



**CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS  
AVANZADOS DEL INSTITUTO POLITÉCNICO NACIONAL**

**UNIDAD ZACATENCO**

**DEPARTAMENTO DE BIOQUÍMICA**

**“ Participación de RNAs intrónicos circulares de longitud completa (flicRNAs) en la regulación de genes del parásito protozoario *Entamoeba histolytica*. ”**

**TESIS**

Que presenta

**Jesús Alberto García Lerena**

Para obtener el grado de

**Maestro en Ciencias**

En la especialidad de

**Bioquímica**

Director de la Tesis

**Dr. Jesús Valdés Flores**

**ESTE TRABAJO SE REALIZÓ EN EL DEPARTAMENTO DE BIOQUÍMICA DEL CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL INSTITUTO POLITÉCNICO NACIONAL, BAJO LA DIRECCIÓN DEL Dr. JESÚS VALDÉS FLORES.**

**DURANTE LA REALIZACIÓN DE ESTE TRABAJO SE CONTÓ CON LA BECA SEP-CINVESTAV 601410 FIDSC2018/56 y Beca CONACYT 485522, BECARIO 633127, POR LO CUAL SE AGRADECE A LAS INSTITUCIONES.**

## **DEDICATORIA**

*A mis padres y hermanos, por ser mi sostén y mi referencia.*

## **AGRADECIMIENTOS**

A:

Dr. Jesús Valdés Flores por las enseñanzas, apoyo y aceptación en su grupo de trabajo que me dio la oportunidad de obtener una excelente preparación profesional.

Dr. Edgar Morales Ríos y Dr. Nicolás Villegas Sepúlveda por asesorarme en este trabajo de tesis y la Dr. Elisa Azuara Liceaga por su participación en el desarrollo experimental.

José Manuel Galindo, por su participación y total disposición en la realización de este trabajo. A Tea y Martín por todo el apoyo brindado. A Gretter por su amistad y apoyo en lo profesional y personal. A mis compañeros de generación Vicente, Nicole, César y Jessica por siempre habernos mantenido unidos. A las secretarias Priscila y Angélica por su apoyo y asesoramiento en todos los trámites y su excelente trabajo.

A los profesores del Departamento de Bioquímica por los cursos impartidos durante la maestría.

A mi toda mi familia, lo más importante, por brindarme todo su amor y comprensión desde la distancia. A Reinier y Cynthia, Francisco Penney, Jenis, Ariday, Renier Garnier, Guillermo, Luis Miguel y Yunior por siempre estar presente y ser tan importantes en mi vida. A mis amigos de Cuba por no dejar morir la amistad después de tantos años. Y a todas las maravillosas personas que he conocido en esta aventura y que han ganado mi corazón Erika, Sergio, Pepe, Alonso, Gabo, Lisset, Zule, Diego, Dayana, Esly y tantos buenos corazones más.

Gracias.

# Índice

<b>Lista de Abreviaturas</b> .....	vi
<b>Resumen</b> .....	ix
<b>Abstract</b> .....	x
<b>I. Introducción</b> .....	1
I.1 Estructura y Función del RNA.....	1
I.2 Procesamiento del mRNA .....	2
I.3 Mecanismo de <i>splicing</i> constitutivo ( <i>forward</i> ) y elementos de regulación en <i>cis</i> . .....	5
I.4 Transcripción y <i>splicing</i> .....	7
I.5 Una enzima dinámica: ensamblaje y actividad del spliceosoma .....	8
I.6 Intrón <i>lariat</i> , eventos post- <i>splicing</i> .....	10
I.7 RNA no codificantes .....	11
I.8 RNAs circulares.....	13
I.9 Biogénesis de los circRNAs .....	14
I.10 Funciones de los circRNAs.....	18
I.10.1 Función nuclear de los circRNAs .....	19
I.10.2 Función citoplasmática de los circRNAs.....	21
I.11 <i>Entamoeba histolytica</i> .....	23
I.11.1 Splicing en <i>Entamoeba histolytica</i> .....	26
<b>II. Antecedentes</b> .....	27
<b>III. Justificación</b> .....	32
<b>IV. Hipótesis</b> .....	33
<b>V. Objetivos</b> .....	34
Objetivo general.....	34
Objetivos específicos .....	34
<b>VI. Estrategia experimental</b> .....	35
<b>VII. Materiales y Métodos</b> .....	36

VII.1 Restricción, construcción de plásmidos y transformación .....	36
VII.2 Transfección de plásmidos en trofozoítos de <i>E. histolytica</i> cepa HM1: IMSS .....	38
VII.3 Cultivos de trofozoitos de <i>Entamoeba histolytica</i> y tratamientos .....	39
VII.4 Extracción de RNA.....	39
VII.5 RT-PCR .....	41
VII.6 PCR en tiempo real (qPCR) .....	42
VII.7 <i>Cross-linking</i> inmunoprecipitación anti-HA (CLIP).....	42
VII.8 Estandarización de ensayo de Inmunoprecipitación de cromatina (ChIP) ....	44
VII.9 Análisis estadístico .....	45
<b>VIII. Resultados</b> .....	47
VIII.1 Transfección en cepas HM1: IMSS.....	47
VIII.2 Identificación de flicRNAs en cepas transfectantes.....	49
VIII.3 <i>Cross-linking</i> Inmunoprecipitación anti-HA (CLIP) en cepas pHA-CTD.....	50
VIII.3.1 Eficiencia del ensayo de inmunoprecipitación.....	51
VIII.3.2 CLIP .....	52
VIII.3.3 Localización de los snRNAs U2 y U6 .....	56
VIII.3.4 Validación de CLIP, tratamiento con RNasa R .....	56
VIII.4 Tratamiento con ácido bórico en cepas transfectantes .....	58
VIII.5 Estandarización de ensayo de Inmunoprecipitación de cromatina (ChIP) en cepas HM1-IMSS .....	61
<b>IX. Discusión</b> .....	64
<b>X. Conclusiones</b> .....	72
<b>XI. Perspectivas</b> .....	73
<b>XII. Bibliografía</b> .....	74
<b>XIII. Anexos</b> .....	84

## Lista de Abreviaturas

3'ss: Sitio de *splicing* 3', del inglés "3' Splice Site"

5'ss: Sitio de *splicing* 5', del inglés "5' Splice Site"

ActD: Actinomicina

AMO: Morfolino-oligonucleótidos antisentido

As: antisentido

ASO: oligonucleótidos antisentido específicos

asRNA: RNA antisentido

BP: Punto de ramificación, del inglés "*Branch Point*"

cDNA: DNA complementario, del inglés "*complementary DNA*"

ChIP: Inmunoprecipitación de cromatina

circRNA: RNA circular

ciRNA: RNA circular intrónico

CLIP: Cross-linking Inmunoprecipitación

CPSF: Factor de Poliadenilación y corte

cRNA: RNA codificante

CstF: Factor de estimulación de la transcripción

CTD: C-terminal de la RNA Polimerasa II

Dbr1: Enzima desramificadora (por sus siglas en inglés: *Debranching*)

Dbr1  $\Delta$ C: Dbr1 catalíticamente deficiente

dNTPs: deoxinucleótidos trifosfatos

ElciRNA: RNA circular exón-intrónico

EMT: Transición epitelial mesenquimal

ESE: Potenciador exónico

flicRNA: RNA intrónico circular de longitud completa

FUS: Proteína de fusión del sarcoma

guanina-N7 MTasa: Guanina-N7 metiltransferasa

GTasa: RNA guanililtransferasa

Gal-Lectin: Galactosa/N-acetil-D-galactosamina

hnRNP: Ribonucleoproteína heterogeneonuclear, del inglés "*heterogeneous RiboNucleoParticle*"

IECs: Células epiteliales intestinales

IL-1 $\beta$ : Interleucina-1

iRNA: RNA de interferencia

IRES: Sitio interno de entrada al ribosoma

ISE: Potenciador Intrónico.

ITS1: Espaciador interno del RNA ribosomal

kb: kilobases

lcrRNA: RNAs largos no codificantes

miRNA: micro RNA

MPEs: Calcineurinas tipo metalofosfodiesterasa

mRNA: RNA mensajero, del inglés “*messenger RNA*”

MUC2: Mucina

ncRNA: RNA no codificante, del inglés “*non coding RNA*”

ncRNA: RNA no codificante

NOS: Especies reactivas del nitrógeno

NSCLC: Células no pequeñas de cáncer de pulmón

nt: nucleótido

ORF: Marco abierto de lectura

pb: Pares de bases

PBS: Solución buffer fosfato

PCNA: Antígeno nuclear de proliferación celular

PCR: Reacción en cadena de la polimerasa, del inglés “*Polymerase Chain Reaction*”

pEhExHA y pHA-V: Plásmido de expresión en *E. histolytica* con etiqueta de hemaglutinina

pHA- ΔGU: Construcción que contiene el minigen mutado en la secuencia 5´ss del intrón del gen RabX13

Pol II: RNA polimerasa II

Poli A: Poli adeninas

pre-mRNA: precursor del RNA mensajero, del inglés “*precursor messenger RNA*”

pre-rRNA: precursor de RNA ribosomal, del inglés “*precursor ribosomal RNA*”

QKI: Proteínas *Quaking* de unión al RNA

qPCR: Reacción en cadena de la polimerasa cuantitativa en tiempo real inglés “*Polymerase Chain Reaction*”

ROS: Especies reactivas del oxígeno

Rpm: Revoluciones por minuto

rRNA: RNA ribosomal, del inglés “*ribosomal RNA*”

RT: Retro transcripción

S: Sentido



SBA: Suero Bovino Adulto

SHPRH: Helicasa del anillo PHD de la histona SNF2

siRNA: RNA pequeños interferentes

sisRNA: secuencias de RNA intrónicas estables

snoRNA: RNA nucleolares pequeños

snRNA U1: RNA pequeño nuclear U1, del inglés "*small nuclear RNA U1*"

snRNA U2: RNA pequeño nuclear U2, del inglés "*small nuclear RNA U2*"

snRNA U4: RNA pequeño nuclear U4, del inglés "*small nuclear RNA U4*"

snRNA U5: RNA pequeño nuclear U5, del inglés "*small nuclear RNA U5*"

snRNA U6: RNA pequeño nuclear U6, del inglés "*small nuclear RNA U6*"

snRNA: RNA pequeño nuclear, del inglés "*small nuclear RNA*"

snRNP U1: Ribonucleoproteína pequeña nuclear U1, del inglés "*small nuclear RiboNucleoProtein U1*"

snRNP U1C: Componente de la ribonucleoproteína pequeña nuclear U1, del inglés "*small nuclear RiboNucleoProtein U1*"

snRNP: Ribonucleoproteína pequeña nuclear, del inglés "*small nuclear RiboNucleoProtein*"

