

# 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  
LICENCIATURA EN CIENCIAS GENÓMICAS

## **Análisis de los genes implicados en la maquinaria de la replicación del DNA en *Entamoeba histolytica***

### **TESIS**

QUE PARA OPTAR POR EL TÍTULO DE LICENCIADO EN:

**CIENCIAS GENÓMICAS**

PRESENTA:

**ALAN JESÉ RODRÍGUEZ ALMARAZ**

DIRECTORA DE TESIS  
**Dra. Elisa Irene Azuara Liceaga**

CODIRECTORA  
**Dra. Patricia Areli Cuellar Silva**

Ciudad de México, Octubre de 2025

## **SISTEMA BIBLIOTECARIO DE INFORMACIÓN Y DOCUMENTACIÓN**



### **UNIVERSIDAD AUTÓNOMA DE LA CIUDAD DE MÉXICO COORDINACIÓN ACADÉMICA**

#### **RESTRICCIONES DE USO PARA LAS TESIS DIGITALES**

#### **DERECHOS RESERVADOS<sup>©</sup>**

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: Dr. Máximo Berto Martínez Benítez

Secretaria: Dra. Elisa Irene Azuara Liceaga

Vocal: Dra. Patricia Areli Cuellar Silva

**Plantel de Adscripción:**

**PLANTEL DEL VALLE, UACM**

**DIRECTORA DE TESIS**

---

Dra. Elisa Irene Azuara Liceaga  
Universidad Autónoma de la Ciudad de México

**CODIRECTORA**

---

Dra. Patricia Areli Cuellar Silva  
Universidad Autónoma de la Ciudad de México

## **LECTORES DE TESIS**

**Dra. Lilia López Canovas**

Posgrado en Ciencias Genómicas  
Universidad Autónoma de la Ciudad de México

**Dr. Máximo Berto Martínez Benítez**

Posgrado en Ciencias Genómicas  
Universidad Autónoma de la Ciudad de México

**Dra. Silvia Castellanos Castro**

Academia de Biología  
Universidad Autónoma de la Ciudad de México

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 codirección con la Dra. Patricia Areli Cuellar Silva. Este proyecto contó con el apoyo de los proyectos UACM-CCYT-2025-CON-05 y CBF-2025-I-3795 proporcionado por el Colegio de Ciencia y Tecnología de la UACM y la SECIHTI, respectivamente, además de la beca de Estancia Postdoctoral SECIHTI otorgada a la Dra. Cuellar (CVU 333090).

## AGRADECIMIENTOS

A la **Universidad Autónoma de la Ciudad de México**, por brindarme la oportunidad de acceder a una educación superior de calidad, permitirme desarrollarme académicamente y concluir mis estudios. Agradezco el apoyo institucional que me permitió llevar a cabo la presente tesis, así como las becas otorgadas: De Nuevas Becas Para Estudios De Licenciatura y Obtención Para La Elaboración De Tesis (Folio: BTL-2024-I-74279), la cual fue fundamental para continuar con mi formación profesional.

A la **Dra. Elisa Irene Azuara Liceaga**, le agradezco por aceptarme en su laboratorio, por brindarme su apoyo, atención, paciencia y enseñanzas a lo largo de mi proyecto de investigación, también por transmitirme siempre confianza, motivación y entusiasmo por la investigación. Su apoyo y guía han sido fundamentales para el desarrollo de este proyecto y para mi formación académica.

A la **Dra. Patricia Areli Cuellar Silva**, por su apoyo en los experimentos, paciencia, enseñanzas y acompañamiento durante este proceso. Su disposición para compartir sus conocimientos, así como su orientación y consejos, enriquecieron mi experiencia en el laboratorio y contribuyeron de manera significativa a mi formación académica. También valoro los buenos momentos y risas compartidas.

A la **Dra. Elizabeth Jacqueline Castañeda Ortiz**, por su apoyo en los experimentos, paciencia, consejos y enseñanzas a lo largo de este proceso, así como por compartir sus conocimientos y experiencias que fueron fundamentales tanto para mi trabajo en el laboratorio como para mi formación académica. También agradezco los buenos momentos, que hicieron de esas horas de trabajo una experiencia más enriquecedora y sobre todo divertida.

A la **Dra. Abigail Betanzos Fernández**, del Departamento de Infectómica y Patogénesis Molecular, por su ayuda en la observación de las muestras por medio de microscopía confocal e interpretación de resultados.

A **Dora Angélica Silva Olivares**, por su ayuda en la observación de las muestras por medio de microscopía confocal, auxiliar de la Dra. Abigail Betanzos del Departamento de Infectómica y Patogénesis Molecular del CInvestav.

Al **Dr. Israel Canela-Pérez** del Departamento de Bioquímica del CInvestav, determinación *in silico* de las estructuras de las proteínas.

Al **M. en C. Carlos Ernesto Martínez Rodríguez** de la Academia de Matemáticas del CCyT, UACM por su ayuda en el análisis de los transcriptomas y la obtención del mapa de calor.

A **Alfredo Padilla Barberi** por su apoyo en el diseño gráfico del póster presentado en XXXIII Congreso Nacional de Parasitología: CONAPAR-2024 "Una sola salud".

A mis compañeros y amigos del grupo de trabajo del laboratorio del posgrado en Ciencias Genómicas, **Claudia, Marlene, Danna, César, Aldair, Jose Luis, Montse,**

**Fernanda, Karen y Dulce**, gracias por su compañía y los buenos momentos divertidos, que hicieron de este proceso una experiencia más enriquecedora.

A mis padres **Josefina Almaraz Marín y Martin Rodríguez Flores**, por su amor, guía y apoyo en cada paso de mi vida. Gracias por enseñarme a creer en mí mismo, por estar a mi lado en los momentos difíciles y por celebrar conmigo cada logro. Este triunfo también es suyo, y siempre será un reflejo de los valores y la fortaleza que me han transmitido.

A mis hermanos **Gabriela Rodríguez Almaraz y Fabricio Rodríguez Almaraz**, por su compañía, apoyo y alegría en cada etapa de mi vida. Gracias por compartir conmigo risas, consejos y momentos difíciles, y por recordarme siempre la importancia de la familia. Este logro también es suyo, y refleja el cariño y la unión que hemos construido juntos.

A mis sobrinos **Sebastian y Valentina**, por llenar mi vida de alegría, ternura y risas. Gracias por enseñarme a disfrutar los pequeños momentos y por recordarme, con su inocencia y entusiasmo, que cada logro es más especial cuando se comparte con quienes se ama.

A **Amaris Luna Ramírez Torres**, por ser mi apoyo incondicional, mi inspiración y mi refugio en cada etapa de este camino. Gracias por tu amor, paciencia y comprensión, por acompañarme en los momentos difíciles, por motivarme con tu cariño cuando más lo necesitaba y por celebrar conmigo cada logro con entusiasmo y alegría. Eres una de las razones más lindas por las que todo ha valido la pena.

A **Samari Hernández Martínez**, a quien conozco desde la preparatoria y que se ha convertido en alguien muy especial en mi vida. Gracias por tu amistad sincera, por el apoyo mutuo en todo momento y por estar siempre presente en las buenas, en las malas y en las peores. Tu compañía y cariño han sido una gran fortaleza para mí, y me siento muy afortunado de contar contigo.

A **Norma Angelica Lira Rodríguez**, por acompañarme a lo largo de la carrera con su alegría, comprensión y compañía constante. Gracias por estar a mi lado en cada momento y por compartir conmigo tantas experiencias, por las risas, por hacer de esta etapa un recuerdo inolvidable y me siento muy afortunado de contar contigo.

A **Mónica Guadalupe Pérez Rebolledo**, por estar siempre a mi lado en cada momento. Gracias por escucharme, por tus consejos y por recordarme con tus palabras que soy capaz, incluso en los momentos en que yo dudaba de mí. Tu apoyo ha sido fundamental en mi vida académica y personal, y me siento muy afortunado de contar contigo.

A mis amigos **Luis, Lin Lin, Yess, Jess, Nadia, Adrian, Montse, Aylin, Mario y Diana** por su apoyo, su compañía, las risas compartidas y los momentos compartidos a lo largo de este camino. Gracias por estar presentes en cada momento y por hacer que cada logro sea más especial con su amistad.

A mis mascotas **Pinky, Wendy, Pichirilo y Tiburoncin** por acompañarme siempre y darme tanto amor.

## **DEDICATORIA**

A mis padres Josefina Almaraz Marín y Martin Rodríguez Flores, por su amor infinito, su paciencia y su guía constante a lo largo de mi vida. Gracias por enseñarme a creer en mí mismo, por sostenerme en los momentos difíciles y por celebrar conmigo cada logro, por pequeño que fuera. Gracias por los valores que me han transmitido, por su ejemplo de fuerza y perseverancia, y por mostrarme que con dedicación y amor todo es posible. Todo lo que soy y he alcanzado se lo debo a ustedes; este triunfo es tan suyo como mío.

## ÍNDICE

|                                                                                                |    |
|------------------------------------------------------------------------------------------------|----|
| ÍNDICE DE TABLAS                                                                               | 3  |
| ÍNDICE DE FIGURAS                                                                              | 4  |
| Abreviaturas                                                                                   | 5  |
| Resumen                                                                                        | 7  |
| Abstract                                                                                       | 8  |
| 1. Introducción                                                                                | 9  |
| 1.1 <i>Entamoeba histolytica</i>                                                               | 9  |
| 1.2 Amebiasis y ciclo de vida de <i>E. histolytica</i>                                         | 9  |
| 1.3 Epidemiología de la amebiasis                                                              | 10 |
| 1.4 Genoma de <i>E. histolytica</i>                                                            | 11 |
| 1.5 Proceso de división celular de <i>E. histolytica</i>                                       | 13 |
| 2. Antecedentes particulares                                                                   | 15 |
| 2.1 Replicación del DNA en organismos eucariontes                                              | 15 |
| 2.2 Proteínas involucradas en la replicación del DNA en otros protozoarios                     | 20 |
| 2.3. Replicación del DNA en <i>E. histolytica</i> , identificación de secuencias Ori           | 20 |
| 2.4. Identificación de proteínas implicadas en la replicación del DNA en <i>E. histolytica</i> | 21 |
| 3. Hipótesis                                                                                   | 27 |
| 4. Justificación                                                                               | 27 |
| 5. Objetivos                                                                                   | 27 |
| 5.1 Objetivo general                                                                           | 27 |
| 5.2 Objetivos particulares                                                                     | 27 |
| 6. Estrategia experimental                                                                     | 28 |
| 7. Materiales y métodos                                                                        | 29 |
| 7.1 Búsqueda y alineamiento de secuencias                                                      | 29 |
| 7.2 Identificación de dominios funcionales y características de las proteínas                  | 29 |
| 7.3 Análisis de las interacciones proteína-proteína                                            | 29 |
| 7.4 Análisis de los patrones de expresión                                                      | 30 |
| 7.5 Diseño de primers                                                                          | 30 |
| 7.6 Cultivo de <i>E. histolytica</i> HM1:IMSS y ensayos de sincronización                      | 30 |
| 7.7 Extracción de DNA genómico                                                                 | 33 |
| 7.8 Extracción de RNA                                                                          | 33 |
| 7.9 Tratamiento con DNAsa                                                                      | 34 |
| 7.10 Síntesis de cDNA                                                                          | 34 |
| 7.11 Amplificación de los genes por RT-PCR                                                     | 34 |

|                                                                                                                  |    |
|------------------------------------------------------------------------------------------------------------------|----|
| 7.12 Electroforesis en gel de agarosa                                                                            | 35 |
| 7.13 Análisis del nivel de expresión relativa de cada gen seleccionado                                           | 35 |
| 7.14 Ensayo de Inmunofluorescencia indirecta (IFI)                                                               | 35 |
| 7.15 Análisis de la intensidad de la fluorescencia nuclear                                                       | 36 |
| 8. Resultados                                                                                                    | 37 |
| 8.1 Identificación de las proteínas que conforman el replisoma de <i>E. histolytica</i>                          | 37 |
| 8.1.1 Etapa Pre-RC                                                                                               | 37 |
| 8.1.2 Etapa Pre-IC                                                                                               | 42 |
| 8.1.3 Etapa de síntesis (S)                                                                                      | 49 |
| 8.1.4 Etapa terminación (T)                                                                                      | 57 |
| 8.2 Interactoma del replisoma de <i>E. histolytica</i>                                                           | 62 |
| 8.3 Expresión diferencial de genes del replisoma durante la depleción y reposición de suero.                     | 66 |
| 8.4 Ensayo de sincronización celular de trofozoítos de <i>E. histolytica</i> mediante RT-PCR                     | 69 |
| 8.4.1 Análisis de la expresión de genes que conforman la etapa Pre-RC                                            | 69 |
| 8.4.2 Análisis de la expresión de genes que conforman la etapa Pre-IC                                            | 71 |
| 8.4.3 Análisis de la expresión de genes que conforman la etapa S                                                 | 72 |
| 8.4.4 Análisis de la expresión de genes que conforman la etapa T                                                 | 74 |
| 8.5 Expresión y localización de EhPCNA durante la sincronización celular de trofozoítos de <i>E. histolytica</i> | 74 |
| 9. Discusión                                                                                                     | 77 |
| 10. Conclusiones                                                                                                 | 88 |
| 11. Bibliografía                                                                                                 | 90 |
| 12. Anexos                                                                                                       | 98 |

## ÍNDICE DE TABLAS

|                                                                                                                                                                                          |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Tabla 1. Características generales del genoma de <i>E. histolytica</i>                                                                                                                   | 12 |
| Tabla 2. Duración del ciclo celular de <i>E. histolytica</i>                                                                                                                             | 14 |
| Tabla 3. Programas utilizados para las características de las proteínas del replisoma de <i>E. histolytica</i>                                                                           | 29 |
| Tabla 4. Oligonucleótidos diseñados para ensayos de RT-PCR                                                                                                                               | 31 |
| Tabla 5. Comparación de los componentes de la etapa Pre-RC de la replicación del DNA de <i>E. histolytica</i> con respecto a los organismos <i>H. sapiens</i> y <i>S. pombe</i>          | 38 |
| Tabla 6. Dominios funcionales y estructuras tridimensionales de las proteínas involucradas en la replicación de la etapa Pre-RC                                                          | 40 |
| Tabla 7. Comparación de los componentes de la etapa Pre-IC de la replicación del DNA de <i>E. histolytica</i> con respecto a los organismos <i>H. sapiens</i> y <i>S. pombe</i>          | 43 |
| Tabla 8. Dominios funcionales y estructuras tridimensionales de las proteínas involucradas en la replicación de la etapa Pre-IC                                                          | 46 |
| Tabla 9. Comparación de los componentes de la etapa síntesis de la replicación del DNA de <i>E. histolytica</i> con respecto a los organismos <i>H. sapiens</i> y <i>S. pombe</i>        | 50 |
| Tabla 10. Dominios funcionales y estructuras tridimensionales de las proteínas involucradas en la replicación de la etapa de síntesis                                                    | 54 |
| Tabla 11. Comparación de los componentes de la etapa de terminación de la replicación del DNA de <i>E. histolytica</i> con respecto a los organismos <i>H. sapiens</i> y <i>S. pombe</i> | 58 |
| Tabla 12. Dominios funcionales y estructuras tridimensionales de las proteínas involucradas en la replicación de la etapa de terminación                                                 | 60 |

## ÍNDICE DE FIGURAS

|                                                                                                                                                                                  |    |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 1. Ciclo de vida de <i>E. histolytica</i>                                                                                                                                 | 10 |
| Figura 2. Modelo de replicación en organismos eucariontes                                                                                                                        | 15 |
| Figura 3. Modelo de replicación en <i>Saccharomyces</i> durante la gemación                                                                                                      | 17 |
| Figura 4. Replisoma en organismos eucariontes                                                                                                                                    | 19 |
| Figura 5. Árbol filogenético de las proteínas ORC en organismos eucariontes                                                                                                      | 22 |
| Figura 6. Estructuras tridimensionales e interacción entre EhOrc1/Cdc6 y EhCdc45                                                                                                 | 23 |
| Figura 7. Estructura tridimensional de CDC45 de <i>E. histolytica</i>                                                                                                            | 24 |
| Figura 8. Estructura tridimensional de EhPCNA de <i>E. histolytica</i>                                                                                                           | 25 |
| Figura 9. Estructura tridimensional de EhDNA ligasa de <i>E. histolytica</i>                                                                                                     | 26 |
| Figura 10. Condiciones del termociclador para la amplificación de genes en ensayos de RT-PCR.                                                                                    | 35 |
| Figura 11. Interactoma <i>in silico</i> de las proteínas del replisoma en <i>E. histolytica</i>                                                                                  | 64 |
| Figura 12. Análisis de enriquecimiento funcional del replisoma de <i>E. histolytica</i>                                                                                          | 65 |
| Figura 13. Mapa de calor de la expresión de genes relacionados con la replicación del DNA en <i>E. histolytica</i>                                                               | 68 |
| Figura 14. Análisis de la expresión de genes relacionados con componentes del complejo Pre-RC en <i>E. histolytica</i> mediante RT-PCR semicuantitativa                          | 70 |
| Figura 15. Análisis de la expresión de genes relacionados con componentes del complejo Pre-IC en <i>E. histolytica</i> mediante RT-PCR semicuantitativa                          | 71 |
| Figura 16. Análisis de la expresión de genes relacionados con componentes de la etapa de síntesis en <i>E. histolytica</i> mediante RT-PCR semicuantitativa                      | 73 |
| Figura 17. Análisis de la expresión de genes relacionados con componentes de la etapa de terminación de la replicación en <i>E. histolytica</i> mediante RT-PCR semicuantitativa | 75 |
| Figura 18. Detección de EhPCNA mediante inmunofluorescencia indirecta en trofozoítos de <i>E. histolytica</i> sincronizados                                                      | 76 |
| Figura 19. Modelo de replicación en <i>E. histolytica</i>                                                                                                                        | 89 |

## Abreviaturas

|                       |                                                                                                                                   |
|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------|
| <b>aa</b>             | Aminoácido                                                                                                                        |
| <b>AMP</b>            | Adenosín monofosfato                                                                                                              |
| <b>ARS</b>            | Del inglés Autonomous Replication Sequences                                                                                       |
| <b>AT</b>             | Alto contenido de bases nitrogenadas adenina (A) y timina (T)                                                                     |
| <b>ATP</b>            | Trifosfato de adenosina                                                                                                           |
| <b>ATPasa</b>         | Enzima que cataliza ATP a ADP                                                                                                     |
| <b>BER</b>            | Escisión de bases                                                                                                                 |
| <b>°C</b>             | Grados centígrados                                                                                                                |
| <b>CDC</b>            | Del inglés, Centers for Disease Control and Prevention                                                                            |
| <b>CDC45</b>          | Del inglés, Cell Division Cycle 45 Protein                                                                                        |
| <b>CDC6</b>           | Del inglés, Cell Division Cycle 6 Protein                                                                                         |
| <b>CDK</b>            | Cinasa dependiente de ciclina, del inglés Cyclin dependent Kinase                                                                 |
| <b>cDNA</b>           | DNA de cadena simple                                                                                                              |
| <b>DAPI</b>           | 4',6-diamino-2-fenilindol diclorhidrato                                                                                           |
| <b>DBD</b>            | Dominio de unión a DNA, del inglés "DNA-binding domain"                                                                           |
| <b>DDK</b>            | Cinasa dependiente de DBF4, del inglés DBF4-dependent kinase                                                                      |
| <b>Dmc1</b>           | Recombinasa asociada a recombinación 1                                                                                            |
| <b>DNA</b>            | DNA, Ácido desoxirribonucleico, del inglés deoxyribonucleic acid                                                                  |
| <b>DNAg</b>           | DNA genómico                                                                                                                      |
| <b>DNALig1</b>        | DNA ligasa 1                                                                                                                      |
| <b>DNAPol</b>         | DNA polimerasas                                                                                                                   |
| <b>DNAr</b>           | DNA ribosomal                                                                                                                     |
| <b>dNTPs</b>          | Desoxinucleótidos trifosfato                                                                                                      |
| <b>Fen1</b>           | Endonucleasa Flap 1                                                                                                               |
| <b>GINS</b>           | Complejo conformado por 4 subunidades: Sld5, Psf1, Psf2 y Psf3 que corresponden a los números 5-1-2-3, del japonés Go-Ichi-Ni-San |
| <b>h</b>              | Horas                                                                                                                             |
| <b>H<sub>2</sub>O</b> | Agua                                                                                                                              |
| <b>IFI</b>            | Inmunofluorescencia indirecta                                                                                                     |
| <b>IPI3_WD</b>        | Proteína IPI3 que contiene repeticiones WD                                                                                        |
| <b>kb</b>             | Kilobases                                                                                                                         |
| <b>kDa</b>            | Kilodaltones                                                                                                                      |
| <b>LB</b>             | Medio Luria-Bertani                                                                                                               |
| <b>Mb</b>             | Megabase                                                                                                                          |
| <b>MCM</b>            | De inglés, Minichromosome Maintenance                                                                                             |
| <b>MCM BP</b>         | Proteína de unión a mantenimiento de minicromosomas                                                                               |
| <b>µg</b>             | Microgramos                                                                                                                       |
| <b>min</b>            | Minutos                                                                                                                           |
| <b>mL</b>             | Mililitros                                                                                                                        |
| <b>µL</b>             | Microlitros                                                                                                                       |
| <b>mM</b>             | Milimolar                                                                                                                         |
| <b>µm</b>             | Micrómetro                                                                                                                        |
| <b>µM</b>             | Micromolar                                                                                                                        |
| <b>mRNA</b>           | Ácido ribonucleico mensajero                                                                                                      |
| <b>OMS</b>            | Organización Mundial de la Salud                                                                                                  |

|                 |                                                                                                                                       |
|-----------------|---------------------------------------------------------------------------------------------------------------------------------------|
| <b>ORC</b>      | Complejo de Reconocimiento de Origen                                                                                                  |
| <b>pb</b>       | Pares de bases                                                                                                                        |
| <b>PBS</b>      | Del inglés, Phosphate Buffered Saline                                                                                                 |
| <b>PCNA</b>     | De inglés, proliferating Cell Nuclear Antigen                                                                                         |
| <b>PCR</b>      | Reacción en cadena de la polimerasa, del inglés Polymerase Chain Reaction                                                             |
| <b>PFGE</b>     | Electroforesis de campos pulsados                                                                                                     |
| <b>PM</b>       | Peso molecular                                                                                                                        |
| <b>Pre-IC</b>   | Complejo de Pre-inicio                                                                                                                |
| <b>Pre-LC</b>   | Complejo de Pre-carga                                                                                                                 |
| <b>Pre-RC</b>   | Complejo de Pre-replicación                                                                                                           |
| <b>RFC</b>      | Factor de Replicación C                                                                                                               |
| <b>RNA</b>      | Ácido ribonucleico, del inglés Ribonucleic acid                                                                                       |
| <b>RPA</b>      | Proteína de Replicación A                                                                                                             |
| <b>rRNA</b>     | RNA ribosomal                                                                                                                         |
| <b>RT-PCR</b>   | Reacción en Cadena de la Polimerasa con Transcriptasa Inversa                                                                         |
| <b>RT-qPCR</b>  | Reacción en cadena de la polimerasa de transcripción inversa cuantitativa, del inglés Reverse transcription polymerase chain reaction |
| <b>S</b>        | Síntesis                                                                                                                              |
| <b>T</b>        | Terminación                                                                                                                           |
| <b>TOP</b>      | Topoisomerasa                                                                                                                         |
| <b>TPM</b>      | Transcritos por millón                                                                                                                |
| <b>TRF</b>      | Del inglés, Telomeric Repeat binding Factor                                                                                           |
| <b>TRF-Like</b> | Proteína similar a TRF                                                                                                                |

## Resumen

*Entamoeba histolytica* es responsable de la amibiasis en humanos, que afecta hasta el 3.5% de la población mundial. Ha sido considerado un eucariota primitivo debido a la ausencia de mitocondria y por su retículo endoplásmico y aparato de Golgi atípicos. En eucariotas, la replicación del DNA inicia con el reconocimiento de la secuencia Ori por el complejo ORC, que posteriormente recluta a las helicasas MCM. Otras proteínas involucradas en la replicación en eucariotas son las DNA polimerasas, el complejo GINS, proteínas CDK, PCNA y DNA ligasas. Aunque la maquinaria de replicación está bien caracterizada en humano y levadura, se sabe poco de su conformación en *E. histolytica*. Es por ello que decidimos identificar *in silico* a la maquinaria de replicación de este parásito, mediante la base de datos de AmoebaDB, empleando como sondas las secuencias de aminoácidos de las proteínas de *Homo sapiens* y *Schizosaccharomyces pombe* involucradas en la replicación del DNA mediante la base de datos AmoebaDB. Además, se analizaron los dominios funcionales y la estructura tridimensional de las proteínas identificadas. Los resultados obtenidos mostraron que en el genoma de *E. histolytica* se encuentran codificadas las proteínas necesarias para la replicación de sus cromosomas; sin embargo, existen diferencias en la composición, la cual parece ser más simplificada, ya que solo codifica en su genoma 43 de las ~63 proteínas que conforman el replisoma humano. Por otro lado, analizamos los cambios de expresión de genes característicos de las diferentes etapas, empleando cultivos de trofozoítos privados de suero para sincronizar su replicación y observamos que algunos de ellos modifican su expresión de manera cíclica, sugiriendo dos eventos de duplicación en un lapso de 24 h. Finalmente, evaluamos la localización de EhPCNA, una proteína clave en la replicación del DNA, por medio de ensayos de inmunofluorescencia y observamos, durante la depleción de suero, una abundante cantidad de proteína en el núcleo de los trofozoítos en comparación con la condición basal, para posteriormente cambiar su nivel de expresión conforme transcurre el tiempo después de la adición de suero. En conclusión, en este trabajo identificamos las proteínas que podrían conformar el replisoma de *E. histolytica*, lo que nos permitirá comprender de una manera más completa el proceso de replicación del DNA en este parásito.

## Abstract

*Entamoeba histolytica* is the causative agent of human amebiasis, affecting up to 3.5% of the global population. It is considered a primitive eukaryote due to the absence of mitochondria and the presence of atypical endoplasmic reticulum and Golgi apparatus. In eukaryotes, DNA replication initiates by the recognition of the origin (Ori) sequence by the origin recognition complex (ORC), followed by the recruitment of MCM helicases. Other essential components of the replication machinery include DNA polymerases, the GINS complex, CDKs, PCNA, and DNA ligases. While the replication machinery has been well characterized in humans and yeast, its composition in *E. histolytica* remains poorly understood. Therefore, in this study, we performed an *in silico* identification of DNA replication-related proteins in *E. histolytica* using amino acid sequences from *Homo sapiens* and *Schizosaccharomyces pombe* as queries in the AmoebaDB database. We also analyzed the functional domains and predicted three-dimensional structures of the identified proteins. Our results indicate that the *E. histolytica* genome encodes homologs of key proteins involved in chromosomal replication. However, notable differences were observed: the *E. histolytica* replisome appears to be simplified, encoding only 43 proteins compared to the approximately 63 components of the human replisome. Furthermore, we analyzed gene expression of several replication-associated genes in normal conditions cultured-trophozoites and subjected to serum deprivation to synchronize replication. The expression of these genes followed a cyclic pattern, suggesting two rounds of DNA replication occur within a 24-hour period. Finally, we assessed the localization of PCNA—a key replication marker—via immunofluorescence. We observed increased nuclear accumulation of PCNA in trophozoites under serum deprivation, with changes in localization following serum replenish. In conclusion, our study identifies putative components of the *E. histolytica* replisome and provides insights into the DNA replication dynamics of this parasite, laying the groundwork for further functional characterization.

## 1. Introducción

### 1.1 *Entamoeba histolytica*

*Entamoeba histolytica* es un protozoario anaerobio, no flagelado, pertenece al filo *Sarcomastigophora* y subfilo Sarcodina, clase Lobosea, orden Amoebida y familia Entamoebidae (Clark y Diamond, 2002). *E. histolytica* presenta dos formas: trofozoíto y quiste. Los trofozoítos miden aproximadamente de 15 a 20  $\mu\text{m}$  de diámetro (en un rango de 10 a 60  $\mu\text{m}$ ) (Sodeman, 1996; Centers for Disease Control and Prevention [CDC], 2019), tienen forma irregular pleomórfica y alargada, son anaerobios facultativos y requieren medios axénicos complejos para su crecimiento. En el citoplasma, tienen de uno a varios núcleos con cariosoma central, cromatina periférica y las vacuolas pueden contener eritrocitos del huésped (Ravdin y Stauffer, 2005). Los trofozoítos son la forma activa, patógena e invasiva del parásito; esta forma se encuentra en el lumen del intestino grueso y puede invadir la mucosa intestinal causando daño tisular (Espinosa y Martínez, 2000). Por otro lado, los quistes, que son la fase infectiva y de resistencia (Fotedar *et al.*, 2007), miden aprox. de 12 a 15  $\mu\text{m}$  de diámetro (CDC, 2019), son redondos, están rodeados por una pared de quitina, contienen cuatro o menos núcleos, tienen ensamblajes ribosómicos llamados cuerpos cromatoidales y tienen una masa de glucógeno cuando son inmaduros (Fleta *et al.*, 2000).

### 1.2 Amebiasis y ciclo de vida de *E. histolytica*

*E. histolytica* es responsable de la amebiasis, una enfermedad intestinal de gran impacto en la incidencia de diarrea grave a nivel mundial. La infección comienza cuando una persona ingiere quistes maduros de *E. histolytica* a través de agua o alimentos contaminados con materia fecal. Una vez que los quistes son ingeridos, llegan al intestino delgado y sufren un proceso llamado desenquistamiento, donde los trofozoítos son liberados. Los trofozoítos migran al intestino grueso y se multiplican por fisión binaria (Cano y López, 1987). En el colon, los trofozoítos pueden permanecer como comensales o bien invadir la mucosa, causando úlceras y disentería e incluso provocando enfermedades gastrointestinales severas (Stanley, 2003; Ximénez *et al.*, 2011).

La amebiasis extraintestinal, ocurre después de que los trofozoítos de *E. histolytica*, atraviesan el epitelio intestinal y, a través de los vasos sanguíneos se disemina a otros órganos como hígado, pulmón y cerebro. El hígado es el órgano más frecuentemente afectado, siendo el absceso hepático amebiano el cuadro clínico más común; el cual suele presentarse en aprox. un 5 % de los pacientes con amebiasis intestinal sintomática. El absceso hepático se manifiesta con síntomas de fiebre y dolor en el hipocondrio derecho (Fleta *et al.*, 2000; Fu *et al.*, 2022). Bajo ciertas condiciones, los trofozoítos pueden enquistarse para ser excretados en las heces, condición en la que pueden sobrevivir en el ambiente durante semanas, esto para reiniciar el ciclo e infectar a nuevos huéspedes (Figura 1) (Ravdin y Stauffer, 2005).

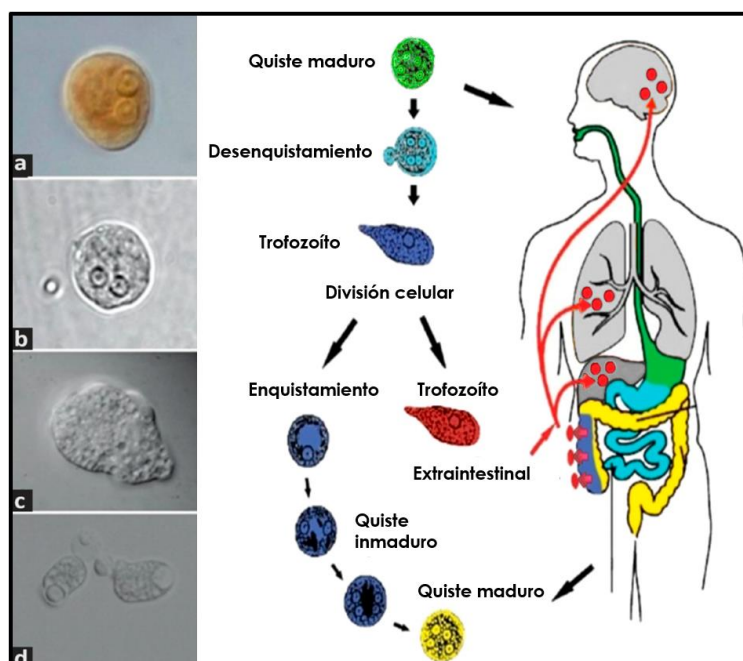


**Figura 1. Ciclo de vida de *E. histolytica*.** El ciclo comienza al ingerir los quistes presentes en agua o alimentos contaminados (a y b; estos se desenquistan (c) y pueden estar como comensales o bien pueden invadir otros órganos o fagocitar eritrocitos (d). Finalmente, algunos trofozoítos pueden enquistar y ser excretados en las heces para volver a comenzar su ciclo de vida. Figura tomada de Ximénez *et al.*, 2011.

### 1.3 Epidemiología de la amebiasis

*E. histolytica* es responsable de un importante problema de salud mundial, ya que es la tercera causa de muerte por infecciones parasitarias que asciende hasta 100,000 muertes al año (Chou y Austin, 2023). El ~90% de las infecciones por *E. histolytica* son asintomáticas; sin embargo, según datos de la Organización Mundial de la Salud (OMS), se estima que, de cada 500 millones de infectados, un 10% aprox. tiene

amebiasis invasiva (Rivero *et al.*, 2021; Shirley *et al.*, 2018). En países desarrollados como Estados Unidos, las infecciones por amebiasis son raras, ya que representan ~5 muertes por año y se observan comúnmente en personas expuestas a áreas endémicas, como inmigrantes o viajeros recientes (Shirley *et al.*, 2018).

La amebiasis intestinal se ha descrito especialmente en regiones tropicales y subtropicales, sobre todo en condiciones sanitarias deficientes (Stanley, 2003). Los países con una alta incidencia de infecciones por amebiasis incluyen regiones de África, India y México, así como también América del Sur (Shirley *et al.*, 2018). En áreas rurales de México, la seroprevalencia de *E. histolytica* ha sido reportada hasta en un 42% (Chou y Austin, 2023). En cuanto a los factores de riesgo de la infección, se relacionan principalmente con la transmisión fecal-oral debido a la mala higiene, la defecación cercana a fuentes de agua como ríos y la cercanía de animales (Gómez *et al.*, 2007).

#### **1.4 Genoma de *E. histolytica***

El genoma de *E. histolytica* ha sido objeto de distintas investigaciones para conocer tanto su información como la regulación de la expresión de sus genes. La secuenciación del genoma se ha logrado mediante distintas aplicaciones y estrategias metodológicas, que métodos como la secuenciación shotgun (Clark *et al.*, 2007), la incorporación de datos de secuenciación de nueva generación, como de la plataforma PacBio, y la técnica de ligación de cromatina Hi-C en el ensamblaje de GenBank (Lieberman-Aiden *et al.*, 2009). La primera secuenciación completa del genoma de *E. histolytica* se realizó utilizando la cepa de referencia HM1:IMSS y fue publicada en 2005 teniendo un ensamble de aprox. 20 Mb (Loftus *et al.*, 2005). Sin embargo, este se reanalizó en 2010 por Lorenzi y col. (2010) determinando que el 49% del genoma corresponde a genes que codifican proteínas (8200), alrededor del 25% de ellos presenta intrones y un 6% de los genes contiene múltiples intrones, lo que sugiere una regulación compleja en este parásito (Tabla 1) (Lorenzi *et al.*, 2010). En cuanto a su contenido de bases, el genoma de *E. histolytica* presenta un porcentaje total de AT del 75% (Lorenzi *et al.*, 2010), lo cual coincide con estimaciones previas que ya se habían reportado (~78%) (Dhar *et al.*, 1996). Actualmente, en el estudio más reciente de Kawano-Sugaya *et al.* (2020), reportó que el genoma es de 26.87 Mb y tiene un total de ~8,734 genes codificantes. Este genoma se organizó en 38 *scaffolds*, los cuales

son estructuras bioinformáticas ensambladas a partir de fragmentos de DNA (contigs) que representan la organización de los cromosomas; ya que existe cierta incertidumbre sobre su ploidía, debido a la dificultad para determinar el número exacto de cromosomas por su falta de condensación (Kawano-Sugaya *et al.*, 2020). Para estimar su número de cromosomas, se ha empleado la electroforesis de campos pulsados (PFGE), permitido determinar que se compone probablemente de entre 31-35 cromosomas (Willhoeft y Tannich, 1999).

En conjunto, estos estudios han contribuido significativamente a la comprensión del genoma de *E. histolytica*. Por otro lado, no se han identificado secuencias canónicas teloméricas, ni el gen de la telomerasa en su genoma (Loftus *et al.*, 2005; Lorenzi *et al.*, 2010). Sin embargo, se han identificado proteínas homólogas a los factores de repetición telomérica (TRF-like), lo que sugiere mecanismos alternativos para la protección y el mantenimiento de los telómeros en este parásito (Rendón-Gandarilla *et al.*, 2018).

**Tabla 1. Características generales del genoma de *E. histolytica***

| <b>Características</b>             | <b>Descripción</b>                                   |
|------------------------------------|------------------------------------------------------|
| Cepa de referencia                 | HM1:IMSS                                             |
| Tamaño del genoma estimado         | Inicial: ~20 MB<br>Más reciente: 26.87 MB            |
| No. estimado de genes codificantes | anterior ~8201 (2005);<br>actualizado a ~8734 (2020) |
| No. estimado de cromosomas         | 31-35 (estimado por PFGE)                            |
| Tamaño de cromosomas               | > 300 kb                                             |
| Organización actual del genoma     | 38 andamios (scaffolds)                              |
| Porcentaje codificante del genoma  | ~49 %                                                |
| Genes con intrones                 | ~25 del total                                        |
| Genes con múltiples intrones       | ~6                                                   |
| Contenido de AT                    | ~75 a 78 %                                           |

Elaborado a partir de Loftus *et al.*, 2005; Lorenzi *et al.*, 2010; Kawano-Sugaya *et al.*, 2020; Dhar *et al.*, 1996; Willhoeft y Tannich, 1999.

### 1.5 Proceso de división celular de *E. histolytica*

Durante la división celular, el DNA debe ser duplicado de manera precisa para asegurar que cada célula hija reciba una copia exacta de la información genética. Este proceso es complejo y está altamente regulado a través de la participación de diversas proteínas (Alberts *et al.*, 2014).

*E. histolytica*, se reproduce de manera asexual mediante fisión binaria; este proceso ocurre en el intestino grueso en humanos, donde los trofozoítos se multiplican antes de transformarse en quistes como parte de su ciclo de vida; cada trofozoíto se divide para poder generar dos células hijas idénticas (Byers *et al.*, 2005; Stanley, 2003). A diferencia de organismos eucariontes, en *E. histolytica* no se ha confirmado la presencia de una mitosis típica; en cambio, sí se ha visto una división nuclear atípica, ya que después del proceso de división celular, las células hijas reciben diferente cantidad de DNA. Por lo tanto, esto podría ser una mitosis cerrada y no convencional con otros organismos (Singh *et al.*, 2013). También se ha visto que, durante el enquistamiento, los núcleos de los trofozoítos pueden dividirse por un proceso mitótico para poder generar quistes con cuatro núcleos. Esto demuestra que la división nuclear es divergente en *E. histolytica* (Byers *et al.*, 2005).

En la actualidad, no se ha reportado meiosis en *E. histolytica*. Sin embargo, estudios previos han identificado la presencia de genes homólogos que están implicados en procesos meióticos. Estos genes codifican para enzimas semejantes a la recombinasa asociada a recombinación (Dmc1) y Spo11 que participan en procesos de meiosis en otros organismos eucariontes. Esto ha generado la hipótesis de que podrían existir procesos ancestrales de recombinación o de meiosis críptica, es decir, una meiosis que ocurre de manera oculta o de manera no convencional e incluso que ocurre bajo ciertas condiciones (Ramesh *et al.*, 2005; Weedall y Hall, 2015). En cuanto a la duración del ciclo celular completo, ha sido estimado aprox. entre 13 a 15 horas (Tabla 2).

**Tabla 2. Duración del ciclo celular de *E. histolytica***

| <b>Fase del ciclo celular</b> | <b>Duración aproximada</b> | <b>Descripción</b>                                                               |
|-------------------------------|----------------------------|----------------------------------------------------------------------------------|
| G1                            | 3 a 4 h                    | Fase previa a la síntesis del DNA; crecimiento celular.                          |
| S                             | 4 a 6 h                    | Síntesis del DNA (replicación)                                                   |
| G2                            | 3 a 5 h                    | Preparación para la división; síntesis de proteínas asociadas.                   |
| M                             | 1 a 2 h                    | Ocurre la mitosis: división nuclear sin desorganización de la envoltura nuclear. |
| <b>Duración total</b>         | 13 a 15 h                  | Ciclo completo                                                                   |

Elaborada a partir de Orozco *et al.*, (1988).

