannotation and analysis of the *Entamoeba histolytica* genome reveal new genomic features and protein content information. *PLOS Neglected Tropical Diseases*, *4*(6), e716.
40. Lambert, S. A., Jolma, A., Campitelli, L. F., Das, P. K., Yin, Y., Albu, M., Rothenberg, M. E., Taipale, J., Hughes, T. R., & Weirauch, M. T. (2018). The human transcription factors. *Cell*, *172*(4), 650-665.
41. Ludwik, K. A., & Lannigan, D. A. (2016). Ribosomal S6 kinase (RSK) Modulators: A patent review. *Expert Opinion on Therapeutic Patents*, *26*(9), 1061-1078.
42. Lv, J., Yang, T., Hong-Yuan, Y., Li, Z., Qin, P., Zhang, X., Liang, X., Li, J., & Chen, Q. (2013). Identification of MYB genes in *euplotes aediculatus* may indicate an early evolutionary process. *Gene*, *530*(2), 266-272.
43. Li, B., Espinal, A., & Cross, G. A. (2005). *Trypanosome* telomeres are protected by a homologue of mammalian TRF2. *Molecular and Cellular Biology*, *25*(12), 5011-5021.
44. Lozano-Amado, D., Herrera-Solorio, A. M., Valdés, J., Alemán-Lazarini, L., De Jesús Almaraz-Barrera, M., Luna-Rivera, E., Vargas, M., & Hernández-Rivas, R. (2016). Identification of repressive and active epigenetic marks and nuclear bodies in *Entamoeba histolytica*. *Parasites & Vectors*, *9*(1).

45. Millard, P. S., Kragelund, B. B., & Burow, M. (2019b). R2R3 MYB transcription Factors – functions outside the DNA-Binding domain. *Trends in Plant Science*, 24(10), 934-946.
46. Meneses, E., Cárdenas, H., Zárate, S., Brieba, L. G., Orozco, E., López-Camarillo, C., Azuara-Liceaga, E. (2010). The R2R3 MYB protein family in *Entamoeba histolytica*. *Gene*, 455(1-2), 32-42.
47. Makarova, O., Makarov, E. M., Urlaub, H., Will, C. L., Gentzel, M., Wilm, M., & Lührmann, R. (2004). A subset of human 35S U5 proteins, including PRP19, function prior to catalytic step 1 of *splicing*. *The EMBO Journal*, 23(12), 2381-2391.
48. Mozgová, I., Schrupfová, P. P., Hofr, C., & Fajkus, J. (2008). Functional characterization of domains in ATTRB1, a putative telomere-binding protein in *Arabidopsis thaliana*. *Phytochemistry*, 69(9), 1814-1819.
49. Martínez, R., Di Geronimo, B., Pastor, M., Zapico, J. M., Coderch, C., Panchuk, R. R., Skorokhyd, N. R., Maslyk, M., Ramos, A., & De Pascual-Teresa, B. (2020). Multitarget anticancer agents based on Histone deacetylase and protein kinase CK2 inhibitors. *Molecules*, 25(7), 1497.
50. Naiyer, S., Bhattacharya, A., & Bhattacharya, S. (2019). Advances in *Entamoeba histolytica* biology through transcriptomic analysis. *Frontiers in Microbiology*, 10.
51. Olivos-García, Alfonso, Saavedra, Emma, Nequiz Avendaño, Mario, Pérez-Tamayo, Ruy. (2011). Amibiasis: mecanismos moleculares de la patogenicidad de *Entamoeba histolytica*. *Revista de la Facultad de Medicina (México)*, 54(2), 10-20.
52. Orphanides, G. M., & Reinberg, D. (2002). A unified theory of gene expression. *Cell*, 108(4), 439-451.