TA: Temperatura Ambiente

TNF- $\alpha$ : Factor de necrosis tumoral alfa

TPasa: RNA trifosfatasa

tRNA: RNA de transferencia

U2AF: Factor auxiliar de splicing de U2, del inglés "*U2 Auxiliary Factor*"

U2AF35: Subunidad pequeña del factor auxiliar de U2 de splicing (humanos)

U2AF65: Subunidad grande del factor auxiliar de U2 de splicing (humanos)

wt: Cepa silvestre

## Resumen

Los RNA no codificantes circulares (circRNAs) son un grupo de transcritos ubicuos en los sistemas biológicos e implicados en la regulación génica. Recientemente en nuestro laboratorio se hallaron por primera vez RNA intrónicos circulares de longitud completa en *Entamoeba histolytica* (flicRNA). Los estudios realizados permitieron determinar que los flicRNA están circularizados por esterificación de sus sitios de *splicing* (ss) 3' y 5' con una secuencia rica en GU al 5' ss esencial para su biogénesis. Mutaciones de este elemento 5'-GU del intron RabX13, causaron una disminución en la producción de flicRX13 que condujo a un aumento en la expresión del transcrito parental Rabx13, implicándolo en la modulación de la expresión de su gen parental. Para ligar aún más los flicRNAs con la regulación transcripcional de sus genes parentales, sobrexpresamos el dominio carboxilo terminal de la RNA Polimerasa II (CTD) fusionado con el antígeno de hemaglutinina en trofozoitos de la cepa HM1-IMSS, y luego de realizar ensayos de cross-linking inmunoprecipitación (CLIP) seguido de purificación de RNA, identificamos mediante RT-PCR circular los flicRNAs correspondientes al gen control RabX13 y a los loci asociados a virulencia EHI\_192510, EHI\_014170 y EHI\_169670 unidos al Pol II-CTD e implicados de este modo en la regulación en *cis* de sus genes parentales. Adicionalmente fue identificado a U2 snRNA como parte del complejo regulador y se descartó la presencia de U6 snRNA. Los flicRNAs correspondientes a los genes de proteínas ribosomales RpS14 y RpL12 fueron localizados en el citosol y proponemos que participan en eventos de regulación post-transcripcional. Con este trabajo se demuestra la presencia de los flicRNAs como elementos del complejo de unión a la RNA Polimerasa II en *E. histolytica*.

**Palabras Clave:** *ncRNA, RNA circular, flicRNA, splicing, regulación génica.*

## Abstract

Non-coding circular RNAs (circRNAs) are a group of RNA transcripts found ubiquitously in eukaryotic organisms and involved in numerous co- and post-transcriptional regulatory systems. Recently, we found full-length intronic circular RNAs (flicRNAs) in *Entamoeba histolytica*. Our studies identified a distinctive 3'ss-5'ss ligation point, and the conserved 5'ss GU-rich elements is essential for its biogenesis. Mutations in the GU-rich 5'ss element of the RabX13 gene test flicRNA (flicRX13) decreased production of flicRX13 which led to an increase in the expression of the Rabx13 parental gene, suggesting that flicRX13 down-regulates RabX13 gene expression. To further link additional flicRNAs with gene expression regulation, here we overexpressed an HA-tagged version of the carboxyl terminal domain of RNA Polymerase II (CTD) in *E. histolytica* trophozoites of strain HM1-IMSS, and we developed cross-linking immunoprecipitation (CLIP) assays followed by RNA purification and flicRNAs identification by RT-PCR, using probes towards the 5'ss-3'ss ligation points. Our results show that the flicRNAs from the positive control RabX13 gene and virulence-associated loci EHI\_192510, EHI\_014170, and EHI\_169670 were bound to the Pol II-CTD suggesting their involvement in the transcriptional regulation of their parental genes in *cis*. Furthermore, U2 snRNA was identified as part of the regulatory complex and the presence of U6 snRNA was ruled out. The flicRNAs from the gene encoding for the ribosomal proteins EpS14 and RpL12 were localized in cytosol and we propose that they are involved in post-transcriptional regulatory events. This work demonstrates the presence of flicRNAs as elements of the Polymerase II binding complex in *E. histolytica*.

Keywords: ncRNA, circular RNA, flicRNA, *splicing*, gene regulation.

# I. Introducción

## I.1 Estructura y función del RNA

El ácido ribonucleico (RNA) está constituido por nucleótidos de ribosa unidos por enlaces fosfodiéster formando un biopolímero de cadena simple que puede variar en longitud. Esta macromolécula es de amplia abundancia en todas sus formas, constituyendo el 1% del peso de las células de mamíferos (Wu *et al.*, 2014). De acuerdo con el dogma central de la biología molecular, descrito por Francis Crick en 1970; la transcripción de la cadena monocatenaria de RNA constituye el primer paso para la expresión génica que culmina con la síntesis proteica. Sin embargo, debido a sus características moleculares, adicionalmente cumple roles esenciales en la regulación del genoma (Morris & Mattick, 2014).

La presencia de un grupo hidroxilo químicamente reactivo en el segundo carbono de los monómeros de ribosa en el RNA, le brinda capacidad de hidrólisis a diferencia del polímero de DNA, el cual contiene monómeros de desoxirribosa y con ello mayor estabilidad (Holley *et al.*, 1965). La complementariedad de secuencias intracatenarias de esta macromolécula de RNA le permite adquirir una estructura tridimensional esencial para su estabilidad, funciones y asociación con proteínas (Miao & Westhof, 2017).

Los RNAs más estudiados son aquellos asociados a la síntesis proteica entre los que encontramos: los RNA ribosomales (rRNA) que forman complejos ribonucleoproteicos y están implicados en la traducción del RNA mensajero (mRNA), los RNA de transferencia (tRNA) encargados de transportar los aminoácidos específicos a cada codón para la síntesis de proteínas y el propio mRNA que constituye la molécula en la cual se transcribe la información contenida en el genoma de la célula (Lewin, 2008).

Por otro lado, tenemos a los RNA reguladores los cuales se clasifican por su tamaño en RNAs largos no codificantes (lncRNA) y RNA antisentido (asRNA) (>200 nt) implicados en la regulación de la transcripción; los RNA de interferencia (iRNA) que a su vez se subdividen en micro RNAs (miRNA) (21-22 nt), RNA nucleolares pequeños (snoRNA) (60-300 nt) involucrados en la modificación del rRNA, RNA pequeños nucleares (snRNA) (150 nt) implicados en el *splicing* y RNA pequeños interferentes (siRNA) (21-25 nt) que regulan la expresión del mRNA (Mattick & Makunin, 2006; Wilusz *et al.*, 2009). Además, se han sido descubierto moléculas de RNA circulares (circRNA) con diversas funciones regulatorias a nivel co- y post- transcripcional (Hsu & Coca-Prados, 1979; Zhang *et al.*, 2013; Chen & Huang, 2018).

Cabe mencionar que esta y otras clasificaciones como la división en RNA codificantes (cRNA) y no codificantes (ncRNA) no resultan del todo precisas ante un análisis detallado de las funciones de cada uno de los tipos de RNA. Por ejemplo, el tRNA adicionalmente cumple funciones como cebador para la replicación de retrovirus y regula la apoptosis por unión al citocromo C (Hopper & Huang, 2015). De igual modo, dentro de los circRNA, que se clasifican como ncRNA, se han asociado a funciones de molde para eventos de traducción independiente del *capping* al 5' (Zhang *et al.*, 2018).

## **I.2 Procesamiento del mRNA**

En eucariontes, la función de transcripción de genes es llevada a cabo por tres RNA Polimerasas (Pol). La Pol I y III transcriben un número limitado de genes: Pol I sintetiza los rRNA grandes y Pol III sintetiza 5S rRNA, 7SL RNA, tRNA y U6 snRNA involucrados en el procesamiento del RNA. En contraste, la Pol II transcribe un gran número de genes codificadores de proteínas (Paule & White, 2000).

El mRNA es sometido a un proceso de maduración cotranscripcional antes de ser exportado del núcleo. El *splicing* o *forward splicing*, así como el *splicing*

alternativo y las modificaciones de bases como el *5'-capping* y la poliadenilación al 3', constituyen las principales fuentes de generación del repertorio de moléculas de RNA en la célula y determinan el destino y función de los transcritos nacientes (Hocine *et al.*, 2010; Bangru & Kalsotra, 2016).

La actividad de la Pol II resulta en la formación de una molécula precursora del mRNA maduro (pre-mRNA) que cuenta con secuencias codificantes (exones) y no codificantes (intrones) que son eliminados en eventos altamente regulados, donde el dominio carboxilo de la Pol II (CTD) es esencial para su desarrollo (Bentley, 2014).