En otros estudios se ha observado una notable heterogeneidad celular en *E. histolytica*; ya que hay trofozoítos multinucleados, anucleados y núcleos con contenido variable de DNA (Mukherjee *et al.*, 2009). Esta variación en el contenido de DNA de este parásito se atribuye a un desacoplamiento temporal y espacial entre la citocinesis y el ciclo nuclear, es decir, la división del citoplasma puede no coincidir con la división del núcleo, lo que da lugar a células hijas desiguales (Mukherjee *et al.*, 2009).

Además, *E. histolytica* presenta múltiples rondas durante su replicación, sin división nuclear ni tampoco citocinesis, lo que resulta en núcleos poliploides, es decir, los núcleos contienen más de dos juegos de cromosomas (Lohia, 2007; Mukherjee *et al.*, 2009).

Por otra parte, *E. histolytica* tiene husos mitóticos multipolares y asimétricos, lo que sugiere una división celular no convencional (Das y Lohia, 2002). La organización del ciclo celular en *E. histolytica* podría conferir una ventaja adaptativa en ambientes cambiantes, para permitir una variación genética sin la necesidad de requerir un ciclo sexual definido (Gangopadhyay *et al.*, 1997; Lohia *et al.*, 2007; Mukherjee *et al.*, 2009).

## 2. Antecedentes particulares

### 2.1 Replicación del DNA en organismos eucariontes

Aunque las proteínas implicadas en la replicación del DNA no han sido bien descritas en *E. histolytica*, este proceso ha sido bien caracterizado en organismos como *H. sapiens* y en levaduras como *Saccharomyces cerevisiae*. En organismos eucariontes, este proceso consta de cuatro etapas principales: 1) pre-replicación (Pre-RC), 2) pre-iniciación (Pre-IC), 3) síntesis de DNA y 4) terminación de la replicación (Figura 2) mientras que en *Saccharomyces*, durante la gemación, el proceso de replicación presenta una etapa adicional después de la Pre-RC denominada Pre-LC (complejo de precarga) (Figura 3) (Li y Araki, 2013). La replicación está estrictamente coordinada con el ciclo celular para poder asegurar la integridad del genoma (Alberts *et al.*, 2014; Watson y Harrison, 2016). En eucariontes, para que ocurra la replicación, se deben reconocer secuencias ricas en AT que funcionan como sitios de inicio de replicación denominados ARS (del inglés Autonomous Replication Sequences) (Stillman, 1996).

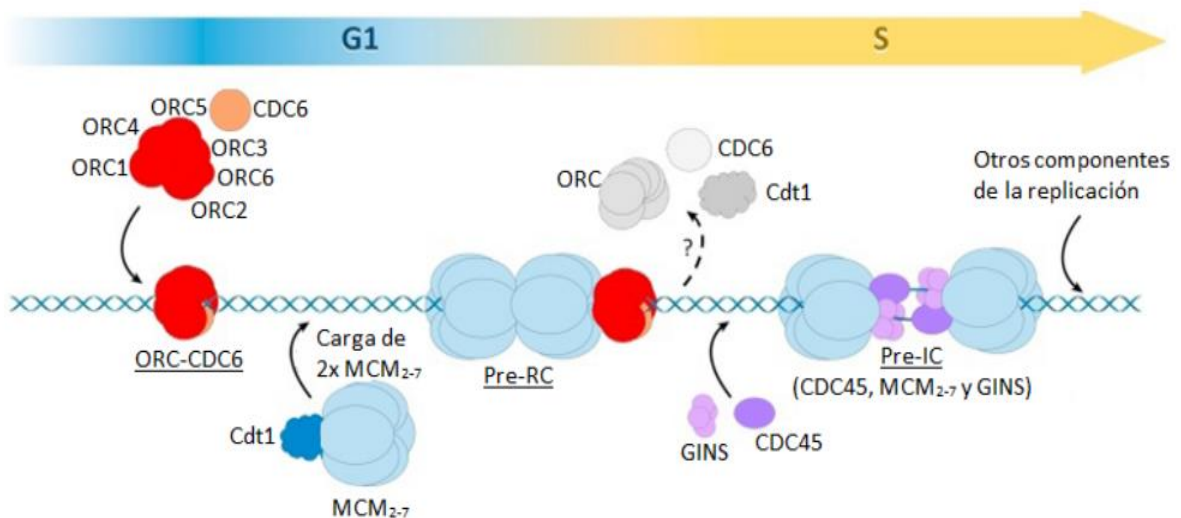


**Figura 2. Modelo de replicación en organismos eucariontes.** Esquema del proceso de replicación en eucariontes. Durante la fase G1 hasta la fase S, el complejo ORC se recluta en la región Ori. ORC interactúa con CDC6, después, la helicasa MCM 2-7 es reclutada por Cdt1 formando el complejo Pre-RC. Al comienzo de la fase S del ciclo celular, CDC6 y Cdt1 se eliminan del pre-RC, entonces CDC45 y GINS se reclutan en el origen, formando el complejo pre-IC que participa en el desenrollamiento del DNA, permitiendo el reclutamiento de los demás componentes de la horquilla de replicación. Figura tomada de M. da Silva *et al.*, 2017.

Estas secuencias son reconocidas durante la etapa Pre-RC en la fase G1 del ciclo celular. Entre las proteínas que forman parte de esta etapa, se encuentra el complejo ORC (del inglés Origin recognition complex), que está conformado por 6 subunidades (ORC 1-6) (Dhar *et al.*, 2001; Bell y Dutta, 2002). Este complejo reconoce las secuencias del origen de replicación (Ori) e interacciona con la proteína CDC6 (del inglés Cell division cycle 6) (Piatti, Hill y Nasmyth, 1996) y se asocia con Cdt1 (del inglés CDC10-dependent transcript 1) (Nishitani y Nurse, 1995). Una vez asociadas, facilitan el reclutamiento de un conjunto de helicasas conocidas como MCM (MCM 2-7) (del inglés minichromosome maintenance protein complex) (Méndez y Stillman 2000).

La etapa Pre-RC abarca desde el reconocimiento del sitio de origen de la replicación hasta el reclutamiento de las helicasas (Watson y Harrison, 2016; Wang y Gao, 2019). Posteriormente, cuando la célula entra en la fase S, inicia la etapa Pre-IC, que se caracteriza por la activación de las helicasas MCM por medio de fosforilación, a través de las proteínas cinasas CDK (del inglés Cyclin dependent kinase) y DDK (del inglés CDC7-DBF4 kinase), que además fosforilan a otros componentes del complejo Pre-RC. Esta activación permite que las helicasas MCM (2-7) se unan a las hebras de DNA y separen a las cadenas de la doble hélice. Algunos de los componentes como ORC, CDC6 y Cdt1 son estrictamente regulados y se inhiben o degradan para evitar que recluten más helicasas y para que la replicación del DNA ocurre solo una vez por ciclo celular (Labib, 2010; Watson y Harrison, 2016).

En cuanto a la etapa Pre-LC de *S. cerevisiae*, después de la activación de las helicasas se deben fosforilar Dpb11 y Sld2 para formar el complejo Pre-LC (Figura 3), el cual se conforma por: Pol  $\epsilon$ , GINS1-4 (del japonés Go-Ichi-Ni-San, conformado por 4 subunidades: Psf1, Psf2, Psf3 y Sld5), Dpb11 y Sld2 ya fosforilados en CDK (Muramatsu *et al.*, 2010). Este complejo Pre-LC funciona como transportador para que se una al complejo Pre-RC para pasar a las demás etapas (Li y Araki, 2013).

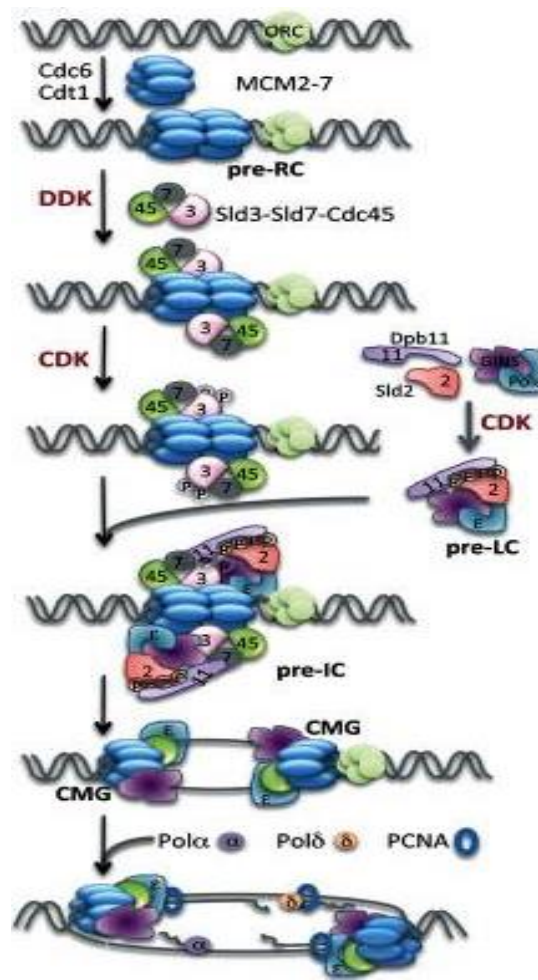


**Figura 3. Modelo de replicación en *Saccharomyces* durante la gemación.** Se muestra el proceso de replicación; el complejo Pre-RC está conformado por ORC, CDC6, Cdt1 y MCM 2-7. Para su activación, las helicasas requieren de DDK y CDK 2 y del complejo Sld3-Sld7-CDC45. Después sigue la etapa Pre-LC, la cual está conformada por Pol  $\epsilon$ , GINS, Dpb11 y Sld2 (ya fosforilados en CDK) y, una vez unidos, el complejo Pre-RC y Pre-LC forman el complejo Pre-IC para pasar a las siguientes etapas de síntesis y terminación. Figura tomada de Li y Araki, 2013.

En la etapa Pre-IC, la proteína Dpb11 se une a Sld2 y Sld3 (del inglés S-phase cyclin-dependent kinase-phosphorylated) para promover su interacción con CDC45 (del inglés, Cell Division Control Protein 45) para regular la activación de MCM y el complejo GINS (Figura 2) (Watson y Harrison, 2016). El complejo GINS es esencial en las etapas Pre-IC y de síntesis durante la replicación (Burgers y Kunkel, 2017). Finalmente, la proteína RPA (del inglés protein replication A) se une a las hebras sencillas de DNA después de su desenrollamiento para 1) evitar la degradación, 2) que se unan otros componentes y 3) que se vuelvan a unir las dos cadenas de DNA (Figura 4) (Pike, Friend y Bell, 2023).

Durante la síntesis, las proteínas Sld2, Sld3 y Dpb11 se liberan del origen de la replicación y se reclutan a las primasas y las DNA polimerasas para la elongación del DNA (Figura 4) (Garcia-Diaz y Bebenek, 2007). Durante la activación de las helicasas, una cadena de DNA debe salir de cada helicasa y se deben romper las interacciones entre los dos complejos de MCM2-7 para que cada complejo de helicasa vaya en dirección contraria al de la otra (Burgers y Kunkel, 2017). Aunque algunas proteínas son reclutadas desde la etapa pre-IC, como la RPA y topoisomerasas, éstas participan más durante la síntesis. Estas proteínas se unen al DNA delante de las helicasas para eliminar los superenrollamientos y permitir el avance de la maquinaria de replicación (Nitiss *et al.*, 2012; Watson y Harrison, 2016).

La separación de las dos cadenas sencillas de DNA crea la horquilla de replicación y, ya separadas, actúan como moldes para la síntesis de las nuevas cadenas de DNA, que va en dirección de 5' a 3' (líder) y la otra de 3' a 5' (rezagada). Para ello, la primasa se une a la cadena principal para generar un fragmento corto de RNA llamado primer o cebador; luego la DNA pol  $\alpha$  se desplaza a lo largo de la cadena y comienza a agregar nucleótidos de DNA en dirección 5' a 3' (Burgers y Kunkel, 2017; Garcia-Diaz y Bebenek, 2007). Sin embargo, debido a su baja progresividad, es rápidamente reemplazada por las DNA polimerasas  $\delta$  o  $\epsilon$ . En cuanto a la cadena rezagada, la primasa sintetiza varios primers en la dirección 5' a 3' a lo largo de la cadena de DNA para que posteriormente la DNA pol  $\delta$  sintetice fragmentos cortos de DNA llamados fragmentos de Okazaki en dirección 5' a 3' (Figura 4) (Garcia-Diaz y Bebenek, 2007; Watson y Harrison, 2016).

En la etapa de terminación, una vez sintetizadas las cadenas nuevas, se reclutan las RNasa H1 y H2 para eliminar los primers de RNA de los fragmentos de Okazaki y reemplazarlos por DNA. La DNA ligasa es responsable de unir los fragmentos de Okazaki mediante la formación de enlaces fosfodiéster (Watson y Harrison, 2016).

Posteriormente, también se recluta a la endo-exonucleasa Fen1 (del inglés Flap endonuclease), que es una enzima que corta enlaces fosfodiéster externos tanto de DNA como de RNA. Esta enzima es multifuncional y participa en distintos mecanismos como la replicación y la reparación del DNA (Tsutakawa *et al.*, 2011; Kimura *et al.*, 2003).

En la replicación del DNA, Fen1 interactúa con proteínas como PCNA (del inglés Proliferating cell nuclear antigen). PCNA pertenece a la familia de abrazaderas deslizantes del DNA y está conformado por un complejo trimérico en forma de anillo (Strzalka y Ziemienowicz, 2011). Este anillo se une a la DNA polimerasa  $\delta$ , sujetándola mientras su forma de anillo rodea al DNA deslizándose a lo largo de la cadena con la finalidad de aumentar la procesividad de la síntesis de la polimerasa (Cardano *et al.*, 2020). Se ha observado que Fen1, también interactúa con el complejo Pre-IC y otros factores que favorecen la procesividad de las DNA polimerasas  $\delta$  y  $\epsilon$ . Cuando hay un error durante la replicación, p21 se une a PCNA para liberar a Fen1 y así permitir la eliminación de los extremos tipo flap, los cuales son estructuras de DNA de hebra sencilla que sobresalen generados durante el desplazamiento de las DNA polimerasas (Hosfield *et al.*, 1998). También se ha observado que Fen1 actúa como cofactor para la DNA Pol  $\delta$  (Matsumoto *et al.*, 1987). Finalmente, para completar la replicación en los extremos terminales de los cromosomas (telómeros) en organismos eucariontes, se requiere de una enzima denominada telomerasa. Esto es necesario debido a que la hebra retrasada no se sintetiza completamente y la telomerasa evita que los extremos terminales se acorten. La telomerasa no requiere de primers, ya que es un complejo ribonucleoproteico con su propio primer de RNA para sintetizar nuevas cadenas en los telómeros de una manera eficiente (Brewer y Fangman, 1987; Pfeiffer y Lingner, 2013).

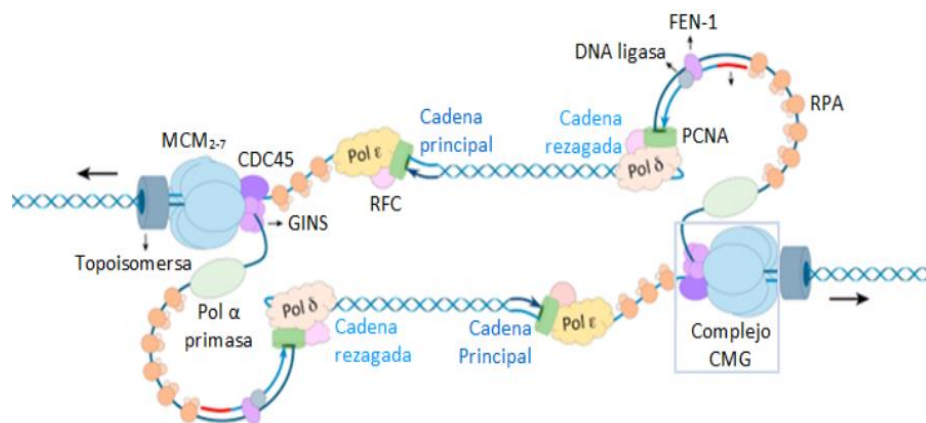


**Figura 4. Replisoma en organismos eucariontes.** Esquema del proceso de replicación en eucariontes durante la síntesis y la terminación. Se observan las proteínas y enzimas implicadas en la replicación del DNA: complejo CMG (Cdc45-MCM-GINS), RPA, topoisomerasas, DNA primasas, RFC (del inglés Replication Factor C), PCNA, DNA polimerasas, Fen1 y DNA ligasa. Se muestran las cadenas, tanto la cadena principal como la rezagada. Las proteínas y enzimas que participan. Figura tomada y editada de M. da Silva *et al.*, 2017).

## **2.2 Proteínas involucradas en la replicación del DNA en otros protozoarios**

La replicación del DNA en protozoarios parásitos tiene ciertas características que reflejan su adaptación a diversos ambientes y en su ciclo de vida. Por ejemplo, en *T. cruzi*, el agente causante de la enfermedad de Chagas, la replicación del DNA ocurre tanto en el núcleo como en una estructura mitocondrial única de los tripanosomátidos denominada cinetoplasto (kDNA) que contiene DNA muy organizado en una red de minicírculos y semicírculos (da Silva *et al.*, 2017). Este protozoario conserva algunos elementos principales de la replicación del DNA que son comunes a otros eucariontes, como el complejo MCM; sin embargo, también tiene algunos elementos distintivos que participan en la replicación del DNA, como la presencia de una única proteína ORC1/CDC6, que combina las dos funciones de los complejos ORC 1-6 y CDC6 sin la necesidad de los demás componentes durante la etapa Pre-RC en tripanosomátidos (Schnauffer *et al.*, 2018).

Por otro lado, para la replicación del kDNA se requiere de un proceso más complejo y específico de los tripanosomátidos. Antes de la replicación, los minicírculos se desacoplan del DNA y se replican en regiones especializadas en la mitocondria para luego reincorporarse en una estructura más ordenada. Este proceso involucra helicasas, polimerasas y topoisomerasas, que trabajan en conjunto para garantizar la replicación del kDNA (Schnauffer *et al.*, 2018). Como se mencionó previamente, en eucariontes, el complejo hexamérico ORC 1-6 es uno de los componentes principales implicados en replicación. Este complejo presenta un dominio ATPasa conservado que se divide en dos subdominios estructurales. En organismos parásitos como *T. brucei* se ha observado una composición más simplificada (Ocaña-Pallarés *et al.*, 2020) además, *T. brucei* presenta una subunidad ORC4 divergente y otras proteínas funcionalmente análogas a ORC2 y ORC5, aunque son altamente divergentes en comparación con otros organismos (Tiengwe *et al.*, 2012; Marques *et al.*, 2016).

## **2.3. Replicación del DNA en *E. histolytica*, identificación de secuencias Ori**

En organismos como las levaduras, los orígenes de replicación están bien definidos y corresponden a secuencias específicas de DNA ricas en AT (Dhar *et al.*, 1996). En cambio, en *E. histolytica* no se han identificado secuencias consenso, aunque se ha observado que ciertas regiones ricas en repeticiones cortas y con sesgo hacia

secuencias AT pueden actuar como sitios Ori cuando se clonan en vectores plasmídicos (Dhar *et al.*, 1996; Ghosh *et al.*, 2003).

Además, se identificaron elementos extracromosómicos, como los plásmidos de DNA ribosomal (DNAr) conocidos como EhR1 y EhR2, los cuales presentan distintos sitios Ori. La activación del Ori depende del estado fisiológico de la célula; por ejemplo, en células en crecimiento exponencial, el Ori principal se localiza en una región espaciadora intergénica (IGS), río arriba del promotor de los genes de rRNA (Ghosh *et al.*, 2003). En cambio, en células depletadas de suero por 24 h, se activan sitios Ori alternativos (Ghosh *et al.*, 2003). Los Ori se identificaron mediante electroforesis bidimensional y microscopía electrónica, donde se observaron burbujas de replicación que corresponden a regiones del DNA desenrollado dentro de una misma molécula, indicando la existencia de múltiples sitios Ori, a diferencia de otros organismos como *Saccharomyces cerevisiae*, que presenta un solo Ori para el plásmido 2 $\mu$ m (Dhar *et al.*, 1996). Interesantemente, en *E. histolytica* las burbujas de replicación aparecen con mayor frecuencia en las regiones menos ricas en AT con un porcentaje aproximado entre 62 y 70%, que en las regiones comunes donde se esperaría con altos contenidos de AT (73 a 85%) (Dhar *et al.*, 1996).

#### **2.4. Identificación de proteínas implicadas en la replicación del DNA en *E. histolytica***

En *E. histolytica*, los análisis filogenéticos sugieren una ausencia de las proteínas ORC2-5 (figura 5), lo que refuerza la idea de que algunos parásitos han simplificado el complejo ORC1-6 a una única proteína ORC1/CDC6 (Ocaña-Pallarès *et al.*, 2020). Adicionalmente, se identificó y caracterizó *in silico* una proteína llamada EhORC1/CDC6 que combina las características de ORC1 y CDC6, como ocurre en *T. cruzi*. Además, se predice que esta proteína podría reclutar los demás componentes de los complejos Pre-RC y Pre-IC sin necesitar de ORC 2-6. En su estructura terciaria *in silico*, se determinaron los dominios conservados e incluso se evaluó la capacidad de formar interacciones proteína-proteína entre esta proteína y CDC45, esencial para la replicación del DNA (figura 6) (Pal y Dam, 2023).

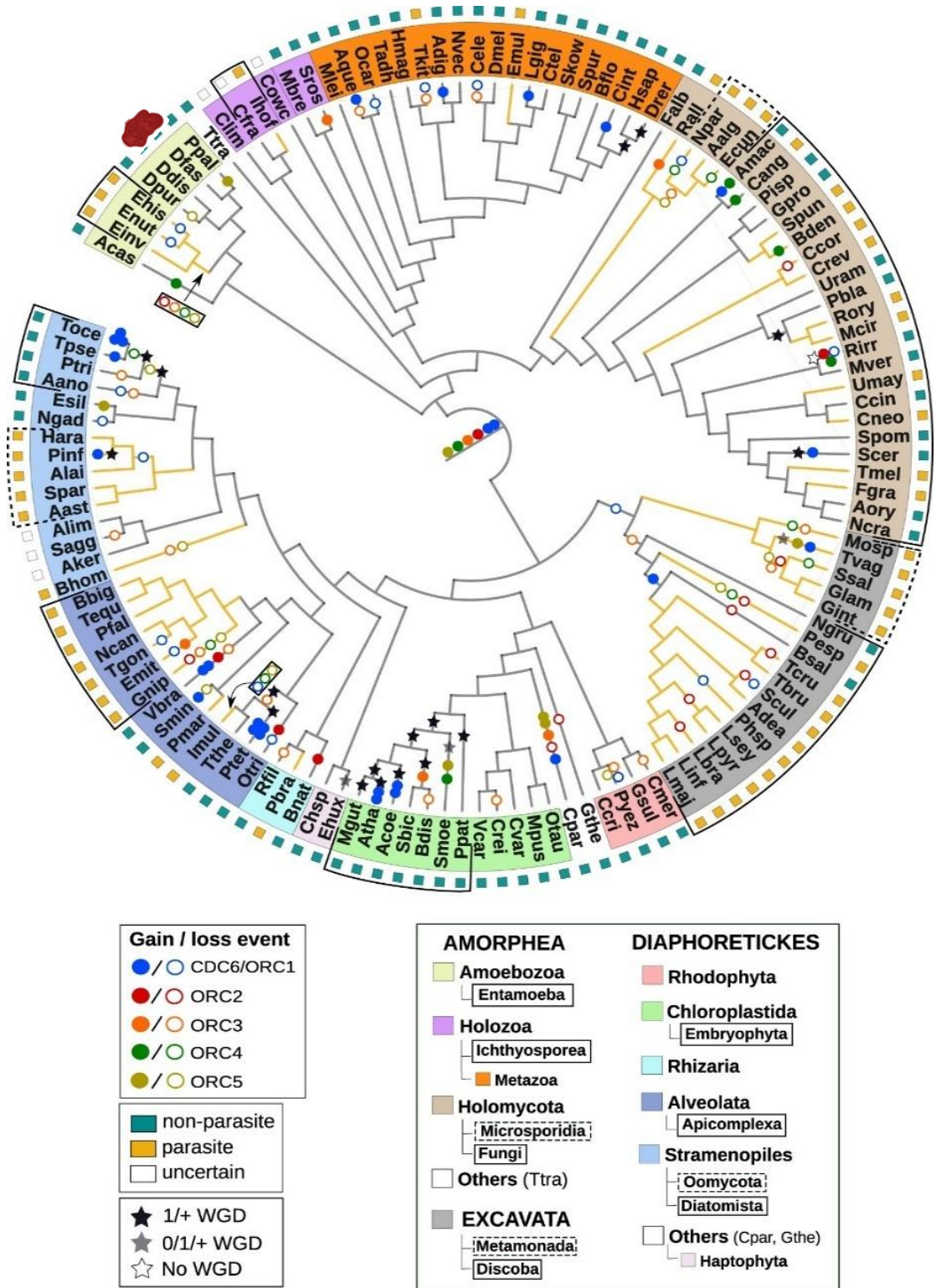


**Figura 5. Árbol filogenético de las proteínas ORC en organismos eucariontes.** Las duplicaciones de las subunidades ORC1-5 y CDC6 se muestran con círculos llenos, mientras que las pérdidas se muestran en círculos vacíos. El género *Entamoeba* se señala con una amiba color rojo; se observa que las tres especies han perdido a las proteínas ORC 2-5. Figura tomada y editada de Ocaña-Pallarès et al., 2020.

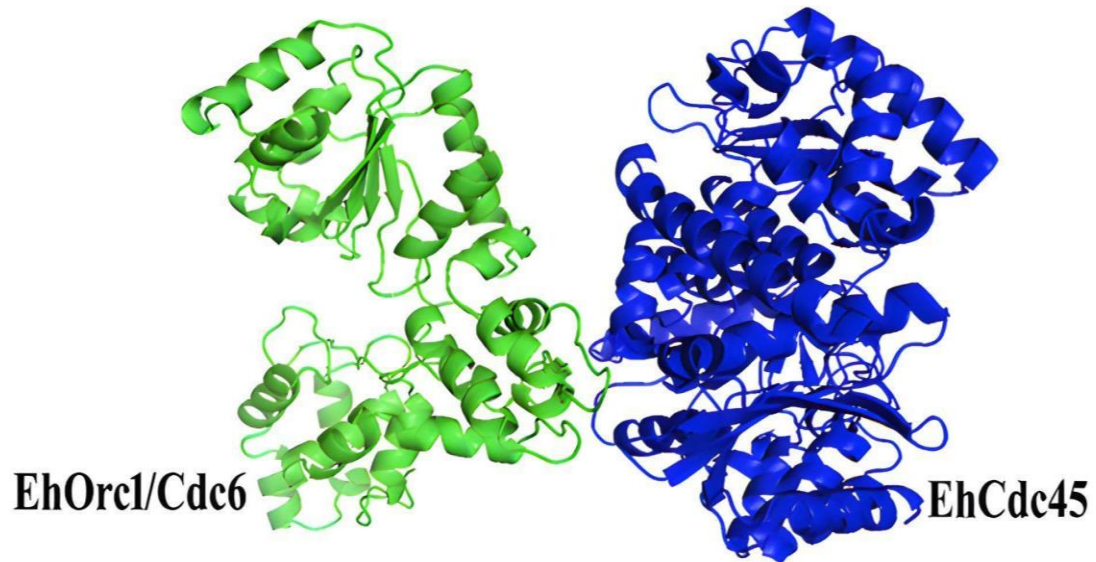


**Figura 6. Estructuras tridimensionales e interacción entre EhOrc1/Cdc6 y EhCdc45.** Se muestra un docking molecular, el cual predice la interacción entre dichas proteínas. En color verde se muestra la estructura tridimensional de EhORC1/CDC6 y en azul EhCDC45. Figura tomada y editada de Pal y Dam, 2023.

Por otra parte, la proteína CDC45 es fundamental en la replicación del DNA en eucariontes, ya que participa durante la etapa pre-IC, en la cual se facilita la activación y mantiene las estructuras y la funcionalidad del complejo MCM 2-7. CDC45 se asocia al complejo GINS y mediante cambios alostéricos, induce un cambio conformacional en el complejo MCM para promover la apertura de su canal central y de esta manera permitir el paso del DNA, formando el complejo CMG (CDC45/MCM2-7/GINS) para que las helicasas puedan desenrollar las cadenas de DNA (Ilves *et al.*, 2015, Simon *et al.*, 2016).

En *E. histolytica*, se han identificado genes homólogos a CDC45, sugiriendo una conservación de este mecanismo en el parásito (Ramírez *et al.*, 2018). De hecho, en 2018, Kurniawan y colaboradores analizaron la estructura cristalográfica y función de la proteína EhCDC45 a través de la proteína recombinante (Figura 7).

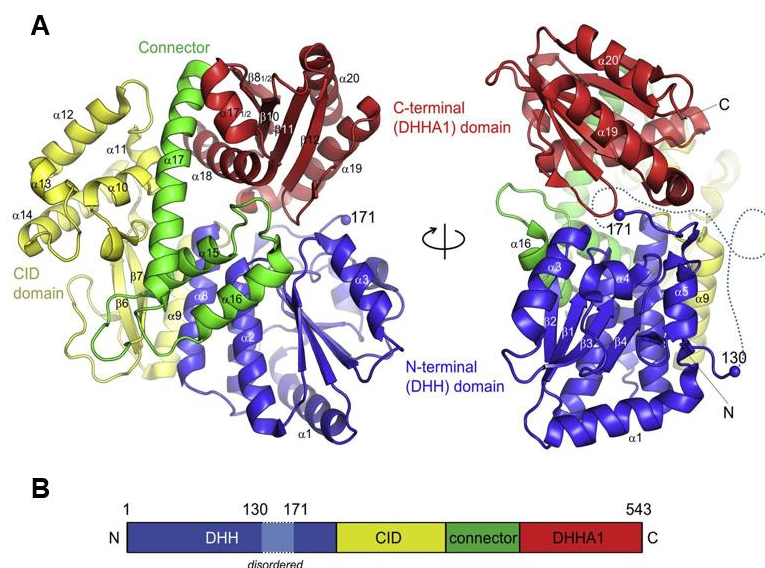


**Figura 7. Estructura tridimensional de CDC45 de *E. histolytica*.** **A)** Estructura tridimensional de EhCDC45, la cual consta de dos dominios: el N-terminal DHH (en color azul) y el C-terminal DHHA1 (en color rojo), a su vez, están conectados por un conector helicoidal  $\alpha$  (en color verde). Además, consta de un dominio CID, el cual interacciona con el complejo CMG (en color amarillo). **B)** Diagrama esquemático que muestra la organización de los dominios de la proteína EhCDC45 los residuos desordenados se muestran en azul claro (130–171). Figura tomada de Kurniawan *et al.*, 2018.

Los resultados revelaron que EhCDC45 presenta dos dominios N-terminal DHH y C-terminal DHHA1 que están conectados a través de un segmento helicoidal  $\alpha$ , típico de las proteínas de CDC45 en otros organismos eucariontes (Kurniawan *et al.*, 2018). Además, identificaron regiones conservadas que son importantes para la interacción con las proteínas de los complejos MCM y GINS. Además, EhCDC45 presenta un dominio CID que podría permitir la formación del complejo CMG (Kurniawan *et al.*, 2018). Por otro lado, PCNA es una proteína que participa en varias etapas, ya que interactúa con proteínas implicadas en las vías de replicación y reparación del DNA. En *E. histolytica*, se identificó y caracterizó estructuralmente esta proteína, conocida como EhPCNA. Cardona-Felix y colaboradores (2011) revelaron que EhPCNA forma un homotrímero funcional, como ocurre en otros eucariontes (Figura 8). EhPCNA conserva la estructura en forma de anillo con un canal central cargado positivamente para la interacción con el DNA bicatenario. Adicionalmente, conserva el dominio bucle conector entre dominios (IDCL, del inglés inter-domain connecting loop) que es crucial para la interacción con proteínas que poseen el motivo PIP box. Además, se demostró que EhPCNA estimula la actividad catalítica de EhDNAIgl durante la unión de fragmentos de DNA (Cardona-Felix *et al.*, 2011). Esto sugiere que los mecanismos bioquímicos mediados por EhPCNA están conservados en este parásito.

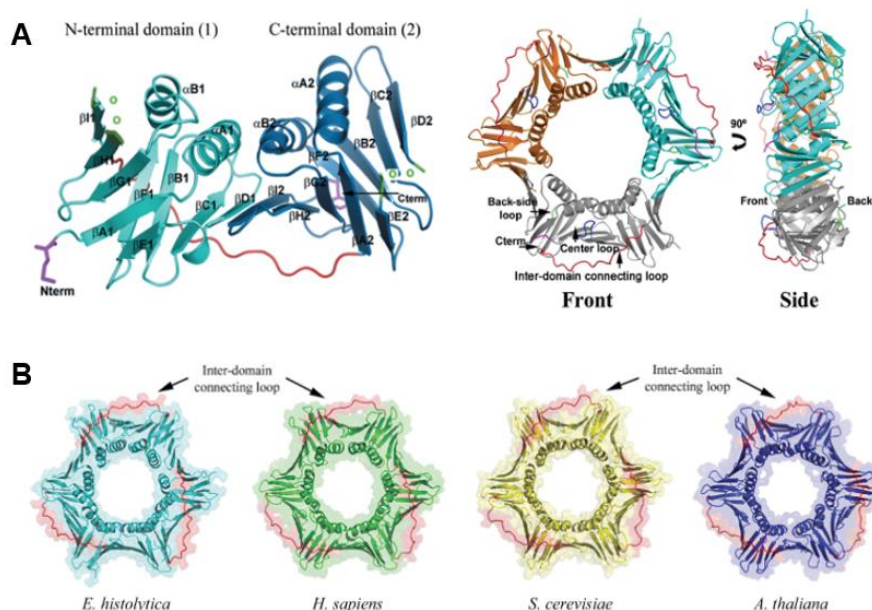


**Figura 8. Estructura tridimensional de EhPCNA de *E. histolytica*.** **A)** Estructura monomérica de EhPCNA que consta de dos dominios, el N-terminal y el C-terminal que están unidos por un bucle de conexión entre dominios (IDCL en color rojo). A la derecha se muestran las tres moléculas monoméricas formando un anillo. **B)** Se muestran las estructuras con forma de anillo tanto de EhPCNA como de otros organismos. Figura tomada y editada de Cardona-Felix *et al.*, 2011.

Otra proteína caracterizada en *E. histolytica* que participa en replicación es la DNA ligasa. La DNA ligasa es una enzima que participa en varios procesos celulares como la replicación, recombinación y reparación del DNA. En replicación, la DNA ligasa participa en la etapa T, donde cataliza la formación de un enlace fosfodiéster entre los extremos 5' y 3' en las cadenas del DNA para unir fragmentos de doble cadena (Watson y Harrison, 2016). Algunos eucariontes poseen múltiples DNA ligasas; en cambio, *E. histolytica* tiene una única DNA ligasa (EhDNAligI) (Figura 9). Esta proteína es dependiente de ATP (trifosfato de adenosina) y similar a las DNA ligasas de tipo I, cuya actividad óptima es a pH 7.5, y utiliza como cofactores a  $Mg^{2+}$  o  $Mn^{2+}$  (Cardona-Félix *et al.*, 2010). Estructuralmente, presenta los dominios DBD (del inglés DNA-binding domain), AD (del inglés adenylation domain) y, OB-fold (del inglés oligonucleotide/oligosaccharide-binding fold). Los dominios AD y OB-fold forman parte del núcleo catalítico de la enzima. Además, EhDNAligI discrimina entre los apareamientos de bases correctos o incorrectos y realiza ligaciones sobre algunos tipos de daño oxidativo como 8-oxoguanina, indicando su participación durante la reparación del DNA mediante las vías de escisión de bases y nucleótidos (BER y NER) (Azuara-Liceaga *et al.*, 2018).

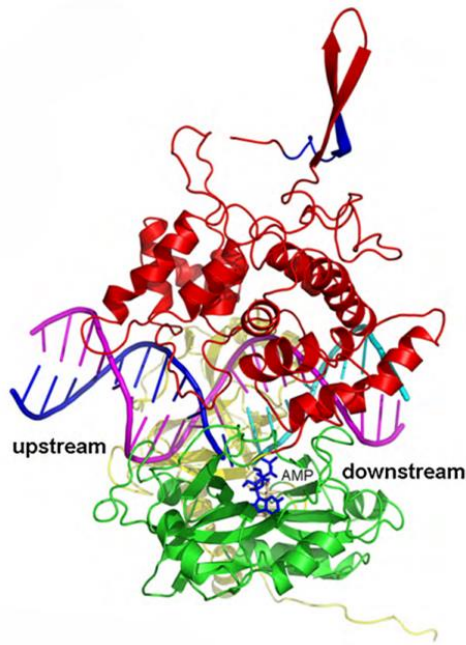


**Figura 9. Estructura tridimensional de EhDNA ligasa de *E. histolytica*.** En la imagen se muestran en la estructura tridimensional de EhDNA ligasa los dominios de unión al DNA (rojo), adenilación (verde) y el pliegue OB (amarillo). Además, se muestra el DNA río arriba en color azul, el DNA río abajo en color celeste y el cofactor AMP (adenosín monofosfato) en azul. Figura tomada de Cardona *et al.*, 2010.

En conclusión, aunque se conocen algunas moléculas de *E. histolytica* implicadas en replicación como EhOrc1, EhCDC45, EhPCNA y EhDNA ligase, es importante caracterizar el replisoma y analizar el proceso de replicación de *E. histolytica*. Este estudio es uno de los primeros trabajos para analizar la composición del replisoma en *E. histolytica* y determinar la expresión de los genes que codifican para estas proteínas durante un proceso de sincronización celular.

### 3. Hipótesis

La maquinaria de replicación del genoma de *E. histolytica* es similar a la que tienen otros organismos eucariontes superiores.

### 4. Justificación

La replicación del DNA es un proceso fundamental en la biología celular que asegura la transmisión precisa de la información genética de una generación a otra. *E. histolytica* es un parásito protozoario causante de la amebiasis, una enfermedad que afecta a millones de personas en todo el mundo, especialmente en regiones con condiciones sanitarias deficientes. A pesar de su importancia médica y su impacto en la salud pública, aún existe un conocimiento limitado sobre los mecanismos moleculares que regulan la replicación del DNA en este organismo.