53. Ong, S., Hsu, H., Liu, H., Chu, C., & Tai, J. (2006). Multifarious transcriptional regulation of adhesion protein gene *AP65 - 1* by a novel MYB1 protein in the protozoan parasite *Trichomonas vaginalis*. *Eukaryotic Cell*, 5(2), 391-399.
54. Pinilla, Análida Elizabeth, López, Myriam Consuelo, & Viasus, Diego Fernando. (2008). Historia del protozoo *Entamoeba histolytica*. *Revista médica de Chile*, 136(1), 118-124.
55. Palm, W., & de Lange, T. (2008). How shelterin protects mammalian telomeres. *Annual review of genetics*, 42, 301–334.
56. Rendón-Gandarilla, F. J., Álvarez-Hernández, V., Castañeda-Ortiz, E. J., Cárdenas-Hernández, H., Cárdenas-Guerra, R. E., Valdés, J., Betanzos, A., Chávez-Munguía, B., Lagunes-Guillén, A., Orozco, E., López-Cánovas, L., & Azuara-Liceaga, E. (2018). Telomeric Repeat-Binding Factor Homologs in *Entamoeba histolytica*: New Clues for Telomeric Research. *Frontiers in Cellular and Infection Microbiology*, 8.
57. Romero Cabello, R. (2008). Amibiasis en el siglo XXI. México: Sociedad Mexicana de Parasitología, A.C.
58. Romeo, Y., Zhang, X., & Roux, P. P. (2011). Regulation and function of the RSK family of protein kinases. *Biochemical Journal*, 441(2), 553-569.
59. Rodríguez-Corona, U., Sobol, M., Rodríguez-Zapata, L. C., Hozák, P., & Castaño, E. (2015). Fibrillarin from archaea to human. *Biology of the Cell*, 107(6), 159-174.
60. Slattery, M., Zhou, T., Yang, L., Machado, A. C. D., Gordân, R., & Rohs, R. (2014). Absence of a simple code: How transcription factors read the genome. *Trends in Biochemical Sciences*, 39(9), 381-399.

61. Suter, D. M. (2020). Transcription factors and DNA play hide and seek. *Trends in Cell Biology*, 30(6), 491-500.
62. Stober-Grässer, U., Brydolf, B., Bin, X., Grässer, F., Firtel, R. A., & Lipsick, J. S. (1992). The Myb DNA-binding domain is highly conserved in *Dictyostelium discoideum*. *Oncogene*, 7(3), 589–596.
63. Schreiber, E., Matthias, P., Müller, M. M., & Schaffner, W. (1989). Rapid detection of octamer binding proteins with ‘mini extracts’, prepared from a small number of cells. *Nucleic Acids Research*, 17(15), 6419.
64. Sun, C., Palm, D., McArthur, A. G., Svärd, S. G., & Gillin, F. D. (2002). A novel MYB-related protein involved in transcriptional activation of encystation genes in *giardia lamblia*. *Molecular Microbiology*, 46(4), 971-984.
65. Shirley, D. T., Farr, L., Watanabe, K., & Moonah, S. (2018). A Review of the Global Burden, New Diagnostics, and Current Therapeutics for Amebiasis. *Open forum infectious diseases*, 5(7), ofy161.
66. Tajrishi, M. M., Tuteja, R., & Tuteja, N. (2011). Nucleolin: The most abundant multifunctional phosphoprotein of nucleolus. *Communicative & Integrative Biology*, 4(3), 267-275.
67. Tökés ZA, Clawson GA. Proteolytic activity associated with the nuclear scaffold. The effect of self-digestion on lamins. *J Biol Chem*. 1989 Sep 5;264(25):15059-65.
68. U.N.A.M. (s. f.). Parásitos y Protozoarios. depa.fquim.unam.mx.
http://depa.fquim.unam.mx/bioseguridad/agentes/parasitos/paras_ehisto.html.

69. Waldmann, T., Scholten, I., Kappes, F., Hu, H., & Knippers, R. (2004). The DEK protein—an abundant and ubiquitous constituent of mammalian chromatin. *Gene*, 343(1), 1-9.
70. Weber, C., Koutero, M., Dillies, M. A., Varet, H., Lopez-Camarillo, C., Coppée, J. Y., Hon, C. C., & Guillén, N. (2016). Extensive transcriptome analysis correlates the plasticity of *Entamoeba histolytica* pathogenesis to rapid phenotype changes depending on the environment. *Scientific reports*, 6, 35852.
71. Wingfield, P. T. (2015). Overview of the purification of recombinant proteins. *Current protocols in protein science*, 80(1).
72. Wahl, M. C., Will, C. L., & Lührmann, R. (2009). The *spliceosome*: design principles of a dynamic RNP machine. *Cell*, 136(4), 701-718.
73. Will, C. L., & Lührmann, R. (2010). Spliceosome structure and function. *Cold Spring Harbor Perspectives in Biology*, 3(7), a003707.
74. Walker, K. W. (2013). Methionyl aminopeptidase type 1. En *Elsevier eBooks* (pp. 1495-1500).
75. Youn, M., Gomez, J. O., Mark, K., & Sakamoto, K. M. (2021). RSK isoforms in acute myeloid leukemia. *Biomedicines*, 9(7), 726.
76. Yang, T., Perasso, R., & Baroin-Tourancheau, A. (2003). MYB genes in ciliates: a common origin with the protooncogene? *Protist*, 154(2), 229-238.