El *capping* es la primera modificación de mRNA y es esencial en la iniciación de la traducción, la protección contra nucleasas 5'-3' y el reclutamiento de factores de *splicing*, poliadenilación y exportación del núcleo. Posterior a la incorporación de los primeros 25-30 nucleótidos al extremo 5' del transcrito naciente, ocurre la formación de la estructura cap. Esta consiste en una guanósina metilada en la posición N7 unida al primer nucleótido del RNA vía esterificación 5'-5' trifosfato. La actividad de tres enzimas es requerida para generar la estructura cap 0: RNA trifosfatasa (TPasa), RNA guanililtransferasa (GTPasa) y la guanina-N7 metiltransferasa (guanina-N7 MTasa). La TPasa elimina el fosfato  $\gamma$  generando un RNA 5'-difosfato. La GTPasa transfiere un grupo GMP de un GTP al RNA 5'-difosfato vía un intermediario covalente Lisina-GMP y finalmente la guanina-N7 MTasa agrega un grupo metilo en la posición N7 de la guanina del cap formando la estructura del cap 0. Posteriormente la 2'O metiltransferasa metila el ribonucleótido +1, subsiguiente al cap 0, generando la estructura cap 1 (Cowling & Cole, 2010; Ramanathan *et al.*, 2016). Las enzimas del *capping* se unen directamente a la CTD fosforilada de la Pol II por su dominio guanililtransferasa permitiendo la modificación del 5' al inicio de la transcripción. Se observa una correlación en la liberación de la enzima *capping* ante la pérdida de la fosforilación del residuo de serina 5 de los repetidos de pentapéptidos en el CTD de la polimerasa

(Hocine *et al.*, 2010). De igual modo, las evidencias indican una función del *capping* en la regulación del escape de la Pol II del promotor durante la síntesis de pre-mRNA (Mandal *et al.*, 2004).

El descubrimiento de que los genes eucarióticos son discontinuos y con segmentos intrónicos no codificantes, fue un hallazgo inesperado en la biología molecular. Un alto porcentaje de genes, principalmente de los reinos *Animalia* y *Plantae*, llevan a cabo *splicing* alternativo para la remoción de los segmentos no codificantes denominados intrones. En humanos, más del 95 % de los genes son procesados por este mecanismo lo cual expande significativamente el proteoma y sus funciones; además coloca al mecanismo de *splicing* como un evento altamente conservado y esencial en el desarrollo celular. Las funciones del *splicing* van desde la determinación del sexo en la mosca de la fruta hasta la variación de las funciones de receptores de membrana en el sistema nervioso de mamíferos (Kelemen *et al.*, 2013; Lee & Rio, 2015).

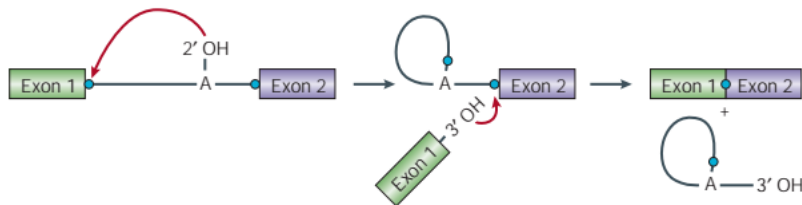
Datos de microarreglos y análisis sistemático han revelado hasta la fecha 7 tipos de *splicing* alternativo. El más prevalente en vertebrados e invertebrados es el tipo cassette exón alternativo o *exon skipping*, en tanto que, en plantas y eucariontes unicelulares, prevalece la retención intrónica, la cual ha sido asociada con sitios de *splicing* débiles, elementos de regulación en *cis* e intrones de longitud corta (Braunschweig *et al.*, 2014; Wang *et al.*, 2014). El mecanismo de *splicing* es mediado por una gran maquinaria ribonucleoproteica y alrededor de 100 factores proteicos auxiliares denominado spliceosoma que incluye a las snRNP (U1, U2, U4, U5 y U6). El spliceosoma funciona en un ciclo complejo y dinámico de ensamblaje-desensamblaje RNA-RNA, RNA-proteína y proteína-proteína que llevan a la ligación de exones y la remoción de intrones en forma de estructura *lariat* (Patel & Steitz, 2003).

El último paso de la transcripción consiste en el corte endonucleolítico, realizado treinta nucleótidos río abajo de la secuencia señal AAUAAA en mamíferos y secuencias homólogas en levaduras seguidas de la adición de una secuencia de poli adeninas (Poli A). La cola de Poli A es esencial para la estabilidad y eficiencia de traducción del mRNA. Para el corte inicial en células de mamíferos son necesarios el factor de poliadenilación y corte (CPSF), el factor de estimulación (CstF) y los factores de corte (CFIm y CFII<sub>m</sub>). Posterior al 3'-OH expuesto la Poli A Polimerasa agrega la cola de Poli A. El mRNA puede ser alternativamente poliadenilado, dando como resultado genes con variaciones en la longitud de sus 3'UTR o que codifican a diferentes proteínas. Al igual que con el *capping*, existe interacción de los factores de poliadenilación con la CTD fosforilada de la polimerasa. Se ha reportado que pérdida de la fosforilación de la serina 2 de la CTD, afecta la poliadenilación por un reclutamiento impropio de factores como CstF (Hocine *et al.*, 2010).

### **I.3 Mecanismo de *splicing* constitutivo (*forward*) y elementos de regulación en *cis***

Joe Sambrook en 1977 descubrió la presencia secuencias no codificantes que separaban las tres secuencias líderes o codificantes de la proteína hexon de los adenovirus (Sambrook, 1977). Estas se nombraron como intrones y exones respectivamente. Adicionalmente, se predijo que diferentes arreglos de exones podían obtenerse mediante *splicing* alternativo para dar isoformas de mRNA por Walter Gilbert (Roy *et al.*, 2013). Este proceso es llevado a cabo por el spliceosoma, el cual cataliza dos reacciones de transesterificación. La primera reacción de transesterificación del pre-mRNA resulta en la liberación del exón flanqueante al 5' del intrón y la formación de un intermediario *lariat*-3' exón. En la segunda reacción ocurre la liberación del intrón *lariat* y la unión de los exones (Figura 1) (Shi, 2017).





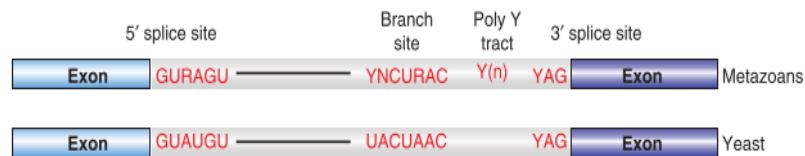
**Figura 1. Reacciones de transesterificación del pre-mRNA durante el *splicing* constitutivo.** Los exones están representados en cajas y el intrón como línea sólida. La adenosina del sitio de ramificación se indica con la letra A. Los ataques nucleofílicos están representados con flechas rojas hacia el enlace fosfodiéster (puntos azules) desde el el hidroxil 2' del sitio de ramificación y el hidroxil 3' del exón rio arriba del intrón. Los productos finales de estas reacciones son los exones ligados y el intrón *lariat* (Tomado de Patel & Steitz, 2003).

Para el *splicing* del pre-mRNA se ha propuesto un mecanismo catalítico mediado por iones metálicos en donde los snRNPs, componentes del espliceosoma, llevan a cabo la coordinación de la catálisis. Experimentos *in vitro* de sustitución de sulfuros en levaduras provocó un spliceosoma catalíticamente inactivo y fue reestablecida su actividad con iones tiofílicos como el manganeso ( $Mn^{2+}$ ). Por otro lado, se pudo ver que U6 snRNA cataliza ambas reacciones de transesterificación mediada por metales divalentes (Yean *et al.*, 2000; Keating *et al.*, 2010).

Para el correcto procesamiento del pre-mRNA es esencial la presencia de los sitios de *splicing* 5' ss y 3' ss a ambos extremos del intrón, además del sitio de ramificación (BP, por sus siglas en inglés: *Branch Point*) y un tracto de poli pirimidinas entre el BP y el 3' ss, todas son secuencias conservadas que garantizan el ensamblaje y funcionamiento del spliceosoma. El BP se encuentra localizado a 18-40 nucleótidos rio arriba del del 3' ss (Figura 2) (Patel & Steitz, 2003). La interacción entre elementos *cis* y factores *trans* de igual modo es esencial para el funcionamiento y la regulación del mecanismo de *splicing*. Los elementos *cis* incluyen potenciadores intrónicos (ISE) y

exónicos (ESE), los cuales se unen a factores *trans* como las proteínas SR (familia de fosfoproteínas ricas en serina/arginina). De igual modo, factores como las ribonucleoproteínas heterogéneas nucleares (hnRNPs) se unen para regular negativamente el spliceosoma (Wang *et al.*, 2014).

Los genes de metazoos presentan dos clases de intrones: tipo U2 y tipo U12, estos últimos no presentan las secuencias canónicas consenso y son procesados por una maquinaria de *splicing* diferente a la anteriormente mencionada, pero de funcionalidad análoga. El spliceosoma tipo U12 está constituido por cuatro snRNPs (U11, U12, U4atac y U6atac) y ambas comparten a U5 snRNPs. Los intrones de tipo U2 presentan terminaciones GT-AG y elementos en *cis* intrónicos conservados (Figura 2). Mientras los intrones tipo U12 presentan terminaciones tanto GT-AG como AT-AC, pero diferentes secuencias consenso 5' ss y BP además de que carecen del tracto de Polipirimidinas río arriba del 3' ss (Patel & Steitz, 2003; Shi, 2017).



**Figura 2. Elementos reguladores en *cis* del *splicing*.** Secuencias conservadas encontradas en los sitios de *splicing* 5' and 3' y en el sitio de ramificación de los intrones tipo U2 en metazoos y levaduras de la especie *Saccharomyces cerevisiae*. Y: pirimidinas, R: purinas y Yn: tracto de pirimidinas (Tomado de Will & Lührmann, 2011).