### 5. Objetivos

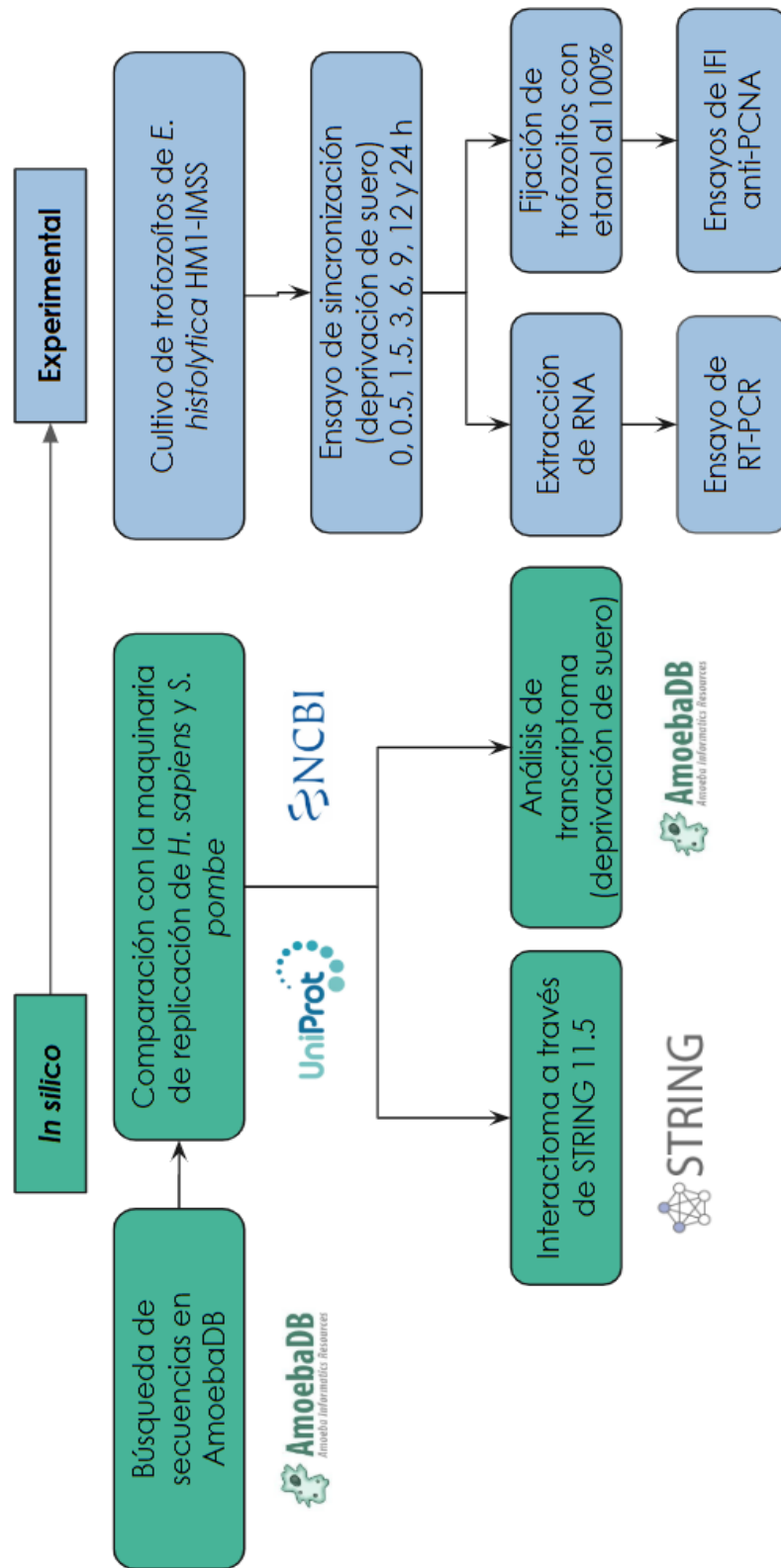
#### 5.1 Objetivo general

Identificar y analizar los genes que codifican para proteínas implicadas en las distintas etapas de la replicación del DNA en *E. histolytica*.

#### 5.2 Objetivos particulares

- Identificar los genes que codifican para las proteínas del replisoma relacionadas con cada una de las etapas de la replicación del DNA de *E. histolytica*.
- Analizar *in silico* el interactoma y la expresión de las proteínas involucradas en las diferentes etapas de la replicación del DNA en *E. histolytica*.
- Analizar *in vitro* los patrones de expresión de algunos genes que codifican para proteínas involucradas en la replicación del DNA durante la sincronización celular.
- Analizar la localización subcelular de EhPCNA durante la sincronización celular.

## 6. Estrategia experimental



## 7. Materiales y métodos

### 7.1 Búsqueda y alineamiento de secuencias

Para la búsqueda de las secuencias de las proteínas relacionadas con la replicación del DNA se emplearon las bases de datos AmoebaDB (<https://amoebadb.org/amoeba/app>) para *E. histolytica*, y UniProt (<https://www.uniprot.org/>) para *H. sapiens* y *S. pombe*. Para el alineamiento de las secuencias se empleó el programa blastp de NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) utilizando el algoritmo DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST) para obtener el porcentaje de identidad y el valor e de las proteínas de *E. histolytica* con respecto a las de *H. sapiens* y *S. pombe*.

### 7.2 Identificación de dominios funcionales y características de las proteínas

Para la identificación de las características de las proteínas implicadas en replicación se emplearon los programas indicados en la tabla 3.

**Tabla 3. Programas utilizados para las características de las proteínas del replisoma de *E. histolytica*.**

| Característica obtenida       | Plataforma utilizada        | Página Web                                                                                                                                       |
|-------------------------------|-----------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|
| Dominios funcionales          | InterPro v.104.0            | <a href="https://www.ebi.ac.uk/interpro/">https://www.ebi.ac.uk/interpro/</a> <sup>1</sup>                                                       |
| PM y pl                       | ExPASy tools                | <a href="https://web.expasy.org/compute_pi/">https://web.expasy.org/compute_pi/</a> <sup>2</sup>                                                 |
| Dominios transmembranales     | TMHMM v.1.0                 | <a href="https://services.healthtech.dtu.dk/services/DeepTMHMM-1.0/">https://services.healthtech.dtu.dk/services/DeepTMHMM-1.0/</a> <sup>3</sup> |
| Señal de localización nuclear | NLStradamus                 | <a href="http://www.moseslab.csb.utoronto.ca/NLStradamus/">http://www.moseslab.csb.utoronto.ca/NLStradamus/</a> <sup>4</sup>                     |
| Anotación y tamaño en pb y aa | AmoebaDB                    | <a href="https://amoebadb.org/amoeba/app">https://amoebadb.org/amoeba/app</a> <sup>5</sup>                                                       |
| Localización subcelular y GO  | AmoebaDB/UniProt            | <a href="https://amoebadb.org/amoeba/app">https://amoebadb.org/amoeba/app</a> <sup>6</sup>                                                       |
| Estructura tridimensional     | Alpha Fold 3 / Pymol v. 3.0 | <a href="https://alphafoldserver.com/welcome">https://alphafoldserver.com/welcome</a> <sup>7</sup>                                               |
| Representación lineal         | DOG 2.0                     | <a href="http://dog.biocuckoo.org/">http://dog.biocuckoo.org/</a> <sup>8</sup>                                                                   |

<sup>1</sup>Blum *et al.*, 2025; <sup>2</sup>Gasteiger *et al.*, 2005; <sup>3</sup>Hallgren *et al.*, 2022; <sup>4</sup>Nguyen *et al.*, 2009; <sup>5</sup>Álvarez *et al.*, 2023; <sup>6</sup>Bateman *et al.*, 2024; <sup>7</sup>Abramson *et al.*, 2024 y <sup>8</sup>Ren *et al.*, 2009

### 7.3 Análisis de las interacciones proteína-proteína

Para determinar las interacciones proteína-proteína se utilizaron los ID de las proteínas que participan en cada etapa en la herramienta bioinformática STRING v.12.0 (<https://string-db.org/>) con nivel de confianza media 0.400 limitado. Fuentes de interacción: base de datos, experimental y coexpresión.

#### 7.4 Análisis de los patrones de expresión

Para analizar los patrones de expresión, se utilizaron los datos del transcriptoma de *E. histolytica* de la base de datos AmoebaDB (<https://amoebadb.org>, consultado el 10 de agosto de 2024) y se tomaron los valores de transcritos por millón (TPM) de los análisis por RNAseq de trofozoítos en condiciones normales y en la condición de depleción de suero (Naiyer *et al.*, 2019). Los datos se normalizaron de acuerdo con la siguiente ecuación: cada valor se normalizó con la siguiente ecuación:  $\log_2(\text{TPM}+1)$ . El mapa de calor se elaboró mediante el paquete mapa de calor del software R (versión 3.4.3.2). El diagrama de caja se construyó utilizando el paquete ggplot2 del mismo software.

#### 7.5 Diseño de primers

Para el diseño de primers de genes involucrados en la replicación, se utilizó la herramienta bioinformática Primer3Plus (<https://www.primer3plus.com/index.html>) y se seleccionaron 4 genes de cada etapa de replicación. Los primers se diseñaron con base a las siguientes características: 1) entre 100 y 200 pb de longitud del amplicón, 2) longitud de los primers entre 20 y 30 pb 3) TM cercana a 55°C. Los primers diseñados se muestran en la Tabla 4.

#### 7.6 Cultivo de *E. histolytica* HM1:IMSS y ensayos de sincronización

Trofozoítos de *E. histolytica* HM1-IMSS se cultivaron en medio TYI-S-33 complementado con 16% de suero de bovino (Equitech-Bio, INC) previamente inactivado con calor, 3% de vitaminas Diamond (Invitrogen) y penicilina/estreptomicina (100 U/mL). Una vez alcanzada la fase logarítmica, se enfriaron a 4°C por 10 min y se centrifugaron a 1800 rpm por 5 min. Para la sincronización celular se utilizaron variaciones de los procedimientos reportados por Das *et al.*, 2004 y Mukherjee *et al.*, 2009. Se cultivaron trofozoítos en cajas T25 para los tiempos 0, 0.5, 1.5, 3, 6, 9, 11 y 24 h. Después, se decantó el medio y se lavaron con 10 mL de medio TYI-S-33 tibio sin suero y se añadieron 40 mL de medio sin suero por caja y se incubó durante 12 h a 37°C (depleción de suero, inducción de arresto celular). Posteriormente, se procesó la primera condición (tiempo 0) y al resto de las cajas se les adicionó suero bovino al 16% y se incubaron durante 0.5, 1.5, 3, 6, 9, 11 y 24 h (condiciones de reposición de suero) y se procesaron. Este ensayo se replicó a partir de una réplica biológica y se hicieron triplicados experimentales.

**Tabla 4. Oligonucleótidos diseñados para ensayos de RT-PCR**

| <b>Etapas</b> | <b>Nombre</b>    | <b>ID*</b>  | <b>Sentido</b> | <b>Secuencia**</b>               | <b>Amplificación pb</b> |
|---------------|------------------|-------------|----------------|----------------------------------|-------------------------|
| Pre-RC        | <i>ehorc1</i>    | EHI_105150  | Forward        | 5'-TGATGATGAAATTGTTGGAAGAGAAG-3' | 118                     |
|               |                  |             | Reverse        | 5'-CCCTGTTCTGCTGGTGTTCAT-3'      |                         |
|               | <i>ehcdc6</i>    | EHI_192770  | Forward        | 5'-TGGAGTAATTGGAATAGGTGAATGGT-3' | 136                     |
|               |                  |             | Reverse        | 5'-ACGTACAACACCTTGAGATGC-3'      |                         |
|               | <i>ehmcm3</i>    | EHI_103600  | Forward        | 5'-TGTGGTGATCCATCTACTGCT-3'      | 101                     |
|               |                  |             | Reverse        | 5'-CCAACTCCAGTTGCACCTCT-3'       |                         |
|               | <i>ehmcm4</i>    | EHI_187720  | Forward        | 5'-AGGCAATTCCAGCAGGTGAA-3'       | 85                      |
|               |                  |             | Reverse        | 5'-CGATCACCTGGTTTTGCGGA-3'       |                         |
| Pre-IC        | <i>ehcdc45</i>   | EHI_049900  | Forward        | 5'-TCGCTGTCCTTGATGCTGAG-3'       | 106                     |
|               |                  |             | Reverse        | 5'-GCTGTTTGAGTGAATTTAGCCCC-3'    |                         |
|               | <i>ehtop2***</i> | EHI_120640  | Forward        | 5'-CAGGATGGAGTACTGCCGTTT-3'      | 120                     |
|               |                  |             | Reverse        | 5'-GTCCTGTCCATCCTCTAAACCA-3'     |                         |
|               | <i>ehpsf1</i>    | EHI_1236990 | Forward        | 5'-CGTGCGTCAATGACTAATGATG-3'     | 182                     |
|               |                  |             | Reverse        | 5'-CTTGATTTTGTGGCAAATCATCC-3'    |                         |
|               | <i>ehrp2***</i>  | EHI_147000  | Forward        | 5'-CTGATTACAAACAAAGTAATGGG-3'    | 281                     |
|               |                  |             | Reverse        | 5'-ATCTACTTCAACAGCATAGC-3'       |                         |
| Síntesis      | <i>ehdnapolε</i> | EHI_141410  | Forward        | 5'-CCCATTAATTGCTCATGCTATTC-3'    | 201                     |
|               |                  |             | Reverse        | 5'-GTTGATTTTGAAGATGCATGTGC-3'    |                         |
|               | <i>ehdnapolδ</i> | EHI_006690  | Forward        | 5'-CACTCCAACAGCAAATACTGGA-3'     | 180                     |

|                  |                           |            |         |                                |     |
|------------------|---------------------------|------------|---------|--------------------------------|-----|
|                  |                           | EHI_038920 | Reverse | 5'-CGATGGTATGCGTTTTCTGA-3      | 160 |
|                  |                           |            | Forward | 5'-GCTTACCGTTCATGGGGTAG-3      |     |
|                  | <i>ehpcna</i> ***         | EHI_128450 | Reverse | 5'-CTCCTTCTCTATCACAGTCAAGC-3   | 142 |
|                  |                           |            | Forward | 5'-TCCAGAACTTCAATCTGACG-3      |     |
| Terminación      | <i>ehfen1</i> ***         | EHI_099740 | Forward | 5'-GAAGGTACTTGTGCTGCATTAG-3    | 218 |
|                  |                           |            | Reverse | 5'-CATCCAAGTAAAATACAAAGG-3     |     |
|                  | <i>ehdnalig1</i> ***      | EHI_111060 | Forward | 5'-GATTAGCCGATACAGTTGAC-3      | 142 |
|                  |                           |            | Reverse | 5'-ATCTCCACTGTCTTCATCTC-3      |     |
|                  | <i>ehtrf-like I</i> ***   | EHI_001090 | Forward | 5'-ACCCTCAGTTTGCAACCCTT-3      | 95  |
|                  |                           |            | Reverse | 5'-AGGACTTTGTTCCATTTTTACTTCA-3 |     |
|                  | <i>ehtrf-like II</i> ***  | EHI_001110 | Forward | 5'-AAAAGAGCATTAGCAACGGTTAT-3   | 127 |
|                  |                           |            | Reverse | 5'-TTTCTGCATTTGGATCTGAACAA-3   |     |
|                  | <i>ehtrf-like III</i> *** | EHI_148140 | Forward | 5'-GCGCGTTCTCAAATACAACC-3      | 70  |
|                  |                           |            | Reverse | 5'-GTTTCAGAGGCTGCAGGAAGT-3     |     |
| Control de carga | <i>ehcoat</i> ***         | EHI_129520 | Forward | 5'-ACTTCCATTTGACGGTGACA-3      | 145 |
|                  |                           |            | Reverse | 5'-TGAAATTCCTTTTTGTCCTGGT-3    |     |

\* De acuerdo a AmoebaDB

\*\* TM de 55°C para todos los primers

Forward y Reverse: orientación de los primers 5' a 3' y 3' a 5', respectivamente

\*\*\* Previamente reportados en Azuara-Liceaga et al., 2018 y Rendón-Gandarilla et al., 2018

## **7.7 Extracción de DNA genómico**

Para la extracción de DNA genómico, se empleó el kit comercial Wizard® Genomic DNA Purification Kit de Promega siguiendo las indicaciones del fabricante. Para ello se cosecharon  $4 \times 10^6$  trofozoítos a 1800 rpm durante 5 min. A la pastilla se le adicionaron 600  $\mu$ L de NLS, pipeteando hasta eliminar agregados; posteriormente, se añadieron 3  $\mu$ L de RNAsa y se incubó la muestra por 30 min a 37°C. Una vez incubado, se adicionaron 200  $\mu$ L de PPS, se agitó en vórtex durante 20 s y se dejó enfriar por 5 min a 4°C. A continuación, se centrifugó a 13000 rpm durante 4 min para adicionar al sobrenadante 600  $\mu$ L de isopropanol, se mezcló sin agitar y se incubó a -20°C durante 10 min. Después, se centrifugó a 13000 rpm durante 2 min y se lavó la pastilla 2 veces con 600  $\mu$ L de etanol al 70%. Finalmente, la pastilla se dejó secar y se resuspendió en 100  $\mu$ L de H<sub>2</sub>O inyectable y se cuantificó en el equipo Nanodrop 2000. La integridad del DNA genómico se visualizó en un gel de agarosa al 1% teñido con RedGel 1x (Biotium).

## **7.8 Extracción de RNA**

La extracción de RNA se realizó utilizando el protocolo basado en TRIzol®. Para cada muestra de  $5 \times 10^6$  trofozoítos, se añadió 1 mL de TRIzol. Las muestras se incubaron en hielo durante 5 min y después se agregaron 200  $\mu$ L de cloroformo a cada tubo, seguido de una mezcla por inversión por 30 s. Posteriormente, se incubó a temperatura ambiente (TA) durante 3 min y se centrifugó a 12,000 rpm durante 15 min a 4°C, separando las fases acuosa y orgánica. La fase acuosa superior se transfirió cuidadosamente a un tubo nuevo de 1.5 mL, mientras que la fase orgánica se almacenó a -20°C para posibles análisis posteriores. Para precipitar el RNA, se agregaron 0.5 mL de isopropanol frío, se mezcló por inversión suave (2-5 veces) y se incubó a TA durante 10 minutos. Después se centrifugó a 11,000 rpm durante 20 min, el pellet se lavó dos veces con 750  $\mu$ L de etanol al 75% (v/v), centrifugando a 10,000 rpm durante 5 min en cada paso. Tras eliminar el etanol y dejar secar el pellet a TA, el RNA se resuspendió en 100  $\mu$ L de agua inyectable estéril. La concentración y pureza se determinaron mediante espectrofotometría (Nanodrop 2000), evaluando la relación de absorbancia. Adicionalmente, la integridad del RNA pudo verificarse mediante electroforesis en gel de agarosa al 1% teñido con RedGel 1x.

### **7.9 Tratamiento con DNAsa**

El tratamiento con DNAsa se realizó de acuerdo con el protocolo de Promega. Para ello se utilizó 1 unidad de la enzima RQ1 RNase-Free DNase por cada  $\mu\text{g}$  de RNA. Se ajustó el volumen de cada muestra para una concentración final de  $100 \text{ ng}/\mu\text{L}$  de RNA tratado. El tratamiento con DNAsa se realizó durante 1.5 h a  $37^\circ\text{C}$ . Tras la incubación, se agregó solución STOP ( $1 \mu\text{L}$  por cada  $1 \mu\text{g}$  de RNA tratado) y se incubó a  $65^\circ\text{C}$  durante 10 min.

### **7.10 Síntesis de cDNA**

Para la síntesis de cDNA se siguió el protocolo de Promega. Para ello, se tomó  $1 \mu\text{g}$  de RNA tratado con DNAsa y se adicionó  $1 \mu\text{L}$  de oligo dT (Promega). Las muestras se incubaron a  $70^\circ\text{C}$  durante 5 min, seguido de una incubación a  $4^\circ\text{C}$  durante 5 min. En un tubo separado, se preparó un mix compuesto por buffer 5x,  $\text{MgCl}_2$  (concentración final 3 mM),  $1 \mu\text{L}$  de dNTPs, RNAsin (1 unidad), y enzima transcriptasa reversa (RT) ( $1 \mu\text{L}$ ) para un volumen final de  $20 \mu\text{L}$ . La rxn de síntesis de cDNA se llevaron a cabo en un termociclador Thermo Scientific (GeneAmp PCR System 9700) con los siguientes pasos de incubación:  $25^\circ\text{C}$  durante 5 min,  $42^\circ\text{C}$  durante 60 min y  $70^\circ\text{C}$  durante 15 min.

### **7.11 Amplificación de los genes por RT-PCR**

La amplificación por RT-PCR se realizó utilizando el kit Master Mix de Roche®, adaptando el protocolo estándar. Se realizaron reacciones de un volumen final de  $12 \mu\text{L}$ ; para ello, cada reacción incluyó el master mix 2x ( $6 \mu\text{L}$ ), primer forward y reverso ( $5 \mu\text{M}$  cada uno) y cDNA o DNAg como control ( $50 \text{ ng}$ ). Para la amplificación del gen *ehorc1* se añadió  $\text{MgCl}_2$  para una concentración final de 3 mM. La amplificación de los diferentes genes se realizó en el termociclador ThermoFisher Scientific 9700 programado con las condiciones descritas en la figura 10. Los productos de RT-PCR se almacenaron a  $-20^\circ\text{C}$  hasta su posterior análisis.

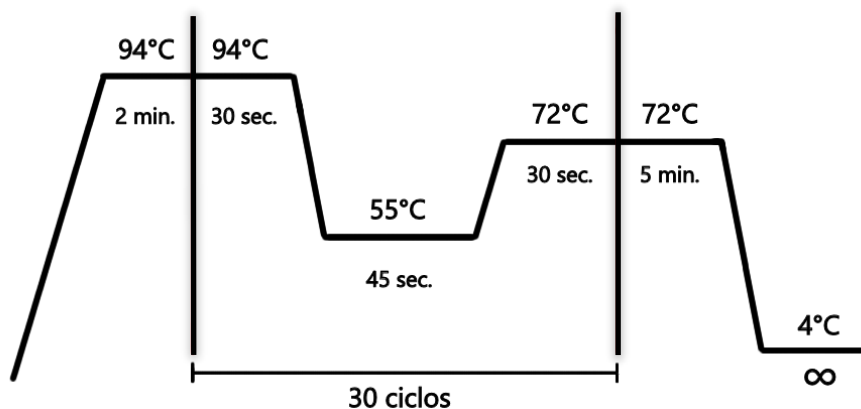


**Figura 10. Condiciones del termociclador para la amplificación de genes en ensayos de RT-PCR.** Las condiciones se emplearon para todos los genes, a excepción de *ehpcna* y del control de carga (*ehcoat*) ambos se amplificaron a 25 ciclos.

### 7.12 Electroforesis en gel de agarosa

Para el corrimiento de las muestras, se prepararon geles de agarosa al 1% y 2% teñidos con Red Gel 1x (Biotium) para el RNA y las RT-PCRs respectivamente. El corrimiento se realizó a 80 V con buffer TAE 1X (1L: Tris 40 mM, ácido acético 20 mM, EDTA 1 mM pH 8.0). Se utilizaron los marcadores de 50 bp y 100 bp DNA Step Ladder de Promega, a la concentración recomendada. Los geles se visualizaron en un transiluminador y se registraron en el fotodocumentador Gel-DOC XR (BIORAD).

### 7.13 Análisis del nivel de expresión relativa de cada gen seleccionado

Las imágenes de los geles de RT-PCR se analizaron mediante el software ImageJ v1.53. Para ello, se convirtieron a escala de grises y se seleccionaron las bandas correspondientes a los amplicones para obtener la intensidad de píxeles. Los valores de los dos ensayos de sincronización que se realizaron se exportaron a Excel, donde se normalizaron utilizando el gen de control de carga correspondiente para graficar el promedio y la desviación estándar de cada condición.

### 7.14 Ensayo de Inmunofluorescencia indirecta (IFI)

Para los ensayos de inmunofluorescencia, trofozoítos sincronizados se adhirieron a cubreobjetos estériles de 22 x 22 mm. Posteriormente, se fijaron con etanol frío al 100% y se incubaron a -20°C por 30 min. Pasado ese tiempo, se retiró el etanol y se realizaron 3 lavados con PBS pH 6.8 (NaCl 136 mM, KCl 2 mM, Na<sub>2</sub>HPO<sub>4</sub> 10 mM, KH<sub>2</sub>PO<sub>4</sub> 1.76 mM). Posteriormente, los cubreobjetos se incubaron con NH<sub>4</sub>Cl 50 mM

a 37°C por 30 min; para después bloquearse con BSA (Albúmina sérica bovina, Sigma-Aldrich) al 1% a 37°C por 1 hr. Los cubreobjetos se incubaron con el anticuerpo primario anti-EhPCNA 1:250 (producido en el laboratorio de la Dra. Elisa) toda la noche a 4°C. Después se hicieron tres lavados con PBS y se decantó y se adicionó el anticuerpo secundario anti-IgG de conejo acoplado a FITC 1:100 (Jackson Immuno Research). Los cubreobjetos se incubaron a TA por 2 h y las muestras se lavaron 3 veces con PBS y se montaron en portaobjetos con el medio de montaje VectaShield® Antifade Mounting Medium con DAPI (Vector Laboratories). Las laminillas se visualizaron en un microscopio confocal Carl Zeiss a 60x en el Departamento de Infectómica y Patogenéesis Molecular del Cinvestav y las imágenes se analizaron a través de los softwares Zen Blue e ImageJ.

#### **7.15 Análisis de la intensidad de la fluorescencia nuclear**

Para los ensayos de IFI, se midió la intensidad de fluorescencia nuclear a partir de las imágenes de inmunofluorescencia mediante el software ImageJ v1.53. Para ello, las imágenes fueron convertidas a 8 bits, en escala de grises y se creó una máscara binaria. Después se seleccionaron manualmente los parámetros con respecto a las regiones de interés (ROI) correspondientes a los núcleos de los trofozoítos. Los datos de intensidad relativa se promediaron y graficaron con su respectiva desviación estándar para cada tiempo evaluado.

## 8. Resultados

### 8.1 Identificación de las proteínas que conforman el replisoma de *E. histolytica*

#### 8.1.1 Etapa Pre-RC

Durante la replicación del DNA se deben de formar diferentes complejos de proteínas y DNA, los cuales están finamente regulados a través del ciclo celular. Estos complejos se han clasificado de acuerdo con la etapa de la replicación en la que se forman. La formación del complejo Pre-RC es el primer paso para la replicación del DNA, el cual involucra el reconocimiento del origen de replicación y el ensamblaje de las helicasas que permiten la apertura de la doble hélice (Bell y Labib, 2016).

Para identificar las proteínas que conforman el complejo Pre-RC en *E. histolytica*, se tomó como referencia la secuencia de las proteínas caracterizadas en *H. sapiens* y *S. pombe* que corresponden a 19 proteínas (Costa y Diffley, 2022). En el genoma de *E. histolytica* solo se identificaron 12 genes que codifican para proteínas de la etapa Pre-RC. Dos de los 12 genes identificados presentan intrones (*ehmcm2* y *ehmcm5*) en general estos 12 genes codifican para proteínas cuyos pesos moleculares oscilan entre 43 y 100 kDa y tienen un pI entre 5.15 y 8.27. La localización subcelular de estas proteínas está predicha en el núcleo de la célula, aunque solo encontramos 4 proteínas (*ehmcm3*, *ehmcm5*, *ehmcm6* y *ehmcm7*) que tienen una señal de localización nuclear respectivamente (Anexo 1). Identificamos un gen que codifica para la proteína ORC1, con un porcentaje de identidad del 23% con respecto a *H. sapiens*, en lugar de las seis proteínas ORC2–6 características de otros organismos eucariontes (Tabla 5). Esta proteína se denominó EhORC1 y presenta un dominio AAA\_22, el cual pertenece a la subfamilia o variante específica dentro de la superfamilia AAA+ ATPasa (del inglés, ATPases Associated with diverse cellular Activities). Esto podría reflejar una especialización funcional o una simplificación estructural del complejo Orc en este organismo (Tabla 6).

**Tabla 5. Comparación de los componentes de la etapa Pre-RC de la replicación del DNA de *E. histolytica* con respecto a los organismos *H. sapiens* y *S. pombe***

| Etapa Pre-RC |                                   |                                |                                                           |                                                         |
|--------------|-----------------------------------|--------------------------------|-----------------------------------------------------------|---------------------------------------------------------|
| Nombre       | <i>E. histolytica</i><br>Proteína | <i>E. histolytica</i><br>(ID)* | <i>H. sapiens</i><br>(identidad % / valor e) <sup>a</sup> | <i>S. pombe</i><br>(identidad % / valor e) <sup>a</sup> |
| ORC1         | EhORC1                            | EH1_105150                     | Q13415<br>(23.94% / 1e-07)                                | P54789<br>(24.48% / 4e-09)                              |
| ORC2         | NI                                | NI                             | Q13416<br>(ORC2_HUMAN)                                    | Q09142<br>(ORC2_SCHPO)                                  |
| ORC3         | NI                                | NI                             | Q9UBD5<br>(ORC3_HUMAN)                                    | Q10067<br>(ORC3_SCHPO)                                  |
| ORC4         | NI                                | NI                             | O43929<br>(ORC4_HUMAN)                                    | Q9Y794<br>(ORC4_SCHPO)                                  |
| ORC5         | NI                                | NI                             | O43913<br>(ORC5_HUMAN)                                    | O43114<br>(ORC5_SCHPO)                                  |
| ORC6         | NI                                | NI                             | Q9Y5N6<br>(ORC6_HUMAN)                                    | O74796<br>(ORC6_SCHPO)                                  |
| CDC6         | EhCDC6                            | EH1_192770                     | Q99741<br>(25.61% / 3e-08)                                | P41411<br>(23.10% / 3e-10)                              |
| Cdt1         | NI                                | NI                             | Q9H211<br>(CDT1_HUMAN)                                    | P40382<br>(CDT1_SCHPO)                                  |
| MCM2         | EhMCM2 <sup>#</sup>               | EH1_117970                     | P49736<br>(41.46% / 0.0)                                  | P40377<br>(42.59% / 0.0)                                |
| MCM3         | EhMCM3 <sup>#</sup>               | EH1_103600                     | P25205<br>(48.02% / 2e-147)                               | P30666<br>(42.22% / 0.0)                                |
| MCM4         | EhMCM4 <sup>#</sup>               | EH1_187720                     | P33991<br>(43.61% / 4e-176)                               | P29458<br>(43.07% / 0.0)                                |
| MCM5         | EhMCM5 <sup>#</sup>               | EH1_069980                     | P33992<br>(42.92% / 5e-163)                               | P41389<br>(44.09% / 7e-170)                             |
| MCM6         | EhMCM6 <sup>#</sup>               | EH1_118870                     | Q14566<br>(44.49% / 2e-178)                               | P49731<br>(45.17% / 0.0)                                |
| MCM7         | EhMCM7 <sup>#</sup>               | EH1_158110                     | P33993<br>(41.35% / 0.0)                                  | O75001<br>(44.89% / 8e-170)                             |
| MCM8         | EhMCM8                            | EH1_140670                     | Q9UJA3<br>(44.24% / 8e-161)                               | NI                                                      |

|        |           |            |                            |                            |
|--------|-----------|------------|----------------------------|----------------------------|
| MCM9   | EhMCM9    | EH1_076880 | Q9NXL9<br>(42% / 1e-155)   | NI                         |
| MCM10  | NI        | NI         | Q7L590<br>(MCM10_HUMAN)    | O42709<br>(MCM10_SCHPO)    |
| MCM BP | EhMCM BP  | EH1_166840 | Q9BTE3<br>(20.75% / 1e-96) | NI                         |
| IPI3   | EhIPI3_WD | EH1_137840 | Q9BV38<br>(19.14% / 2e-14) | Q10272<br>(23.72% / 1e-16) |

\*Número de acceso obtenido de AmoebaDB

a. Número de acceso obtenido de Uniprot y los porcentajes obtenidos mediante un blast de NCBI

NI: no identificado

# Proteínas previamente identificadas en *E. histolytica*

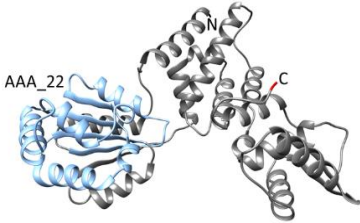
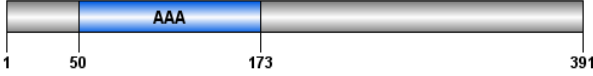
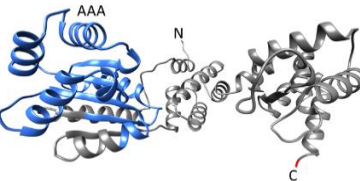
Por otra parte, se identificó al gen que codifica para la proteína CDC6, con un porcentaje de identidad del 25.61% y 23.10% en *H. sapiens* y *S. pombe*, respectivamente (Tabla 5). Esta proteína se denominó EhCDC6 y presenta un dominio AAA es un dominio proteico altamente conservado que forma parte de la superfamilia AAA+ ATPasas. Otras proteínas importantes de la etapa Pre-RC son las MCM 2-7. Identificamos 8 genes que codifican para proteínas MCM (MCM2–9) denominadas EhMCM de 2 a la 9 (Tabla 5). Estas proteínas presentan un alto grado de conservación con respecto a sus homólogas en *H. sapiens* y *S. pombe*, con porcentajes de identidad entre 41% y 48% en *H. sapiens*, y entre 42% y 45% en *S. pombe*. Entre ellas, MCM3 presentó el mayor porcentaje de identidad con *H. sapiens* (48.02%) (Tabla 5). MCM2–7 conforman el complejo MCM, mientras que MCM8 y MCM9 forman un complejo funcional (MCM8–9) que colabora con otras proteínas (RPA y RAD51) implicadas con la reparación; también ayudan para la formación del complejo MCM 2-7. Estructuralmente, se identificaron los dominios de las proteínas MCM2-9 y conservan varios dominios fundamentales para su función helicasa (Tabla 6):

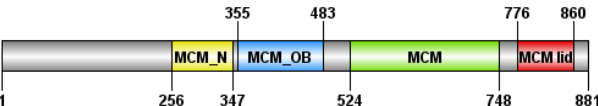
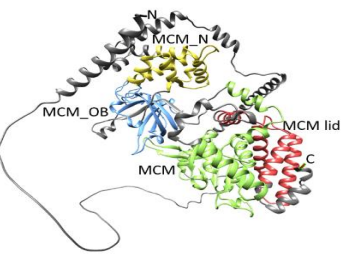
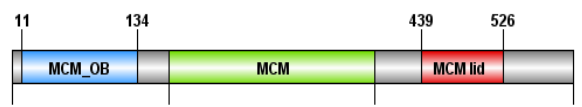
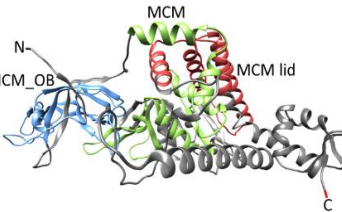
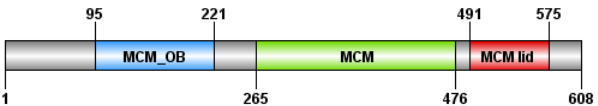
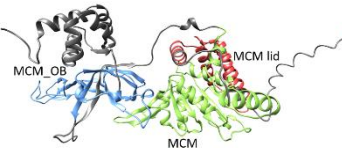
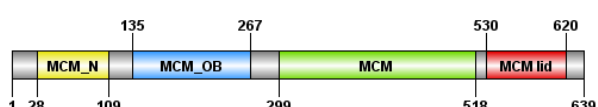
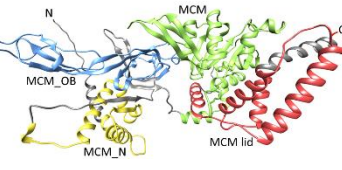
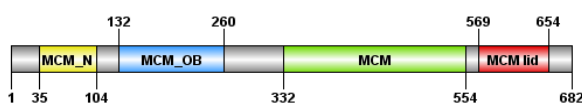
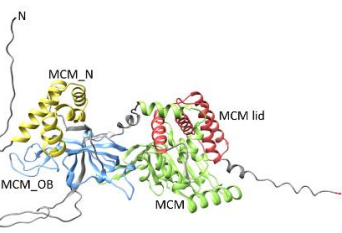
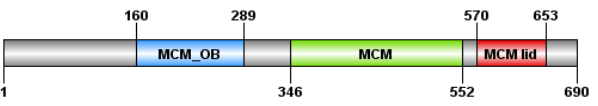
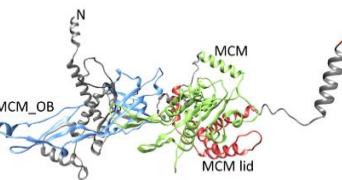
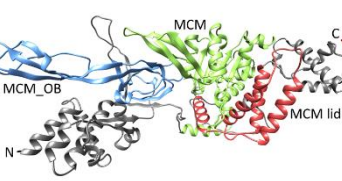
- MCM\_N: Dominio N-terminal con estructura tipo zinc finger implicado en la estabilización del complejo.
- MCM\_OB: pliegue tipo OB (del inglés, oligonucleotide oligosaccharide binding fold) asociado con la unión al DNA.
- MCM (AAA+ ATPasa): dominio central que contiene los motivos Walker A y B necesarios para la hidrólisis de ATP.

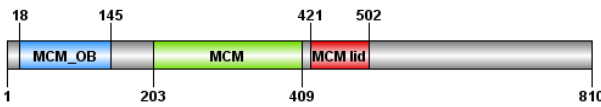
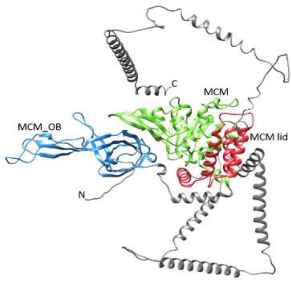
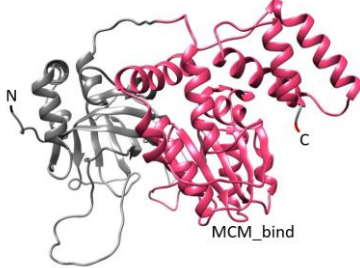
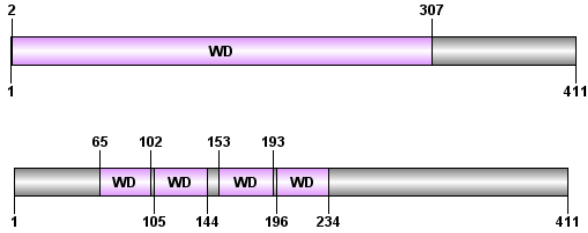
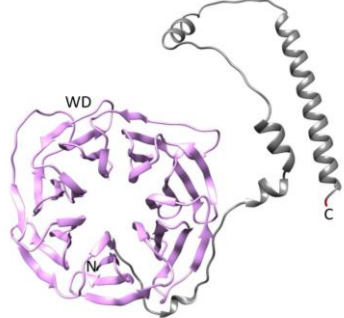
- MCM\_lid: Dominio C-terminal involucrado en el ensamblaje del anillo MCM y en la regulación de la actividad helicasa.

Además, identificamos otras proteínas adicionales que están relacionadas con el complejo Pre-RC, la proteína MCM-BP e IPI3\_WD, previamente descritas en mamíferos y levaduras, respectivamente, que participan durante la replicación del DNA (Saito *et al.*, 2022; Huang *et al.*, 2016). En *E. histolytica*, EhMCM BP y EhIPI3\_WD presentan valores de identidad de 20.75% y 19.14% con respecto a *H. sapiens*, respectivamente. La proteína EhIPI3\_WD presentó una identidad de 23.72% con respecto a *S. pombe* (Tabla 5). En EhMCM BP presenta el dominio MCM\_bind, que participa en la interacción con las MCM2-7 para asociar o disociar el complejo MCM en respuesta al daño celular (Samel *et al.*, 2014; Saito *et al.*, 2022). Por otra parte, en EhIPI3\_WD identificamos el dominio WD, el cual tiene 4 dominios repetidos de tipo WD40, que están involucrados en la interacción proteína-proteína y que contribuyen al reclutamiento del complejo MCM2-7 en los orígenes de replicación (Takahashi *et al.*, 2010; Campbell y Struhl, 2012) (Tabla 6).

**Tabla 6. Dominios funcionales y estructuras tridimensionales de las proteínas involucradas en la replicación de la etapa Pre-RC**

| Etapa Pre-RC                      |                                                                                      |                                                                                       |
|-----------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| Proteína /<br>Número de<br>acceso | Dominios* (aa)                                                                       | Estructura**                                                                          |
| EhORC1 /<br>EHI_105150            |  |  |
| EhCDC6 /<br>(EHI_192770)          |  |  |

|                                  |                                                                                      |                                                                                       |
|----------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| <p>EhMCM2 /<br/>(EHI_117970)</p> |    |    |
| <p>EhMCM3 /<br/>(EHI_103600)</p> |    |    |
| <p>EhMCM4 /<br/>(EHI_187720)</p> |    |    |
| <p>EhMCM5 /<br/>(EHI_069980)</p> |   |   |
| <p>EhMCM6 /<br/>(EHI_118870)</p> |  |  |
| <p>EhMCM7 /<br/>(EHI_158110)</p> |  |  |
| <p>EhMCM8 /<br/>(EHI_140670)</p> |  |  |

|                                     |                                                                                     |                                                                                      |
|-------------------------------------|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|
| <p>EhMCM9 /<br/>(EHI_076880)</p>    |   |   |
| <p>EhMCM BP /<br/>(EHI_166840)</p>  |   |   |
| <p>EhIP13_WD /<br/>(EHI_137840)</p> |  |  |

\* Dominios identificados con Interpro

\*\* Estructuras predichas con el programa Alpha fold

# Proteína previamente identificada

C: carboxilo terminal

N: amino terminal

### 8.1.2 Etapa Pre-IC

La etapa Pre-IC se caracteriza por la activación del complejo de replicación mediante la incorporación de factores esenciales como CDC45, el complejo GINS y CDKs (Moyer *et al.*, 2006). Aunque en esta etapa participan 13 a 15 proteínas (da Silva *et al.*, 2017), interesantemente en el genoma de *E. histolytica* identificamos 11 genes que codifican para proteínas de esta etapa Pre-IC, aunque también encontramos varios, dando un total de 20 genes codificantes para las proteínas de esta etapa.