13. ANEXOS

EHMYB10 TRANSCRIPTION FACTOR INTERACTOME: EVIDENCE OF COTRANSCRIPTIONAL REGULATION

Danna Paola Aguirre Casimiro¹, Patricia Cuellar², Claudia Castillo Mercado³,
Elizabeth J. Castañeda-Ortiz², Landy Marlene Rojas Avilés¹, Jesús Valdés⁴, Elisa Azuara Liceaga¹

¹Licenciatura en Ciencias Genómicas, UACM; ²Posgrado en Ciencias Genómicas, UACM,
³Programa de Maestría en Ciencias Genómicas; ⁴Departamento de Bioquímica-Cinvestav
elisa.azuara@uacm.edu.mx

Abstract:

Gene expression is a highly dynamic process that involves many regulating proteins to enable RNA pol II to access promoters. Therefore, several transcription factors (TFs) are needed to activate or inhibit mRNA transcription. During its synthesis, immature mRNA must be processed by co-transcriptional regulation, to allow its exportation to the cytoplasm, where it is translated. Splicing is a complex process that involves almost 150 proteins termed spliceosomal factors (SF); however, additional proteins have been recently identified with a non-splicing function. The Myb transcription factors are proteins recently implicated in transcription regulation and splicing. *Entamoeba histolytica* is an enteric protozoan possessing 8,201 genes, and interestingly almost 30% contain at least one intron, suggesting the presence of co-transcriptional regulation. In addition, EhMyb proteins are the most abundant TFs in this protozoan, which allow us to hypothesize that could regulate some virulence factors needed for host invasion. We previously characterized EhMyb10 and determined the DNA-binding domain (DBD) through EMSA assays. In this work, we performed an *in silico* interactome between EhMyb10 and the *E. histolytica* proteome to determine putative factors binding to this TF. Through the STRING website, we observed that EhMyb10 binds to both transcription and splicing-related proteins. The interactome showed 31 nodes and 260 edges with an e-value of 1.0e-16. Interestingly, one of the proteins bound to EhMyb10 is EhCDC5-like. This protein belongs to the Myb family and has two imperfect R2/R3 repeats, although in *Schizosaccharomyces pombe* its function has been related to spliceosome activation rather than transcription. We also analyzed the primary structure of EhCDC5-like and found a disordered region, as well as, many predicted phosphorylation residues that may be involved in the regulation of the spliceosome active state. In the EhCDC5-like interactome, we found 41 nodes and 756 edges, indicating that the interaction of putative proteins is more likely than those of EhMyb10 binding proteins. *In silico*, EhCDC5-like binds to many SFs and interestingly to chromatin remodeling proteins. In conclusion, through this bioinformatic analysis, we propose a role for EhMyb10 and EhCDC5-like as a complex during co-transcriptional regulation in this parasite.

This work is funded by CONACYT (grant CF-2019-194163)

EhMyb10 Transcription Factor Interactome: evidence of cotranscriptional regulation

Danna Paola Aguirre Casimiro¹, Patricia Cuellar², Claudia Castillo Mercado², Elizabeth J. Castañeda-Ortiz³, Landy Mariene Rojas Áviles¹, Jesús Valdés³, Elisa Azuara Liceaga^{1,2*}

¹Licenciatura en Ciencias Genómicas; ²Posgrado en Ciencias Genómicas, Universidad Autónoma de la Ciudad de México, UACM, Piantel del Valle; Calle San Lorenzo 290, Col del Valle Sur, Benito Juárez, 03100 Ciudad de México, CDMX. 55 5488 6661

³Departamento de bioquímica CINVESTAV-IPN
*Corresponding author: elisa.azuara@uacm.edu.mx



Every amoeba has a story to tell...