#### 1.4 Transcripción y *splicing*

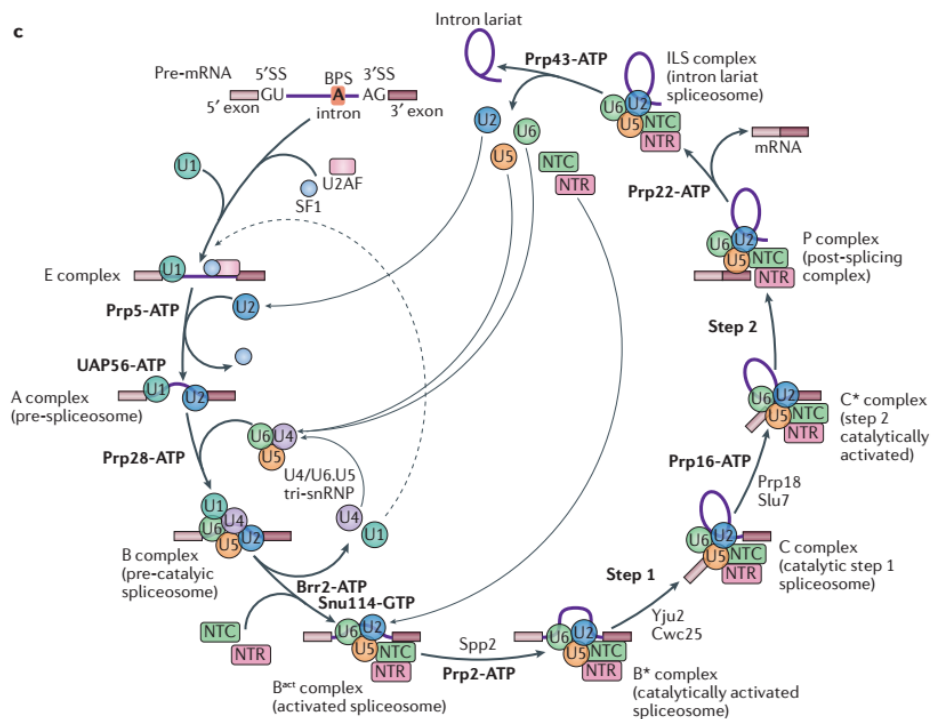
El mecanismo de *splicing* es un evento co-transcripcional y se encuentra estrechamente relacionado a la transcripción génica. Estudios donde se truncó la CTD de la Pol II causaron un *splicing* ineficiente e inhibición en la colocalización de factores de *splicing* con sitios de inicio de la transcripción en

células de mamíferos. (Misteli & Spector 1999). Se ha descrito la asociación de factores de *splicing* a la Pol II, los cuales son reclutados por el CTD fosforilado. De igual modo existe relación *splicing*-transcripción mediada por proteínas. Un ejemplo es el coactivador transcripcional p52, el cual interacciona con SF2/ASF lo cual sugiere que la maquinaria transcripcional modula el reclutamiento de factores de *splicing*. La presencia de intrones en los transcritos incrementa a su vez la eficiencia transcripcional de 10 a 100 veces. De los efectos provocados por los intrones sobre la transcripción tenemos: el potenciamiento de la procesividad de la Pol II y la iniciación de la transcripción debido a la asociación a través de U1 snRNP, reclutamiento de factores de iniciación de la transcripción como TFIIH o el reclutamiento de snRNPs por el factor de elongación de la transcripción TAT-SF1 que se asocia con P-TEFb y provoca la activación de la elongación por fosforilación del CTD (Le Hir *et al.*, 2003; Hocine *et al.*, 2010).

### **I.5 Una enzima dinámica: ensamblaje y actividad del spliceosoma**

El *splicing* del pre-mRNA catalizado por el spliceosoma comienza con el reclutamiento de U1 snRNP al 5' ss y los factores SF1, U2AF65 y U2AF35 que interaccionan con el BP, el tracto de polipirimidinas y el 3' ss respectivamente, formando el complejo temprano o complejo E. En genes eucariontes, las secuencias consenso de *splicing* están poco conservadas, por lo que la unión de los snRNP requiere factores auxiliares como los antes mencionados SF1 y U2AF entre otros. La subunidad mayor del heterodímero U2AF35/U2AF65, contiene una región homóloga a dominios de unión a RNA e interacciona con el 3' ss en tanto que U2AF65 que se une directamente al tracto de polipirimidinas, interacciona con U2AF35 y presenta un dominio rico en arginina-serina hacia su extremo amino. La proteína SF1 se une cooperativamente a U2AF65 para la formación del complejo E. (Figura 3) (Will & Lührmann, 2011; Jenkins *et al.*, 2013; Wu & Fu, 2015).

Posteriormente el U2 snRNP reconoce el BP en una reacción dependiente de ATP mediado por las helicasas Prp5 y Sub2. Con la unión estable de U2 snRNP se forma el complejo A. El tri-snRNP U4/U6 U5 preensamblado es reclutado en una reacción catalizada por Prp28 para formar el complejo B pre-catalítico. Las RNA helicasas DEAH-box Brr2 y Snu114 garantizan la desestabilización de complejo y la liberación de U1 y U4 snRNPs para la formación del complejo B activado (U2, U5 y U6) que con la entrada de la RNA helicasa Prp2 conllevan a la formación del complejo B\*, activado catalíticamente. Con la entrada de los factores Yju2 y Cwc25 el complejo C lleva a cabo la primera reacción de transesterificación en la cual se obtiene el exón libre al 5' del intrón y el intermediario *lariat*-3'exón (Lee & Rio, 2015).



**Figura 3. Ensamblaje y ciclo catalítico del *splicing*.** Ensamblaje y activación del spliceosoma en levaduras. Los complejos del spliceosoma son nombrados según la nomenclatura para metazoos. Los exones e intrones son representados con cuadros y líneas respectivamente. Los elementos en *cis* del intrón presentes en levaduras son mostrados. Las helicasas son nombradas en negrita. (Tomado de Shi, 2017).

La incorporación de la helicasa Prp16 forma el complejo catalíticamente activado C\* donde ocurren una serie de re-arreglos conformacionales que conlleva al desarrollo de la segunda reacción de transesterificación para la remoción del intrón *lariat* y la unión de los exones. La segunda reacción catalítica da lugar al complejo P post-spliceosoma que contiene unidos el mRNA procesado y el intrón *lariat*. Helicasas como la Prp22 y la Prp43 garantizan la liberación del mRNA maduro, del intrón *lariat* y de los snRNPs U2, U5 y U6 los cuales se reciclan para un nuevo evento (Figura 3) (Will & Lührmann, 2011; Matera & Wang, 2014). Como puede observarse, los componentes del spliceosoma son distintos en cada paso de la reacción, incluyendo las dos reacciones de transesterificación.

### **I.6 Intrón *lariat*, eventos post-splicing**

Después de ser procesado el pre-mRNA, el intrón-*lariat* es liberado del spliceosoma en asociación a factores spliceosomales entre los que se encuentran las snRNPs U2, U5, U6 y proteínas del *nineteen complex*. El desensamblaje de este complejo es llevado a cabo por las DEAH-box ATPasas Prp43 y Brr2, así como la GTPasa Snu114. Posteriormente la enzima de-ramificadora (Dbr1 por sus siglas en inglés: *Debranching*) reconoce el sitio de ramificación y linealiza el intrón *lariat* promoviendo su rápida degradación por exonucleasas o su empleo para la obtención de lncRNAs, microRNAs y snoRNAs (Small *et al.*, 2006; Fourmann *et al.*, 2013).

La enzima Dbr1 fue identificada en un tamizaje genético de retrotransposones modificados en extractos de células HeLa y se observó su actividad 2'-5'fosfodiesterasa. Esta es una proteína de la familia de las calcineurinas tipo metalofosfodiesterasa (MPEs) con centro metalo-iónico binuclear, con preferencia por iones manganeso como cofactor y que hidroliza enlaces mono, di y trifosfato (Ruskin *et al.*, 1985; Valdés *et al.*, 2018). Dbr1 presenta alto grado de conservación en diferentes organismos. El *knockout* de esta proteína en

eucariontes pluricelulares como *Arabidopsis thaliana* resulta letal, mayormente, por la ausencia de RNAs pequeños esenciales en la embriogénesis de esta especie. Mientras que en levaduras como *S. cerevisiae* afecta el crecimiento y provoca acumulación de *lariats* no procesados (Wang *et al.*, 2004; Hesselberth, 2013; Katolik *et al.*, 2017).

En ciertos casos, intrones de longitudes cortas con una citocina en el sitio de ramificación pudieran no ser procesados por la Dbr1, la cual reconoce preferentemente una adenina como nucleótido canónico en el BP (Talhouarne & Gall, 2018). Estos intrones pueden permanecer estables en la célula y afectar directamente el procesamiento del pre-mRNA en *cis*. En la mayoría de los casos los intrones estables liberados del *splicing* se mantienen unidos a una serie de hnRNPs. Los niveles de hnRNPs son esenciales para el correcto procesamiento de transcritos y el secuestro de estas proteínas puede afectar rondas posteriores de *splicing*. Por el tamaño y abundancia de los intrones, pudieran provocar un desbalance en la reserva de nucleótidos disponible para eventos de transcripción subsecuentes si no son procesados (Hesselberth, 2013).

## **I.7 RNA no codificantes**

Las investigaciones referentes a los RNAs no codificantes de las últimas décadas demuestran la gran importancia de este grupo de transcritos en la regulación génica y otros eventos celulares. La mayor parte de la información genética de mamíferos y otros organismos complejos es transcrita a ncRNAs, mucho de los cuales son procesados mediante *splicing* alternativo para la obtención de productos más pequeños. Esta clasificación de ncRNAs incluye a microRNAs, snoRNAs, lncRNAs, siRNAs, piRNAs y circRNAs, grupo de RNAs circulares que se abordará detalladamente en el siguiente acápite. Estos transcritos pueden ser derivados tanto de secuencia intrónicas como exónicas (Mattick & Makunin, 2006).

Los rRNAs, snRNAs y los tRNAs entran de igual modo en la clasificación de no codificantes y son de los más ampliamente estudiados. Se considera que cumplen funciones genéricas en la célula, como la traducción del mRNA, el *splicing* en el caso de los snRNAs y la modificación de rRNAs. Sin embargo, también llevan a cabo funciones regulatorias en *cis*. Por ejemplo, además de su participación en el *splicing*, U1 snRNA está involucrado en la regulación del inicio de la transcripción por unión a la RNA polimerasa II en asociación con el factor transcripcional TFIIH (Mattick & Makunin, 2006). Los ElciRNAs son un tipo de circRNA que regulan la expresión de sus genes parentales por unión a la Pol II mediada igualmente por U1 snRNA (Li *et al.*, 2015). Por otro lado, el U1 snRNA interacciona con la ciclina H, lo cual sugiere su implicación en la regulación del ciclo celular (O'Gorman *et al.*, 2005).

El snRNA 7SK tiene actividad inhibitoria sobre el complejo cinasa CDK9/ciclina T, esto reduce la fosforilación de la Pol II llevada a cabo por este complejo y con ello disminuyen los eventos de transcripción (Mattick & Makunin, 2006).