Diez de los 20 genes codificantes presentan intrones (*ehpsf1*, *ehsld5*, *ehcdk1-2*, *ehcdk2-1*, *ehcdk2-2*, *ehcdk2-3*, *ehcdk2-4*, *ehrpa1-3*, *ehrpa2* y *ehrpa3*) y en general estos 20 genes codifican para proteínas cuyos pesos moleculares oscilan entre 12.39 y 155.44 kDa y con un pI entre 4.07 y 8.78. Para esta etapa solo encontramos 2 proteínas con señales de localización nuclear (EhTOP2 y EhTOP3 $\beta$ ) (Anexo 1). La proteína EhCDC45, caracterizada por Kurniawan y *et al.*, (2018) presenta una identidad superior al 20% con respecto a sus homólogos en *S. pombe* y *H. sapiens* (Tabla 7). Kurniawan y *et al.* (2018) identificaron por cristalografía a los dominios DHH N-, un dominio C-terminal y un dominio de interacción con CMG (CID), en esta proteína, así mismo encontraron un segmento conector  $\alpha$ -helicoidal (Tabla 8).

**Tabla 7. Comparación de los componentes de la etapa Pre-IC de la replicación del DNA de *E. histolytica* con respecto a los organismos *H. sapiens* y *S. pombe***

| Etapa Pre-IC    |                                   |                                |                                                         |                                                       |
|-----------------|-----------------------------------|--------------------------------|---------------------------------------------------------|-------------------------------------------------------|
| Nombre          | <i>E. histolytica</i><br>Proteína | <i>E. histolytica</i><br>(ID)* | <i>H. sapiens</i><br>(identidad %/valor e) <sup>a</sup> | <i>S. pombe</i><br>(identidad %/valor e) <sup>a</sup> |
| CDC45           | EhCDC45                           | EHI_049900                     | O75419<br>(22.39% / 2e-138)                             | O74113<br>(21.92% / 0.0)                              |
| GINS1<br>(Psf1) | EhPsf1                            | EHI_136990                     | Q14691<br>(28.10% / 4e-38)                              | Q9P7X6<br>(22.55% / 3e-30)                            |
| GINS2<br>(Psf2) | EhPsf2                            | EHI_069340                     | Q9Y248<br>(27.78% / 9e-66 )                             | O94329<br>(30.77% / 7e-55)                            |
| GINS3<br>(Psf3) | EhPsf3                            | EHI_069960                     | Q9BRX5<br>(20.80% / 7e-39 )                             | Q9UTC3<br>(18.14% / 2e-32 )                           |
| GINS4<br>(Sdl5) | EhSdl5                            | EHI_048430                     | Q9BRT9<br>(19.61% / 2e-49)                              | Q9P7C8<br>(21.95% / 5e-45)                            |
| CDK1            | EhCDK1-1                          | EHI_068070                     | P06493<br>(39.93% / 3e-73)                              | P04551<br>(39.25% / 7e-77)                            |
|                 | EhCDK1-2                          | EHI_105300                     | P06493<br>(44.29% / 3e-109)                             | P04551<br>(41.02% / 1e-79)                            |
|                 | EhCDK1-3                          | EHI_048710                     | P06493<br>(35.25% / 5e-107)                             | P06493<br>(35.10% / 4e-50)                            |
| CDK2            | EhCDK2-1                          | EHI_065280                     | P24941<br>(52.40% / 2e-106)                             | NI                                                    |

|            |                    |            |                             |                            |
|------------|--------------------|------------|-----------------------------|----------------------------|
|            | EhCDK2-2           | EHI_153140 | P24941<br>(46.67% / 1e-89 ) | NI                         |
|            | EhCDK2-3           | EHI_152430 | P24941<br>(47.37% / 2e-93)  | NI                         |
|            | EhCDK2-4           | EHI_054090 | P24941<br>(31.68% / 2e-38)  | NI                         |
| DDK (Cdc7) | NI                 | NI         | O00311<br>(CDC7_HUMAN)      | P41892<br>(CDC7_SCHPO)     |
| DDK (Dbf4) | NI                 | NI         | Q9UBU7<br>(DBF4A_HUMAN)     | O59836<br>(HIM1_SCHPO)     |
| Sld2       | NI                 | NI         | NI                          | O14216<br>(SLD2_SCHPO)     |
| Sld3       | NI                 | NI         | NI                          | Q09761<br>(SLD3_SCHPO)     |
| TOP2       | EhTOP2             | EHI_120640 | P11388<br>(47.90% / 0.0)    | P08096<br>(49.65% / 0.0)   |
| RPA1       | EhRPA1-1<br>70 kDa | EHI_062980 | P27694<br>(28.79% / 1e-32)  | Q92372<br>(24.94% / 8e-32) |
|            | EhRPA1-2<br>70 kDa | EHI_127070 | P27694<br>(18.18% / 2e-38)  | Q92372<br>(25.77% / 0.022) |
|            | EhRPA1-3<br>70 kDa | EHI_067930 | P27694<br>(23.73% / 0.007)  | Q92372<br>(23.23% / 1e-12) |
| RPA2       | EhRPA2<br>17 kDa   | EHI_147000 | P15927<br>(23.94% / 0.021)  | Q92373<br>(29.47% / 8e-05) |
| RPA3       | EhRPA3<br>12 kDa   | EHI_187270 | NI                          | Q92374<br>(25.81% / 0.018) |

\*Número de acceso obtenido de AmoebaDB

a. Número de acceso obtenido de Uniprot y los porcentajes obtenidos mediante un blast de NCBI

NI no identificado

En cuanto al complejo de GINS identificamos a las cuatro subunidades que la conforman codificadas en su genoma: EhPsf1, EhPsf2, EhPsf3 y EhSld5, con un rango de identidad entre 19% y 30%.

Estas subunidades del complejo GINS conservan el dominio GINS\_A correspondiente al dominio  $\alpha$ -helicoidal, el cual interactúa con las MCM, mientras que EhPsf2 y EhSld5 presentan un GINS\_B en la región C-terminal que interactúa con CDC45.

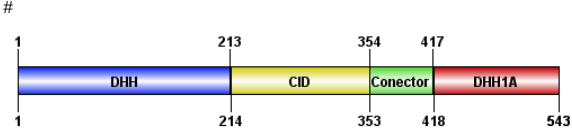
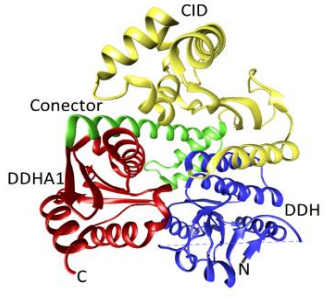
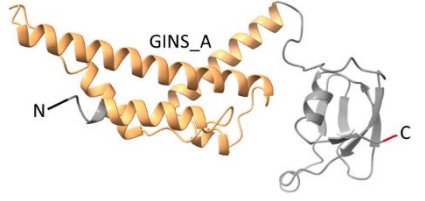
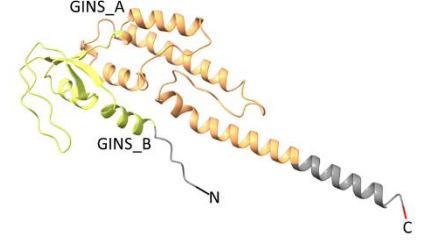
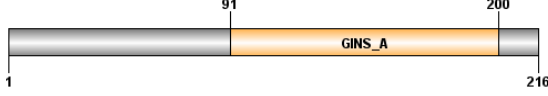
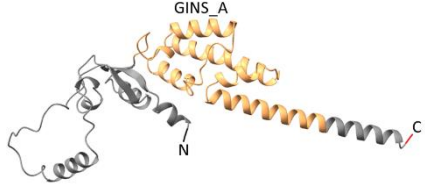
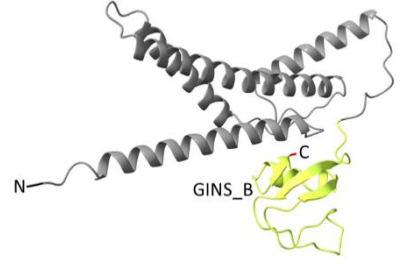
Por otro lado, identificamos múltiples isoformas de las proteínas CDK1 y CDK2 en *E. histolytica*, con identidades que van desde el 31.68% hasta el 52.40% en *H. sapiens*. CDK2 I, presentó el porcentaje más alto de identidad de 52.40%, lo que sugiere una buena conservación de estos reguladores durante el ciclo celular. Todas las isoformas que identificamos conservan el dominio característico Pkinase (Tabla 8) que contiene la función catalítica de las proteínas cinasas (Hanks *et al.*, 1988).

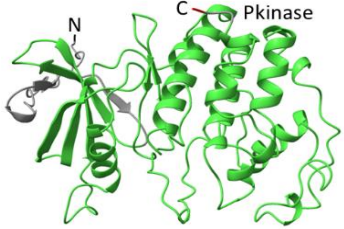
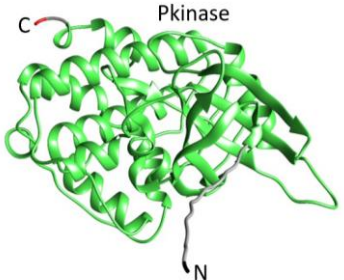
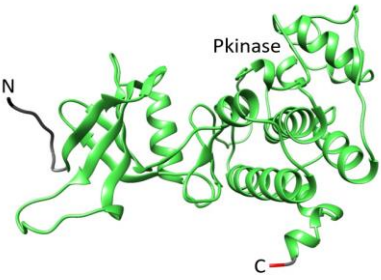
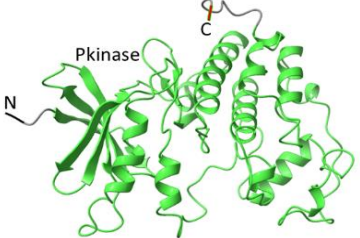
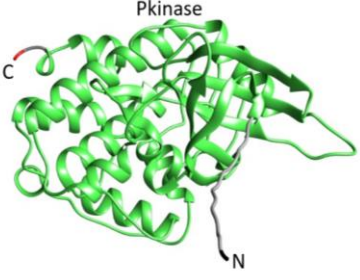
También identificamos a las topoisomerasas tipo II, con porcentajes 47.90 y 49.65% de identidad con respecto a *H. sapiens* y *S. pombe* (Tabla 7). En el análisis de los dominios identificamos 5 dominios conservados para la EhTOP II (Tabla 8); el dominio HATPase\_c es un dominio de hidrólisis de ATP, permitiendo un cambio conformacional que permite cortar y religar el DNA, el dominio DNA\_gyrase B es un dominio relacionado con la subunidad B de girasa bacteriana y que en eucariontes funciona en la unión de ATP y en el ensamblaje de complejos catalíticos (Oliveira-Costa *et al.*, 2010; Champoux, 2001).

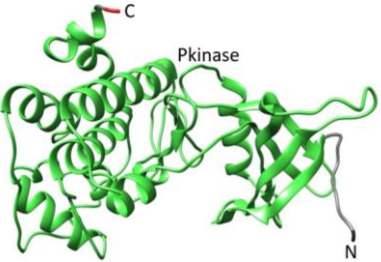
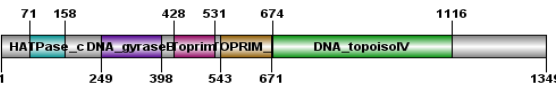
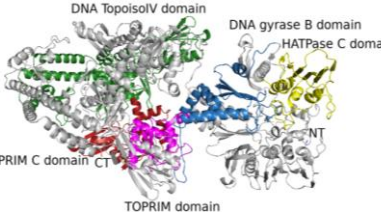
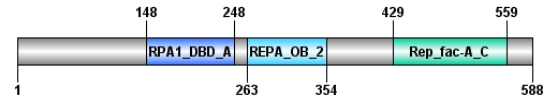
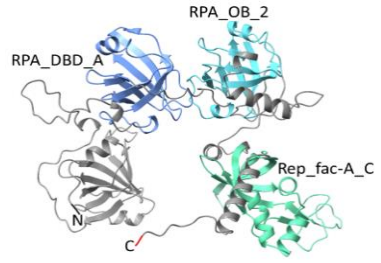
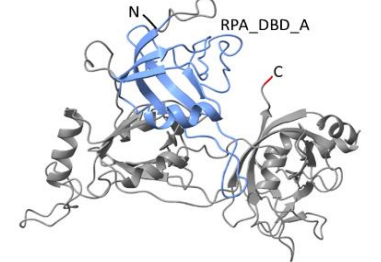
Finalmente, identificamos los genes que codifican para proteínas RPA y encontramos 3 isoformas para EhRPA1, las identidades de EhRPA1 a 3 están en un rango de 18.18% a 29.47%. En cuanto a los dominios, en las RPA de *E. histolytica* identificamos la mayoría de los dominios característicos, sin embargo, en algunas proteínas, como RPA3, no identificamos ningún dominio (Tabla 8).

En este trabajo no logramos identificar a los componentes del complejo DDK codificados en el genoma de *E. histolytica* (Cdc7 y Dbf4); sin embargo, encontramos 3 isoformas para la CDK 1 y 4 isoformas para la CDK 2. Tampoco identificamos a las proteínas Sld2 y Sld3 que forman parte de esta etapa en levadura.

**Tabla 8. Dominios funcionales y estructuras tridimensionales de las proteínas involucradas en la replicación de la etapa Pre-IC**

| Etapa Pre-IC                      |                                                                                      |                                                                                       |
|-----------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| Proteína /<br>Número de<br>acceso | Dominios* (aa)                                                                       | Estructura**                                                                          |
| EhCDC45# /<br>(EHI_049900)        |    |    |
| EhPsf1 /<br>(EHI_136990)          |    |   |
| EhPsf2 /<br>(EHI_069340)          |  |  |
| EhPsf3 /<br>(EHI_069960)          |  |  |
| EhSdl5 /<br>(EHI_048430)          |  |  |

|                                    |                                                                                      |                                                                                       |
|------------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| <p>EhCDK1-1 /<br/>(EHI_068070)</p> |    |    |
| <p>EhCDK1-2 /<br/>(EHI_105300)</p> |    |    |
| <p>EhCDK1-3 /<br/>(EHI_048710)</p> |   |   |
| <p>EhCDK2-1 /<br/>(EHI_065280)</p> |  |  |
| <p>EhCDK2-2 /<br/>(EHI_153140)</p> |  |  |

|                                    |                                                                                      |                                                                                       |
|------------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| <p>EhCDK2-3 /<br/>(EHI_152430)</p> |    |    |
| <p>EhCDK2-4 /<br/>(EHI_054090)</p> |    |    |
| <p>EhTOP2 /<br/>(EHI_120640)</p>   |   |   |
| <p>EhRPA1-1 /<br/>(EHI_062980)</p> |  |  |
| <p>EhRPA1-2 /<br/>(EHI_127070)</p> |  |  |

|                            |  |  |
|----------------------------|--|--|
| EhRPA1-3 /<br>(EHI_067930) |  |  |
| EhRPA2 /<br>(EHI_147000)   |  |  |
| EhRPA3 /<br>(EHI_187270)   |  |  |

\* Dominios identificados con Interpro

\*\* Estructuras predichas con el programa Alpha fold

# Proteína previamente identificada (Kurniawan *et al.*, 2018)

\* NI No identificado

C: carboxilo terminal

N: amino terminal

### 8.1.3 Etapa de síntesis (S)

Durante la etapa síntesis, el replisoma se ensambla por completo permitiendo la elongación de las nuevas cadenas de DNA. Esta etapa se caracteriza por la actividad coordinada de primasas, polimerasas, proteínas deslizantes como PCNA y los factores de carga que aseguran la alta procesividad del complejo de replicación (Stillman, 2008). En esta etapa participan 15 a 22 proteínas (da Silva *et al.*, 2017), sin embargo, en el genoma de *E. histolytica* identificamos solo 13 genes que codifican para proteínas de la etapa de síntesis, aunque también encontramos varios genes duplicados para la misma proteína que podrían corresponder a isoformas de la misma, dando un total de 19 genes codificantes para esta etapa.

Cuatro de los 19 genes codificantes presentan intrones (*ehdnprimase1s-1*, *ehdnprimase1s-2*, *ehrfc1-2* y *ehrfc5*). Y en general, estos 19 genes codifican para proteínas cuyos pesos moleculares oscilan entre 18.31 y 130.17 kDa con un pI entre 4.56 y 8.92.

En cuanto a la localización subcelular de estas proteínas, está predicha en el núcleo de la célula, aunque solo 3 proteínas (EhDNAprimase1S-1, EhRFC1-1 y EhDNApol $\alpha$ ) presentaron señal de localización nuclear (Anexo 1). En esta etapa, identificamos 3 isoformas para cada subunidad pequeña (S) y grande (L) de las DNA primasas (EhDNAprimase 1 y 2) estas presentan entre un 9.17% y un 35.10% de identidad con respecto a *H. sapiens* y *S. pombe* (Tabla 9). EhDNAprimase1S-2 presentó la identidad más baja de 9.17% contra *S. pombe*, mientras que EhDNAprimase1S-1 mostró el valor más alto de 35.10% contra *H. sapiens* (Tabla 10).

**Tabla 9. Comparación de los componentes de la etapa síntesis de la replicación del DNA de *E. histolytica* con respecto a los organismos *H. sapiens* y *S. pombe***

| Etapa síntesis (S) |                                   |                                |                                                          |                                                        |
|--------------------|-----------------------------------|--------------------------------|----------------------------------------------------------|--------------------------------------------------------|
| Nombre             | <i>E. histolytica</i><br>Proteína | <i>E. histolytica</i><br>(ID)* | <i>H. sapiens</i><br>(identidad %/ valor e) <sup>a</sup> | <i>S. pombe</i><br>(identidad %/ valor e) <sup>a</sup> |
| DNAprimase S       | EhDNAprimase1<br>S-1              | EHI_103530                     | P49642<br>(35.10% / 1e-65)                               | O14215<br>(31.84% / 4e-84)                             |
|                    | EhDNAprimase1<br>S-2              | EHI_129630                     | P49642<br>(31.82% / 4e-19)                               | O14215<br>(9.17% / 5e-19)                              |
|                    | EhDNAprimase1<br>S-3              | EHI_010140                     | P49642<br>(31.74% / 5e-59)                               | O14215<br>(32.45% / 3e-63)                             |
| DNAprimase L       | EhDNAprimase1<br>L-1              | EHI_127060                     | P49643<br>(22.37% / 2e-06)                               | O74761<br>(20.05% / 1e-20)                             |
|                    | EhDNAprimase1<br>L-2              | EHI_147530                     | P49643<br>(25.26% / 7e-27)                               | O74761<br>(20.66% / 3e-73)                             |
|                    | EhDNAprimase1<br>L-3              | EHI_008590                     | P49643<br>(22.87% / 1e-04)                               | O74761<br>(17.66% / 9e-15)                             |
| TOP3 $\alpha$      | EhTOP3 $\alpha$                   | EHI_038920                     | Q13472<br>(44.30% / 2e-180)                              | O60126<br>(40.84% / 3e-153)                            |

|                     |                     |            |                             |                             |
|---------------------|---------------------|------------|-----------------------------|-----------------------------|
| TOP3 $\beta$        | EhTOP3 $\beta$      | EHI_042880 | O95985<br>(38.56% / 0.0)    | O60126<br>(28.20% / 2e-161) |
| RFC1                | EhRFC1-1            | EHI_158610 | P35251<br>(26.64% / 3e-35)  | O60182<br>(22.93% / 2e-107) |
|                     | EhRFC1-2            | EHI_175240 | P35251<br>(22.78% / 1e-13)  | O60182<br>(12.84% / 2e-12)  |
| RFC2                | EhRFC2              | EHI_039320 | P35250<br>(24.18% / 1e-15)  | O14003<br>(25.00% / 2e-16)  |
| RFC3                | EhRFC3              | EHI_081730 | P40938<br>(18.12% / 3e-05)  | O14003<br>(26.69% / 3e-43)  |
| RFC4                | EhRFC4-1            | EHI_167150 | P35249<br>(42.72% / 6e-86)  | O94449<br>(44.76% / 4e-92)  |
|                     | EhRFC4-2            | EHI_086540 | P35249<br>(36.96% / 2e-81)  | O94449<br>(36.71% / 9e-78)  |
| RFC5                | EhRFC5              | EHI_099850 | P40937<br>(43.26% / 9e-89)  | O94697<br>(22.19% / 3e-65)  |
| PCNA                | EhPCNA              | EHI_128450 | P12004<br>(34.62% / 5e-64)  | Q03392<br>(39.38% / 4e-68)  |
| DNApol $\alpha$ 1   | EhDNApol $\alpha$   | EHI_151520 | P09884<br>(25.86% / 1e-169) | P28040<br>(25.66% / 1e-169) |
| DNApol $\alpha$ 2   | NI                  | NI         | Q14181<br>(DPOA2_HUMAN)     | O74946<br>(DPOA2_SCHPO)     |
| DNApol $\delta$ 1   | EhDNApol $\delta$   | EHI_006690 | P28340<br>(45.87% / 0.0)    | P30316<br>(45.98% / 0.0)    |
| DNApol $\delta$ 2   | NI                  | NI         | P49005<br>(DPOD2_HUMAN)     | P87324<br>(DPOD2_SCHPO)     |
| DNApol $\delta$ 3   | NI                  | NI         | Q15054<br>(DPOD3_HUMAN)     | P30261<br>(DPOD3_SCHPO)     |
| DNApol $\delta$ 4   | NI                  | NI         | Q9HCU8<br>(DPOD4_HUMAN)     | O59835<br>(DPOD4_SCHPO)     |
| DNApol $\epsilon$ 1 | NI                  | NI         | Q07864<br>(DPOE1_HUMAN)     | P87154<br>(DPOE_SCHPO)      |
| DNApol $\epsilon$ 2 | EhDNApol $\epsilon$ | EHI_141410 | P56282<br>(18.06% / 1e-17)  | O94263<br>(13.00% / 1e-24)  |

|           |    |    |                         |                        |
|-----------|----|----|-------------------------|------------------------|
| DNApolε 3 | NI | NI | Q9NRF9<br>(DPOE3_HUMAN) | C6Y4D0<br>(DPB3_SCHPO) |
| DNApolε 4 | NI | NI | Q9NR33<br>(DPOE4_HUMAN) | P87174<br>(DPB4_SCHPO) |

\*Número de acceso obtenido de AomebaDB

a. Número de acceso obtenido de Uniprot y los porcentajes obtenidos mediante un blast de NCBI  
NI no identificado

En cuanto a los dominios de las DNA primasas en *E. histolytica*, identificamos los dominios DNA\_primase\_S en las tres isoformas de EhDNAprimase1S (1, 2 y 3) y el dominio DNA\_primase\_lrg en las tres isoformas de EhDNAprimase2L (1, 2 y 3) (Tabla 10). Por otro lado, en las proteínas que constituyen al complejo RFC, se observó una amplia variación en los porcentajes de identidad, que van desde 18.12% hasta 43.26%. Para las TOP3α y TOP3β presentaron un porcentaje entre 28 a 44% para *H. sapiens* y *S. pombe*. En eucariontes presentan las dos regiones Toprim\_c y Toprim\_IV que conforman el dominio TOPRIM implicado en la coordinación catalítica con iones metálicos y la activación del sitio de corte (Oliveira-Costa *et al.*, 2010). También presentan el dominio DNA\_topoiso IV es el dominio central que realiza el corte y religación del DNA (Champoux, 2001). Las topoisomerasas III A y B que identificamos en *E. histolytica*, conservan los dominios Toprim y Topoisom\_bac, pero adicionalmente encontramos un dominio Zn\_ribbon\_TOP3B con respecto a EhTOPIIIβ, implicado en la catálisis del corte y religación del DNA (Champoux, 2001).

Las subunidades EhRFC4 y EhRFC5 fueron las más conservadas, con valores de identidad superiores al 40%. Respecto a los dominios, las subunidades EhRFC1-1, EhRFC4-1, EhRFC4-2 y EhRFC5 conservan el dominio AAA es un dominio proteico altamente conservado que forma parte de la superfamilia AAA+ ATPasas. Además, EhRFC1-1 contiene los dominios BRC1 y el dominio RFC1, característico de la subunidad grande del complejo clamp loader (Bunz *et al.*, 1998). En cuanto a las proteínas EhRFC4-1, EhRFC4-2 y EhRFC5, también presentan el dominio Rep\_factorC\_C, una región C-terminal involucrada en el ensamblaje estable del complejo RFC y su unión con PCNA (Neuwald *et al.*, 1999). EhRFC1-2 y RFC2 contienen el dominio P-loop\_NTPase, típico de proteínas que se unen a nucleótidos y catalizan la hidrólisis de ATP, función clave para la actividad dinámica del complejo RFC (Walker *et al.*, 1982).

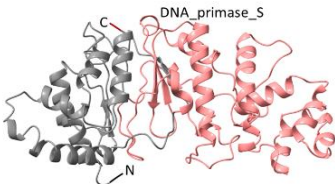
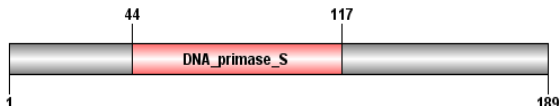
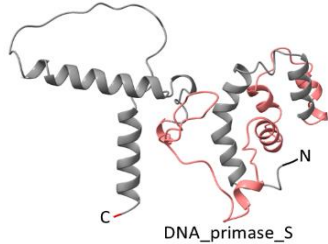
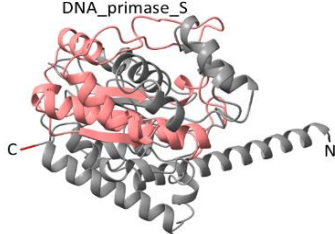
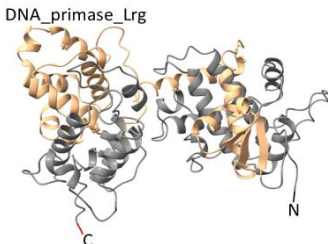
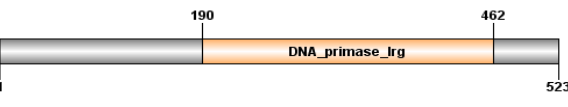
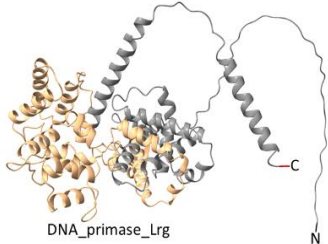
Adicionalmente, la proteína EhRFC1-2 presenta el dominio DNA\_pol3\_clamp-load\_cptx\_C, implicado en el reconocimiento del DNA y la carga de PCNA (Bunz *et al.*, 1998). Finalmente, EhRFC3 presenta dos dominios particulares: DNA\_pol\_3\_delta, que está asociado estructuralmente a las polimerasas tipo III, y RFC\_C3, que está implicado en la interacción con otras subunidades del complejo RFC y posiblemente con el DNA (Bunz *et al.*, 1998).

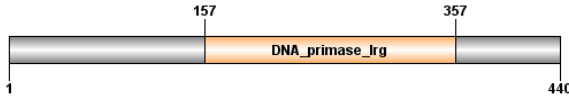
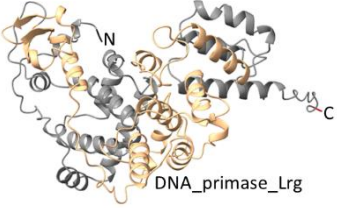
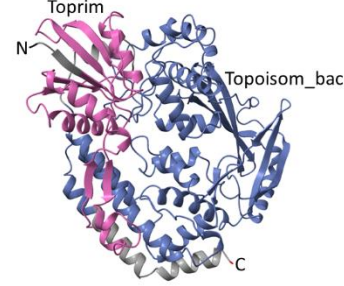
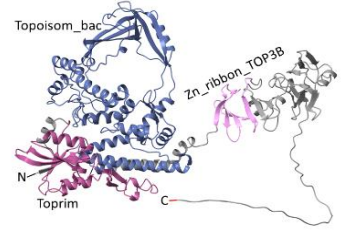
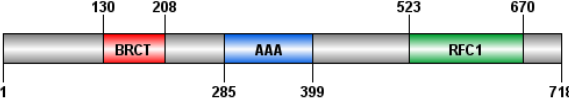
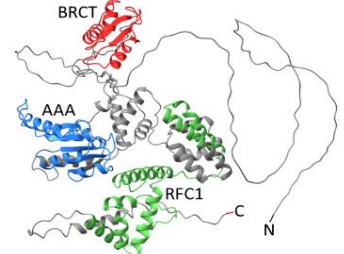
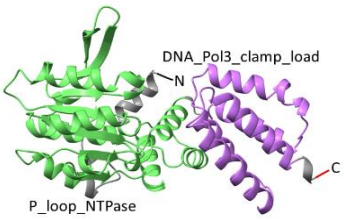
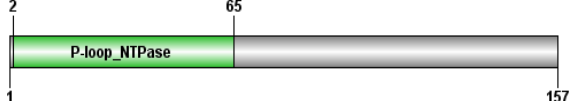
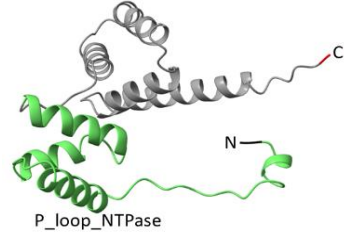
Por otra parte, determinamos los porcentajes de identidad de la proteína PCNA previamente reportada en *E. histolytica* por Cardona-Felix y *et al.* (2011). EhPCNA presenta una identidad entre el 34.62% y 39.38% contra *H. sapiens* y *S. pombe*, respectivamente. Con respecto a los dominios de EhPCNA, Cardona-Felix y *et al.* (2011) identificaron 2 dominios: el dominio PCNA\_N que corresponde al dominio N-terminal y el dominio PCNA\_C que corresponde a C-terminal, que a su vez, están conectados por un bucle de conexión interdominios (Tabla 10). En cuanto a las DNA polimerasas, en el genoma de *E. histolytica* encontramos genes que codifican para las proteínas EhDNApol  $\alpha$ ,  $\delta$  y  $\epsilon$ . La proteína EhDNApol $\delta$  fue la más conservada, con un porcentaje cercano al 46%, mientras que la subunidad 2 de EhDNApol $\epsilon$  mostró los niveles más bajos de identidad (13.00% y 18.06% con respecto a *H. sapiens* y *S. pombe*).

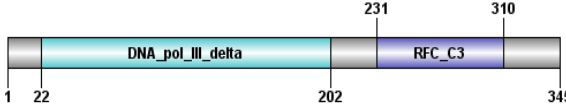
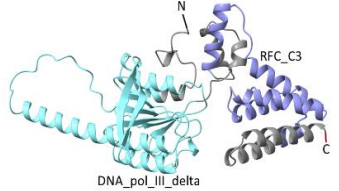
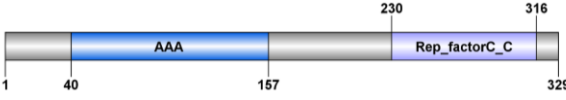
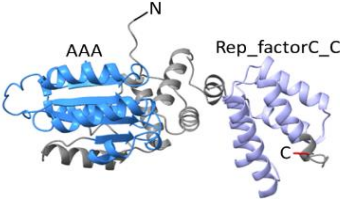
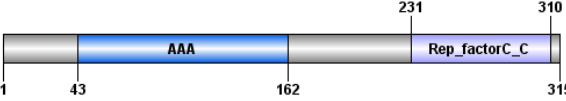
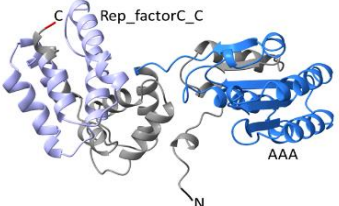
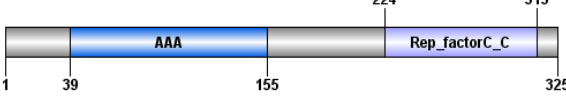
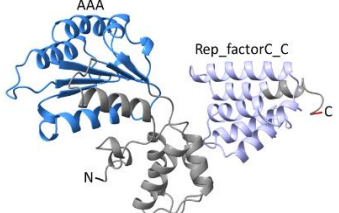
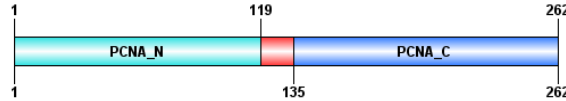
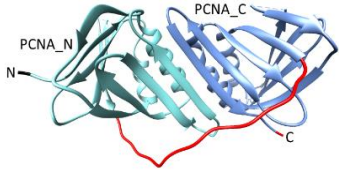
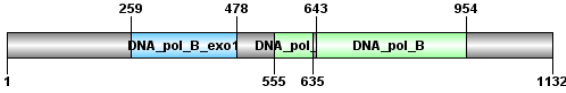
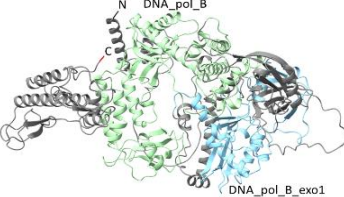
Esta proteína conserva los dominios DNA\_pol\_B y DNA\_pol\_B\_exo1 que presentan actividad de polimerasa y de exonucleasa 3' a 5' para corrección de errores, respectivamente (Aravind y Koonin, 1998). En la proteína EhDNApol $\delta$  identificamos un dominio denominado zf-C4pol, tipo dedos de zinc, que está implicado en la interacción con ácidos nucleicos y proteínas accesorias (Aravind y Koonin, 1998).

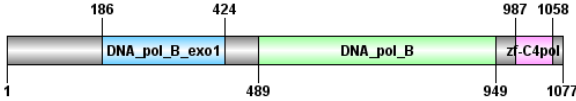
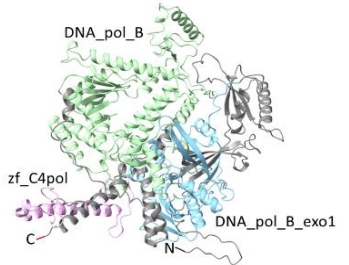
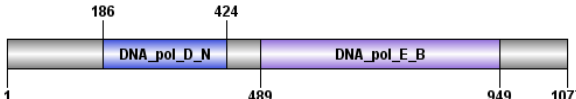
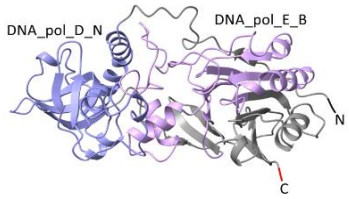
En cuanto a la proteína EhDNApol $\epsilon$ , los dominios que identificamos son: DNA\_pol\_D\_N y DNA\_pol\_E\_B están involucrados en la síntesis de la cadena líder y en la regulación del ciclo celular (Henninger y Pursell, 2014). Además, identificamos 3 isoformas para las 2 subunidades de DNA primase S y L, 2 isoformas para RFC1, 2 isoformas para RFC4 y no encontramos las subunidades que conforman DNApol $\delta$  (2, 3 y 4) al igual que DNApol $\epsilon$  (2, 3 y 4), solo identificamos una única proteína tanto para EhDNApol $\delta$  como para EhDNApol $\epsilon$ .

**Tabla 10. Dominios funcionales y estructuras tridimensionales de las proteínas involucradas en la replicación de la etapa de síntesis**

| Etapa síntesis (S)                    |                                                                                      |                                                                                       |
|---------------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| Proteína / Número de acceso           | Dominios* (aa)                                                                       | Estructura**                                                                          |
| EhDNAprimase1S-1<br>/<br>(EHI_103530) |    |    |
| EhDNAprimase1S-2<br>/<br>(EHI_129630) |    |    |
| EhDNAprimase1S-3<br>/<br>(EHI_010140) |  |   |
| EhDNAprimase1L-1 /<br>(EHI_127060)    |  |  |
| EhDNAprimase1L-2 /<br>(EHI_147530)    |  |  |

|                                            |                                                                                      |                                                                                       |
|--------------------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| <p>EhDNAprimase1L-3 /<br/>(EHI_008590)</p> |    |    |
| <p>EhTOP3α /<br/>(EHI_038920)</p>          |    |    |
| <p>EhTOP3β /<br/>(EHI_042880)</p>          |    |   |
| <p>EhRFC1-1 /<br/>(EHI_158610)</p>         |  |  |
| <p>EhRFC1-2 /<br/>(EHI_175240)</p>         |  |  |
| <p>EhRFC2 /<br/>(EHI_039320)</p>           |  |  |

|                                     |                                                                                      |                                                                                       |
|-------------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| <p>EhRFC3 /<br/>(EHI_081730)</p>    |    |    |
| <p>EhRFC4-1 /<br/>(EHI_167150)</p>  |    |    |
| <p>EhRFC4-2 /<br/>(EHI_086540)</p>  |    |    |
| <p>EhRFC5 /<br/>(EHI_099850)</p>    |  |  |
| <p>EhPCNA# /<br/>(EHI_128450)</p>   |  |  |
| <p>EhDNApolα /<br/>(EHI_151520)</p> |  |  |

|                                                 |                                                                                    |                                                                                     |
|-------------------------------------------------|------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|
| <p>EhDNApolδ<sup>#</sup> /<br/>(EHI_006690)</p> |  |  |
| <p>EhDNApolε /<br/>(EHI_141410)</p>             |  |  |

\* Dominios identificados con Interpro

\*\* Estructuras predichas con el programa Alpha fold

# Proteína previamente identificada (Pastor-Palacios *et al.*, 2010; Cardona-Felix *et al.*, 2011)

C: Carboxilo terminal

N: amino terminal

#### 8.1.4 Etapa terminación (T)

En la etapa de terminación se llevan a cabo los eventos finales del proceso de replicación del DNA como 1) la eliminación de cebadores de RNA, 2) el sellado de los fragmentos de Okazaki y 3) la protección de los extremos cromosómicos, asegurando así la integridad genómica y la correcta finalización del ciclo replicativo (Burgers y Kunkel, 2017; Carty *et al.*, 2003). Durante la etapa de terminación participan entre 6 y 10 proteínas aprox. (Dewar y Walter, 2017). En el genoma de *E. histolytica*, identificamos 7 genes que codifican para proteínas de la etapa de terminación (Tabla 11), aunque también encontramos varios genes duplicados para la misma proteína que podrían corresponder a isoformas de esta, dando un total de 8 genes codificantes para esta etapa. Dos de los 8 genes identificados presentan intrones (*ehRNAaseh2-1* y *ehRNAaseh2-2*). Los 8 genes codifican proteínas con pesos moleculares entre 21.97 y 78.16 kDa y con un *pI* entre 5.13 y 8.90.

La localización subcelular de estas proteínas está predicha en el núcleo de la célula y en congruencia se encontró señal de localización nuclear en 6 proteínas (EhRNAaseH2-2, EhDNALig1, EhFen1, EhTRF-Like1, EhTRF-Like2 y EhTRF-Like3) (Anexo 1). Además, identificamos 2 isoformas para la subunidad de RNAaseH2.

La proteína EhTRF-Like III, previamente caracterizada por Rendón-Gandarilla y *et al.*, (2018) no se identificó en *H. sapiens* y *S. pombe*.