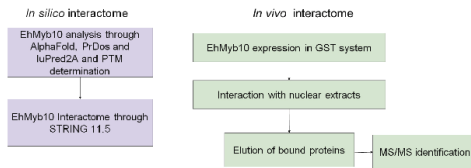
INTRODUCTION

Gene expression is a highly dynamic process that involves many regulating proteins that enable RNA pol II access to the promoter. Therefore, several proteins or transcription factors (TFs) are needed to activate or inhibit transcription of mRNA. TFs recognize a specific DNA sequence due to their secondary structure that can be zinc fingers, helix-turn-helix among others. While synthesis, immature mRNA must be processed in a process termed cotranscriptional regulation, to allow its exportation to the cytoplasm for its translation. Splicing is a complex process that involves almost 150 proteins termed spliceosomal factors (SF); however, additional proteins have been recently identified with a non-splicing function. Myb TFs are proteins that have been recently implicated in a variety of functions such as telomere protection, transcription regulation and splicing. Interestingly *E. histolytica* is an enteric protozoan that has 8,201 genes where almost 30% of the genes have at least one intron, suggesting the presence of a cotranscriptional regulation. *E. histolytica* is responsible for amebiasis, the third leading cause of death among protozoan parasitic diseases that affects approximately 500 million people could be infected worldwide. In addition, Myb TFs are the most abundant transcription factors in this protozoan and therefore might regulate many of the virulence properties needed for invasion. We previously characterized the TF EhMyb10 and determined the DNA Binding domain (DBD) through EMSA assays. However, besides binding to DNA, TFs must bind also to proteins to allow transcription.

OBJECTIVE

Determine the factors or proteins that bind to EhMyb10 to contribute to the regulation of gene expression

METHODOLOGY



RESULTS: In silico interactome

1 EhMyb10 has a R2R3 Myb DBD and disordered regions
Structure predictors showed disordered regions in the N- and C- terminal of EhMyb10 that may contribute to protein binding, it is also subject to many posttranslational modifications.

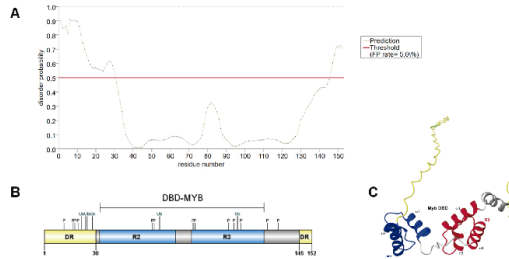


Figure 1. EhMyb10 contains disordered regions. A) EhMyb10 analysis through PiDoss 2.0 indicating the disordered regions. B) Schematic representation of the regions present in EhMyb10. Yellow box indicates the position of the disordered region, blue boxes represent the R2/R3 domain. C) 3-D structure of MYB-DBD obtained by AlphaFold, yellow indicates the disordered regions, R2 and R3 indicate the 2 MYB repeats and their three-helix structure

2 EhMyb10 interactome
STRING predicted that EhMyb10 binds to transcription, splicing proteins and proteins involved in chromatin organization.

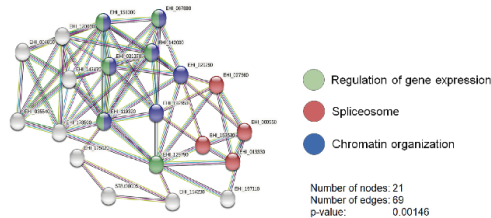


Figure 2. EhMyb10 interacts with splicing involved proteins. Interactome of EhMyb10 through STRING 11.5 with *E. histolytica* proteins, in green, blue and red are indicated proteins involved in transcription, splicing and chromatin organization respectively.

RESULTS: In vivo interactome

3 Nuclear proteins bind to EhMyb10
Recombinant EhMyb10 was used in pulldown assays using nuclear extracts from *E. histolytica* in which differential bands were observed and analyzed by IMS-D/A-MS.

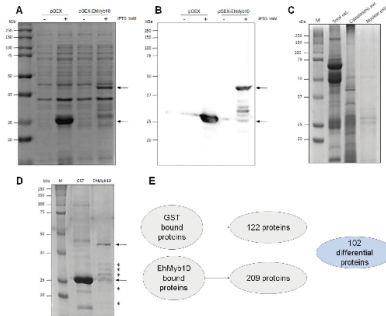


Fig. 3. EhMyb10 interactome *in vivo*. A) EhMyb10 recombinant protein was expressed in *E. coli* BL21 rosettas with IPTG 1mM. GST was used as negative control. B) A duplicate from A was transferred to nitrocellulose and was used for WB with anti-GST antibody. Arrowheads indicate the identification of GST and EhMyb10-GST at 26 and 44 kDa respectively. C) *E. histolytica* trophozoites were harvested and cellular extracts were obtained. D) Proteins bound to GST or EhMyb10 were eluted and analyzed through SDS-PAGE 12%. Asterisks indicate the bands observed only in the EhMyb10 fraction.