Existen muchos ejemplos bien caracterizados de secuencias de RNA regulatorias en los UTR y en zonas conservadas alrededor del *splice junction* de los mRNAs. Estos elementos en *cis* actúan como sitios de unión para elementos reguladores, los cuales son mayormente proteínas de unión a RNA. Sin embargo, existe evidencia de la unión de ncRNAs a estas secuencias, como el caso de los miRNAs que se unen a los UTRs. (Lewis *et al.*, 2005). Los SnoRNAs se han visto involucradas de igual modo en el control del *splicing* del mRNA del receptor de serotonina 5-HT(2C) por unión a estos elementos en *cis* (Kishore *et al.*, 2006). Estos son también componentes esenciales de la telomerasa, garantizando el mantenimiento y segregación de los cromosomas (Bernstein & Allis 2005).

Los snoRNAs tienen una longitud de 60 a 300 nucleótidos y se dividen en dos clases: los de caja C/D que guían la metilación de la ribosa en la posición 2'-O y los de caja H/ACA las cuales guían la pseudouridilación de RNAs dianas.

De este modo, llevan a cabo funciones en la modificación sitio específica de nucleótidos en rRNAs, mRNA y snRNAs vía complementariedad de bases (Eddy, 2001; Bachellerie *et al.*, 2002; Meier *et al.*, 2005).

Los siRNAs y miRNAs se derivan de horquillas y cadenas dobles del pre-mRNAs. Estos últimos, suprimen la traducción por apareamiento con el mRNA y activación del complejo de degradación RISC constituyendo un eficiente mecanismo de control de la expresión de genes presente en 1/3 de los genes codificadores de proteínas humanas. Por su parte, los siRNAs participan en el silenciamiento de la cromatina (heterocromatina) (Mattick & Makunin, 2006).

## **I.8 RNA circulares**

Los RNA circulares han sido descritos en prácticamente todos los dominios de la vida incluyendo eucariontes, archaea, bacteria y virus. Fueron descritos por primera vez por Sanger *et al.*, (1976) los cuales mediante estudios de microscopia electrónica y sedimentación determinaron la estructura circular de la cadena simple de viroides patógenos de plantas superiores. Posteriormente Hsu & Coca-Prados, (1979) hallaron circRNAs exónicos en el citoplasma de células HeLa. Entre las características que presentan estas moléculas se encuentra la resistencia a exonucleasas, restricción en la formación de estructuras secundarias y capacidad de unión a proteínas y otros RNAs con evidencia funcional (Lasda & Parker, 2014).

Se han descrito varios tipos de RNA circulares como los provenientes de intrones del grupo I y grupo II de bacterias y virus que llevan a cabo autoprosesamiento. Estos RNA circulares han sido relacionados con eventos de reinserción y movilidad de elementos génicos. El genoma de RNA de cadena sencilla circular de 1.75 kb del virus de la hepatitis delta es otro ejemplo de RNA circular. Igualmente, en *Caenorhabditis elegans*, *Drosophila melanogaster*, algas rojas como *Cyanidioschyzon merolae* y la especie de archaea *Sulfolobus acidocaldarius* se han identificado circRNAs provenientes



del procesamiento de pre-tRNAs y pre-rRNAs. También en la archaea *Sulfolobus solfataricus*, fueron descritos circRNAs exónicos asociados a 37 genes. Estos hallazgos denotan la posible conservación evolutiva de la circularización de moléculas de RNA (Danan *et al.*, 2011; Zhang *et al.*, 2013; Holdt *et al.*, 2018).

Avances en transcriptómica y secuenciación profunda han revelado miles de genes que expresan RNAs no codificantes circulares estables y funcionales producto del procesamiento del pre-mRNA en células eucarióticas. Los mayores estudios se han realizado en células humanas donde han sido encontrados alrededor de 25000 isoformas de circRNAs en fibroblastos, sugiriendo que el 20% de la expresión de genes activos en la célula están asociados a la producción de circRNAs. A nivel genómico se estima que más del 50% de los eventos de circularización por *backsplicing* involucra eventos de circularización exónica y un 20% corresponde a círculos con secuencias intrónicas. Debido a su abundancia y presencia prácticamente ubicua en los sistemas biológicos, los circRNAs son considerados a la fecha un importante mecanismo de control celular altamente conservado (Salzman, 2014; Holdt *et al.*, 2018; Zhang *et al.*, 2018).

### **I.9 Biogénesis de los circRNAs**

Independientemente del mecanismo de biogénesis analizado, han sido descritos diversos factores que promueven la biogénesis de los circRNAs en eucariontes. En aproximadamente el 70% de los eventos de circularización se utilizan señales del spliceosoma tipo U2 y elementos en *cis* y *trans* relacionados con la regulación del *splicing*. (Chen, 2016). Se han descrito tres mecanismos de biogénesis dependiente del spliceosoma para los circRNAs: 1) circularización post-*splicing* de intrones *lariat*, 2) *backsplicing* co-transcripcional de pre-mRNA y 3) *backsplicing* post-transcripcional (Holdt *et al.*, 2018).

Topológicamente el *lariat* es una molécula circular con unión covalente 2´-5´ posterior a la degradación de la cola 3´. Los intrones *lariats* tienen un tiempo de vida muy corto desde que son liberados del spliceosoma siendo linealizados por la enzima Dbr1 y posteriormente degradados dentro del núcleo en cuestión de minutos. Sin embargo, pueden ser estables y cumplir funciones en la célula. Zhang *et al.* (2013) identificaron círculos intrónicos estables (ciRNAs) con unión covalente 2´-5´ en células humanas (Figura 4A). A partir de la construcción de minigenes se estudiaron los elementos *in cis* que garantizaban el escape a la acción de la Dbr1, es decir, implicados en su biogénesis. Observaron la presencia de una secuencia conservada de 7 nucleótidos rica en GU en el 5´ y una secuencia rica en citocinas de alrededor de 11 nucleótidos río arriba del BP.

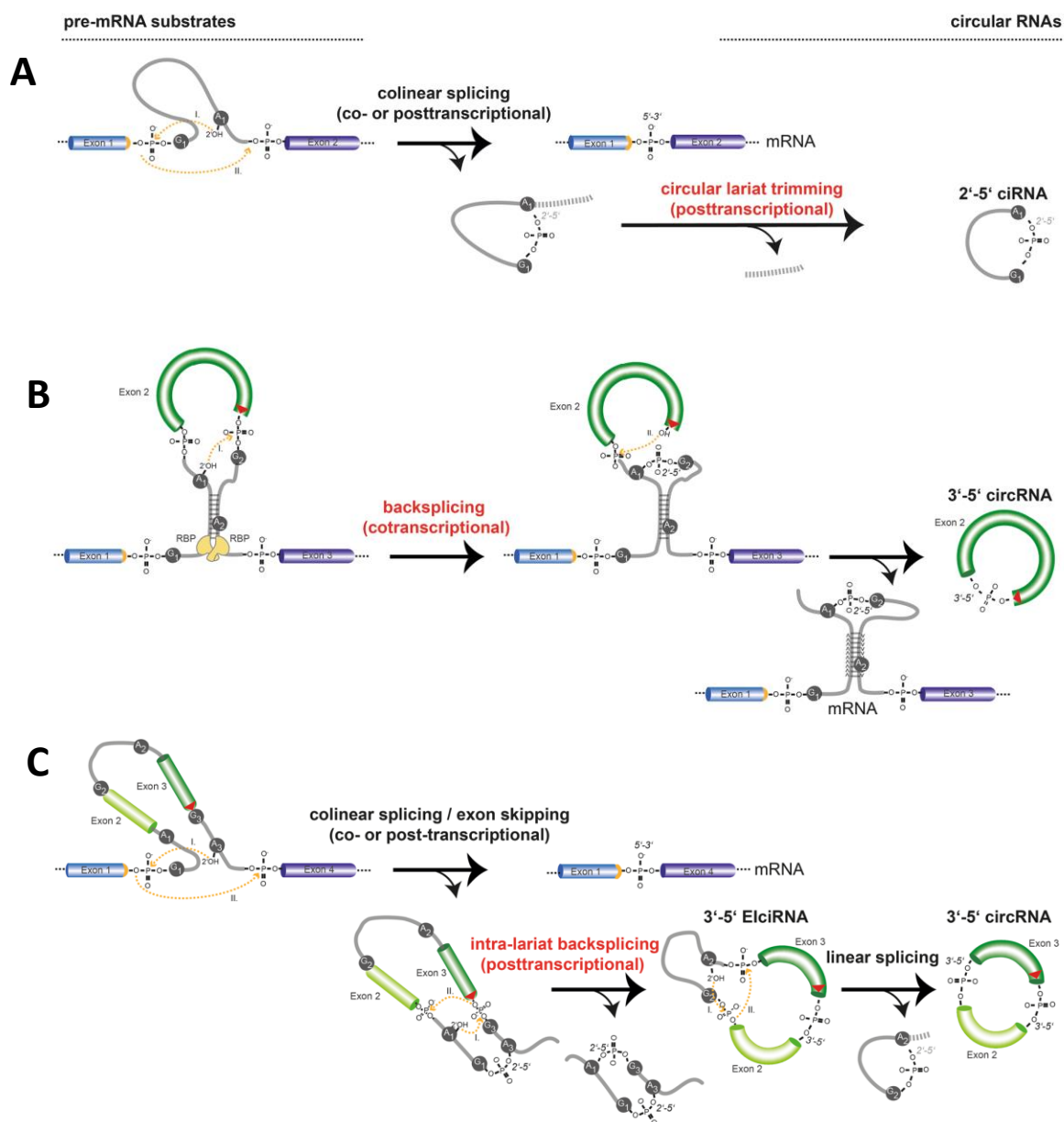
La biogénesis de los circRNAs exónicos o intrón-exónicos transcurre por el mecanismo de *backsplicing* donde un donador 5´ss río abajo se une a un aceptor 3´ss río arriba generando una molécula circular que puede ser multi-exónica, exón-intrónica o de un solo exón. Este mecanismo puede ocurrir co-transcripcionalmente (Figura 4B), o postranscripcional en un exón que contiene unido un *lariat* producido por el *splicing* canónico para dar lugar los denominados ElciRNAs. (Figura 4C) (Chen, 2016). Los intrones en los ElciRNAs pueden ser eliminados en eventos post-*splicing* resultando en un circRNA exónico (Holdt *et al.*, 2018).