**Tabla 11. Comparación de los componentes de la etapa de terminación de la replicación del DNA de *E. histolytica* con respecto a los organismos *H. sapiens* y *S. pombe***

| Etapa de terminación (T) |                                   |                                |                                                         |                                                        |
|--------------------------|-----------------------------------|--------------------------------|---------------------------------------------------------|--------------------------------------------------------|
| Nombre                   | <i>E. histolytica</i><br>Proteína | <i>E. histolytica</i><br>(ID)* | <i>H. sapiens</i><br>(identidad %/valor e) <sup>a</sup> | <i>S. pombe</i><br>(identidad %/ valor e) <sup>a</sup> |
| RNAaseH1                 | EhRNAaseH1                        | EHI_042240                     | O60930<br>(31.11% / 1e-16)                              | Q9UST8<br>(35.56% / 6e-19)                             |
| RNAaseH2                 | EhRNAaseH2-1                      | EHI_084700                     | O75792<br>(36.12% / 9e-85)                              | Q10236<br>(38.39% / 6Le-69)                            |
|                          | EhRNAaseH2-2                      | EHI_134360                     | O75792<br>(36,98% / 4e-60)                              | Q10236<br>(38.46% / 1e-56)                             |
| DNALig1                  | EhDNALig1                         | EHI_111060                     | P18858<br>(36.19% / 0.0)                                | P12000<br>(35.29% / 0.0)                               |
| Fen1                     | EhFen1                            | EHI_099740                     | P39748<br>(49.47% / 1e-171)                             | P39750<br>(45.24% / 1e-168)                            |
| TRF-Like I               | EhTRF-Like I                      | EHI_001090                     | P54274<br>(40.00% / 2e-14)                              | Q6E434<br>(27.69% / 6e-10)                             |
| TRF-Like II              | EhTRF-Like II                     | EHI_001110                     | Q15554<br>(37.93% / 6e-15)                              | P79005 (Taz1)<br>(30.99% / 9e-10)                      |
| TRF-Like III             | EhTRF-Like III                    | EHI_148140                     | NI                                                      | NI                                                     |

\* Número de acceso obtenido de AomebaDB

a. Número de acceso obtenido de Uniprot y los porcentajes obtenidos mediante un blast de NCBI

NI no identificado

Las proteínas EhRNAasaH1 y dos variantes de EhRNAasaH2 mostraron porcentajes de identidad entre 31.11% y 38.46%. En cuanto a los dominios, identificamos que EhRNAasaH1 contiene el dominio Ribonuclease\_H1, responsable de su capacidad catalítica para degradar los primers de RNA. Por otro lado, EhRNAasa H2 I y II conservan el dominio RNase\_H\_2 que participa en la eliminación de cebadores y en la vigilancia del genoma (Burgers y Kunkel, 2017) (Tabla 12).

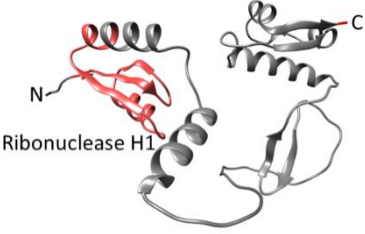
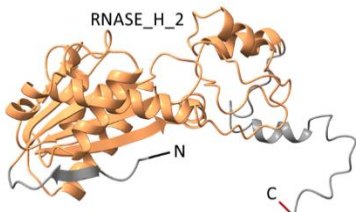
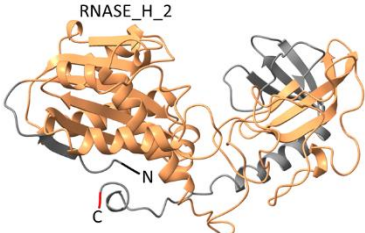
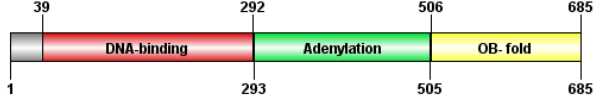
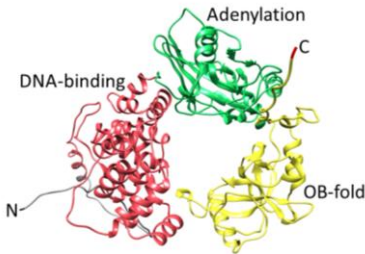
De manera interesante, *E. histolytica* solamente presenta una proteína de tipo ligasa previamente caracterizada por Cardona-Felix y *et al.* (2010).

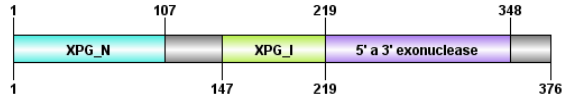
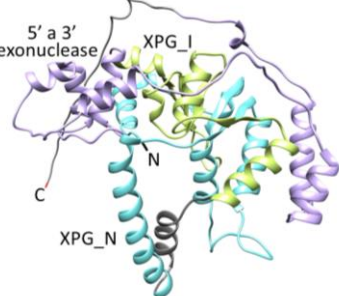
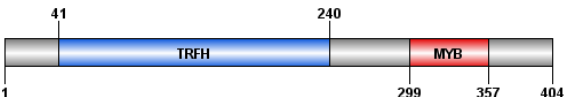
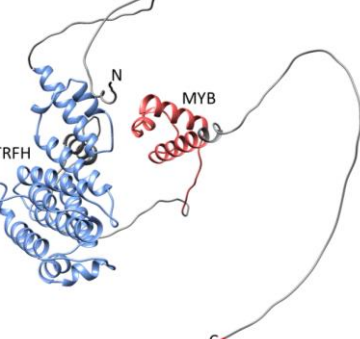
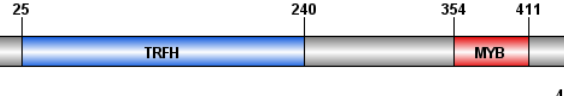
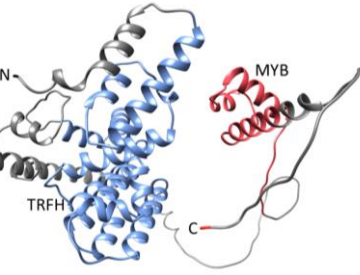
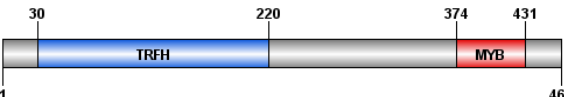
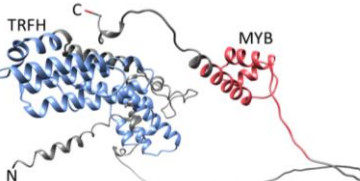
Esta proteína presentó valores de identidad ligeramente superiores al 35% en ambos organismos de referencia (*H. sapiens* y *S. pombe*), lo que indica una conservación funcional relevante, ya que cataliza la unión covalente de los fragmentos de Okazaki (Cardona-Felix *et al.*, 2010). Se ha reportado previamente que, a nivel estructural, la EhDNALig1 presenta tres dominios: un dominio de unión al DNA (DNA binding), un dominio de adenilación, y un dominio OB fold. Estos dominios corresponden a las regiones N-terminal, central y C-terminal de la proteína, respectivamente (Cardona-Felix *et al.*, 2010). Identificamos también en *E. histolytica* a la proteína EhFen1 que presenta un valor de identidad de 49.47% y 45.24% para *H. sapiens* y *S. pombe*. En EhFen1 identificamos dos dominios: XPG\_N y XPG\_I y un subdominio C-terminal con pliegue de exonucleasa 5'→3', los cuales son comunes en nucleasas de tipo XPG y están relacionados con la actividad de escisión de estructuras en “flap”, que aparecen durante la replicación y la reparación por escisión de bases (González-Corrochano *et al.*, 2020).

Por otra parte, Rendón-Gandarilla y *et al.* (2018) previamente caracterizó 3 proteínas TRF-like de *E. histolytica* y determinó que podrían estar relacionadas con la protección de los extremos teloméricos en este parásito, ya que presentan los dominios de repetición telomérica (TRF) asociados con el mantenimiento de la estabilidad de los telómeros al evitar su reconocimiento como rupturas de DNA (Smogorzewska y de Lange, 2004). Estas proteínas mostraron valores entre 27.69% y 40% de identidad (Tabla 11). Las proteínas EhTRF-like presentan además el dominio TRFH, que permite que las proteínas TRF1 y TRF2 se asocien entre sí formando homodímeros o multímeros (Palm y De Lange, 2008).

El otro dominio que conservan es el dominio de unión a DNA MYB (DBD-MYB, del inglés DNA binding domain), que está conservado en varios eucariontes y la característica común de este dominio es su ubicación cerca del N-terminal y consta de un repetido de aprox. 50 aa de longitud que a su vez forman 3 hélices alfa involucradas en el reconocimiento de la secuencia telomérica de DNA. (Arratia y Aguirre, 2013) (Tabla 12).

**Tabla 12. Dominios funcionales y estructuras tridimensionales de las proteínas involucradas en la replicación de la etapa de terminación**

| Etapa terminación                 |                                                                                      |                                                                                       |
|-----------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| Proteína /<br>Número de<br>acceso | Dominios* (aa)                                                                       | Estructura**                                                                          |
| EhRNAaseH1 /<br>(EHI_042240)      |    |    |
| EhRNAaseH2-1 /<br>(EHI_084700)    |    |    |
| RNAsa H2 II /<br>(EHI_134360)     |  |  |
| EhDNALig1# /<br>(EHI_111060)      |  |  |

|                                        |                                                                                      |                                                                                       |
|----------------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| <p>EhFen1 /<br/>(EHI_099740)</p>       |    |    |
| <p>EhTRF-Like1# /<br/>(EHI_001090)</p> |    |    |
| <p>EhTRF-Like2# /<br/>(EHI_001110)</p> |  |   |
| <p>EhTRF-Like3# /<br/>(EHI_148140)</p> |  |  |

\* Dominios identificados con InterPro

\*\* Estructuras predichas con el programa AlphaFold.

# Proteína previamente identificada (Cardona-Félix *et al.*, 2010; Rendón-Gandarilla *et al.*, 2018)

C: carboxilo terminal

N: amino terminal

En resumen, de las ~63 proteínas identificadas en *H. sapiens* y *S. pombe* que participan en replicación, en el genoma de *E. histolytica* identificamos 43 genes que codifican para las mismas (Pre-RC 19 vs 12; Pre IC 15 vs 13; S 19 vs 22 T 10 vs 7), sin considerar las isoformas de las mismas proteínas, que nos darían un total de 57 proteínas.

Esto nos sugiere que *E. histolytica* presenta una maquinaria de replicación simplificada donde EhORC1, EhDNAIlg y las DNA polimerasas podrían ser críticas para la replicación del genoma de este organismo.

## **8.2 Interactoma del replisoma de *E. histolytica***

Para determinar si las proteínas identificadas y codificadas en el genoma de este parásito interactúan entre sí conformando el replisoma de *E. histolytica*, se hizo un interactoma *in silico* empleando la plataforma STRING. Este análisis se hizo incluyendo las 57 proteínas que participan en las cuatro etapas de la replicación del DNA: Pre-RC (azul), Pre-IC (celeste), síntesis (verde) y terminación (amarillo).

El interactoma obtenido presenta un total de 57 nodos correspondientes a las proteínas analizadas. La red de interacciones contiene 807 interacciones (*edges*), con un grado promedio por nodo de 28.3, lo que indica una alta conectividad entre las proteínas (Figura 11). El coeficiente promedio de agrupamiento local fue de 0.787, lo cual sugiere una fuerte tendencia a formar clústeres funcionales. Comparado con el número esperado de interacciones (56), el número observado es considerablemente mayor, lo que indica que las proteínas no están conectadas al azar. Esta diferencia es estadísticamente significativa, con un valor de *p* para el enriquecimiento en interacciones proteína-proteína (PPI) menor a  $1.0e-16$ . En este análisis observamos que 3 proteínas no presentaron ninguna interacción, (EhRPA1-2, EhRPA3 y EhRNAasaH1) las cuales se eliminaron del interactoma obtenido. El interactoma obtenido muestra que las proteínas que conforman la etapa Pre-RC (ORC1, CDC6, MCM2-7, MCM8-9, MCMBP e IPI3) están interconectadas entre sí e interactúan con las proteínas de las otras etapas (Pre-IC, S y T). Los nodos de esta etapa son centrales, a excepción de los que corresponden a EhIPI3\_WD y EhMCM BP, las cuales están alejadas, pero mantienen sus interacciones e interactúan con otros nodos de diferentes etapas. En el caso de EhIPI3\_WD, mantiene interacciones con las EhCDK 1 y 2, en cuanto a EhMCM BP, está interactuando con las proteínas del complejo EhMCM y con otras proteínas de otras etapas como S y T (EhDNAprimase y EhRNAase respectivamente). En cuanto a la etapa Pre-IC, se conforma de las proteínas EhCDK1 y 2, EhCDC45, el complejo de GINS (Psf1, Psf2, Psf3 y Sld5), EhRPA (1, 2 y 3).

Es interesante que las isoformas de las proteínas EhCDK 1 y 2 están muy cercanas entre sí e interactuando con otras proteínas de otras etapas, lo cual refleja su papel importante de regular el ciclo celular; de igual manera pasa con el complejo de GINS.

Con respecto a la etapa síntesis, observamos que las proteínas EhDNAprimase, EhRFCs, EhPCNA y DNA polimerasas interactúan tanto con las proteínas de la propia etapa como también de otras etapas. Finalmente, para la etapa de terminación, conformada por EhRNAsaH 1 y 2, EhFen1, EhDNA ligasa y EhTRF-like (I, II y III) observamos que presentan múltiples interconexiones, lo cual concuerda con su papel crucial en la etapa tardía de la replicación, pero coordinada con el resto del replisoma. Se observó que las EhTRF-like presentan pocas interacciones, lo que sugiere que, al participar en las etapas terminales de mantenimiento telomérico, posiblemente interaccionen con menos proteínas relacionadas con el proceso de replicación, mientras que EhFen1 y EhDNALig1 están más cercanas a las proteínas de la etapa síntesis (Figura 11).

El análisis de enriquecimiento funcional identificó procesos biológicos fuertemente representados entre las proteínas identificadas. El proceso más enriquecido fue “Replicación del DNA”, que incluyó ~40 proteínas y mostró una tasa de descubrimiento falso (FDR) muy baja ( $\approx 1.0e-57$ ), lo que indica una alta significancia estadística. Otros términos con un enriquecimiento destacado fueron “Inicio de la replicación del DNA” y “Elongación del DNA durante la replicación”, con valores de FDR inferiores a  $1.0e-48$  y  $1.0e-39$ , respectivamente. Estos procesos están estrechamente relacionados con la replicación del DNA y el mantenimiento de la integridad genómica, lo que refuerza la hipótesis del replisoma de *E. histolytica* (Figura 12 A).



Le siguen en relevancia términos estrechamente relacionados como “división nuclear” y “segregación cromosómica”, con FDR menores a  $1.0e-19$  y  $1.0e-17$  respectivamente. Estos resultados muestran que las proteínas identificadas se encuentran relacionadas con procesos como la mitosis, la segregación de cromátidas hermanas, la organización del huso mitótico, y la condensación cromosómica, todos ellos representados con significancia estadística considerable ( $FDR < 1.0e-14$ ) (Figura 12 B).

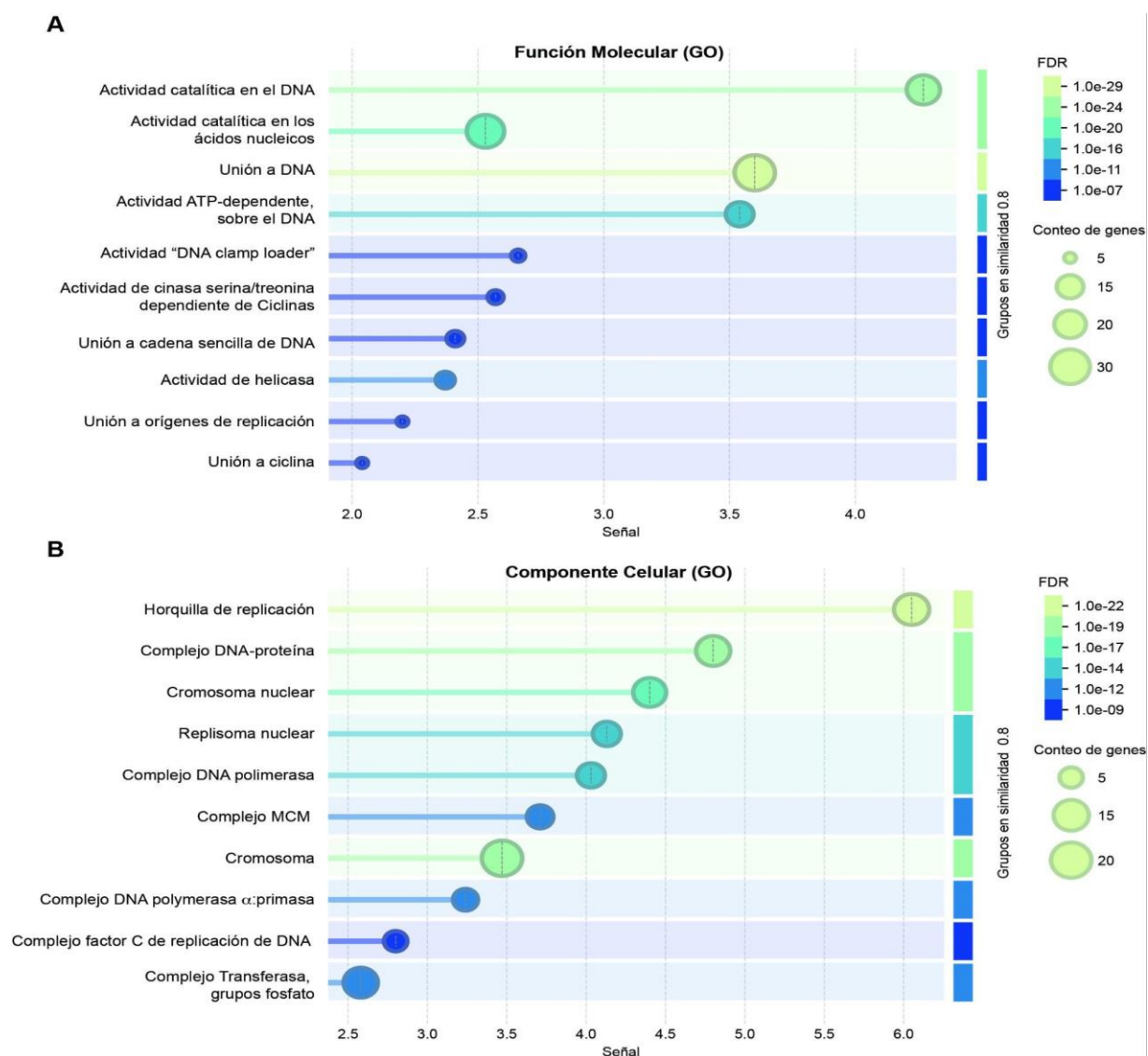


**Figura 12. Análisis de enriquecimiento funcional del replisoma de *E. histolytica*.** Los gráficos de burbujas muestran los términos significativamente enriquecidos: (A) Función molecular y (B) Componente celular. El tamaño de las burbujas representa el número de proteínas asociadas a cada término, mientras que el color indica la significancia estadística del enriquecimiento (FDR).

Este análisis *in silico* reveló una red altamente conectada entre las proteínas asociadas a la replicación del DNA en *E. histolytica*, lo que proporciona evidencia de un replisoma en este parásito. Así mismo, estos resultados permiten proponer que *E. histolytica* posee un conjunto conservado y funcionalmente organizado de proteínas que integran un replisoma activo, conectado además con mecanismos clave del ciclo celular.

### **8.3 Expresión diferencial de genes del replisoma durante la depleción y reposición de suero.**

Para explorar la dinámica de expresión de genes que codifican proteínas del replisoma en *Entamoeba histolytica*, analizamos los estudios transcriptómicos reportados en la base de datos AmoebaDB, específicamente el reportado por Naiyer y *et al.* (2019). Este transcriptoma se obtuvo a partir de trofozoítos de *E. histolytica* de las siguientes condiciones experimentales: 1) trofozoítos cultivados en condiciones normales de cultivo, 2) sometidos a depleción de suero durante 24 h y 3) condición de reposición de suero por 2 h. Aunque Naiyer *et al.*, 2019 estudiaron el efecto en el transcriptoma de la depleción de suero como una condición de estrés, a nosotros nos interesó ya que previamente se reportó que la privación de suero provoca un arresto reversible en el ciclo celular de trofozoítos de *E. histolytica*, emulando así una sincronización celular (Austin y Warren, 1983). Esto se debe a que la falta de factores de crecimiento y nutrientes que están presentes en el suero impiden que las células avancen en el ciclo celular y así mismo se detiene la síntesis de proteínas y se bloquea la progresión de la fase S.

Se obtuvieron los transcritos por millón (TPM) de las condiciones antes descritas (basal, privación de suero y reposición de suero), y se visualizaron los patrones de expresión de los datos normalizados por medio de un gráfico tipo mapa de calor, el cual muestra los niveles de expresión de los genes asociados al replisoma bajo las tres condiciones experimentales. Los colores representan la magnitud del cambio en la expresión (escala de -1 a 1), donde el rojo indica sobreexpresión relativa y el azul, subexpresión relativa (Figura 13). De manera interesante, se observan diferentes agrupamientos (*clusters*): 1) dependiendo de la condición y 2) por grupo de genes.

En el caso del agrupamiento por condición, el mapa de calor muestra que en la condición basal, la mayoría de los genes se expresan poco (tonos azules), lo que sugiere una actividad replicativa baja, probablemente debido a que el cultivo se encuentra asincrónico. En la depleción de suero, se observa un aumento importante en la expresión de muchos genes, lo cual podría reflejar una activación que permita a los trofozoítos prepararse para la replicación. En la reposición de suero, se detecta una sobrerregulación más intensa de un subconjunto de genes, lo que sugiere que el retorno de nutrientes promueve la reactivación del ciclo celular y la replicación del DNA (Figura 13). Para el agrupamiento por grupo de genes, estos se agruparon según su patrón de expresión en seis grupos señalados en recuadros de colores (rojo, naranja, azul, morado, amarillo y verde) (Figura 13).

Grupo rojo: se conforma de 13 genes que mostraron muy alta expresión en la condición basal. Sin embargo, muestran una drástica disminución durante la depleción de suero y su expresión permanece baja tras la reposición. Este patrón sugiere que estos genes podrían estar activos durante fases específicas del ciclo celular (Fig. 13, recuadro rojo).

Grupo naranja: agrupa 8 genes que presentan en la condición basal una expresión alta y muy alta; con fuerte represión tras la depleción de suero y aumento en su expresión después de la reposición de suero. Este comportamiento podría corresponder a genes involucrados en la fase S/G2, los cuales se están expresando constitutivamente para mantener la replicación del DNA en condiciones normales y una reactivación progresiva del ciclo celular.

Grupo azul: incluye 4 genes que, en la condición basal, presentan una baja expresión y que en la condición de reposición de suero presentaron una muy alta expresión. Lo que sugiere que podrían ser genes clave en el inicio de la replicación, activados por señales de entrada a la fase S.

Grupo morado: En este grupo se encuentran 5 genes de baja expresión en la condición basal, curiosamente, en la depleción de suero hubo un gran incremento en su expresión, seguido de una disminución tras la reposición de suero. Este comportamiento sugiere que estos genes podrían estar implicados en la reconfiguración de la maquinaria replicativa durante la fase G1.

Grupo amarillo: es el grupo que integra el mayor número de genes (16), los cuales incrementan moderadamente su expresión durante la depleción, con regulación variable tras la reposición de suero. Lo cual podría implicar una preparación temprana de la maquinaria replicativa durante el arresto celular inducido por la depleción.

Grupo verde: compuesto por 11 genes que mostraron muy baja expresión en la condición basal, aumento progresivo con la depleción y máxima activación tras la reposición de suero. Este comportamiento sugiere una activación dependiente de fase S, en donde estos genes estarían directamente involucrados en la síntesis y finalización de la replicación del DNA.

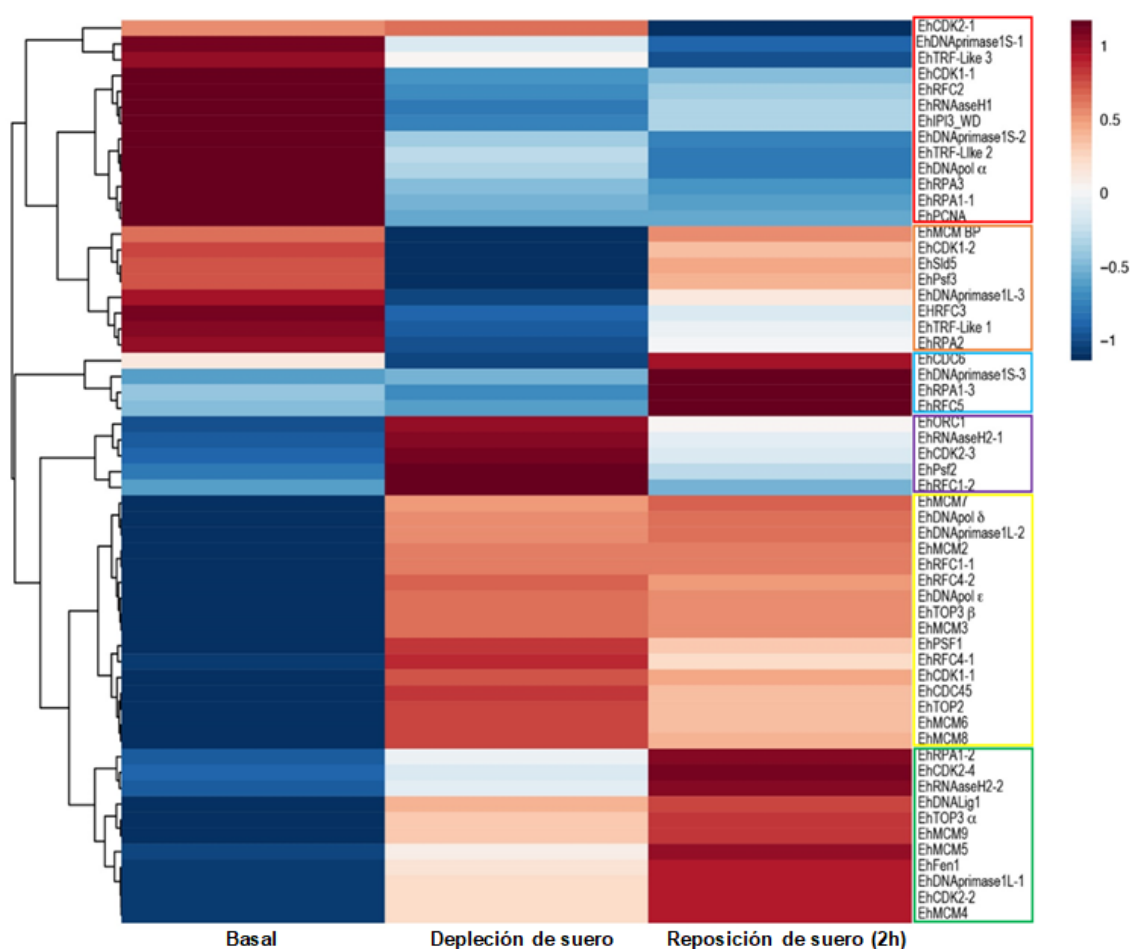


**Figura 13. Mapa de calor de la expresión de genes relacionados con la replicación del DNA en *E. histolytica*.** mapa de calor de 57 genes implicados en la replicación del DNA de trofozoítos de *E. histolytica* en condiciones de cultivo basal, depleción de suero y reposición de suero. Los datos corresponden a valores de TPM obtenidos de AmoebaDB y fueron normalizados para su representación. Los niveles de expresión están por color: rojo (alta expresión), azul (baja expresión) y blanco (expresión intermedia).

En resumen, el análisis transcriptómico muestra que la replicación del DNA está regulada por la disponibilidad de nutrientes. La depleción de suero provoca un bloqueo del ciclo celular, mientras que su reposición induce la expresión de genes del replisoma, lo que indica una activación coordinada de la replicación y la progresión hacia fase S.

#### **8.4 Ensayo de sincronización celular de trofozoítos de *E. histolytica* mediante RT-PCR**

Ya que la depleción y reposición de suero modifica la expresión de los genes involucrados en la replicación del DNA (posiblemente relacionado a un proceso de sincronización celular), realizamos un análisis de expresión de algunos genes de las diferentes etapas de la replicación del DNA. Para ello se obtuvieron cultivos de trofozoítos a los cuales se les depletó el suero durante 12 h, seguido de una reposición de suero, y se cosecharon a diferentes intervalos de tiempo: 0, 0.5, 1.5, 3, 6, 9, 11 y 24 h, además de una condición basal (sin depleción de suero). Para ello, a partir de los primers diseñados para amplificar genes representativos de cada etapa (tabla 4), se hicieron ensayos de RT-PCR semicuantitativa utilizando el gen *ehcoact* como control de carga, ya que codifica para un coactivador transcripcional de expresión constitutiva que no presenta cambios en su expresión durante la sincronización celular. Adicionalmente, como control positivo se utilizó DNAg y como control negativo se utilizó RNA tratado con DNAasa al cual no se le agregó retrotranscriptasa.

##### **8.4.1 Análisis de la expresión de genes que conforman la etapa Pre-RC**

En la etapa Pre-RC, se analizó la expresión de los genes representativos *ehorc1*, *ehcdc6*, *ehmcm3* y *ehmcm4* (Figura 14). Este ensayo mostró que todos los genes se expresan en condiciones basales de cultivo. En este ensayo observamos a las 0 horas una disminución de la expresión de los genes *ehorc1*, *ehmcm3* y *ehmcm4* con respecto a la condición basal, lo cual concuerda con el estado de arresto o bloqueo celular. Observamos también un aumento paulatino en la expresión de estos mismos genes durante los tiempos de reposición de suero (0.5 a 3 h). Posteriormente, a las 6 h, se evidenció una disminución en la expresión, lo que podría reflejar la finalización de la fase de síntesis o la terminación de la replicación del DNA. A las 9 h, la expresión vuelve a incrementar de manera gradual hasta las 24 h.

El único que no presentó este efecto fue el gen *ehcdc6*, ya que se expresó abundantemente a lo largo de todos los tiempos analizados, lo cual pudiera estar relacionado con una expresión constitutiva.

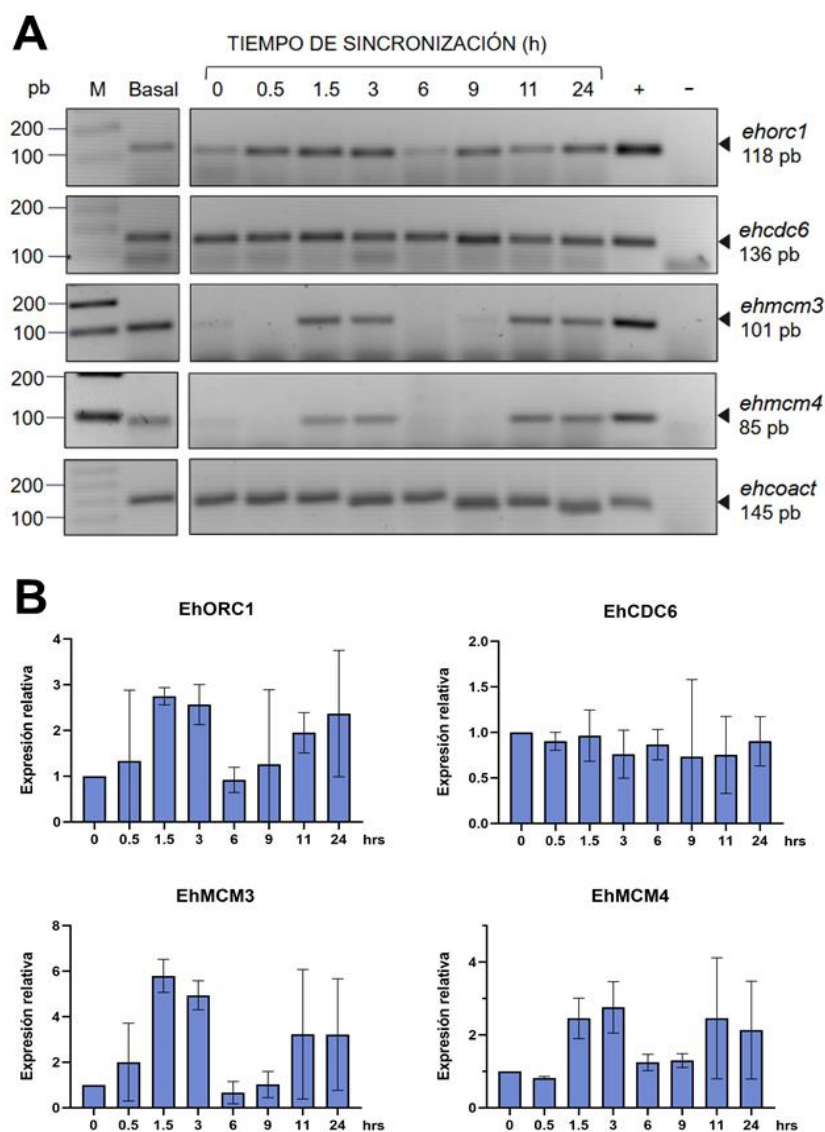


**Figura 14. Análisis de la expresión de genes relacionados con componentes del complejo Pre-RC en *E. histolytica* mediante RT-PCR semicuantitativa.** A) Se analizaron los genes *ehorc1*, *ehcdc6*, *ehmcm3* y *ehmcm4* en trofozoítos de *E. histolytica* sincronizados por depleción de suero durante 24 h y posteriormente reposición de suero. Los tiempos post-reposición evaluados fueron: 0, 0.5, 1.5, 3, 6, 9, 11 y 24 h. La condición basal corresponde a trofozoítos no sincronizados, y los controles positivos (+) y negativos (–) corresponden a DNA genómico y RNA tratado sin transcriptasa inversa, respectivamente. El marcador (M) indica pesos moleculares. El gen *ehcoact* se utilizó como control de carga. B) Gráficos de la expresión relativa de genes representativos, obtenidos a partir de duplicados independientes del ensayo de RT-PCR semicuantitativa. Expresión relativa a la expresión del gen control (coactivador).

### 8.4.2 Análisis de la expresión de genes que conforman la etapa Pre-IC

En la etapa Pre-IC se analizaron los genes *ehpsf1*, *ehcdc45*, *ehtop II* y *ehrpall* (Figura 15). En la condición basal, todos los genes se expresan abundantemente, con excepción de *ehtopII*, cuya expresión fue baja. El comportamiento de *ehpsf1* a lo largo del ensayo fue semejante al observado para los genes de la etapa Pre-RC (*ehorc-1*, *ehmcm3*, *ehmcm4*). El gen *ehcdc45* presenta un patrón de expresión semejante a *ehcdc6*, abundante y aparentemente de expresión constitutiva.

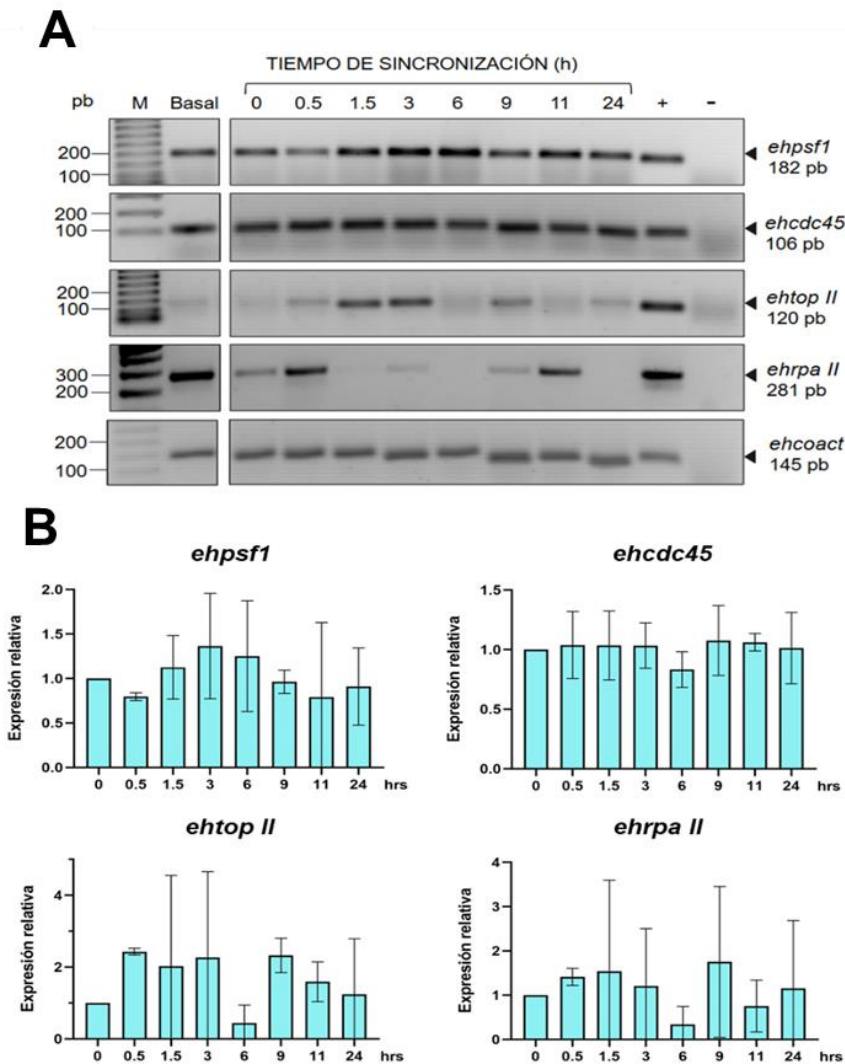


**Figura 15. Análisis de la expresión de genes relacionados con componentes del complejo Pre-IC en *E. histolytica* mediante RT-PCR semicuantitativa.** A) Se evaluó la expresión de los genes *ehpsf1*, *ehcdc45*, *ehtopII*, y *ehrpall* durante un experimento de sincronización celular por depleción y reposición de suero en trofozoítos de *E. histolytica*. Las muestras se recolectaron en los tiempos: 0, 0.5, 1.5, 3, 6, 9, 11 y 24 h post-reposición. Se incluyen la condición basal (B), basal con mayor densidad (B+), y controles positivos (+), negativos (-) y marcador de peso molecular (M). Se utilizó el gen *ehcoact* como control de carga. B) Gráficos de la expresión relativa de genes representativos, obtenidos a partir de duplicados independientes del ensayo de RT-PCR semicuantitativa.

Por su parte, *ehtopII* mantuvo baja expresión a las 0 h, pero mostró un aumento progresivo hasta las 3 h, seguido de una disminución a las 6 h. En el caso de *ehrpII*, alcanza su mayor expresión a las 0.5 h, y disminuye su expresión. De manera interesante, a las 11 h la expresión de *ehrpII* se elevó nuevamente a niveles similares a los de las 0.5 h.

#### 8.4.3 Análisis de la expresión de genes que conforman la etapa S

Para esta etapa se analizaron los genes *ehtop III*, *ehpcna*, *ehdnapol δ* y *ehdnapol ε* (Figura 16). En la etapa Pre-IC, se analizó el gen *ehtopII*, ya que su participación es en la apertura del DNA y la formación del complejo CMG, donde alivia la tensión generada por la helicasa MCM al desenrollar la doble hélice (Wang, 2002). En cambio, el gen analizado *ehtopIII* fue incluido en la etapa de S debido a su participación en la resolución de estructuras intermedias durante la elongación del DNA; así mismo, participa en la etapa T, ya que también colabora con FEN1 para la maduración de los fragmentos de Okazaki y en la reparación (Champoux, 2001; Plank y Hsieh, 2009). En la condición basal, la mayoría de los genes se expresa de forma abundante con excepción de *ehtop III* que se expresa menos. También observamos que la expresión de *ehtop III* pareciera ser constante y abundante a lo largo del ensayo, a excepción de las 9 horas, donde se observa una disminución de la expresión.

En cuanto a *ehpcna*, este gen se expresa tempranamente desde las 0 y hasta las 3 h. Posteriormente, disminuye drásticamente hasta las 6 h y se observa una recuperación de la expresión a las 9 y 11 h, seguido de una reducción a las 24 h. En cuanto a las polimerasas, para *ehdnapolδ*, la expresión se mantuvo moderada a lo largo del ensayo, mientras que para *ehdnapolε*, se observó una disminución a partir de las 6 horas.

En conjunto, se observó que los genes analizados presentan una regulación dinámica a lo largo del tiempo. La mayoría mostró una disminución de la expresión a partir de las 6 y 11 h, similar para las etapas anteriores, lo que podría estar asociado a la finalización de la fase S o al inicio del otro nuevo ciclo de replicación.