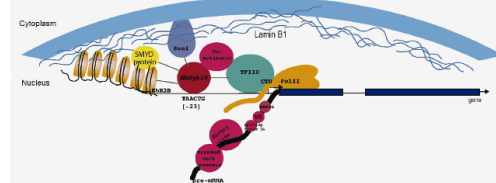
4 EhMyb10 may participate in cotranscriptional events

Table 1. Proteins identified by MS that interact with EhMyb10

Probable function	Number of proteins
RNA binding	13
Chromatin organization	6
Transcription related	6
Protein binding	6
Nuclei structure	1
Splicing	1

Protein identification was realized in the Laboratorio Nacional de Servicios Experimentales, Unidad de Genómica, Proteómica y Metabómica (CINVESTAV-IPN).

EhMyb10 protein-protein interaction model



REFERENCES

Meneses, E., Cárdenas, H., Zárate, S., Brieba, L. G., Orozco, E., López-Camantillo, C. & Azuara-Liceaga, E. (2010). The R2R3 Myb protein family in *Entamoeba histolytica*. *Gene*, 455(1-2), 32-42. <https://doi.org/10.1016/j.gene.2010.02.004>

Haque, F., Mazzeo, D., Patel, J. T., Smallwood, D. T., Ellis, J. A., Shanahan, C. M., & Shackleton, S. (2010). Mammalian SUN protein interaction networks at the inner nuclear membrane and their role in laminopathy disease processes. *The Journal of biological chemistry*, 285(5), 3487-3498. <https://doi.org/10.1074/jbc.M109.071910>

Szkarczyk, D., Gabie, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N. T., Legeay, M., Fang, T., Bork, P., Jensen, L. J., & von Mering, C. (2021). The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic acids research*, 49(D1), D605-D612. <https://doi.org/10.1093/nar/nkaa1074>

Acknowledgements

This work was supported by CONACYT CF-2019-194163 and PC is grateful to CONACYT Estancias Posdoctorales por México. The authors thank the Universidad Autónoma de la Ciudad de México (UACM) for the constant support.



Sociedad Mexicana de
Bioquímica A.C.



SACRAMENTO NO. 413
COL. INSURGENTES BORJA
DEL. BENITO JUÁREZ
CP. 03100
CDMX
TEL. (55)56225742
WWW.SMB.ORG.MX
FACEBOOK:
facebook.com/RedSMB

MESA DIRECTIVA 2021-2023

PRESIDENTA
Dra. Soledad María Teresa Hernández
Sotomayor

VICEPRESIDENTE
Dr. Agustín Guerrero Hernández

SECRETARIA TESORERA
Dra. Bertha María Josefina González Fedrajo

SUBSECRETARIA TESORERA
Dra. María de Lourdes Girard Cuzey

SOCIOS FUNDADORES

Dr. Barbarín Arreguín Lozano
Dr. Edmundo Calva Cuadrilla
Dr. Guillermo Carvajal Sandoval
Dr. Joaquín Cravioto
Dr. Carlos del Río Estrada
Dr. Silvestre Frenk Freund
Dr. Mario García Hernández
Dr. Jesús Guzmán García
Dr. Jesús Kumate Rodríguez
Dr. José Laguna García
Dr. Guillermo Massieu Higuera
Dr. Raúl Ondarza Vidaurreta
Dr. Efraín G. Pardo Codina
Dr. Guillermo Soberón Acevedo

Se otorga la presente

CONSTANCIA a:

Danna Paola Aguirre Casimiro

Quien asistió y presentó el trabajo:

**EhMyb10 Transcription Factor Interactome: evidence
of cotranscriptional regulation**

Por: Danna Paola Aguirre Casimiro, Patricia Cuellar,
Claudia Castillo Mercado, Elizabeth J. Castañeda-Ortiz,
Landy Marlene Rojas Avilés, Jesús Valdés, Elisa Azuara Liceaga

En la modalidad de cartel durante el
XXXIII Congreso Nacional de Bioquímica
16 - 21 de octubre de 2022, Mérida, Yucatán