El *backsplicing* puede ocurrir mediado por elementos adicionales al spliceosoma. Secuencias complementarias (repeticiones *Alu* inversamente orientadas) en los intrones flanqueantes al exón o exones a circularizar constituyen elementos *in cis* que permiten la formación de un lazo que garantiza la cercanía de los sitios 5´ss y 3´ss para su circularización. Las complementariedades de 30-40 nucleótidos son suficientes para garantizar este mecanismo (Pasman *et al.*, 1996; Liang *et al.*, 2014). Investigaciones de edición de genes con CRISPR/Cas9 fueron usadas para la delección de

secuencias repetidas implicadas en la formación de círculos, lo cual fue suficiente para eliminar la circularización (Zhang Y *et al.*, 2016).

La interacción proteína-proteína constituye un mecanismo alternativo para la formación de circRNAs. Las proteínas de unión al RNA *Quaking* (QKI), el factor de *splicing* MBL (Muscleblind) y la proteína de fusión del sarcoma (FUS), han sido relacionadas con el incremento de la circularización. Estas proteínas de unión al RNA forman homodímeros que unen intrones flanqueantes a sitios de formación de círculos (Ashwal-Fluss *et al.*, 2014; Conn *et al.*, 2015).

La proteína QKI es un factor de biogénesis responsable de la circularización de aproximadamente un tercio de todos los circRNAs inducidos durante la transición epitelial mesenquimal (EMT) en humanos. Los resultados sugieren que estas proteínas regulan el balance entre la biogénesis de los circRNAs y el *splicing* canónico como mecanismos excluyentes en competencia (Conn *et al.*, 2015). Otro ejemplo son las proteínas FUS, las cuales son conocidas por regular tanto el inicio de la transcripción por unión a la RNA Pol II, como el *splicing* alternativo por unión al spliceosoma. Durante la diferenciación de células madres de embión de ratón a motoneuronas en cultivo, se pudo observar que las proteínas FUS se unieron a secuencias exón-intrón circularizadas y regularon de este modo la biogénesis de circRNAs (Errichelli L *et al.*, 2017).



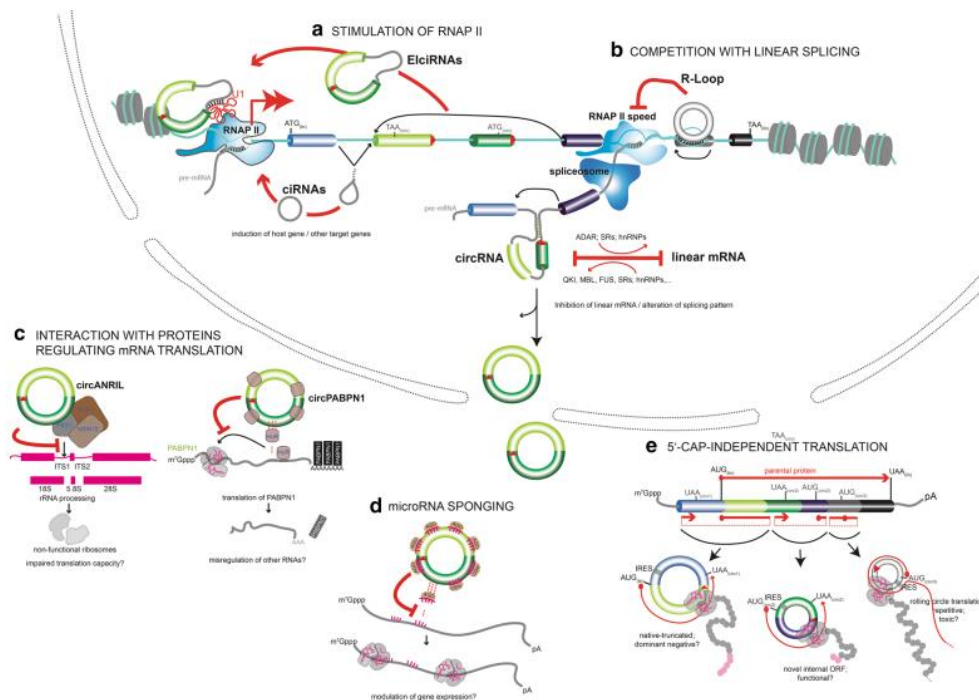
**Figura 4. Biogénesis de circRNAs en eucariotes.** Los sustratos de la reacción están a la izquierda y el producto de la reacción a la derecha. **(A):** *Splicing* co-lineal o canónico da lugar a mRNA procesado y un intrón en estructura *lariat* el cual puede ser procesado para la obtención de ciRNAs con unión covalente 2'-5'. **(B):** Formación de circRNAs por *backsplicing* cotranscripcional. Esta reacción ocurre en el pre-mRNA y puede ser mediada por la maquinaria de *splicing*, secuencias complementarias flanqueantes al sitio de circularización o por dimerización de proteínas de unión a RNA que se unen de igual modo a los intrones flanqueantes (proteínas en amarillo). **(C):** Formación de circRNAs por *backsplicing* posttranscripcional. En el primer paso ocurre el mecanismo de *splicing* da lugar a la liberación de

un exón que contiene un intrón *lariat* el cual se convierte en sustrato para un evento de *backsplicing* intra-*lariat*. La posición de los donadores de *splicing* se representa con puntos naranjas. Las uniones de *backsplicing* se representan en triángulos rojos. La reacción de transesterificación y su dirección se representa con flechas naranja y la cabeza de flecha representa la dirección del ataque nucleofílico. Secuencias flanqueantes representadas con líneas punteadas (Tomado de Holdt *et al.*, 2018).

## I.10 Funciones de los circRNAs

Hasta la fecha son pocas las investigaciones relacionadas con las funciones de los circRNAs que han sido publicadas, pero la evidencia experimental indica un alto potencial funcional para estas moléculas. Dependiendo de su localización, se han asignado cinco funciones de como los circRNAs pueden regular la expresión de genes a diferentes niveles. Los circRNAs pueden estimular la iniciación y elongación de la transcripción por unión a la RNA Pol II como elemento de regulación en trans (Figura 5A).

De igual modo, el proceso de biogénesis en si, puede significar un mecanismo de inhibición “pasiva” de la expresión del mRNA lineal por competencia entre el *splicing* canónico y el *backsplicing* ante la formación de un híbrido DNA-RNA que afecta la función de la RNA Pol II (Figura 5B). Estos círculos pueden interactuar con proteínas en citoplasma para la regulación de la traducción del mRNA (Figura 5C), así como actuar de esponjas de microRNAs controlando la función de estos ncRNAs sobre sus mRNA lineales blanco (Figura 5D) (Holdt *et al.*, 2018). Otros estudios han demostrado que los circRNAs pueden codificar para péptidos funcionales por la presencia, en su secuencia, de un marco abierto de lectura y sitios de entrada al ribosoma, garantizando eventos de traducción independientes de capping al 5´ (Figura 5E) ( Zhang *et al.*, 2013; Ashwal-Fluss *et al.*, 2014; Li *et al.*, 2015; Starke *et al.*, 2015; Holdt *et al.*, 2016; Zhang *et al.*, 2018).



**Figura 5. Función de los circRNAs en eucariotes. (A, B):** Función nuclear de los circRNAs. **(A):** ElciRNAs y ciRNAs estimulan la RNA Pol II para la regulación del inicio de la transcripción de genes codificadores de proteínas. **(B):** Competencia entre el *splicing* canónico y el *backsplicing* por la formación de un híbrido DNA-RNA o por la participación de elementos trans que garantizan una competencia antagónica entre ambos mecanismos. **(C-E):** Funciones citoplasmáticas de los circRNAs. **(C):** Interacción de circRNAs con proteínas y regulación de funciones relacionadas a la traducción de genes. Los casos mostrados son la unión de circANRIL a PES1 para la inhibición del complejo PeBoW durante el procesamiento del rRNA y las funciones de esponja de proteínas Hur por circPABPN1. **(D):** Los circRNAs pueden actuar como esponjas de microRNAs inhibiendo el bloqueo de la traducción de mRNA blancos de esto ncRNAs. **(E):** Traducción de ORFs codificados en circRNAs por iniciación independiente de capping 5' utilizando IRES (Tomado de Holdt *et al.*, 2018).

### I.10.1 Función nuclear de los circRNAs

Los circRNAs con secuencias intrónicas aparecen localizados mayoritariamente en el núcleo y se han relacionado con la regulación génica. Li *et al.* (2015) encontraron círculos de RNA exón-intrónicos (ElciRNAs) unidos a la RNA Pol II regulando sus genes parentales en células HeLa (Figura 5A).

Mediante estudios de inmunoprecipitación contra la RNA Polimerasa II y secuenciación de RNA (CLIP), determinaron que la interacción era dependiente de U1 snRNA, que como fue mencionado anteriormente, se ha relacionado además a sus funciones como elemento del spliceosoma, con la regulación génica por unión al sitio de inicio de la transcripción de diversos genes en complejo con factores transcripcionales como TFIIH (Mattick & Makunin, 2006). El dímero U1-EIciRNA acompañado con la RNA Pol II estimula la actividad transcripcional en *cis* por mecanismos aún no descritos. El *knockdown* de los EIciRNAs así como el bloqueo de la interacción U1-circRNA mediante el uso de morfolino-oligonucleótidos antisentido (AMO) contra U1, causaron la disminución de la transcripción del gen parental (Li *et al.*, 2015).

A pesar de presentar mayormente localización citoplasmática, se ha reportado un círculo exónico de localización nuclear producido a partir del gene UBR5 que codifica para la proteína ubiquitin-ligasa E3, componente de reconocimiento N-5. Los resultados mostraron que la expresión de este circRNA estuvo significativamente disminuido en células no pequeñas de cáncer de pulmón (NSCLC). Circ-UBR5 se relaciona con la unión a factores reguladores del *splicing* como las proteínas QK1, NOVA1 y U1 snRNA en el núcleo y se correlacionó positivamente con células tumorales menos agresivas en este tejido (Quin *et al.*, 2018).