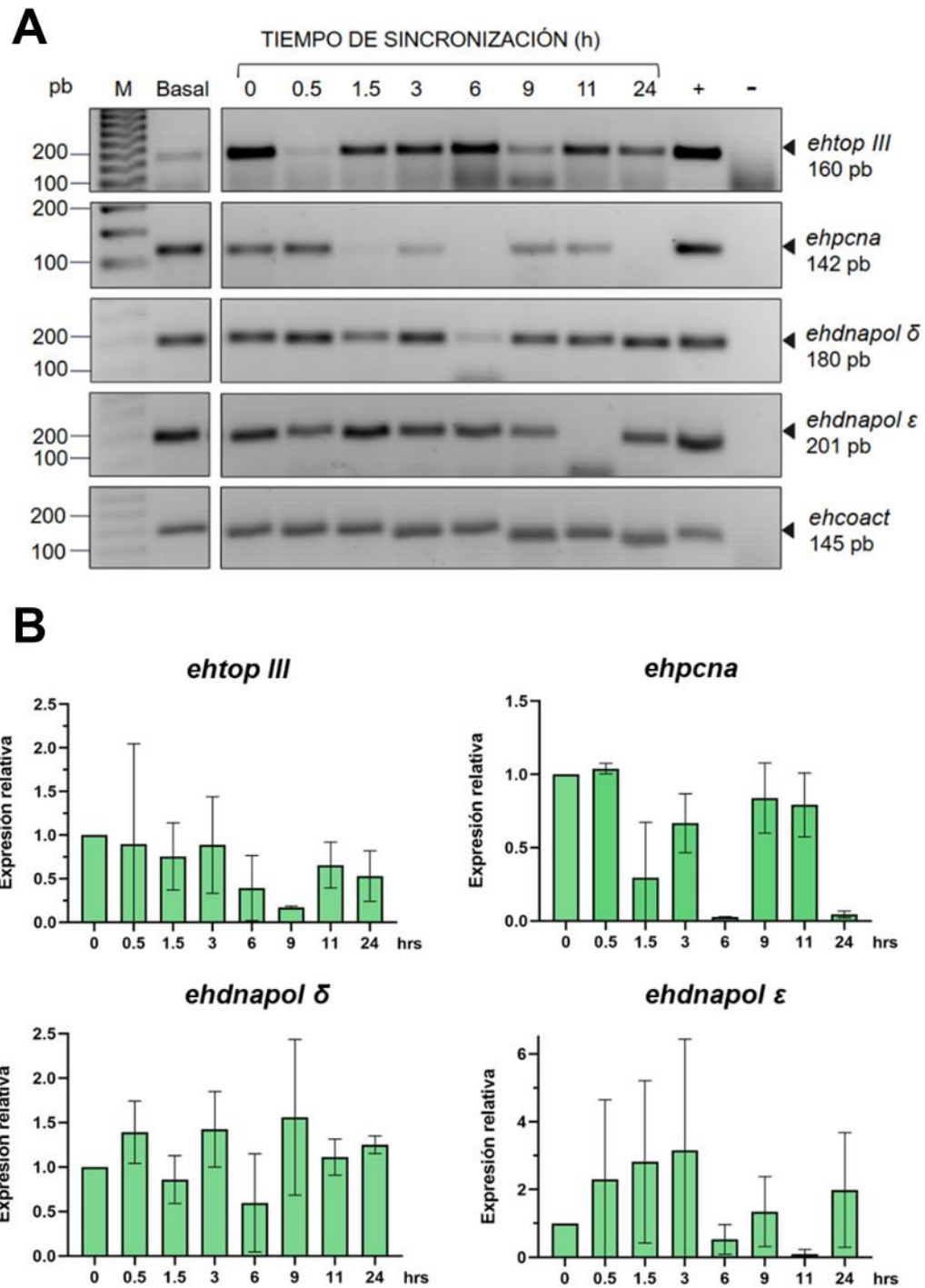


**Figura 16. Análisis de la expresión de genes relacionados con componentes de la etapa de síntesis en *E. histolytica* mediante RT-PCR semicuantitativa.** A) Se analizó la expresión de los genes *ehtop III*, *ehpcna*, *ehdnapolδ* y *ehdnapolε* en trofozoítos sincronizados. Las condiciones experimentales incluyeron depleción de suero por 24 h seguida de reposición, con recolección de muestras en los tiempos indicados (0 a 24 h), así como la condición basal (B), control positivo (+), negativo (-) y marcador de peso molecular (M). El gen *ehcoact* fue utilizado como control de carga para normalizar la expresión. B) Gráficos de la expresión relativa de genes representativos, obtenidos a partir de duplicados independientes del ensayo de RT-PCR semicuantitativa.

#### **8.4.4 Análisis de la expresión de genes que conforman la etapa T**

Se analizaron los genes *ehdnalig1*, *ehfen1* y *ehtrf-like I*, *ehtrf-like II* y *ehtrf-like III*, los cuales están implicados en el sellado de fragmentos de Okazaki, procesamiento y mantenimiento de los telómeros, respectivamente (Figura 17). En la condición basal, todos los genes mostraron niveles bajos de expresión. Para los genes *ehdnalig 1*, *ehfen1* y *ehtrf-like I* observamos que la expresión inicia a partir de las 0.5 h y continúa hasta las 3 h. A las 6 h, estos genes presentaron una disminución en su expresión, seguida por una recuperación a las 9 h. En el caso de *ehtrf-like II* y *ehtrf-like III*, se observó un patrón similar entre ellas, donde la expresión es constante en todas las condiciones.

En resumen, el análisis de expresión génica evidenció una activación temporal y coordinada de los genes involucrados en la replicación del DNA tras la privación y la reposición de suero. Estos resultados sugieren que la primera activación replicativa, inducida por la reposición, podría estar ocurriendo de 0.5 hasta las 6 h, empezando un nuevo ciclo de replicación a las 9 h, lo cual es consistente con estudios previos que señalan que *E. histolytica* puede replicar su DNA cada 8 a 10 h (Mukherjee *et al.*, 2009).

#### **8.5 Expresión y localización de EhPCNA durante la sincronización celular de trofozoítos de *E. histolytica***

Para observar lo que ocurre con EhPCNA durante la progresión del ciclo celular, se realizó un ensayo de inmunofluorescencia indirecta en trofozoítos de *E. histolytica*, cultivados sin suero por 12 y posteriormente con suero por diferentes tiempos (0, 0.5, 1.5, 3, 6, 9, 11 y 24 h). Para ello, se utilizó un anticuerpo primario anti-EhPCNA, seguido de un anticuerpo secundario conjugado con FITC (verde), mientras que el núcleo fue marcado con DAPI (azul) (Figura 18; Anexo 3).

En la condición basal, se observó que EhPCNA está localizada principalmente en el núcleo de los trofozoítos; esto es consistente con la función coordinar múltiples procesos asociados con la síntesis del DNA como incrementar la procesividad de las DNA polimerasas (Moldovan *et al.*, 2007).



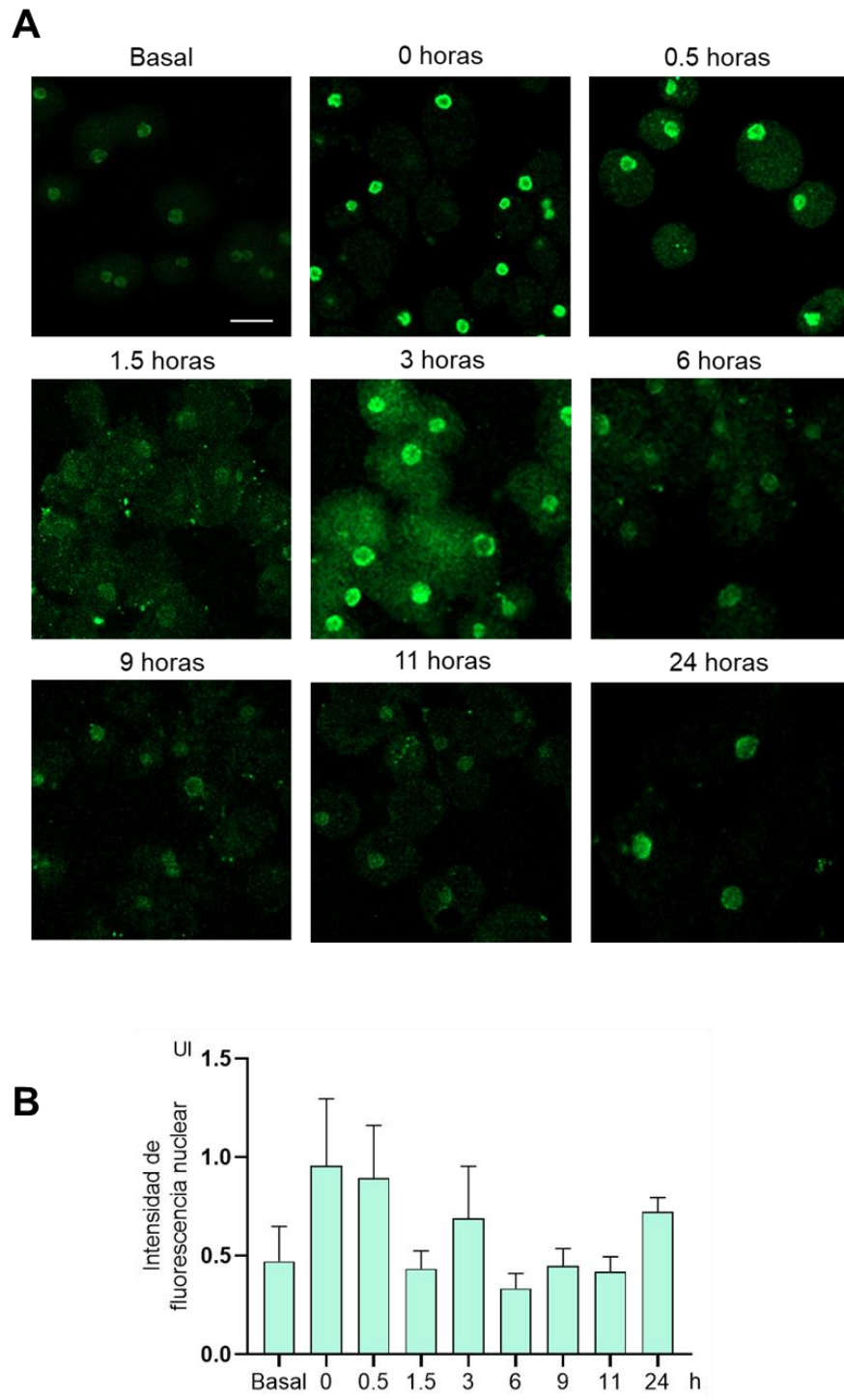


**Figura 18. Detección de EhPCNA mediante inmunofluorescencia en trofozoítos de *E. histolytica* sincronizados.** A) Trofozoítos de *E. histolytica* incubados sin suero durante 12 horas, seguida de reposición de suero, se fijaron y procesaron a diferentes tiempos. Las muestras se incubaron con anti-EhPCNA (1:250) y se visualizaron por medio de microscopía confocal. B) Cuantificación de la intensidad de fluorescencia nuclear empleando el ImageJ de los diferentes tiempos analizados.

## 9. Discusión

En *E. histolytica*, se han identificado previamente algunas proteínas relacionadas con el proceso de replicación del DNA. Algunas de ellas se han analizado de manera *in silico* y de otras se ha caracterizado su actividad funcional (Das y Lohia, 2000; Das *et al.* 2004; Cardona-Félix *et al.*, 2010; 2011; Pastor-Palacios *et al.*, 2010; Azuara *et al.*, 2018; Kurniawan *et al.*, 2018; Rendón-Gandarilla *et al.*, 2018; Pal y Dam, 2023). Sin embargo, hasta la fecha no se han identificado todos los componentes que conforman el replisoma de este parásito. Por ello, en este estudio, se llevó a cabo la identificación *in silico* de las proteínas correspondientes a las cuatro etapas que conforman la replicación del DNA: Pre-RC, Pre-IC, síntesis (S) y terminación (T) y se analizaron la presencia de los dominios conservados y sus patrones de expresión.

Identificamos en *E. histolytica*, 12 genes que codifican para proteínas de la etapa Pre-RC, de los cuales algunos presentan señal de localización nuclear y conservan dominios funcionales e identidad con proteínas homólogas de *H. sapiens* y *S. pombe* (Anexo 1). El complejo ORC, compuesto por 6 subunidades, se une a los orígenes de replicación de forma dependiente del ATP y funciona de andamiaje para el ensamblaje de otros factores de la iniciación. Sin embargo, en *E. histolytica* no se han identificado los genes que codifican para las subunidades ORC2–6, solo se identificó un gen que codifica para una subunidad denominada EhOrc1, lo que sugiere una simplificación estructural del complejo ORC en este organismo (Grewal y Lohia, 2015). Sería interesante analizar si EhOrc1 es capaz de formar hexámeros durante el reconocimiento del sitio Ori. Recientemente, la proteína EhOrc1 se analizó *in silico* y se propuso que presenta una función dual ORC1/CDC6 (Pal y Dam, 2023). No obstante, en este trabajo identificamos dos genes independientes que codifican para EhORC1 y EhCDC6 (Anexo 2) que se expresan diferente durante el ayuno y la reposición de suero. Esto confirma que en el genoma de *E. histolytica* están presentes ambos genes, a diferencia de lo que se había reportado por Pal y Dam (2023) sobre una sola proteína con función dual EhORC1/CDC6. Además, EhORC1 y EhCDC6 presentan una identidad del 23% y 25% con respecto a *H. sapiens*, EhORC1 contiene al dominio AAA\_22, el cual es una variante de la superfamilia AAA+ ATPasa, mientras que EhCDC6 presenta el dominio AAA. Ambos dominios comparten una arquitectura central de aprox. 250 aa.

Dentro de estos dominios se encuentran los motivos tipo Walker A (P-loop) Walker B, necesario para la hidrólisis de ATP (Neuwald *et al.*, 1999; Iyer *et al.*, 2004), lo que indica que conservan su capacidad de hidrolizar y unir ATP, respectivamente, función esencial en el inicio de la replicación del DNA. En estas proteínas no se identificaron señales de localización nuclear (NLS), a diferencia de otros eucariontes como *S. cerevisiae*, donde ORC y CDC6 tienen múltiples NLS (Anexo 1), así como dominios de interacción con MCM (Bell y Dutta, 2002; Kragelund *et al.*, 2006).

En relación con el complejo MCM, identificamos a las proteínas EhMCM 2–9, de las cuales las subunidades MCM 2–7 conforman el complejo helicasa, encargado del desenrollamiento del DNA. Algunos de los genes *mcm* han sido previamente identificados en el genoma de este organismo (Gangopadhyay *et al.*, 1997; Grewal y Lohia *et al.*, 2015). Adicionalmente, analizamos la secuencia de proteínas y observamos que todas conservan los dominios MCM\_N, MCM\_OB, MCM/AAA+ ATPasa y el dominio MCM\_lid, los cuales son importantes para su función de helicasa, lo que sugiere que *E. histolytica* conserva las características estructurales necesarias para el funcionamiento del complejo MCM.

Las helicasas EhMCM8 y EhMCM9 no forman parte del complejo MCM2–7, pero están implicadas en el reclutamiento de CDC6, factor clave para la carga de las proteínas MCM en los orígenes de replicación (Helderman *et al.*, 2023; Weng *et al.*, 2023). Finalmente, identificamos dos proteínas adicionales: EhMCM-BP y EhIPI3-WD. En mamíferos, MCM-BP participa en el ensamblaje, estabilización y disociación del complejo MCM, interactuando con MCM3 y MCM5 a través del dominio MCM\_bind (Samel *et al.*, 2014; Saito *et al.*, 2022; Hatoyama y Kanemaki, 2023). La proteína EhMCM-BP presenta este dominio, lo cual sugiere una conservación funcional. La proteína EhIPI3-WD presenta un dominio WD con 4 regiones o subdominios repetidos, similares a los dominios WD40 observados en *H. sapiens* (que presenta cinco repetidos). Estos dominios están implicados en interacciones proteína–proteína y, en levaduras, IPI3 participa en el reclutamiento del complejo MCM al sitio Ori (Takahashi *et al.*, 2010; Campbell y Struhl, 2012). En mamíferos, estos dominios sirven como plataformas de ensamblaje multiproteico (Huang *et al.*, 2016), lo que sugiere que en *E. histolytica*, EhIPI3-WD podría participar tanto en la estabilidad del complejo MCM como en la interacción con otras proteínas de la replicación como EhORC1 y EhCDC6.

En cuanto a la expresión de estos genes, *ehmcm2-7* aumentan su expresión de forma similar en las condiciones de depleción y reposición de suero, siendo los genes *ehmcm4* y *5* los de mayor incremento (Naiyer *et al.*, 2019). Notoriamente, observamos un comportamiento similar para *ehmcm3* y *4*, en nuestras condiciones experimentales.

Nuestros resultados evidencian que *E. histolytica* conserva un repertorio básico de proteínas para formar el complejo Pre-RC, aunque con una organización simplificada. Esto puede deberse a adaptaciones evolutivas propias de organismos unicelulares parásitos, como se ha observado en *T. cruzi*, donde algunas subunidades del complejo ORC no están presentes (da Silva *et al.*, 2017). La identificación de proteínas como MCM-BP e IPI3-WD y la conservación de dominios funcionales refuerzan la idea de que, aunque reducido en número, el complejo Pre-RC de *E. histolytica* podría ser funcional y participar en el inicio de la replicación del DNA de manera eficiente bajo su contexto biológico.

Con respecto a la etapa Pre-IC, identificamos 20 genes que codifican para proteínas relacionadas con esta etapa, de los cuales algunos se encuentran duplicados. Esto podría indicar que pertenecen a familias de genes o bien que codifican para isoformas de la misma proteína que surgieron durante la adaptación a la vida parásita. De estos genes, la mitad presentan intrones, y solo dos codifican para proteínas que con NLS (Anexo 1). Sin embargo, la mayoría conservan tanto sus dominios funcionales y presentan altos porcentajes de identidad con proteínas homólogas de *H. sapiens* y *S. pombe*.

La proteína EhCDC45 previamente se identificó en *E. histolytica* y mostró una identidad superior al 20% con respecto a sus ortólogos en humanos y levaduras (Kurniawan *et al.*, 2018). Esta proteína conserva dominios clave como DHH en el extremo N-terminal, DHHA1 C-terminal, el dominio CID y un segmento conector  $\alpha$ -helicoidal; dichos dominios son fundamentales para su interacción con las proteínas MCM, el complejo de GINS y el complejo CMG (Cdc45- Mcm2-7- GINS). De manera interesante, este parásito tiene todas las proteínas para formar los complejos GINS y CMG. Las proteínas EhPSF1-3 y EhSLD5 presentan el dominio GINS\_A, responsable de la interacción con las helicasas MCM y alguna también presentan el dominio GINS\_B que facilita la interacción con CDC45.

Las proteínas mantienen la arquitectura  $\alpha$ -helicoidal característica, que estabiliza la horquilla de replicación. Aunque el complejo GINS fue identificado previamente por Grewal y Lohia (2015), hasta ahora no se habían analizado en detalle sus dominios conservados ni su estructura tridimensional. Por su parte, el complejo CMG tiene una función en la activación de la helicasa replicativa, y facilitando la apertura de la doble hélice de DNA (Moyer *et al.*, 2006).

Por otro lado, en *E. histolytica* encontramos siete genes que codifican para cinasas tipo CDK1 y CDK2, con identidades de hasta 52.4%, mientras que en *H. sapiens* se han identificado 20 genes que codifican para CDKs, agrupados en familias de acuerdo con su función (Fisher, 2011). Todas las proteínas identificadas conservan el dominio catalítico Pkinase, característico de serina/treonina cinasas (Hanks *et al.*, 1988). Las CDKs promueven la conversión de los complejos de pre-replicación (pre-RC) en complejos de iniciación activos mediante la fosforilación de los factores Sld2 y Sld3. La cinasa dependiente de Dbf4 (DDK) regula la activación de los orígenes de replicación mediante fosforilación. De manera interesante, en *E. histolytica* no se identificaron genes que codifican para DDK, lo que podría indicar que la fosforilación de ORC, CDC6 y MCM está mediada únicamente por CDKs (Anexo 2). Esto sugiere diferencias importantes en cuanto a la regulación de esta etapa en este parásito.

En cuanto a las topoisomerasas (TOP), las enzimas encargadas de aliviar la tensión provocada por el desenrollamiento del DNA (Pommier *et al.*, 2016), en este trabajo identificamos a las topoisomerasas de tipo II y III (isoformas A y B) que mantienen una alta identidad con los ortólogos de *H. sapiens* y levadura. La topoisomerasa II en *E. histolytica* conserva los cinco dominios clave asociados con la hidrólisis de ATP, el corte reversible de DNA y el mantenimiento de la estructura de la doble hélice durante la replicación. Las topoisomerasas III A y B presentaron los dominios Toprim, Topoisom\_bac y Zn\_ribbon\_TOP3B que respaldan su función en el alivio de superenrollamientos y la resolución de estructuras. En *E. histolytica*, la topoisomerasa tipo II A se expresa abundantemente durante el estrés y el enquistamiento. De manera interesante, cuando se silenció este gen se redujo la viabilidad y la eficiencia del enquistamiento (Varghese y Ghosh, 2020).

Finalmente, la proteína de replicación A (RPA) es un complejo heterotrimérico compuesto por tres subunidades (RPA1, RPA2 y RPA3) de 70, 32 y 14 kDa, respectivamente esencial en eucariotas para los procesos de replicación, recombinación y reparación del DNA que además interactúa con PCNA. En *E. histolytica*, identificamos los genes que codifican para las tres subunidades del complejo RPA (EhRPA1, EhRPA2 y EhRPA3), con identidades entre 18.18 % y 29.47 % respecto a sus ortólogos humanos. De manera interesante, se identificaron tres genes que codifican isoformas de RPA1 con un peso molecular aproximado de 70 kDa similar al de RPA1 de *H. sapiens*. Estas isoformas conservan los tres dominios OB-fold fundamentales para la unión al DNA monocatenario (Broderick et al., 2010; Krasikova et al., 2016). En el caso de EhRPA2, se identificó un dominio OB-fold y secuencias de fosforilación también presentes en el ortólogo humano, aunque presenta un peso molecular menor (16 vs 32 kDa). Por su parte, EhRPA3 muestra un peso molecular similar al de la proteína humana y aunque no se identificaron dominios en su secuencia la estructura tridimensional predicha mediante AlphaFold revela una organización muy similar a la de un dominio OB-fold (Tabla 8). De manera interesante, los tres genes disminuyen su expresión tras la remoción y reposición de suero por lo que podrían responder ante estímulos similares. La conservación estructural sugiere que dichas proteínas podrían conservar funciones clave en la estabilización de hebras monocatenarias durante la replicación del DNA.

Para la etapa de síntesis del DNA, en *E. histolytica* identificamos 19 genes que codifican para proteínas que conservan sus dominios funcionales, sin embargo, muy pocas presentaron señales de localización nuclear (Tabla 9, Anexo 1). La etapa S en eucariotas depende de la actividad de una RNA primasa y la DNA-pol  $\alpha$  formando el complejo pol-prim. En *E. histolytica* se identificaron tres isoformas para la subunidad pequeña y grande de la primasa (PriS y PRIM2, respectivamente). Al comparar con sus ortólogos de humano y levadura, llama la atención que EhDNAprimase1S-2 mostró un valor reducido (~9 %), lo que sugiere una divergencia evolutiva. Además, se predice un peso molecular menor (22.48 kDa), en comparación con las demás isoformas (44–45 kDa) y con su ortólogo humano (~49 kDa). En el caso de las subunidades grandes, EhDNAprimase1L-2 mostró un peso semejante a la del ortólogo humano (~58 kDa).

Sin embargo, todas las isoformas conservan los dominios funcionales característicos, lo que sugiere que podrían estar involucradas tanto en la replicación como en la respuesta a estrés replicativo o procesos de reparación. La región C-terminal de p180 (p180C), altamente conservada, desempeña un papel estructural clave al unir funcionalmente las subunidades p70 y la primasa con la polimerasa catalítica. Estas proteínas podrían además tener funciones alternativas, incluyendo posibles roles en el mantenimiento de telómeros, ya que se ha demostrado que el complejo pol-prim puede interactuar con TRF1 y TRF2, actuando como una segunda maquinaria para el mantenimiento telomérico (Zinder *et al.*, 2022). Los genes *ehdnprimase1s-1* y *2* y *ehdnprimase1l-3* se expresan abundantemente en la condición basal y su expresión disminuye tras la depleción y reposición de suero, mientras que para *ehdnprimase1s-3* y *ehdnprimase1l-1* y *2* la expresión ocurre de manera inversa (expresión baja en la condición basal y alta tras depleción y reposición de suero) indicando una posible activación durante la replicación del DNA.

Las topoisomerasas EhTOP3  $\alpha$  y  $\beta$  las incluimos en la etapa S del ciclo celular debido a que participan en conjunto con FEN1, en la resolución de intermediarios de replicación durante la elongación y en la maduración de fragmentos de Okazaki (Plank y Hsieh, 2009). EhTOP3  $\alpha$  tiene un peso molecular significativamente menor que su ortólogo humano (69.6 vs 112.3 kDa), mientras que EhTOP3  $\beta$  es similar en tamaño y presentan una NLS (Anexo 1), lo que podría indicar diferencias en su regulación y transporte nuclear. Ambas topoisomerasas conservan los dominios Toprim y Topoisom\_bac, esenciales para la actividad catalítica, pero EhTOP3  $\beta$  contiene adicionalmente un dominio Zn\_ribbon\_TOP3B, implicado en la catálisis del corte y religación del DNA (Champoux, 2001), lo que sugiere una especialización funcional. En cuanto a la expresión, *ehtop3  $\alpha$*  pareciera un gen constitutivo, mientras que *ehtop3  $\beta$*  aumenta tras la reposición de suero, sugiriendo una regulación diferencial durante la replicación.

Por otro lado, el complejo RFC en *H. sapiens* está compuesto por cinco subunidades (RFC1–RFC5) y actúa facilitando la interacción entre PCNA y el DNA para conferir alta procesividad a las DNA polimerasas  $\delta$  y  $\epsilon$  (Yao y O'Donnell, 2012).

En *E. histolytica* identificamos los cinco genes que codifican para las subunidades del complejo RFC, incluyendo dos genes que codifican para isoformas de EhRFC1. EhRFC 3-5 presentan pesos moleculares similares a los observados en *H. sapiens*, lo que sugiere una posible conservación estructural y funcional del complejo en este parásito. Sin embargo, las isoformas EhRFC1-1 y EhRFC1-2 presentan un peso molecular de 80.9 y 38.4 kDa respectivamente, que es considerablemente menor a la proteína de *H. sapiens* (128 kDa). También, EhRFC2 tiene un tamaño reducido en comparación con su ortólogo humano (18.3 vs ~39 kDa). A nivel estructural, EhRFC1-1 conserva los dominios BRCT y RFC1, este último característico de la subunidad mayor. EhRFC1-2 y EhRFC2 conservan el dominio P-loop\_NTPase, relacionado con la unión de nucleótidos, mientras que EhRFC3 contiene dominios asociados a la interacción con DNA y otras subunidades del complejo. Además, estas proteínas conservan dominios esenciales como AAA+ ATPase, necesarios para la hidrólisis de ATP, y Rep\_factorC\_C, implicado en el ensamblaje e interacción con PCNA. Estos hallazgos sugieren que *E. histolytica* mantiene la funcionalidad básica del complejo RFC, aunque con adaptaciones estructurales propias y dominios especializados que podrían reflejar ajustes con respecto a sus necesidades.

En cuanto a la proteína homotrimérica PCNA, esta actúa en *H. sapiens* como plataforma, reclutando y coordinando la actividad de múltiples proteínas como las DNA polimerasas, para aumentar su procesividad durante la replicación (Zhang *et al.*, 1998). En *E. histolytica*, se ha identificado y caracterizado a EhPCNA y su estructura cristalográfica muestra una organización muy similar a la de *H. sapiens* (Cardona-Félix *et al.*, 2011). El gen *ehpcna* muestra una expresión basal alta que disminuye tras la depleción y reposición de suero (Naiyer *et al.*, 2019). En congruencia, nosotros observamos el mismo comportamiento en nuestras condiciones experimentales. Adicionalmente, observamos cambios en el patrón de expresión a lo largo del tiempo que sugieren que la regulación de *ehpcna* podría ser sensible a la sincronización celular inducida, lo que podría explicar la presencia de señal tanto nuclear como citoplasmática observada en los ensayos de IFI (Anexo 3). Además, en los ensayos de IFI observamos que EhPCNA se localiza predominantemente en el núcleo, coherente con su función en la replicación del DNA.

Estos datos sugieren que EhPCNA podría ser un marcador útil para monitorear la progresión de la replicación en *E. histolytica*. Por otro lado previamente, en *E. histolytica* se identificaron los genes codificantes para las 4 subunidades catalíticas de las DNA polimerasas  $\alpha$ ,  $\delta$  y  $\epsilon$  todas ellas miembros de la familia B. Esta familia de polimerasas está conformada por diferentes subunidades las cuales pueden tener actividad de polimerasa y de exonucleasa entre otras (Pastor-Palacios *et al.*, 2015; Nozaki y Bhattacharya, 2015). Sin embargo, debido a las actualizaciones y reanotaciones en las bases de datos, los números de acceso previamente reportados han cambiado. Por ello, realizamos de nuevo la búsqueda de estos genes en el genoma y sólo identificamos 1 subunidad para cada polimerasa. En el caso de la EhDNApol  $\delta$ , la subunidad catalítica presenta una identidad de 45.98%, mientras que la subunidad de EhDNApol  $\epsilon$  mostró identidades bajas (13-18%) con respecto a *S. pombe* y *H. sapiens*, lo que sugiere una mayor divergencia evolutiva. En cuanto a los pesos moleculares las subunidades catalíticas de EhDNApol  $\alpha$  y EhDNApol  $\epsilon$ , son más pequeñas con respecto a su ortólogo en *H. sapiens*, mientras que el peso molecular de EhDNApol  $\delta$  es muy similar.

Desde el punto de vista estructural, EhDNApol  $\alpha$  y EhDNApol  $\delta$  presentan dominios típicos de DNA polimerasas de la familia B, como DNA\_pol\_B\_exo1 y DNA\_pol\_B, mientras que EhDNApol  $\alpha$  contiene además un dominio zf-C4pol, relacionado con la unión al DNA y la estabilidad estructural. Por otro lado, EhDNApol  $\epsilon$  posee los dominios distintivos, DNA\_pol\_D\_N y DNA\_pol\_E\_B esenciales para la función catalítica y la interacción con el DNA durante la replicación (García-Díaz *et al.*, 2007). La presencia conjunta en EhDNApol  $\epsilon$  de los dominios DNA\_pol\_D\_N (característico de la pol  $\delta$ ) y DNA\_pol\_E\_B (propio de la pol  $\epsilon$ ) sugiere que esta proteína podría desempeñar una función dual o combinada, integrando actividades que en eucariotes superiores están separadas en dos enzimas diferentes. Esto podría representar una adaptación evolutiva para optimizar recursos en un organismo unicelular. Por otro lado, los análisis de expresión muestran que EhDNA pol  $\alpha$  está más activa durante condiciones de crecimiento normal y se regula negativamente bajo estrés nutricional o condiciones adversas, mientras que EhDNApol  $\delta$  y EhDNApol  $\epsilon$  tienen un patrón similar: baja expresión en condición basal, con un ligero aumento durante la depleción y mantenimiento tras la reposición de suero (Naiyer *et al.*, 2019).

Este comportamiento sugiere diferencias en la regulación temporal y funcional de las polimerasas durante el ciclo celular.

Finalmente, en la etapa terminal de la replicación del DNA, identificamos 8 genes que codifican proteínas que conservan sus dominios y cuya mayoría presentaron NLS predicha (Anexo 1, Tabla 12). Las RNAasas H en *H. sapiens*, son endonucleasas que degradan los cebadores de RNA sintetizados por las primasas (Hyjek y Nowotny, 2019). En *E. histolytica* encontramos 4 genes que codifican para las RNAasas H1 y H2 que tienen un porcentaje de identidad moderado (31-36.9%) con respecto a *H. sapiens*, reflejando una conservación funcional parcial. En cuanto a los pesos moleculares, estas proteínas son ligeramente más pequeñas en comparación con *H. sapiens* (21 vs 30 kDa), mientras que EhRNAasaH2-2 es 10 kDa más grande con respecto a su homólogo en *H. sapiens*, lo que podría indicar variaciones en la estructura ya que aunque conservan los dominios RNASE\_H1 y RNASE\_H2, no encontramos el dominio HBD (dominio de unión híbrido), indispensable para la eliminación de los cebadores de RNA, presente en *H. sapiens* pero menos conservado entre especies (Hyjek y Nowotny, 2019), lo cual es consistente con su baja conservación evolutiva. En cuanto a la expresión, el gen *ehrnaaseh1* mostró alta expresión en condiciones basales, disminuyendo tras la depleción y reposición de suero, lo que sugiere que su actividad está asociada principalmente a condiciones de crecimiento activo y replicación normal, mientras que las isoformas *ehrnaaseh2* aumentan su expresión durante la depleción y reposición de suero, indicando que podrían desempeñar un papel más relevante en condiciones de estrés o en fases específicas de la replicación del DNA.

Adicionalmente, en *E. histolytica*, previamente se identificó y caracterizó una única proteína EhDNALig1 (Cardona-Félix *et al.*, 2010; Azuara-Liceaga *et al.*, 2018), lo que sugiere una enzima centralizada para estas funciones, mientras que en humanos, se conocen varias isoformas (ligasa I, III y IV) responsables de sellar los fragmentos de Okazaki (Pascal *et al.*, 2004), reparar el daño a DNA por escisión de bases (BER) y reparar rupturas de doble hebra por unión de extremos no homólogos (NHEJ) (Ellenberger y Tomkinson, 2008), respectivamente.

El porcentaje de identidad de EhDNALig1 es de 36.19% con respecto a *H. sapiens*, sugiriendo una conservación evolutiva. Además, presenta NLS coherente con su función nuclear. Esta proteína es más pequeña que su homólogo humano (78 vs 101 kDa), lo que podría reflejar una estructura más compacta ya que conserva los tres dominios fundamentales característicos: el dominio de unión a DNA (DBD), el dominio de adenilación (NTasa), y el dominio de pliegue OB (Cardona-Félix et al., 2010). En cuanto a la expresión génica, el gen *ehdnalig1* aumenta tras la depleción y reposición de suero. Este patrón también se observó en nuestras condiciones experimentales lo que sugiere que la expresión de *ehdnalig1* está fuertemente correlacionada con los periodos de síntesis activa de la replicación del DNA.

Por otro lado, *E. histolytica* codifica en su genoma para la proteína FEN1, una nucleasa que participa en el procesamiento de los fragmentos de Okazaki y en la reparación del DNA. Su principal función es reconocer y escindir estructuras de tipo "flap" (Liu et al., 2004). EhFEN1 presenta un porcentaje de identidad cercano a 50% con respecto a *H. sapiens* y un peso molecular muy similar y contiene los dominios característicos XPG\_N, XPG\_I y un subdominio C-terminal con pliegue de exonucleasa 5'→3' que sugieren una alta conservación funcional. Además, se ha observado que los dominios XPG N y I participan en la reparación del daño al DNA causado por la luz ultravioleta (Habracken et al., 1994; González-Corrochano et al., 2020) por lo que sería relevante determinar la expresión de EhFEN1 durante estas condiciones de estrés. En cuanto a la expresión génica, observamos un patrón muy similar al de *ehdnalig1* lo que sugiere que ambas proteínas están coordinadamente reguladas en respuesta a la sincronización celular, y la replicación del DNA.

Finalmente, las TRF-Like en *H. sapiens* son esenciales para el mantenimiento, protección y regulación de los telómeros (Fairall et al., 2001). En *E. histolytica* previamente se identificaron y caracterizaron 3 genes que codifican para estas proteínas y se denominaron EhTRF-like I, II y III (Rendón-Gandarilla et al., 2018). Estas proteínas presentan un porcentaje de identidad del 27-40% con respecto a *H. sapiens* y *S. pombe*, reflejando una conservación funcional parcial. Respecto a los pesos moleculares, estas proteínas son muy similares en comparación con *H. sapiens*, lo que podría sugerir una estructura ligeramente simplificada.

En lo estructural, las 3 isoformas conservan los dominios característicos de las TRFs, TRFH y el dominio MYB, que permiten la multimerización y la unión al DNA telomérico, respectivamente (Palm y De Lange, 2008). Estas proteínas, reconocen y se unen a sondas teloméricas de *H. sapiens* y se localizan en el núcleo de los trofozoítos (Rendón-Gandarilla *et al.*, 2018). Además, la expresión de los genes *ehtrf-like II* y *ehtrf-like III* mostró un patrón de expresión semejante durante el ensayo de sincronización celular lo que sugiere que ambos genes se expresan de manera coordinada en respuesta a la replicación del DNA. Esta conservación estructural sugiere que la amiba podría presentar un mecanismo de protección y regulación de telómeros.

Finalmente, este trabajo muestra lo que sabemos acerca del proceso de replicación del DNA y de las proteínas que constituyen el replisoma de *E. histolytica* y también, describen aspectos importantes que nos podrían ayudar a determinar proteínas relevantes en este proceso que afecten al desarrollo del parásito dentro del hospedero. Este trabajo nos podría ayudar a identificar proteínas clave que podrían emplearse con fines terapéuticos para el desarrollo de vacunas o fármacos contra la amibiasis.

## 10. Conclusiones

- Se identificaron genes que codifican para proteínas involucradas en las cuatro etapas de la replicación del DNA en *E. histolytica*: Pre-RC, Pre-IC, síntesis (S) y terminación (T).
- Se observó la ausencia de genes y la duplicación de otros lo cual podría representar isoformas relacionadas a mecanismos alternativos de replicación propios de este parásito.
- Se propuso un replisoma de *E. histolytica* que conserva muchas de las proteínas de organismos eucariontes superiores, pero con una organización más simplificada.
- Los análisis de expresión génica revelaron una regulación dinámica durante la sincronización celular que sugiere 2 replications en el transcurso de 24 h tras la privación y reposición de suero, lo cual podría ayudar a explicar los estados de poliploidía de este parásito.
- Estos hallazgos sugieren que *E. histolytica* posee una maquinaria de replicación funcional, compacta y adaptada a su biología parasit, lo que podría representar una oportunidad para identificar blancos terapéuticos específicos para su control.