Atentamente
Por el Comité Organizador

Dra. Teresa Hernández Sotomayor
Presidenta





LA UNIVERSIDAD AUTÓNOMA DE LA CIUDAD DE MÉXICO
OTORGA LA PRESENTE

C O N S T A N C I A

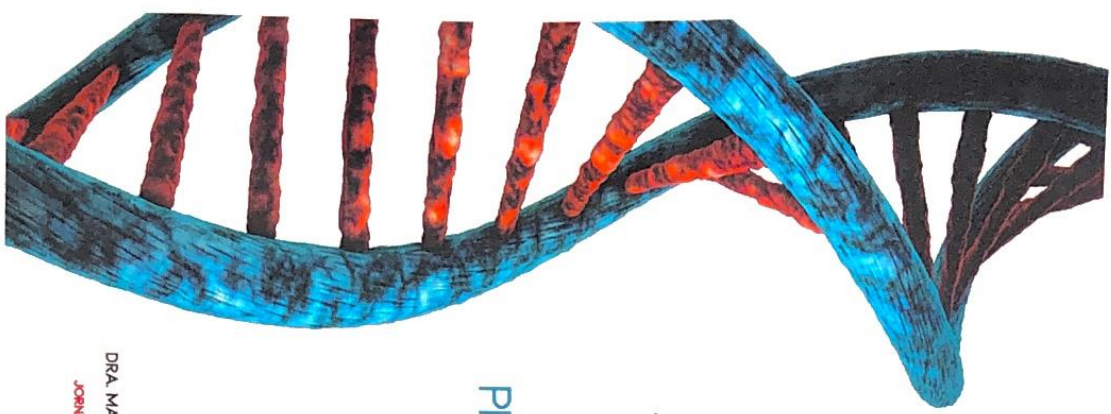
A: Aguirre Casimiro Danna Paola Ixyalit

POR SU PARTICIPACIÓN EN LAS

**PRIMERAS JORNADAS DE DIVULGACIÓN DE LA CIENCIA
DE LA LICENCIATURA EN CIENCIAS GENÓMICAS**

DIRIGIDA AL PÚBLICO EN GENERAL PARA PROMOVER
LA DIVULGACIÓN Y EL ACCESO UNIVERSAL AL CONOCIMIENTO

MAYO 23 DEL 2023




DRA. MARÍA ELIZBETH ALVAREZ SANCHEZ
COORDINADORA DE LAS
JORNADAS DE DIVULGACIÓN DE LA CIENCIA


C. FRANCISCO DÍAZ BAUTISTA
ENCARGADO GENERAL DE LA
COORDINACIÓN DEL PLANTEL DEL VALLE


DRA. SELENE ZARATE GUERRA
ENLACE DE LA LICENCIATURA EN
CIENCIAS GENÓMICAS

SISTEMA BIBLIOTECARIO DE INFORMACIÓN Y DOCUMENTACIÓN



UNIVERSIDAD AUTÓNOMA DE LA CIUDAD DE MÉXICO COORDINACIÓN ACADÉMICA

RESTRICCIONES DE USO PARA LAS TESIS DIGITALES

DERECHOS RESERVADOS[©]

La presente obra y cada uno de sus elementos está protegido por la Ley Federal del Derecho de Autor; por la Ley de la Universidad Autónoma de la Ciudad de México, así como lo dispuesto por el Estatuto General Orgánico de la Universidad Autónoma de la Ciudad de México; del mismo modo por lo establecido en el Acuerdo por el cual se aprueba la Norma mediante la que se Modifican, Adicionan y Derogan Diversas Disposiciones del Estatuto Orgánico de la Universidad de la Ciudad de México, aprobado por el Consejo de Gobierno el 29 de enero de 2002, con el objeto de definir las atribuciones de las diferentes unidades que forman la estructura de la Universidad Autónoma de la Ciudad de México como organismo público autónomo y lo establecido en el Reglamento de Titulación de la Universidad Autónoma de la Ciudad de México.

Por lo que el uso de su contenido, así como cada una de las partes que lo integran y que están bajo la tutela de la Ley Federal de Derecho de Autor, obliga a quien haga uso de la presente obra a considerar que solo lo realizará si es para fines educativos, académicos, de investigación o informativos y se compromete a citar esta fuente, así como a su autor ó autores. Por lo tanto, queda prohibida su reproducción total o parcial y cualquier uso diferente a los ya mencionados, los cuales serán reclamados por el titular de los derechos y sancionados conforme a la legislación aplicable.