Los ciRNAs son otros de los RNA circulares relacionados a la regulación de genes en *cis*. Estos pertenecen a la categoría de intrones de secuencia estable (sisRNAs), los cuales incluyen una serie de moléculas intrónicas con elementos en *cis* que le garantizan la evasión de la actividad 2'-5' fosfodiesterasa de la Dbr1 (Talhouarne & Gall, 2018). Los estudios más completos referentes a la funcionalidad de los ciRNAs fueron llevados a cabo por Zhang *et al.* (2013) para el caso de ci-*ankrd52* y se estudió parcialmente ci-*sirt7*. De acuerdo a lo reportado por este trabajo los ciRNAs presentan asociación directa con la cromatina y regulan la expresión de su gen parental

por mecanismos aún no determinados (Figura 5A). Se observó asociación de *ci-ankrd52* con la RNA Pol II fosforilada, como indicativo de etapas de elongación de la transcripción, lo cual fue determinado mediante ensayos de inmunoprecipitación contra la CTD de la polimerasa fosforilada en la serina 2. Adicionalmente se observó que los *ciRNAs* pueden asociarse con loci diferentes a su gen parental por lo que pudieran influir en la expresión génica global de la célula (Zhang *et al.*, 2013).

Se han establecido dos conceptos referentes a como la circularización del RNA puede afectar el *splicing*: la inhibición “pasiva” del *splicing* linear por eventos de *backsplicing* en la misma molécula de pre-mRNA y la inhibición “activa” por interacción del círculo naciente con el gen parental afectando la procesatividad de la RNA Pol II (Figura 5B). Estudios en el gene *Mbl* permitieron determinar que ambos mecanismos compiten por el procesamiento del pre-mRNA de este gen. En el caso de la formación de híbridos DNA-circRNAs como mecanismo “activo” de inhibición, se describe la posibilidad de que los circRNAs puedan unirse a la hebra molde de DNA por complementariedad de bases ocasionando la formación de un Loop R, el cual podría afectar la organización de la cromatina, ocasionar descompactación del nucleosoma que conlleva a la modificación de histonas así como causar el estancamiento de RNA Pol II (Ashwal-Fluss *et al.*, 2014; Chedín, 2016).

#### I.10.2 Función citoplasmática de los circRNAs

El mayor número de circRNAs aparecen localizados en citoplasma y se han descrito diversas funciones asociadas a estos (Figura 5C-E). Varios estudios reportan circRNAs implicados en la regulación de la traducción del mRNA, entre ellos tenemos los *circPABPN1* que secuestran e inhiben proteínas implicadas en la regulación de la traducción del mRNA. De igual modo, *circANRIL* se une a proteínas implicadas en la maduración del complejo ribosomal regulando hacia abajo la traducción. Entre las proteínas de unión descritas, se encuentra el complejo PeBoW que actúa en el procesamiento del



pre-rRNA durante la maduración de la subunidad 60S, estabilizando nucleasas que remueven los espaciadores internos (ITS1) del pre-rRNA. Se estima que la sobreexpresión de *circANRIL* provoca la disminución de la interacción de PES1, proteína del complejo PeBow, con el intermediario pre-rRNA durante su procesamiento debido a una posible interacción entre *circ-ANRIL* con el complejo PeBoW (Holdt *et al.*, 2018).

Una de las características de los circRNAs es la gran estabilidad que presentan en comparación con su contraparte lineal. Debido a esto, se sugiere que muchos circRNAs compiten en el citoplasma por la unión a miRNAs y de este modo regulen la expresión génica (Chen, 2016). El ciRS-7 es un ejemplo de circRNA esponja de miRNAs. Se forma a partir del gen CDR1 y se ha reportado en cerebro de humanos. Este presenta mas de 60 sitios de unión a miR-7. De acuerdo a su abundancia y sus numerosos sitios de unión, podrían ocurrir alrededor de 20000 uniones de miR-1 por célula lo cual constituye un efecto considerable en la regulación de la traducción (Hansen *et al.*, 2013).

A pesar de clasificarse como ncRNAs se han reportado circRNAs que codifican para proteínas funcionales entre los que se encuentran los *circMbl*, *circZNF609* y *circSHPRH*. La identificación de circRNA codificantes resulta difícil debido a las pocas bases de datos disponibles para el análisis de marcos abiertos de lectura (ORFs) conservados entre especies, además el hecho de que los circRNAs codificantes se generen a partir de exones, durante la identificación, se podrían solapar los ORFs con sus homólogos del mRNA (Zhang *et al.*, 2018). En todos los casos, los circRNAs codificantes surgieron ante circunstancias muy especiales como en el caso de *circZNF609* donde el *backsplicing* permitió la inclusión de parte del 5'-UTR el cual presenta propiedades similares al sitio interno de entrada al ribosoma (IRES) lo cual le permite a este circRNA participar en eventos de traducción (Pamudurti *et al.*, 2017).

Circ*SHPRH* constituye la forma circular del gen que codifica para la helicasa del anillo PHD de la histona SNF2 (SHPRH). Este circRNA presenta altos niveles de expresión en glioblastoma de cerebro humano. Su capacidad codificante se da por el solapamiento del código genético durante su formación por *backsplicing*, que genera un tandem UGAUGA que funciona como codón de inicio (AUG) y parada (UGA) al mismo tiempo y da lugar a una proteína trunca de 17 kDa. Se estima que esta proteína protege a su forma homóloga completa de la degradación mediada por el proteosoma. La proteína SHPRH completa es una bien caracterizada ubiquitin ligasa E3, la cual tiene como blanco el antígeno nuclear de proliferación celular (PCNA), permitiendo la inhibición de la proliferación celular y la tumorigénesis. Por tanto, la función de circ-*SHPRH* constituye un mecanismo supresor de tumores en glioblastoma de cerebro humano (Zhang *et al.*, 2018).

### **I.11 *Entamoeba histolytica***

*Entamoeba histolytica* es el parásito protozoario causante de la amebiasis, el cual, continúa estando entre las tres principales causas de mortalidad por infección parasitaria a nivel mundial. Se estima que el 10% de la población mundial es infectada cada año resultando en alrededor de 100,000 muertes. La cepa HM1-IMSS de *E.histolytica* (ATCC 30459) fue aislada a partir de una úlcera rectal de un adulto masculino humano con disentería amebiana en la Ciudad de México en 1967. Esta cepa presenta características bioquímica y genéticamente que le permiten ser altamente invasiva, por lo cual resulta de gran importancia clínica el estudio de sus mecanismos de virulencia y regulación génica (Solaymani-Mohammadi *et al.*, 2006; Kantor *et al.*, 2018).

La infección ocurre ante la ingesta de quistes amebianos en agua o alimentos contaminados. Al llegar a nivel del intestino delgado ocurre el desenquistamiento y la eclosión a trofozoitos, la forma infectiva de este parásito (Kantor *et al.*, 2018). La colonización del lumen intestinal ocurre de

manera asintomática en un 90% de los casos mientras que en el 10-20% de los individuos se observa amebiasis invasiva donde el parásito logra sobrepasar la mucosa intestinal mediante el reconocimiento y destrucción de la mucina (MUC2), glicoproteínas de la mucosa intestinal secretadas por las células caliciformes, mediante actividad cistein proteasa sobre el carboxilo terminal de MUC2 (Pearson & Sing, 2010; Begum *et al.*, 2015). La invasión del tejido produce disentería ambiana, colitis y en casos mas graves ocasiona abscesos entraintestinales en hígado, pulmones y cerebro (Cornick & Chadee, 2017).

La manifestación invasiva tiene alta relación con el hospedero. Se sugiere que la composición de carbohidratos y factores de histocompatibilidad, así como la microbiota intestinal condicionan el cuadro clínico diferencial observado (Olivos, 2011). Ante la ingesta, los quistes infectivos logran evadir la acción de los ácidos estomacales y pasar al intestino delgado donde eclosionan en un ambiente mas alcalino. La barrera mucosa constituye el primer mecanismo innato de defensa contra este protozoo. Posterior a la eclosión, los trofozoitos se adhieren a células epiteliales, eritrocitarias y MUC2 de la mucosa intestinal mediante el receptor galactosa/N-acetil-D-galactosamina (Gal-Lectin) de la familia de las lectinas así como lipofosfoglicanos de membrana de la amoeba (Morf & Singh. 2012; Uribarren, 2017). Los péptidos formadores de canales iónicos denominados ameboporos y las secreción de proteasas contribuyen además a la destrucción rápida de la matriz extracelular y estimulan la respuesta inflamatoria. (Moonah *et al.*, 2013). Una vez alcanzadas las células epiteliales intestinales (IECs) por la infección, estas producen una serie de mediadores proinflamatorios entre los que se encuentran la interleucina-1 (IL-1 $\beta$ ), TNF- $\alpha$  e interleucina-8. Estos mediadores reclutan células del sistema inmune como neutrófilos y macrófagos al sitio de invasión y se desarrollan robustos mecanismos amebicidas que culminan con la liberación de especies

reactivas del oxígeno (ROS) y nitrógeno (NOS) para la defensa del hospedero (Lidell *et al.*, 2006; Begum *et al.*, 2015).

Sin embargo, *E. histolytica* logra evadir los mecanismos de defensa gracias al desarrollo de una serie de factores de virulencia contra la inmunidad innata y adaptativa. Entre ellos, la excreción de cistein-proteasas constituye una de sus mayores defensas. La enzima amebiana arginasa permite a este parásito convertir la L-arginina en L-ornitina en macrófagos, lo cual disminuye la producción de NOS (Elnekave *et al.*, 2003). Adicionalmente *E. histolytica* logra fagocitar de manera efectiva los neutrófilos infiltrados además de interferir en la actividad oxidasa del mismo mediante la acción de la peroxiredoxina, una enzima de superficie celular con actividad antioxidante que inhibe la producción de ROS (Lin *et al.*, 1994; Sim *et al.*, 2005).