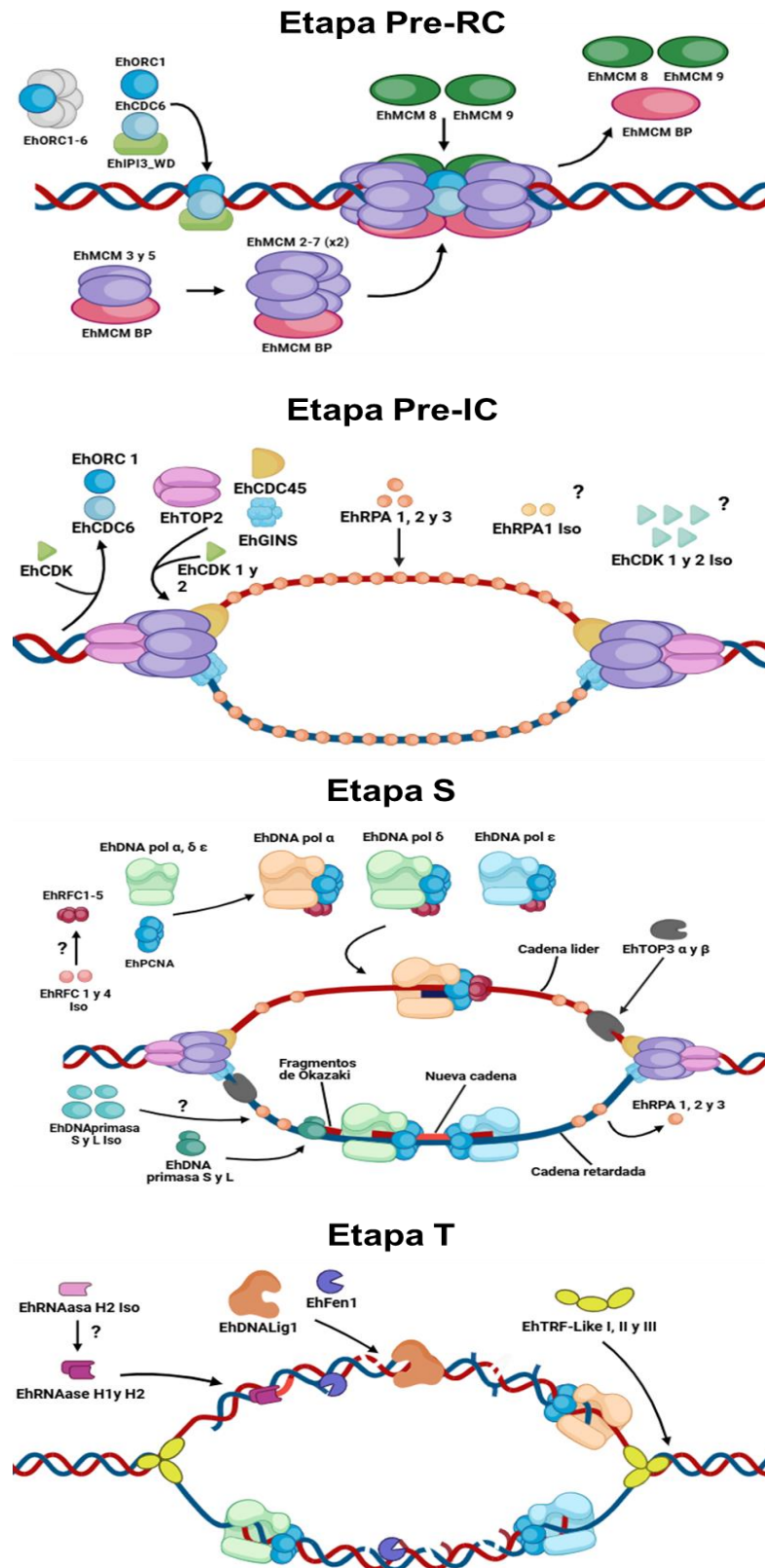


**Figura 18. Modelo de replicación del DNA en *E. histolytica*.** Esquema del proceso de replicación durante las cuatro etapas de la replicación del DNA en *E. histolytica* (Pre-RC, Pre-IC, S y T). Creado en <https://BioRender.com>

## 11. Bibliografía

1. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, et al., (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630(8016), 493–500. <https://doi.org/10.1038/s41586-024-07487-w>
2. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2014). *Molecular Biology of the Cell* (6th ed.). Garland Science.
3. Ali, I. K., Clark, C. G., & Petri, W. A. (2003). Molecular epidemiology of amebiasis. *Infection, Genetics and Evolution*, 3(1), 123-133.
4. Alvarez-Jarreta, J. et al. (2023) 'VEuPathDB: The eukaryotic pathogen, vector and host Bioinformatics Resource Center in 2023', *Nucleic Acids Research*, 52(D1). <https://doi.org/10.1093/nar/gkad1003>
5. Aravind, L., & Koonin, E. V. (1999). DNA primases and related RNA polymerases. *Nucleic Acids Research*, 27(15), 3389–3401. <https://doi.org/10.1093/nar/27.15.3389>[https://doi.org/10.1016/S0968-0004\(00\)01765-1](https://doi.org/10.1016/S0968-0004(00)01765-1)
6. Arezi, B., & Kuchta, R. D. (2000). Eukaryotic DNA primase. *Trends in Biochemical Sciences*, 25(11), 572–576. [https://doi.org/10.1016/S0968-0004\(00\)01762-7](https://doi.org/10.1016/S0968-0004(00)01762-7)
7. 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. Revista especializada en ciencias químico-biológicas*, 16(2), 98-108.
8. Austin, C. J., & Warren, L. G. (1983). Induced division synchrony in *Entamoeba histolytica*. Effects of hydroxyurea and serum deprivation. *The American journal of tropical medicine and hygiene*, 32(3), 507–511. <https://doi.org/10.4269/ajtmh.1983.32.507>
9. Azuara-Liceaga, E., Betanzos, A., Cardona-Felix, C. S., Castañeda-Ortiz, E. J., Cárdenas, H., Cárdenas-Guerra, R. E., Pastor-Palacios, G., García-Rivera, G., Hernández-Álvarez, D., Traviña-Arenas, C. H., Díaz-Quezada, C., Orozco, E., & Briebe, L. G. (2018). The Sole DNA Ligase in *Entamoeba histolytica* Is a High-Fidelity DNA Ligase Involved in DNA Damage Repair. *Frontiers in cellular and infection microbiology*, 8, 214. <https://doi.org/10.3389/fcimb.2018.00214>
10. Balakrishnan, L., & Bambara, R. A. (2013). Okazaki fragment metabolism. *Cold Spring Harbor Perspectives in Biology*, 5(2), a010173. <https://doi.org/10.1101/cshperspect.a010173>
11. Baranovskiy, A. G., Zhang, Y., Suwa, Y., Babayeva, N. D., Gu, J., Pavlov, Y. I., & Tahirov, T. H. (2015). Crystal structure of the human primase. *The Journal of biological chemistry*, 290(9), 5635–5646. <https://doi.org/10.1074/jbc.M114.624742>
12. Bateman, A. et al. (2024) 'Uniprot: The Universal Protein Knowledgebase in 2025', *Nucleic Acids Research*, 53(D1). <https://doi.org/10.1093/bioinformatics/btac793>
13. Bell, S. P., & Dutta, A. (2002). DNA replication in eukaryotic cells. *Annual Review of Biochemistry*, 71(1), 333-374.
14. Bell, S. P., & Labib, K. (2016). Chromosome duplication in *Saccharomyces cerevisiae*. *Genetics*, 203(3), 1027–1067. <https://doi.org/10.1534/genetics.115.186452>
15. Bell, S. P., & Stillman, B. (1992). ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature*, 357(6377), 128-134.
16. Blum, M. et al. (2024) 'InterPro: The Protein Sequence Classification Resource in 2025', *Nucleic Acids Research*, 53(D1). <https://doi.org/10.1101/2022.04.08.487609>
17. Blum, M., Andreeva, A., Florentino, L. C., Chuguransky, S. R., Grego, T., Hobbs, E., Pinto, B. L., Orr, A., Paysan-Lafosse, T., Ponamareva, I., Salazar, G. A., Bordin, N., Bork, P., Bridge, A., Colwell, L., Gough, J., Haft, D. H., Letunic, I., Llinares-López, F., Marchler-Bauer, A., ... Bateman, A. (2025). InterPro: the protein sequence classification resource in 2025. *Nucleic acids research*, 53(D1), D444–D456. <https://doi.org/10.1093/nar/gkae1082>
18. Bochman, M. L., & Schwacha, A. (2009). The Mcm complex: unwinding the mechanism of a replicative helicase. *Microbiology and Molecular Biology Reviews*, 73(4), 652–683. <https://doi.org/10.1128/MMBR.00019-09>
19. Bowman, G. D., O'Donnell, M., & Kuriyan, J. (2004). Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. *Nature*, 429(6993), 724–730. <https://doi.org/10.1038/nature02585>
20. Brewer, B. J., & Fangman, W. L. (1987). The localization of replication origins on ARS plasmids in *S. cerevisiae*. *Cell*, 51(3), 463–471. [https://doi.org/10.1016/0092-8674\(87\)90642-8](https://doi.org/10.1016/0092-8674(87)90642-8)

21. Broderick, S., Rehmet, K., Concannon, C., & Nasheuer, H. P. (2010). Eukaryotic single-stranded DNA binding proteins: central factors in genome stability. *Sub-cellular biochemistry*, 50, 143–163. [https://doi.org/10.1007/978-90-481-3471-7\\_8](https://doi.org/10.1007/978-90-481-3471-7_8)
22. Burgers, P. M. J., & Kunkel, T. A. (2017). Eukaryotic DNA Replication Fork. *Annual review of biochemistry*, 86, 417–438. <https://doi.org/10.1146/annurev-biochem-061516-044709>
23. Byers, J. M., Faubert, G., & Eichinger, D. J. (2005). Cell division in *Entamoeba histolytica*. *International Journal for Parasitology*, 35(8), 903–911. <https://doi.org/10.1016/j.ijpara.2005.03.011>
24. Cai, S. W., & de Lange, T. (2023). CST-Pola/Primase: the second telomere maintenance machine. *Genes & development*, 37(13-14), 555–569. <https://doi.org/10.1101/gad.350479.123>
25. Cano, R., & López, R. (1987). The biology of *Entamoeba histolytica*: an update. *Archives of Medical Research*, 18(1), 13-19.
26. Cardano, M., Tribioli, C., & Prosperi, E. (2020). Targeting Proliferating Cell Nuclear Antigen (PCNA) as an Effective Strategy to Inhibit Tumor Cell Proliferation. *Current cancer drug targets*, 20(4), 240–252. <https://doi.org/10.2174/1568009620666200115162814>
27. Cardona-Felix, C. S., Lara-Gonzalez, S., & Briebe, L. G. (2011). Structure and biochemical characterization of proliferating cellular nuclear antigen from a parasitic protozoon. *Acta crystallographica. Section D, Biological crystallography*, 67(Pt 6), 497–505. <https://doi.org/10.1107/S0907444911010547>
28. Cardona-Felix, C., Pastor-Palacios, G., Cardenas, H., Azuara-Liceaga, E., & Briebe, L. (2010). Biochemical characterization of the DNA ligase I from *Entamoeba histolytica*. *Molecular And Biochemical Parasitology*, 174(1), 26-35. doi: 10.1016/j.molbiopara.2010.06.010
29. Carty, M. P., Zernik-Kobak, M., McGrath, S., & Dixon, K. (2003). UV light-induced DNA damage results in a reorganization of nuclear replication complexes. *Nucleic Acids Research*, 31(19), 291–299. <https://doi.org/10.1093/nar/gkg900>
30. CDC. (2019, Febrero 11). *Amebiasis*. Recuperado el 18 de marzo del 2025 de: <https://www.cdc.gov/dpdx/amebiasis/index.html>
31. Cerritelli, S. M., & Crouch, R. J. (2009). Ribonuclease H: the enzymes in eukaryotes. *FEBS Journal*, 276(6), 1494–1505. <https://doi.org/10.1111/j.1742-4658.2009.06908.x>
32. Cerritelli, S. M., & Crouch, R. J. (2009). Ribonuclease H: The enzymes in eukaryotes. *The FEBS Journal*, 276(6), 1494–1505. <https://doi.org/10.1111/j.1742-4658.2009.06884.x>
33. Champoux, J. J. (2001). DNA topoisomerases: structure, function, and mechanism. *Annual Review of Biochemistry*, 70(1), 369–413.
34. Chou, A., & Austin, R. L. (2023). *Entamoeba histolytica* Infection. In *StatPearls*. StatPearls Publishing.
35. Clark, C. G., & Diamond, L. S. (2002). The *Entamoeba histolytica* complex. *International Journal for Parasitology*, 32(12), 1759-1771.
36. Clark, C.G. et al. (2007) 'Structure and content of the *Entamoeba histolytica* genome', *Advances in Parasitology*, 65, pp. 51–190. doi:10.1016/s0065-308x(07)65002-7.
37. Costa, A., & Diffley, J. F. X. (2022). The Initiation of Eukaryotic DNA Replication. *Annual review of biochemistry*, 91, 107–131. <https://doi.org/10.1146/annurev-biochem-072321-110228>
38. Crosby M. E. (2007). Cell Cycle: Principles of Control. *The Yale Journal of Biology and Medicine*, 80(3), 141–142.
39. da Silva, M. S., Pavani, R. S., Damasceno, J. D., Marques, C. A., McCulloch, R., Tosi, L., & Elias, M. C. (2017). Nuclear DNA Replication in Trypanosomatids: There Are No Easy Methods for Solving Difficult Problems. *Trends in parasitology*, 33(11), 858–874. <https://doi.org/10.1016/j.pt.2017.08.002>
40. Das, S. et al. (2004) "Constitutive association of mcm2-3-5 proteins with chromatin in *Entamoeba histolytica*", *Cellular Microbiology*, 7(2), pp. 259–267. <https://doi.org/10.1111/j.1462-5822.2004.00456.x>
41. Das, S., & Lohia, A. (2000). MCM proteins of *Entamoeba histolytica*. *Archives of medical research*, 31(4 Suppl), S269–S270. [https://doi.org/10.1016/s0188-4409\(00\)00136-3](https://doi.org/10.1016/s0188-4409(00)00136-3)
42. Das, S., & Lohia, A. (2002). Delinking of S phase and cytokinesis in the protozoan parasite *Entamoeba histolytica*. *Cellular Microbiology*, 4(1), 55–60. <https://doi.org/10.1046/j.1462-5822.2002.00165.x>
43. DePamphilis M. L. (2016). Genome Duplication: The Heartbeat of Developing Organisms. *Current topics in developmental biology*, 116, 201–229. <https://doi.org/10.1016/bs.ctdb.2015.10.002>

44. Dewar, J. M., & Walter, J. C. (2017). Mechanisms of DNA replication termination. *Nature reviews. Molecular cell biology*, 18(8), 507–516. <https://doi.org/10.1038/nrm.2017.42>
45. Dhar, S. K., Choudhury, N. R., Mittal, V., Bhattacharya, A., & Bhattacharya, S. (1996). Replication initiates at multiple dispersed sites in the ribosomal DNA plasmid of the protozoan parasite *Entamoeba histolytica*. *Molecular and cellular biology*, 16(5), 2314–2324. <https://doi.org/10.1128/MCB.16.5.2314>
46. Dhar, S. K., Delmolino, L., & Dutta, A. (2001). Architecture of the human origin recognition complex. *The Journal of biological chemistry*, 276(31), 29067–29071. <https://doi.org/10.1074/jbc.M103078200>
47. Ellenberger, T., & Tomkinson, A. E. (2008). Eukaryotic DNA ligases: structural and functional insights. *Annual Review of Biochemistry*, 77, 313–338. <https://doi.org/10.1146/annurev.biochem.77.061306.123941>
48. Espinosa, M., & Martínez, A. (2000). Pathogenesis of intestinal amebiasis: from molecules to disease. *Clinical Microbiology Reviews*, 13(2), 318–331.
49. Essers, J., Theil, A. F., Baldeyron, C., van Cappellen, W. A., Houtsmuller, A. B., Kanaar, R., & Vermeulen, W. (2005). Nuclear dynamics of PCNA in DNA replication and repair. *Molecular and cellular biology*, 25(21), 9350–9359. <https://doi.org/10.1128/MCB.25.21.9350-9359.2005>
50. Fairall, L., Chapman, L., Moss, H., de Lange, T., & Rhodes, D. (2001). Structure of the TRFH dimerization domain of the human telomeric proteins TRF1 and TRF2. *Molecular cell*, 8(2), 351–361. [https://doi.org/10.1016/s1097-2765\(01\)00321-5](https://doi.org/10.1016/s1097-2765(01)00321-5)
51. Fleta, J., Bueno, O., & Bueno, M. (2000). *Amebiasis intestinal Y Absceso Hepático Amebiano*, Medicina Integral. Recuperado el 20 de noviembre del 2023 de: <https://www.elsevier.es/es-revista-medicina-integral-63-articulo-amebiasis-intestinal-absceso-hepatico-amebiano-15325>.
52. Fotedar, R., Stark, D., Beebe, N., Marriott, D., Ellis, J., & Harkness, J. (2007). Laboratory diagnostic techniques for *Entamoeba* species. *Clinical Microbiology Reviews*, 20(3), 511–532.
53. Frick, D. N., & Richardson, C. C. (2001). DNA primases. *Annual review of biochemistry*, 70, 39–80. <https://doi.org/10.1146/annurev.biochem.70.1.39>
54. Fu, B., Wang, J., & Fu, X. (2022). A rare case of extraintestinal amebiasis. *BMC infectious diseases*, 22(1), 364. <https://doi.org/10.1186/s12879-022-07348-9>
55. Gambus, A., & Blow, J. J. (2013). Mcm8 and Mcm9 form a dimeric complex in *Xenopus laevis* egg extract that is not essential for DNA replication initiation. *Cell cycle (Georgetown, Tex.)*, 12(8), 1225–1232. <https://doi.org/10.4161/cc.24310>
56. Gangopadhyay, S. S., Ray, S. S., Kennady, K., Pande, G., & Lohia, A. (1997). Heterogeneity of DNA content and expression of cell cycle genes in axenically growing *Entamoeba histolytica* HM1:IMSS clone A. *Molecular and biochemical parasitology*, 90(1), 9–20. [https://doi.org/10.1016/s0166-6851\(97\)00156-4](https://doi.org/10.1016/s0166-6851(97)00156-4)
57. Garcia-Diaz, M., & Bebenek, K. (2007). Multiple functions of DNA polymerases. *Critical reviews in plant sciences*, 26(2), 105–122. <https://doi.org/10.1080/07352680701252817>
58. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (2005). *Protein Identification and Analysis Tools on the ExPASy Server*. Humana Press. 571–607 pp.
59. Ghosh, S., Satish, S., Tyagi, S., Bhattacharya, A., & Bhattacharya, S. (2003). Differential use of multiple replication origins in the ribosomal DNA episome of the protozoan parasite *Entamoeba histolytica*. *Nucleic acids research*, 31(8), 2035–2044. <https://doi.org/10.1093/nar/gkg320>
60. González-Corrochano, R., Ruiz, F. M., Taylor, N. M. I., Huecas, S., Drakulic, S., Spínola-Amilibia, M., & Fernández-Tornero, C. (2020). The crystal structure of human XPG, the xeroderma pigmentosum group G endonuclease, provides insight into nucleotide excision DNA repair. *Nucleic acids research*, 48(17), 9943–9958. <https://doi.org/10.1093/nar/gkaa688>
61. Gulbis, J. M., Kelman, Z., Hurwitz, J., O'Donnell, M., & Kuriyan, J. (1996). Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell*, 87(2), 297–306. [https://doi.org/10.1016/S0092-8674\(00\)81348-8](https://doi.org/10.1016/S0092-8674(00)81348-8)
62. Habraken, Y., Sung, P., Prakash, L., & Prakash, S. (1994). Human xeroderma pigmentosum group G gene encodes a DNA endonuclease. *Nucleic acids research*, 22(16), 3312–3316. <https://doi.org/10.1093/nar/22.16.3312>
63. Hallgren, J. et al. (2022). DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks [Preprint]. doi:10.1101/2022.04.08.487609.

64. Hanks, S. K., Quinn, A. M., & Hunter, T. (1988). The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science (New York, N.Y.)*, 241(4861), 42–52. <https://doi.org/10.1126/science.3291115>
65. Hatoyama, Y., & Kanemaki, M. T. (2023). The assembly of the MCM2-7 hetero-hexamers and its significance in DNA replication. *Biochemical Society transactions*, 51(3), 1289–1295. <https://doi.org/10.1042/BST20221465>
66. Hattacharya S. (2023). Episomal and chromosomal DNA replication and recombination in *Entamoeba histolytica*. *Frontiers in molecular biosciences*, 10, 1212082. <https://doi.org/10.3389/fmolb.2023.1212082>
67. Helderma, N. C., Terlouw, D., Bonjoch, L., Golubicki, M., Antelo, M., Morreau, H., van Wezel, T., Castellvi-Bel, S., Goldberg, Y., & Nielsen, M. (2023). Molecular functions of MCM8 and MCM9 and their associated pathologies. *iScience*, 26(6), 106737. <https://doi.org/10.1016/j.isci.2023.106737>
68. Henninger, E. E., & Pursell, Z. F. (2014). DNA polymerase  $\epsilon$  and its roles in genome stability. *IUBMB life*, 66(5), 339–351. <https://doi.org/10.1002/iub.1276>
69. Hosfield, D. J., Mol, C. D., Shen, B., & Tainer, J. A. (1998). Structure of the DNA repair and replication endonuclease and exonuclease FEN-1: coupling DNA and PCNA binding to FEN-1 activity. *Cell*, 95(1), 135–146. [https://doi.org/10.1016/S0092-8674\(00\)81789-4](https://doi.org/10.1016/S0092-8674(00)81789-4)
70. Huang, Y., Amin, A., Qin, Y., Wang, Z., Jiang, H., Liang, L., et al. (2016) A Role of hIP13 in DNA Replication Licensing in Human Cells. *PLoS ONE* 11(4): e0151803. <https://doi.org/10.1371/journal.pone.0151803>
71. Hyjek, M., Figiel, M., & Nowotny, M. (2019). RNases H: Structure and mechanism. *DNA repair*, 84, 102672. <https://doi.org/10.1016/j.dnarep.2019.102672>
72. Ilves, I., Petojevic, T., Pesavento, J. J., & Botchan, M. R. (2010). Activation of the MCM2–7 helicase by association with Cdc45 and GINS proteins. *Molecular Cell*, 37(2), 247–258. <https://doi.org/10.1016/j.molcel.2009.12.030>
73. Iyer, L. M., et al. (2004). Evolutionary classification of AAA+ ATPases. *Journal of Structural Biology*, 146(1–2), 11–31. <https://doi.org/10.1016/j.jsb.2003.10.010>
74. Kawano-Sugaya, T., Izumiyama, S., Yanagawa, Y., Saito-Nakano, Y., Watanabe, K., Kobayashi, S., Nakada-Tsukui, K., & Nozaki, T. (2020). Near-chromosome level genome assembly reveals ploidy diversity and plasticity in the intestinal protozoan parasite *Entamoeba histolytica*. *BMC genomics*, 21(1), 813. <https://doi.org/10.1186/s12864-020-07167-9>
75. Kelch, B. A. (2016). Review: The Lord of the rings: Structure and mechanism of the sliding clamp loader. *Biopolymers*, 105(8), 532–546. <https://doi.org/10.1002/bip.22847>
76. Kelman Z. (1997). PCNA: structure, functions and interactions. *Oncogene*, 14(6), 629–640. <https://doi.org/10.1038/sj.onc.1200886>
77. Kimura, S., Furukawa, T., Kasai, N., Mori, Y., Kitamoto, H. K., Sugawara, F., Hashimoto, J., & Sakaguchi, K. (2003). Functional characterization of two flap endonuclease-1 homologues in rice. *Gene*, 314, 63–71. [https://doi.org/10.1016/S0378-1119\(03\)00694-2](https://doi.org/10.1016/S0378-1119(03)00694-2)
78. Klingbeil, M. M., Motyka, S. A., & Englund, P. T. (2011). Replication of the kinetoplast DNA network. *Nucleus*, 2(4), 424–432. <https://doi.org/10.1002/nuc.127095>
79. Kragelund, B. B., et al. (2006). Order by disorder in plant signaling. *Trends in Plant Science*, 11(11), 524–530. <https://doi.org/10.1016/j.tplants.2006.09.003>
80. Krasikova, Y. S., Rechkunova, N. I., & Lavrik, O. I. (2016). *Molekuliarnaia biologiya*, 50(5), 735–750. <https://doi.org/10.7868/S0026898416030083>
81. Kuchta, R. D., & Stengel, G. (2010). Mechanism and evolution of DNA primases. *Biochimica et biophysica acta*, 1804(5), 1180–1189. <https://doi.org/10.1016/j.bbapap.2009.06.011>
82. Kurniawan, F. et al. (2018) 'Crystal structure of *Entamoeba histolytica* cdc45 suggests a conformational switch that may regulate DNA replication', *iScience*, 3, pp. 102–109. [doi:10.1016/j.isci.2018.04.011](https://doi.org/10.1016/j.isci.2018.04.011)
83. Labib, K. (2010). How do Cdc7 and cyclin-dependent kinases trigger the initiation of chromosome replication in eukaryotic cells?. *Genes & Development*, 24(12), 1208–1219.
84. Li, Y., & Araki, H. (2013). Loading and activation of DNA replicative helicases: the key step of initiation of DNA replication. *Genes to cells : devoted to molecular & cellular mechanisms*, 18(4), 266–277. <https://doi.org/10.1111/gtc.12040>
85. Lieberman-Aiden, E., Van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., ... & Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *science*, 326(5950), 289–293.

86. Liu, Y., Kao, H. I., & Bambara, R. A. (2004). Flap endonuclease 1: a central component of DNA metabolism. *Annual Review of Biochemistry*, 73, 589–615. <https://doi.org/10.1146/annurev.biochem.73.011303.073918>
87. Loftus, B., Anderson, I., Davies, R., Alsmark, U. C., Samuelson, J., Amedeo, P., ... & Hall, N. (2005). The genome of the protist parasite *Entamoeba histolytica*. *Nature*, 433(7028), 865–868.
88. Lohia, A. (2003). The cell cycle of *Entamoeba histolytica*. *Molecular and Biochemical Parasitology*, 126(1), 1–10. [https://doi.org/10.1016/S0166-6851\(02\)00243-0](https://doi.org/10.1016/S0166-6851(02)00243-0)
89. Lohia, A., Mukherjee, C., Majumder, S., & Dastidar, P. G. (2007). Genome re-duplication and irregular segregation occur during the cell cycle of *Entamoeba histolytica*. *Bioscience reports*, 27(6), 373–384. <https://doi.org/10.1007/s10540-007-9058-8>
90. Lorenzi, H. A., Puiu, D., Miller, J. R., Brinkac, L. M., Amedeo, P., Hall, N., ... & Kissinger, J. C. (2010). New assembly, reannotation and analysis of the *Entamoeba histolytica* genome reveal new genomic features and protein content information. *PLoS Neglected Tropical Diseases*, 4(7), e716.
91. Marques, C. A., Tiengwe, C., Lemgruber, L., Damasceno, J. D., Scott, A., Paape, D., Marcello, L., & McCulloch, R. (2016). Diverged composition and regulation of the *Trypanosoma brucei* origin recognition complex that mediates DNA replication initiation. *Nucleic acids research*, 44(10), 4763–4784. <https://doi.org/10.1093/nar/gkw147>
92. Mendez, J., & Stillman, B. (2000). Chromatin association of the DNA replication factors Cdc45 and Mcm2-7 during S phase and their redistribution at the end of mitosis. *Molecular and Cellular Biology*, 20(21), 6716–6729.
93. Moldovan, G. L., Pfander, B., & Jentsch, S. (2007). PCNA, the maestro of the replication fork. *Cell*, 129(4), 665–679. <https://doi.org/10.1016/j.cell.2007.05.003>
94. Moyer, S. E., Lewis, P. W., & Botchan, M. R. (2006). Isolation of the Cdc45/Mcm2–7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proceedings of the National Academy of Sciences*, 103(27), 10236–10241. <https://doi.org/10.1073/pnas.0602400103>
95. Mukherjee, C., Majumder, S., & Lohia, A. (2009). Inter-cellular variation in DNA content of *Entamoeba histolytica* originates from temporal and spatial uncoupling of cytokinesis from the nuclear cycle. *PLoS neglected tropical diseases*, 3(4), e409. <https://doi.org/10.1371/journal.pntd.0000409>
96. Muramatsu, S., Hirai, K., Tak, Y. S., Kamimura, Y., & Araki, H. (2010). CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol (epsilon), and GINS in budding yeast. *Genes & development*, 24(6), 602–612. <https://doi.org/10.1101/gad.1883410>
97. Muzi-Falconi, M., Giannattasio, M., Foiani, M., & Plevani, P. (2003). The DNA polymerase alpha-primase complex: multiple functions and interactions. *TheScientificWorldJournal*, 3, 21–33. <https://doi.org/10.1100/tsw.2003.05>
98. Naiyer, S., Kaur, D., Ahamad, J., Singh, S. S., Singh, Y. P., Thakur, V., Bhattacharya, A., & Bhattacharya, S. (2019). Transcriptomic analysis reveals novel downstream regulatory motifs and highly transcribed virulence factor genes of *Entamoeba histolytica*. *BMC genomics*, 20(1), 206. <https://doi.org/10.1186/s12864-019-5570-z>
99. Neer, E. J., et al. (1994). The ancient regulatory-protein family of WD-repeat proteins. *Nature*, 371(6495), 297–300. <https://doi.org/10.1038/371297a0>
100. Neuwald, A. F., et al. (1999). AAA+: A class of chaperone-like ATPases. *Genome Research*, 9(1), 27–43. <https://doi.org/10.1101/gr.9.1.27>
101. Nishimura, K., Ishiai, M., Horikawa, K., Fukagawa, T., Takata, M., & Takisawa, H. (2012). MCM8 and MCM9 form a complex that functions in homologous recombination repair induced by DNA double-strand breaks. *Molecular Cell*, 47(4), 511–522. <https://doi.org/10.1016/j.molcel.2012.05.047>
102. Nishitani, H., & Nurse, P. (1995). The Cdt1 protein is required for the initiation of DNA replication in fission yeast. *Nature*, 374(6518), 656–659.
103. Nitiss, J. L., Soans, E., Rogojina, A., Seth, A., & Mishina, M. (2012). Topoisomerase assays. *Current protocols in pharmacology*, Chapter 3, Unit3.3–3.3.. <https://doi.org/10.1002/0471141755.ph0303s57>
104. Noguchi, K., Vassilev, A., Ghosh, S., Yates, J. L., & DePamphilis, M. L. (2006). The BAH domain facilitates the ability of human Orc1 protein to activate replication origins in vivo. *The EMBO journal*, 25(22), 5372–5382. <https://doi.org/10.1038/sj.emboj.7601396>
105. Nozaki, T. and Bhattacharya, A. (2015) Amebiasis: Biology and pathogenesis of *Entamoeba*. Tokyo: Springer. 393-410.

106. Ocaña-Pallarès, E. et al. (2020) "Origin recognition complex (ORC) evolution is influenced by global gene duplication/loss patterns in eukaryotic genomes", *Genome Biology and Evolution*, 12(2), pp. 3878–3889. doi:10.1093/gbe/evaa011.
107. Oliveira-Costa, J. P., Zanetti, J., Oliveira, L. R., Soares, F. A., Ramalho, L. Z., Silva Ramalho, F., Garcia, S. B., & Ribeiro-Silva, A. (2010). Significance of topoisomerase III $\beta$  expression in breast ductal carcinomas: strong associations with disease-specific survival and metastasis. *Human pathology*, 41(11), 1624–1630.
108. Orozco, E., Báez-Camargo, M., Riveron, A. M., Gharibeh, R., Gariglio, P., de la Cruz Hernández, F., & Chávez, P. (1997). A model for unscheduled DNA replication in *Entamoeba histolytica* trophozoites. *Archives of medical research*, 28 Spec No, 24–26.
109. Orozco, E., Solís, F. J., Domínguez, J., Chávez, B., & Hernández, F. (1988). *Entamoeba histolytica*: cell cycle and nuclear division. *Experimental parasitology*, 67(1), 85–95. [https://doi.org/10.1016/0014-4894\(88\)90011-2](https://doi.org/10.1016/0014-4894(88)90011-2)
110. Pal, S. & Dam, S. (2023) "Identification and characterisation of a novel EHORC1/Cdc6 from the human pathogen *Entamoeba histolytica*: An *in silico* approach", *Journal of Biomolecular Structure and Dynamics*, 43(4), pp. 1883–1892. doi:10.1080/07391102.2023.2293264.
111. Palm, W., & De Lange, T. (2008). How shelterin protects mammalian telomeres. *Annual review of genetics*, 42(1), 301–334. <https://doi.org/10.1146/annurev.genet.41.110306.130350>
112. Park, J. H., Park, E. J., Hur, S. K., Kim, S., & Kwon, J. (2013). Mammalian MCM8–9 complex promotes replication fork progression during homologous recombination repair. *Cell Cycle*, 12(10), 1617–1624. <https://doi.org/10.4161/cc.24601>
113. Pascal, J. M., et al. (2004). Crystal structure of a DNA ligase DNA–adenylate intermediate: Insights into adenylation and the architecture of DNA ends. *Molecular Cell*, 14(5), 757–769. <https://doi.org/10.1016/j.molcel.2004.05.021>
114. Pastor-Palacios, G., Azuara-Liceaga, E., and Briebe, L. G. (2010). A nuclear family A DNA polymerase from *Entamoeba histolytica* bypasses thymine glycol. *PLoS Negl. Trop. Dis.* 4:e786. doi: 10.1371/journal.pntd.0000786
115. Pfeiffer, V., & Lingner, J. (2013). Replication of telomeres and the regulation of telomerase. *Cold Spring Harbor perspectives in biology*, 5(5), a010405. <https://doi.org/10.1101/cshperspect.a010405>
116. Piatti, S., Hill, C. M., & Nasmyth, K. (1996). Cdc6p is required for the association of Mcm proteins with origins of DNA replication in yeast. *Cell*, 86(2), 311–319.
117. Pike, A. M., Friend, C. M., & Bell, S. P. (2023). Distinct RPA functions promote eukaryotic DNA replication initiation and elongation. *Nucleic acids research*, 51(19), 10506–10518. <https://doi.org/10.1093/nar/gkad765>
118. Plank, J. L., & Hsieh, T. S. (2009). Roles of topoisomerases in maintaining genome stability during DNA replication. *Genes & Development*, 23(23), 2780–2793.
119. Pommier, Y., Sun, Y., Huang, S., et al. Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nat Rev Mol Cell Biol* 17, 703–721 (2016). <https://doi.org/10.1038/nrm.2016.111>
120. Ramesh, M. A., Malik, S. B., & Logsdon, J. M. (2005). A phylogenomic inventory of meiotic genes: Evidence for sex in *Giardia* and an early eukaryotic origin of meiosis. *Current Biology*, 15(2), 185–191. <https://doi.org/10.1016/j.cub.2005.01.003>
121. Ramírez-Moreno, M., Martínez-Robles, M. L., & Mendoza-Hernández, G. (2018). Structural insights into the *Entamoeba histolytica* CDC45 protein. *iScience*, 1(15), 45–56. DOI: 10.1016/j.isci.2018.04.003
122. Ravdin, J. I., & Stauffer, W. (2005). *Entamoeba histolytica* (Amebiasis). In Mandell, Douglas, and Bennett's *Principles and Practice of Infectious Diseases* (pp. 3097–3111).
123. Ren, J., Wen, L., Gao, X., Jin, C., Xue, Y., & Yao, X. (2009). DOG 1.0: illustrator of protein domain structures. *Cell research*, 19(2), 271–273. <https://doi.org/10.1038/cr.2009.6>
124. Rendón, F. J., Álvarez, V., Castañeda, E. J., Cárdenas, H., Cárdenas, R. E., Valdés, J., ... & Azuara, E. (2018). Telomeric Repeat-Binding Factor Homologs in *Entamoeba histolytica*: New Clues for Telomeric Research. *Frontiers in Cellular and Infection Microbiology*, 8, 341. <https://doi.org/10.3389/fcimb.2018.00341>
125. Rivero, R. M., Escobedo, A. A., & González, M. E. (2021). Amebiasis. *Infectious Disease Clinics*, 35(1), 1–10.
126. Saito, Y., Santosa, V., Ishiguro, K. I., & Kanemaki, M. T. (2022). MCMBP promotes the assembly of the MCM2–7 hetero-hexamers to ensure robust DNA replication in human cells. *eLife*, 11, e77393. <https://doi.org/10.7554/eLife.77393>

127. Samel, S. A., et al. (2014). A unique DNA entry gate serves for regulated loading of the eukaryotic replicative helicase MCM2–7 onto DNA. *Genes & Development*, 28(15), 1653–1666. <https://doi.org/10.1101/gad.242156.114>
128. Schnauffer, A., Domingo-Sananes, M. R., & Stuart, K. (2018). The evolution of kinetoplast DNA: More than a junkyard of repeated sequences. *Trends in Parasitology*, 34(2), 102–114. DOI: 10.1016/j.pt.2017.11.007
129. 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. <https://doi.org/10.1093/ofid/ofy161>
130. Simon, A. C., Sannino, V., Costanzo, V., & Pellegrini, L. (2016). Structure of human Cdc45 and implications for CMG helicase function. *Nature communications*, 7, 11638. <https://doi.org/10.1038/ncomms11638>
131. Singh, U., Banerjee, S., & Sinha, P. (2013). Mitosis in *Entamoeba histolytica*: Structure and dynamics of mitotic spindles. *Protistology*, 7(2), 77–86.
132. Smogorzewska, A., & de Lange, T. (2004). Regulation of telomerase by telomeric proteins. *Annual Review of Biochemistry*, 73, 177–208. <https://doi.org/10.1146/annurev.biochem.73.011303.073800>
133. Sodeman, W. A., Jr (1996). Intestinal Protozoa: Amebas. In S. Baron (Ed.), *Medical Microbiology*. (4th ed.) Chapter 79. University of Texas Medical Branch at Galveston.
134. Stanley, S. L. (2003). Amoebiasis. *The Lancet, Elsevier* 361(9362), 1025–1034. [https://doi.org/10.1016/S0140-6736\(03\)12830-9](https://doi.org/10.1016/S0140-6736(03)12830-9)
135. Stillman, B. (1996). Cell cycle control of DNA replication. *Science*, 274(5293), 1659–1664. <https://doi.org/10.1126/science.274.5293.1659>
136. Stillman, B. (2008). DNA polymerases at the replication fork in eukaryotes. *Molecular Cell*, 30(3), 259–260. <https://doi.org/10.1016/j.molcel.2008.04.018>
137. Stirpe, M., et al. (2019). The role of the WD40 domain in the regulation of the eukaryotic cell cycle. *Cell Cycle*, 18(1), 1–15. <https://doi.org/10.1080/15384101.2018.1557491>
138. Strzalka, W., & Ziemienowicz, A. (2011). Proliferating cell nuclear antigen (PCNA): a key factor in DNA replication and cell cycle regulation. *Annals of botany*, 107(7), 1127–1140. <https://doi.org/10.1093/aob/mcq243>
139. The PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.
140. Tiengwe, C. et al. (2012) “Identification of orc1/cdc6-interacting factors in *Trypanosoma brucei* reveals critical features of Origin Recognition Complex architecture”, *PLoS ONE*, 7(3). doi:10.1371/journal.pone.0032674.
141. Tsutakawa, S. E., Classen, S., Chapados, B. R., Arvai, A. S., Finger, L. D., Guenther, G., Tomlinson, C. G., Thompson, P., Sarker, A. H., Shen, B., Cooper, P. K., Grasby, J. A., & Tainer, J. A. (2011). Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. *Cell*, 145(2), 198–211. <https://doi.org/10.1016/j.cell.2011.03.004>
142. UniProt Consortium (2025). UniProt: the Universal Protein Knowledgebase in 2025. *Nucleic acids research*, 53(D1), D609–D617. <https://doi.org/10.1093/nar/gkae1010>
143. Varghese, S. S., & Ghosh, S. K. (2020). Stress-responsive *Entamoeba* topoisomerase II: a potential antiamebic target. *FEBS letters*, 594(6), 1005–1020. <https://doi.org/10.1002/1873-3468.13677>
144. Walker, J. E., et al. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *The EMBO Journal*, 1(8), 945–951. <https://doi.org/10.1002/j.1460-2075.1982.tb01276.x>
145. Wang, J. C. (2002). Cellular roles of DNA topoisomerases: a molecular perspective. *Nature Reviews Molecular Cell Biology*, 3(6), 430–440.
146. Watson, J. and Harrison, S., 2016. *Biología molecular del gen*. 7th ed. México, D.F.: Médica Panamericana, pp.256-268.
147. Weedall, G. D., & Hall, N. (2015). Evolutionary genomics of *Entamoeba*. *Research in Microbiology*, 166(10), 942–951. <https://doi.org/10.1016/j.resmic.2015.10.003>
148. Weng, Z., Zheng, J., Zhou, Y., Lu, Z., Wu, Y., Xu, D., Li, H., Liang, H., & Liu, Y. (2023). Structural and mechanistic insights into the MCM8/9 helicase complex. *eLife*, 12, RP87468. <https://doi.org/10.7554/eLife.87468>

149. Willhoeft, U. and Tannich, E. (1999) "The electrophoretic karyotype of *Entamoeba histolytica*", *Molecular and Biochemical Parasitology*, 99(1), pp. 41–53. doi:10.1016/s0166-6851(98)00178-9.
150. Wohlschlegel, J. A., Dhar, S. K., Prokhorova, T. A., Dutta, A., & Walter, J. C. (2002). Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. *Science*, 295(5554), 2456–2459. <https://doi.org/10.1126/science.1068664>
151. Wold, M. S. (1997). Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annual Review of Biochemistry*, 66(1), 61–92. <https://doi.org/10.1146/annurev.biochem.66.1.61>
152. Ximénez, C., Morán, P., Rojas, L., Valadez, A., Gómez, A., Miranda, E., & Valenzuela, O. (2011). Novelties on amoebiasis: a neglected tropical disease. *Journal of global infectious diseases*, 3(2), 164-170. <https://doi.org/10.4103/0974-777X.81695>
153. Yao, N. Y., & O'Donnell, M. (2012). The RFC clamp loader: structure and function. *Sub-cellular biochemistry*, 62, 259–279. [https://doi.org/10.1007/978-94-007-4572-8\\_14](https://doi.org/10.1007/978-94-007-4572-8_14)
154. Zhang, P., Sun, Y., Hsu, H., Zhang, L., Zhang, Y., & Lee, M. Y. (1998). The interdomain connector loop of human PCNA is involved in a direct interaction with human polymerase delta. *The Journal of biological chemistry*, 273(2), 713–719. <https://doi.org/10.1074/jbc.273.2.713>
155. Zinder, J. C., Olinares, P. D. B., Svetlov, V., Bush, M. W., Nudler, E., Chait, B. T., Walz, T., & de Lange, T. (2022). Shelterin is a dimeric complex with extensive structural heterogeneity. *Proceedings of the National Academy of Sciences of the United States of America*, 119(31), e2201662119. <https://doi.org/10.1073/pnas.2201662119>

## 12. Anexos

**Anexo 1. Características de las proteínas que participan en la replicación del DNA en *Entamoeba histolytica***

| No | Nombre de la proteína (ID) | Anotación <sup>a</sup>                     | GO <sup>b</sup>    | Tamaño DNA/mRNA (pb) | Tamaño (aa) | PM kDa | pl   | Secuencia de Localización Nuclear (NLS) <sup>c</sup> | Localización |
|----|----------------------------|--------------------------------------------|--------------------|----------------------|-------------|--------|------|------------------------------------------------------|--------------|
| 1  | EhOrc1 (EHI_105150)        | Hypothetical protein                       | Replicación de DNA | 1116/1116            | 371         | 43.27  | 8.27 | N                                                    | Nu           |
| 2  | EHCDC6 (EHI_192770)        | Hypothetical protein                       | Replicación de DNA | 1176/1176            | 391         | 44.76  | 8.60 | N                                                    | Nu           |
| 3  | EhMCM2 (EHI_117970)        | DNA replication licensing factor           | Replicación de DNA | 2659/2743            | 881         | 100.03 | 5.15 | N                                                    | Nu           |
| 4  | EhMCM3 (EHI_103600)        | DNA replication licensing factor MCM3      | Replicación de DNA | 1806/1806            | 601         | 66.92  | 5.74 | 474- ERLKRIKHNFK -484                                | Nu           |
| 5  | EhMCM4 (EHI_187720)        | DNA replication licensing factor, putative | Replicación de DNA | 1827/ 1827           | 608         | 68.42  | 7.21 | N                                                    | Nu           |
| 6  | EhMCM5 (EHI_069980)        | DNA replication licensing factor           | Replicación de DNA | 1920/ 1965           | 639         | 71.05  | 8.25 | 625- AVRQKEKHKKRRDDK -639                            | Nu           |
| 7  | EhMCM6 (EHI_118870)        | DNA replication licensing factor, putative | Replicación de DNA | 2049/2049            | 682         | 76.49  | 5.30 | 279- EKKR -282                                       | Nu           |
| 8  | EhMCM7 (EHI_158110)        | DNA replication licensing factor, putative | Replicación de DNA | 2073/2073            | 690         | 77.55  | 8.17 | 600- KKRSKPKNK - 608                                 | Nu           |
| 9  | EhIPI3_WD (EHI_137840)     | Hypothetical protein                       | Replicación de DNA | 1236/ 1236           | 411         | 46.08  | 7.86 | N                                                    | Nu           |
| 10 | EhMCM8 (EHI_140670)        | DNA replication licensing factor, putative | Replicación de DNA | 2202/2202            | 733         | 81.91  | 6.80 | N                                                    | Nu           |

Etapa Pre-RC

|              |    |                                   |                                                              |                                          |            |     |       |      |   |       |
|--------------|----|-----------------------------------|--------------------------------------------------------------|------------------------------------------|------------|-----|-------|------|---|-------|
| Etapa Pre-IC | 11 | EhMCM9<br>(EHI_076880)            | DNA replication<br>licensing factor,<br>putative             | Replicación de<br>DNA                    | 2433/ 2433 | 810 | 92.79 | 5.40 | N | Nu    |
|              | 12 | EhMCMBP<br>(EHI_166840)           | Hypothetical<br>protein                                      | Replicación de<br>DNA                    | 1470/ 1470 | 489 | 57.11 | 5.56 | N | Nu    |
|              | 12 | EhCDC45<br>(EHI_049900)           | Cell division control<br>protein 45 CDC45,<br>putative       | Replicación de<br>DNA                    | 1632/ 1632 | 543 | 62.37 | 6.17 | N | Nu    |
|              | 13 | EhPSF1<br>(GINS)<br>(EHI_136990)  | DNA replication<br>complex GINS<br>protein PSF1,<br>putative | Replicación de<br>DNA                    | 570/619    | 189 | 21.57 | 5.17 | N | Nu    |
|              | 14 | EhPSF2<br>(GINS)<br>(EHI_069340)  | Hypothetical<br>protein, conserved                           | Replicación de<br>DNA                    | 600/600    | 199 | 23.33 | 4.68 | N | Nu    |
|              | 15 | Eh PSF3<br>(GINS)<br>(EHI_069960) | Hypothetical<br>protein                                      | Replicación de<br>DNA                    | 651/ 651   | 216 | 24.85 | 5.56 | N | Nu    |
|              | 16 | Eh SDL5<br>(GINS)<br>(EHI_048430) | Hypothetical<br>protein                                      | Replicación de<br>DNA                    | 615/ 716   | 204 | 24.33 | 6.07 | N | Nu    |
|              | 17 | EhCDK1 I<br>(EHI_068070)          | cyclin-dependent<br>kinase protein,<br>putative              | Actividad de<br>cinasa                   | 933/ 933   | 310 | 35.54 | 8.64 | N | C, Nu |
|              | 18 | EhCDK1 II<br>(EHI_105300)         | Cell division protein<br>kinase, putative                    | Fosforilación<br>/Actividad de<br>cinasa | 912/968    | 303 | 35.52 | 8.58 | N | C, Nu |
|              | 19 | EhCDK1 III<br>(EHI_048710)        | Protein kinase<br>domain containing<br>protein               | Fosforilación<br>/Actividad de<br>cinasa | 882/ 882   | 293 | 33.67 | 6.0  | N | C, Nu |

|    |                            |                                                |                                                                                              |           |      |        |      |                                                                                                                                                                                                                                                                                             |       |
|----|----------------------------|------------------------------------------------|----------------------------------------------------------------------------------------------|-----------|------|--------|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| 20 | EhCDK2 I<br>(EHI_065280)   | Cell division protein<br>kinase 2, putative    | Fosforilación<br>/Actividad de<br>cinasa                                                     | 876/955   | 291  | 33.83  | 7.62 | N                                                                                                                                                                                                                                                                                           | C, Nu |
| 21 | EhCDK2 II<br>(EHI_153140)  | Cell division protein<br>kinase 2, putative    | Fosforilación<br>/Actividad de<br>cinasa                                                     | 933/996   | 310  | 35.82  | 7.21 | N                                                                                                                                                                                                                                                                                           | C, Nu |
| 22 | EhCDK2 III<br>(EHI_152430) | Cell division protein<br>kinase 2, putative    | Fosforilación<br>/Actividad de<br>cinasa                                                     | 879/932   | 292  | 33.57  | 7.08 | N                                                                                                                                                                                                                                                                                           | C, Nu |
| 23 | EhCDK2 IV<br>(EHI_054090)  | Protein kinase<br>domain containing<br>protein | Fosforilación<br>/Actividad de<br>cinasa                                                     | 879/930   | 292  | 33.77  | 7.01 | N                                                                                                                                                                                                                                                                                           | Nu    |
| 24 | EhTOP II<br>(EHI_120640)   | DNA<br>topoisomerase II,<br>putative           | DNA<br>topoisomerasa<br>tipoll<br>(hidrólisis de<br>ATP/ corte de<br>DNA de doble<br>cadena) | 4050/4050 | 1349 | 155.44 | 6.64 | 312-<br>EKMKKLNKKGAEIK<br>-325<br>401- GKK -403<br>1040- PKKANKSKP<br>-1048<br>1133-<br>LKKGDKSTKRVIMGK<br>KGKSTKKTSRKVA<br>-1160<br>1195-<br>EIKKKTIKKVTTTRKRK<br>DDSSMSGDEKVKRK<br>RTTKKSTTKKTTKKRT<br>TKASKEEENDTVDTP<br>AKTSRKKKTKKVTVK<br>-1272<br>1322-<br>DPKPKRKTRSSKKEK<br>S -1337 | Nu    |

|         |    |                                       |                                             |                                                         |            |     |       |      |                                                    |    |
|---------|----|---------------------------------------|---------------------------------------------|---------------------------------------------------------|------------|-----|-------|------|----------------------------------------------------|----|
|         | 25 | EhTOP III A<br>(EHI_038920)           | DNA<br>topoisomerase III,<br>putative       | Desenrollamie<br>nto del DNA<br>/replicación            | 1824/1824  | 607 | 69.63 | 7.17 | N                                                  | Nu |
|         | 26 | EhTOP III B<br>(EHI_042880)           | DNA<br>topoisomerase III,<br>putative       | Desenrollamie<br>nto del DNA<br>/replicación            | 2523/2523  | 840 | 95.24 | 8.78 | 820- RGG -822<br>825-<br>SHSKNNSRKHSSRTG<br>K -840 | Nu |
|         | 27 | EhRPA 1 I<br>(EHI_062980)             | Replication factor A<br>protein 1, putative | Unión a ácidos<br>nucleicos                             | 1767/1767  | 588 | 66.69 | 5.78 | N                                                  | Nu |
|         | 28 | EhRPA 1 II<br>(EHI_127070)            | Hypothetical<br>protein                     | NI                                                      | 1353/1353  | 450 | 52.72 | 7.41 | N                                                  | Nu |
|         | 29 | EhRPA 1 III<br>(EHI_067930)           | Replication factor<br>A, putative           | Factor de<br>replicación A                              | 1710/1763  | 569 | 65.49 | 5.47 | N                                                  | Nu |
|         | 30 | EhRPA 2<br>(EHI_147000)               | Hypothetical<br>protein                     | Unión a ácidos<br>nucleicos                             | 450/512    | 149 | 16.74 | 8.39 | N                                                  | Nu |
|         | 31 | EhRPA 3<br>(EHI_187270)               | Hypothetical<br>protein                     | NI                                                      | 336/386    | 111 | 12.39 | 4.07 | N                                                  | Nu |
| Etapa S | 32 | EhDNA prim 1<br>S I<br>(EHI_103530)   | DNA primase small<br>subunit, putative      | Replicación de<br>DNA / síntesis<br>de primer de<br>RNA | 1143/1202  | 380 | 45.19 | 5.68 | 61- LKRKLKEKVP -70                                 | Nu |
|         | 33 | EhDNA prim 1<br>S II<br>(EHI_129630)  | DNA primase small<br>subunit, putative      | Replicación de<br>DNA / síntesis<br>de primer de<br>RNA | 570/691    | 189 | 22.48 | 5.21 | N                                                  | Nu |
|         | 34 | EhDNA prim 1<br>S III<br>(EHI_010140) | DNA primase small<br>subunit, putative      | Replicación de<br>DNA / síntesis<br>de primer de<br>RNA | 1119/1179  | 372 | 44.50 | 7.96 | N                                                  | Nu |
|         | 35 | EhDNA prim 2<br>L I<br>(EHI_127060)   | DNA primase large<br>subunit, putative      | Replicación de<br>DNA / síntesis<br>de primer de<br>RNA | 1374/ 1374 | 457 | 54.37 | 7.22 | N                                                  | Nu |

|         |    |                                |                                              |                                                |           |     |       |      |                                                                                                                    |    |
|---------|----|--------------------------------|----------------------------------------------|------------------------------------------------|-----------|-----|-------|------|--------------------------------------------------------------------------------------------------------------------|----|
| Etapa S | 36 | EhDNA prim 2 L II (EHI_147530) | DNA primase large subunit, putative          | Replicación de DNA / síntesis de primer de RNA | 1572/1572 | 523 | 61.15 | 6.16 | N                                                                                                                  | Nu |
|         | 37 | DNA prim 2 III (EHI_008590)    | Hypothetical protein                         | Replicación de DNA / síntesis de primer de RNA | 1323/1323 | 440 | 52.38 | 8.77 | N                                                                                                                  | Nu |
|         | 38 | RFC 1 I (EHI_158610)           | Activator 1 140 kDa subunit, putative        | Replicación de DNA                             | 2157/2157 | 718 | 80.99 | 8.92 | 14-<br>KKENKVQNNDKEVKK<br>PKMPSSKTSKKKINK<br>QIKSV -48<br>87-<br>NKQKTSKPIESKK -<br>100<br>681- KGSSKKTCK -<br>689 | Nu |
|         | 39 | RFC 1 II (EHI_175240)          | Activator 1 subunit, putative                | Replicación de DNA                             | 996/1050  | 331 | 38.43 | 8.32 | N                                                                                                                  | Nu |
|         | 40 | RFC 2 (EHI_039320)             | Hypothetical protein                         | Factor C de replicación de DNA                 | 474/ 474  | 157 | 18.31 | 8.53 | N                                                                                                                  | Nu |
|         | 41 | RFC 3 (EHI_081730)             | Activator 1 subunit, putative                | Replicación de DNA                             | 1038/1038 | 345 | 39.46 | 8.87 | N                                                                                                                  | Nu |
|         | 42 | RFC 4 I (EHI_167150)           | Replication factor C subunit 4, putative     | Factor C de replicación de DNA                 | 990/990   | 329 | 36.89 | 5.64 | N                                                                                                                  | Nu |
|         | 43 | RFC 4 II (EHI_086540)          | Replication factor C subunit 4, putative     | Factor C de replicación de DNA                 | 948/948   | 315 | 35.65 | 6.62 | N                                                                                                                  | Nu |
|         | 44 | RFC 5 (EHI_099850)             | Replication factor C family protein          | Replicación de DNA                             | 978/1030  | 325 | 35.99 | 6.35 | N                                                                                                                  | Nu |
|         | 45 | PCNA (EHI_128450)              | Proliferating cell nuclear antigen, putative | Regulación de la replicación de DNA            | 789/789   | 262 | 28.50 | 4.56 | N                                                                                                                  | Nu |

|         |    |                                    |                                                                     |                                                         |           |      |        |      |                                                                                           |    |
|---------|----|------------------------------------|---------------------------------------------------------------------|---------------------------------------------------------|-----------|------|--------|------|-------------------------------------------------------------------------------------------|----|
| Etapa T | 46 | DNA pol $\alpha$<br>(EHI_151520)   | DNA polymerase<br>alpha catalytic<br>subunit, putative              | Replicación de<br>DNA                                   | 3399/3399 | 1132 | 130.17 | 8.70 | 96 -<br>SKEKAPPIKMPKKF -<br>109<br>790 - RKKRYAA - 796<br>799 -<br>ILDGGKKKKELKG -<br>811 | Nu |
|         | 47 | DNA pol $\delta$<br>(EHI_006690)   | DNA polymerase<br>delta catalytic<br>subunit, putative              | Replicación de<br>DNA                                   | 3234/3234 | 1077 | 124.38 | 8.30 | N                                                                                         | Nu |
|         | 48 | DNA pol $\epsilon$<br>(EHI_141410) | DNA polymerase<br>epsilon subunit B<br>domain-containing<br>protein | Replicación de<br>DNA                                   | 1188/1188 | 395  | 44.64  | 5.68 | N                                                                                         | Nu |
|         | 49 | RNAsa H1<br>(EHI_042240)           | Ribonuclease,<br>putative                                           | Replicación de<br>DNA /<br>remoción de<br>primer de RNA | 567/567   | 188  | 21.97  | 8.89 | N                                                                                         | Nu |
|         | 50 | RNAsa H2 I<br>(EHI_084700)         | RNase H2 large<br>subunit, putative                                 | Replicación de<br>DNA /<br>remoción de<br>primer de RNA | 813/880   | 270  | 30.57  | 8.30 | N                                                                                         | Nu |
|         | 51 | RNAsa H2 II<br>(EHI_134360)        | Ribonuclease H2<br>subunit A, putative                              | Replicación de<br>DNA /<br>remoción de<br>primer de RNA | 1065/1132 | 354  | 40.33  | 5.13 | 217 -<br>GRKVMKKEKIKG -<br>228                                                            | Nu |
|         | 52 | DNALigI<br>(EHI_111060)            | DNA ligase,<br>putative                                             | Replicación de<br>DNA                                   | 2094/2094 | 685  | 78.16  | 6.78 | 37 - KK - 38<br>400 - K - 400<br>101 - RK - 102<br>125 - KKLMKRT - 131<br>320 -           | Nu |
|         | 53 | Fen 1<br>(EHI_099740)              | Flap nuclease,<br>putative                                          | Reparación de<br>DNA,<br>replicación                    | 1131/1131 | 376  | 42.58  | 8.90 | RVSKGIEKLKNVSKK<br>AQGRLDSEFFNVKKVP<br>LSKSEAASGVKRKKP<br>TTKAKESRKKK - 376               | Nu |

|    |                              |                         |                                              |           |     |       |      |                                                                                                                                                                                                                 |    |
|----|------------------------------|-------------------------|----------------------------------------------|-----------|-----|-------|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 54 | TRF-like I<br>(EHI_001090)   | Hypothetical<br>protein | Unión a DNA<br>telomérico de<br>doble cadena | 1215/1215 | 404 | 46.79 | 5.84 | 276 -<br>KVNKKKNEKVNKPIR<br>QPVDENGQVIKKKKR<br>RF - 307<br>326 - KEEKSKKPSLKL<br>- 338<br>351 -<br>ENNKLKKKKRRF -<br>362<br>6 - KK - 7<br>339 -<br>PKIKQEKEARNIKKKA<br>LNSKSQDATPKMED<br>DEHSLKKKKRRFS -<br>383 | Nu |
| 55 | TRF-like II<br>(EHI_001110)  | Hypothetical<br>protein | Unión a DNA<br>telomérico de<br>doble cadena | 1326/1326 | 441 | 51.48 | 6.68 |                                                                                                                                                                                                                 | Nu |
| 56 | TRF-like III<br>(EHI_148140) | Hypothetical<br>protein | Unión a DNA<br>telomérico de<br>doble cadena | 1383/1383 | 460 | 53.10 | 8.39 |                                                                                                                                                                                                                 | Nu |

<sup>a</sup> AmoebaDB

<sup>b</sup> UniProt

<sup>c</sup> <https://www.moseslab.csb.utoronto.ca/NLStradamus/>

NI, no identificada

N, No

C, Citoplasma; Nu, nucleo

## Anexo 2

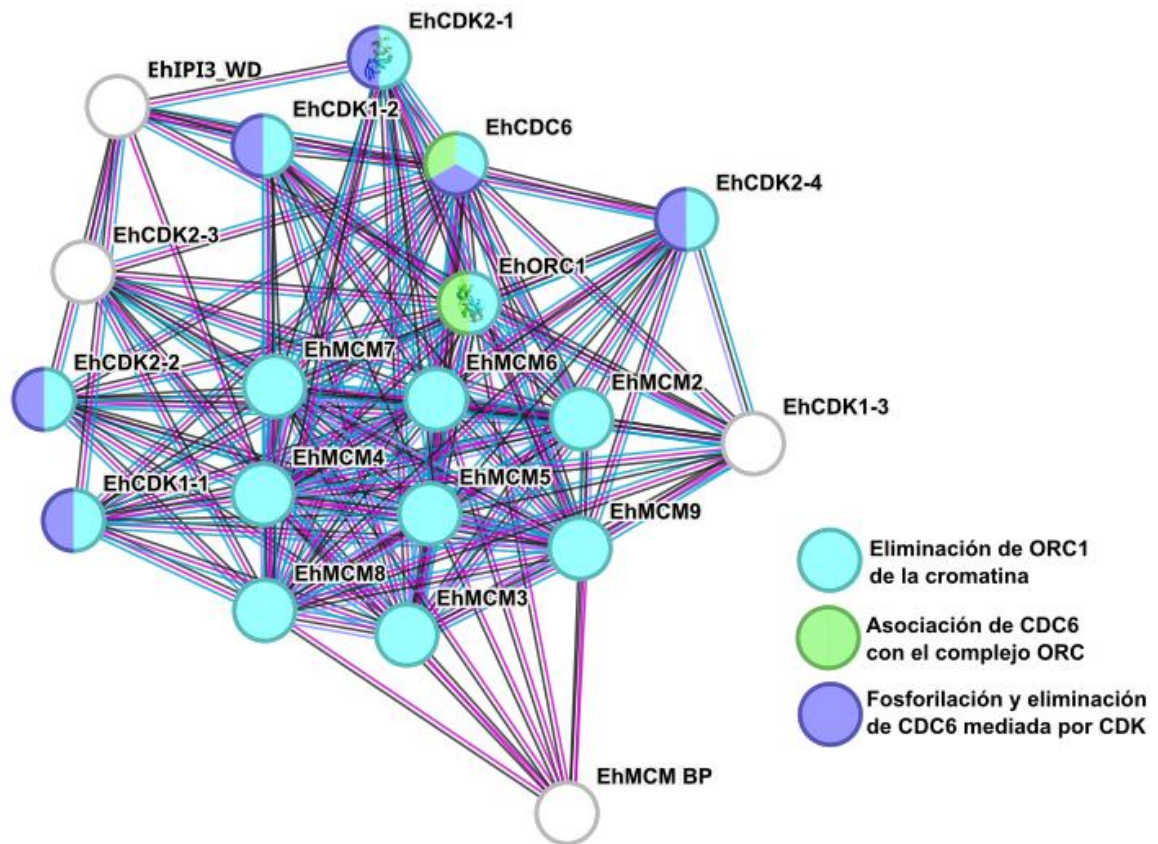


**Figura 1a. Interacciones funcionales predichas entre proteínas del complejo Pre-RC en *E. histolytica*.** En el análisis se incluyó a la etapa Pre-RC y las CDK para confirmar que corresponden a proteínas independientes de EhORC1 y EhCDC6, en contraste con el reporte previo que las considera una única proteína dual EhORC1/CDC6 (Pal y Dam, 2023). Los nodos resaltados representan procesos funcionales específicos definidos por rutas Reactome: en verde claro, asociación de EhCDC6 con el complejo ORC; en celeste, la eliminación de EhORC1 de la cromatina; y en morado, fosforilación y eliminación de EhCDC6 mediada por CDK. Las conexiones de interacción sugieren que EhORC1 y EhCDC6 son proteínas funcionalmente distintas, con roles complementarios en la regulación de la pre-replicación en *E. histolytica* y solo 5 CDK fosforilan y eliminan a EhORC1 y EhCDC6.

### Anexo 3

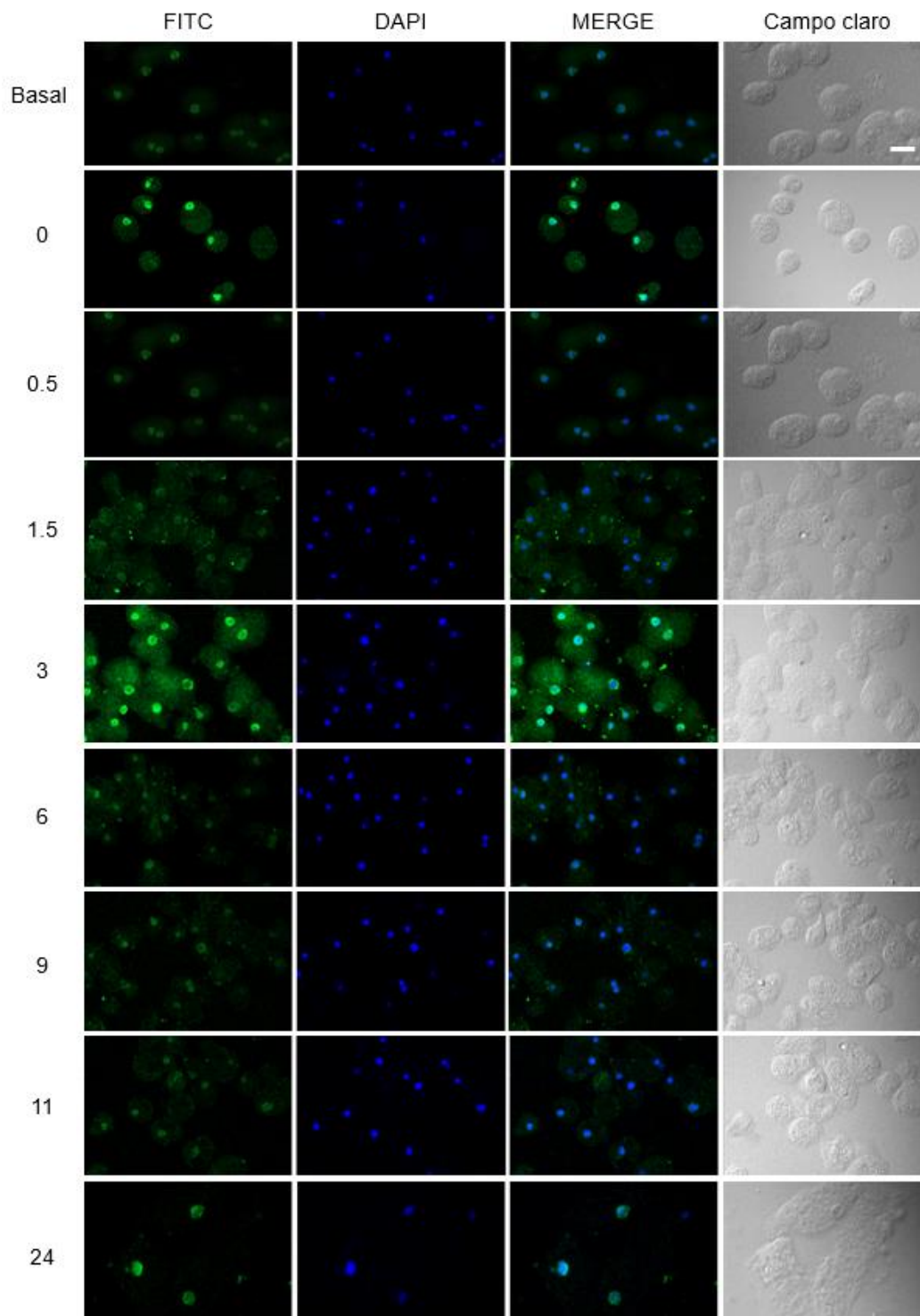


**Figura 2a. Detección de EhPCNA mediante inmunofluorescencia indirecta en trofozoítos de *E. histolytica* sincronizados.** Trofozoítos de *E. histolytica* incubados sin suero durante 12 horas, seguida de reposición de suero, se fijaron y procesaron a diferentes tiempos. Las muestras se incubaron con anti-EhPCNA (1:250) y se visualizaron por medio de microscopía confocal. El núcleo se visualizó con DAPI y se muestra el campo claro. La barra indica 20  $\mu$ m.

## Anexo 4

Este trabajo fue presentado en los siguientes eventos científicos:

- XXXIII Congreso Nacional de Parasitología

### Biología Molecular



#### Las características únicas de la maquinaria de replicación de *Entamoeba histolytica* PBM-15

Rodríguez-Almaraz Alan Jesé<sup>1</sup>, Cuellar-Silva Patricia<sup>2</sup>, Canela-Pérez Israel<sup>3</sup>, Valdés Jesús<sup>3</sup>, Martínez Rodríguez Carlos<sup>4</sup>, Azuara-Liceaga Elisa<sup>1,2</sup>

<sup>1</sup>Licenciatura en Ciencias Genómicas, Universidad Autónoma de la Ciudad de México, UACM, Plantel del Valle, CDMX, México.

<sup>2</sup>Posgrado en Ciencias Genómicas, Universidad Autónoma de la Ciudad de México, UACM, Plantel del Valle, CDMX, México.

<sup>3</sup>Departamento de Bioquímica, CINVESTAV-IPN, Zacatenco, CDMX, México.

<sup>4</sup>Academia de Matemáticas, Universidad Autónoma de la Ciudad de México, UACM, Plantel del Valle, CDMX, México.

*Entamoeba histolytica* es responsable de la amibiasis en humanos, que afecta hasta el 3.5% de la población mundial. Ha sido considerado un eucariota primitivo debido a la ausencia de mitocondria y por su retículo endoplásmico y aparato de Golgi atípicos. En eucariotas, la replicación del DNA inicia con el reconocimiento de la secuencia Ori por el complejo ORC (6 proteínas) que posteriormente recluta a las helicasas MCM. Otras proteínas involucradas en el proceso de replicación en eucariotas son las DNA polimerasas, el complejo GINS, proteínas CDK, PCNA, DNA ligasas, entre otras. Aunque la maquinaria de replicación está bien caracterizada en *Homo sapiens*, se sabe poco de su conformación en *E. histolytica*, debido a ello decidimos identificar in silico las proteínas que conforman su maquinaria de replicación. Para la identificación de los genes que codifican para proteínas involucradas en la replicación del DNA de *E. histolytica* se realizaron búsquedas en la base de datos AmoebaDB, utilizando como sondas las secuencias de aminoácidos de las proteínas de *H. sapiens* y

*Schizosaccharomyces pombe*. Los resultados obtenidos mostraron que en el genoma de *E. histolytica* se encuentran codificadas las proteínas necesarias para realizar la replicación de sus cromosomas; sin embargo, existen diferencias en la composición, la cual parece ser más simplificada. Además, se analizaron los dominios funcionales y la estructura tridimensional de las proteínas identificadas. Finalmente, empleando cultivos de trofozoitos privados de suero, analizamos los cambios de expresión de algunos de los genes identificados que participan en las diferentes etapas de la replicación durante diferentes intervalos de tiempo. Este trabajo es uno de los primeros pasos para la comprensión del proceso de replicación del DNA en *E. histolytica*.

CONAHCYT (CF-2019-194163)

E-mail: alan.rodriguez.almaraz@estudiante.uacm.edu.mx

E-mail: elisa.azuara@uacm.edu.mx

128





**Sociedad Mexicana de  
Parasitología, A.C.**



Universidad Veracruzana

Otorga la presente **CONSTANCIA A:**

**Rodríguez-Almaraz Alan Jesé**

**CO-AUTORES**

**Cuellar-Silva Patricia, Canela-Pérez Israel,  
Valdés Jesús, Martínez Rodríguez Carlos,  
Azuara-Liceaga Elisa**

**Por la presentación del Trabajo en  
la modalidad Poster**

**Las características únicas de la maquinaria  
de replicación de *Entamoeba histolytica***

**Dentro del marco XXIII Congreso Nacional de Parasitología:  
CONAPAR-2024 "Una sola salud".  
Del 28-30 de agosto de 2024.  
Orizaba, Ver., México**

  
**Dra. Gloria León-Ávila**  
Comité Académico SMP

  
**Dra. Aracely López Monteon**  
Comité Científico Local

  
**Dr. Angel Ramos Ligonio**  
Presidente SMP

## B. Conferencia BioPhys Mexico City 2025





# Characterization of the replisome of *Entamoeba histolytica*

**UACM**  
Universidad Autónoma  
de la Ciudad de México  
MEXICO - DF 06702 ADO

Rodríguez-Almaraz Alan Jesé<sup>1</sup>, Cuellar-Silva Patricia<sup>2</sup>, Canela-Pérez Israel<sup>3</sup>, Valdés Jesús<sup>4</sup>, Castañeda-Ortiz Elizabeth J.<sup>2</sup>, Betanzos Abigail<sup>2</sup>, Azuara-Liceaga Elisa<sup>1,2</sup>

E-mail: alan.rodriguez.almaraz@estudiante.uacm.edu.mx

<sup>1</sup> Licenciatura en Ciencias Genómicas, <sup>2</sup> Posgrado en Ciencias Genómicas, Universidad Autónoma de la Ciudad de México, UACM, Plantel del Valle, <sup>3</sup> Departamento de Bioquímica, CINVESTAV, <sup>4</sup> Departamento de Infectómica y Patogénesis Molecular, CINVESTAV.

## 1. INTRODUCTION

DNA replication is an essential process in eukaryotic organisms that must ensure the precise genomic duplication during cell division. This process is divided into four stages: pre-replication (Pre-RC), pre-initiation (Pre-IC), synthesis (S) and termination (T). The replication machinery of *Homo sapiens* has been well described, but its composition and regulation in *Entamoeba histolytica*, the protozoan that causes human amebiasis, has been scarcely studied. *E. histolytica* is a pathogenic eukaryote that lacks mitochondria and has an atypical endoplasmic reticulum. Studying the replication process of *E. histolytica* will allow us to understand how its genome is replicated and also determine important differences regarding its host, which could lead to the identification of therapeutic targets as an alternative to current treatment against amebiasis. Therefore, the aim of this work is to identify the molecular components of the replisome of *E. histolytica* to contribute to the understanding of its DNA replication.

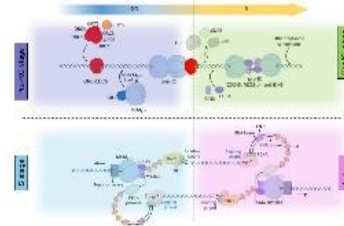


Figure 1. DNA replication stages in eukaryotes

## 2. METHODOLOGY



## 3. RESULTS

### 3.1 Essential domains and 3D-structure of the main proteins involved in DNA replication in *E. histolytica*

|              | PROTEIN                        | DOMAINS | STRUCTURE |
|--------------|--------------------------------|---------|-----------|
| Pre-RC stage | EhORC1                         |         |           |
|              | EhCDC6                         |         |           |
|              | EhMCM2, 3, 4, 5, 6, 7, 8 and 9 |         |           |
|              | EhCDC45                        |         |           |
| Pre-IC stage | EhGINS, PSF1, PSF2, and PSF3   |         |           |
|              | EhCDK1 and 2 (7 isoforms)      |         |           |
|              | EhTOPBII and III (A and B)     |         |           |
|              | EhRPA1 (3 isoforms)            |         |           |
| S stage      | EhDNA prim 1S (3 isoforms)     |         |           |
|              | EhRFC1 (2 isoforms)            |         |           |
|              | EhPCNA                         |         |           |
|              | EhDNApol α, β, γ               |         |           |
| T stage      | EhRNAse H1, H2 (2 isoforms)    |         |           |
|              | EhDNALig I                     |         |           |

## 4. CONCLUSIONS

*E. histolytica* has a more simplified replication machinery than other eukaryotes, which allows us to identify key differences with its host that could be exploited for the development of new therapeutic targets.

## 6. ACKNOWLEDGMENTS

This work was supported by UACM-CCYT-2025-CON-05 and PC is grateful to SICIHTI Estancias Posdoctorales por México (CVU 333090). The authors thank the Universidad Autónoma de la Ciudad de México (UACM) for its support.

## 5. REFERENCES

- Burgess, H., & Kunkel, L. A. (2017). Eukaryotic DNA Replication Fork. Annual review of biochemistry, 86, 477-498.
- da Silva, M. S. et al. (2017). Nuclear DNA Replication in Trypanosomids: There Are No Easy Methods for Solving Difficult Problems. Trends in parasitology, 33(11), 858-874.

**C. Constancias**



El Posgrado en Ciencias Genómicas  
otorga la presente

Constancia

a

**Rodríguez Amaraz Alan Jesé**

por su asistencia al

**1er. Encuentro de Egresados del  
Programa de Maestría en Ciencias Genómicas**

que se llevó a cabo el 28 de Noviembre de 2019  
en la UACM Plantel Del Valle.

Universidad Autónoma de la Ciudad México

*Nada humano me es ajeno*



28 NOV 2019



  
Dra. Martha Yocupicio Monroy

  
Dra. Selene Zárate Guerra

Coordinadoras del 1er. Encuentro de Egresados del  
Programa de Maestría en Ciencias Genómicas



**UACM**  
Universidad Autónoma  
de la Ciudad de México  
NADA HUMANO HE ES AJENO

**POSGRADO  
EN CIENCIAS  
GENÓMICAS**

LA UNIVERSIDAD AUTÓNOMA DE LA CIUDAD DE MÉXICO  
OTORGA LA PRESENTE

## CONSTANCIA

A: **Rodríguez Almaraz Alan Jese**

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 ÁLVAREZ SÁNCHEZ  
COORDINADORA DE LAS  
JORNADAS DE DIVULGACIÓN DE LA CIENCIA

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

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

**UACM**  
Universidad Autónoma  
de la Ciudad de México  
NADA HUMANO HE ES AJENO

LA UNIVERSIDAD AUTÓNOMA DE LA CIUDAD DE MÉXICO

OTORGA LA PRESENTE

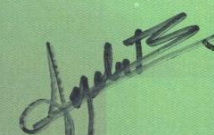
## CONSTANCIA

A: **ALAN JESE RODRÍGUEZ ALMARAZ**

POR SU PARTICIPACIÓN EN LAS  
**SEGUNDAS 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

9 DE MAYO DEL 2024

  
DRA. ANGELES C. TECALCO CRUZ  
COORDINADORAS DE LAS SEGUNDAS JORNADAS DE  
DIVULGACIÓN DE LA CIENCIA DE LA LICENCIATURA EN CIENCIAS GENÓMICAS

  
DRA. HELENA SOLLEIRO VILLAVICENCIO  
COORDINADORAS DE LAS SEGUNDAS JORNADAS DE  
DIVULGACIÓN DE LA CIENCIA DE LA LICENCIATURA EN CIENCIAS GENÓMICAS

otorgan la presente

# CONSTANCIA

a: **RODRÍGUEZ ALMARAZ ALAN JESÉ**

Por su participación en el Simposio Internacional  
Genómica: de lo molecular a lo traslacional.

24 y 25 de Septiembre del 2024



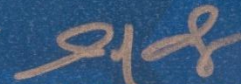
Dra. Elizabeth Álvarez Sánchez  
Coordinadora del Simposio UACM



Mtro. César Enrique Fuentes Hernández  
Coordinador Académico UACM



Dr. Armando Raúl Esparza Ávila  
Subdirector de Enseñanza e Investigación



Dra. Sofía Lizeth Alcaraz Estrada  
Jefa de División de Medicina Genómica



**UACM**

Universidad Autónoma  
de la Ciudad de México  
NADA HUMANO ME ES AJENO



La Universidad Autónoma de la Ciudad de México  
y la Sociedad Mexicana de Virología  
otorgan la presente

**Constancia** a

**Alan Jesé Rodríguez Almaraz**

por su asistencia a la

**Jornada Académica**

**Virus, Antivirales y Vacunas de  
Importancia para la Salud Pública**

realizada en el plantel Del Valle  
el 9 de noviembre de 2024.

*Nada humano me es ajeno*

**Dra. María Elizabeth  
Álvarez Sánchez**  
Coordinadora  
Académica-UACM

**Mtra. Erika Lorena  
Álvarez Ramírez**  
Coordinadora del Colegio  
de Ciencia y Tecnología-UACM

**Dra. Martha  
Yocupicio Monroy**  
Coordinadora de la  
Jornada Académica

**Dra. Sofía Lizeth  
Alcaraz Estrada**  
Sociedad Mexicana  
de Virología

VEN Y CUELGA TU POSTER



La Universidad Autónoma de la Ciudad de México  
y el El Posgrado en Ciencias Genómicas

otorgan la presente

# CONSTANCIA

a:

**Alan Jesé Rodríguez Almaraz**

Por su participación en el evento "Ven y cuelga tu poster"  
con el trabajo:

**Las características únicas de la maquinaria  
de replicación de *Entamoeba histolytica***

Celebrado el 21 de noviembre de 2024 en el patio central del  
Posgrado en Ciencias Genómicas de la UACM, plantel del valle.

  
Dra. María Elizabeth Álvarez Sánchez  
Coordinadora Académica  
UACM

  
Mtra. Erika Lorena Álvarez Ramírez  
Coordinadora del  
Colegio de Ciencia y Tecnología  
UACM

  
Dra. Eliza Azuara Lieceaga  
Profesora Investigadora del PCG  
Coordinadora del PMCG, UACM

**UACM**

Universidad Autónoma  
de la Ciudad de México

NADA HUMANO ME ES AJENO



**POSGRADO  
EN CIENCIAS  
GENÓMICAS**

El Posgrado en Ciencias Genómicas otorga el siguiente reconocimiento a:

**ALAN JESÉ RODRÍGUEZ ALMARAZ**

Por su participación con el trabajo

**Detective parásito**

En la modalidad: "Conociendo a las Ciencias Ómicas con una Dinámica"

en la Tercera Jornada de Divulgación de la Licenciatura  
en Ciencias Genómicas de la  
Universidad Autónoma de la Ciudad de México

Ciudad de México, a 13 de mayo de 2025

**Nada Humano me es ajeno**



**Dra. Mireya Camacho Nuez**

Coordinadora de la Tercera Jornada de Divulgación

**Dra. Martha Rocapinto Monroy**

Coordinadora de la Tercera Jornada de Divulgación

**UACM**

Universidad Autónoma  
de la Ciudad de México

NADA HUMANO ME ES AJENO



**POSGRADO  
EN CIENCIAS  
GENÓMICAS**

El Posgrado en Ciencias Genómicas otorga el

**TERCER LUGAR**

a:

**Luna Ibarra Dulce Maria Concepción  
Rodríguez Almaraz Alan Jesé  
Herrera Téllez Karen Elizabeth**

por el trabajo titulado:

**“Detective parásito”**

en la modalidad de

***Dinámica de acercamiento al conocimiento y aplicaciones genómicas***

en la Tercera Jornada de Divulgación de la Licenciatura

en Ciencias Genómicas de la

Universidad Autónoma de la Ciudad de México

Ciudad de México, a 13 de mayo de 2025

***Nada Humano me es ajeno***



**Dra. Minerva Camacho Nuez**

Coordinadora de la Tercera Jornada de Divulgación

**Dra. Martha Yocupicio Monroy**

Coordinadora de la Tercera Jornada de Divulgación