Análisis de transcriptoma comparativo entre cepas virulenta y no virulentas permitieron identificar a una serie de genes que se encontraban diferencialmente expresados en cepas virulentas entre los que se encuentran genes implicados en el metabolismo del RNA, transporte celular, factores de *splicing*, de la maquinaria de traducción y del citoesqueleto de actina así como genes que codifican para proteínas de choque térmico como Hsp70 y Hsp90. De igual modo fueron identificados con expresión diferencial, loci que codifican para proteínas hipotéticas, muchos de los cuales han sido relacionados a la producción de circRNAs (EHI\_144490, EHI\_169670, EHI\_050490, EHI\_192510, EHI\_014170 y EHI\_062080) (Mendoza-Figueroa *et al.*, 2018). De este modo se puede concluir que las condiciones del microambiente conlleva a la una modulación severa del transcriptoma amebiano que le garantiza la adaptabilidad, supervivencia y comportamiento invasivo a este parásito (Weber *et al.*, 2016; Mayer *et al.*, 2016).

### I.11.1 Splicing en *Entamoeba histolytica*

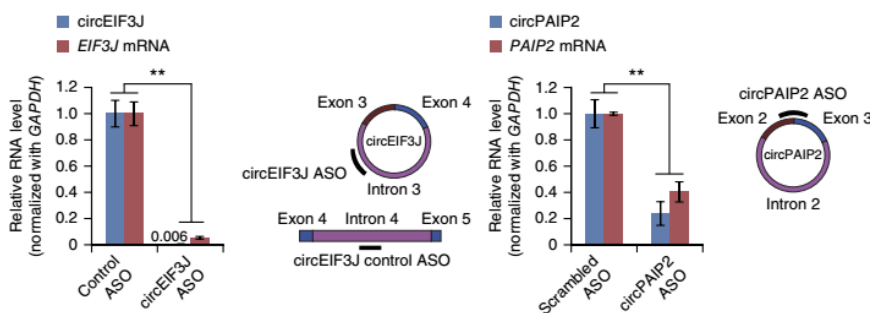
Se han reportado alrededor de 4000 intrones en los 8333 genes anotados de *E. histolytica*, lo cual resulta peculiar para un organismo antiguo, el cual pertenece a un linaje de eucariontes, que ramificaron justo antes de los eventos endosimbióticos de proteobacterias que dieron lugar a la formación de la mitocondria (Morf & Singh, 2012; Valdés *et al.*, 2014). El genoma de *E. histolytica* presenta 75.9% en contenido de adeninas y timinas con un tamaño promedio de genes de 1260 pares de bases con intrones de 74 pares de bases aproximadamente (Loftus, 2005; Lorenzi *et al.*, 2010) y a partir de datos de secuenciación de RNA poliadenilado se estima que alrededor del 50% de los transcritos presentan uniones exónicas con sitios 5´ss y 3´ss alternativos. La mayoría de las secuencias intrónicas flanqueantes presentan alto grado de conservación con la secuencia GUUUGU al 5´ss y UAG al 3´ss, mientras que no presenta tales grados de conservación en las secuencias del BP ni en el tracto de polipirimidinas (Wilihoeft *et al.*, 2001; Hon *et al.*, 2013) .

La mayoría de los componentes del spliceosoma amebiano han sido identificados. Se ha encontrado confirmación molecular de la existencia de homólogos a los snRNAs de U2, U4, U5 y U6 pero no existe evidencia de la presencia del homólogo para U1 snRNA aunque la presencia de factores de la U1 snRNP (U1-A, U1-C y U1-70k) fueron predichos mediante estudios *in silico* y la expresión funcional *in vivo* en ensayos de CLIP con U1-A marcada. Estos experimentos permitieron identificar alrededor de 32 factores de *splicing* en trofozoitos y permitió realizar una descripción mas detallada de los eventos de procesamiento del pre-mRNA por el complejo spliceosomal. Los hallazgos sumados a datos de RNA-seq y predicciones bioinformáticas indican que la retención intrónica es la principal ruta para el *splicing* alternativo en *E. histolytica*, donde prevalecen pre-mRNAs monointrónicos y el *splicing* tiene gran importancia en la regulación de genes y la expansión del proteoma (Valdés *et al.*, 2014).

## II. Antecedentes

Los RNA circulares han sido descritos recientemente como una nueva clase de RNA no codificantes, clasificados dentro de los lncRNAs que existen de manera prácticamente ubicua (Burd *et al.*, 2010; Hansen *et al.*, 2013; Zhang *et al.*, 2013; Ashwal-Fluss *et al.*, 2014; Li *et al.*, 2015; Boeckel *et al.*, 2015; Holdt *et al.*, 2016; Chedín, 2016; Pamudurti *et al.*, 2017; Quin *et al.*, 2018; Zhang *et al.*, 2018). Su conformación circular brinda a estos RNAs propiedades únicas, como la resistencia a exonucleasas, restricción en la formación de estructuras secundarias, y capacidad de unión a proteínas y otros RNA que le permite participar en eventos de regulación de la transcripción génica (Lasda & Parker, 2014, Qu *et al.*, 2015).

Los RNA circulares con secuencias intrónicas como los ElciRNAs y los ciRNAs han sido vinculados a la regulación en *cis* de sus genes parentales. Los ElciRNAs: circ-EIF3J y circ-PAIP2 son un ejemplo. Li *et al.* (2015) llevaron a cabo el knockdown de estos ElciRNAs utilizando oligonucleótidos antisentido específicos (ASO) en células HeLa y observaron una disminución de los genes parentales en ambos casos (Li *et al.*, 2015) (Figura 6).



**Figura 6. Regulación en cis de los genes EIF3J y PAIP2 mediada por ElciRNAs:** Disminución en los niveles de mRNA de sus genes parentales después de Knockdown de circEIF3J y circPAIP2 con ASOs oligoespecíficos (Tomado de Li *et al.*, 2015).

Los ciRNAs fueron descritos por primera vez por Zhang *et al.* (2013) y posteriormente se clasificaron como intrones de secuencia estable (sisRNAs) los cuales son intrones con elementos en *cis* altamente conservados que permiten la formación de estructuras circulares en su mayoría con unión covalente 2'-5' por circularización del intron *lariat*. Estos sisRNAs fueron identificados en grandes cantidades en ovocitos de ranas del género *Xenopus*, así como en glóbulos rojos de ratones y humanos. Estudios de secuenciación del BP permitieron observar que el 80% de los siRNAs presentaban una citocina en el BP en lugar de un adenina lo cual podría garantizar a este grupo de intrones evadir la acción de la Dbr1 para su linealización (Talhouarne & Gall, 2018).

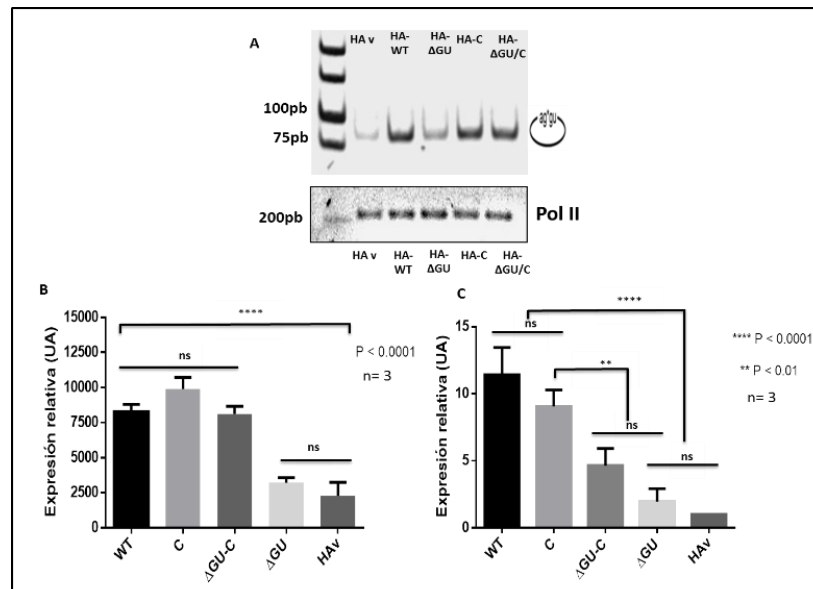
De acuerdo a lo reportado por Zhang *et al.* (2013) los ciRNAs se unen a la cromatina y participan en eventos de regulación génica en diferentes loci por mecanismos aún no determinados influyendo así en la expresión global de la célula (Figura 5A). Observaron asociación de *ci-ankrd52* con la RNA Pol II fosforilada, como indicativo de etapas de elongación de la transcripción, lo cual fue determinado mediante ensayos de inmunoprecipitación contra la CTD de la polimerasa fosforilada en la serina 2 (Zhang *et al.*, 2013).

Recientemente, en nuestro laboratorio se reportó la existencia de RNAs intrónicos circulares de longitud completa (flicRNAs) en diferentes genes de *Entamoeba histolytica*. Estos círculos difieren de los hallados por Zhang *et al.* (2013) en que presentan unión covalente entre los sitios 3' y 5' *splicing* y se originan por eventos *post-splicing* producto del procesamiento del intron *lariat* por mecanismos aún no descritos (Mendoza-Figueroa *et al.*, 2018).

Estudios de secuenciación de los intrones formadores de flicRNAs, permitieron determinar la presencia de una secuencia altamente conservada rica en GU al 5' ss, la cual se vinculó estrechamente a la biogénesis de los mismos mediante la transfección de trofozoitos con minigenes en los cuales se mutó la secuencia conservada (HA-ΔGU) y se observó una disminución significativa en la

producción de flicRX13 producto del gen RabX13 que codifica para una GTPasa pequeña (Figura 7) (Mendoza-Figueroa *et al.*, 2018).

Por otro lado, se pudo observar la ausencia del tracto de polipirimidinas río abajo del BP reportado en metozoos. Respecto a la estructura *ariat* de este grupo de intrones, se observó que la distancia entre el sitio de ramificación 2´-5 y el sitio de *splicing* 3´ss eran de menos de 10 nucleótidos como característica común, lo cual sumado a la baja conservación del BP en intrones de *E. histolytica* podrían ser factores determinantes en los eventos de circularización (Mendoza-Figueroa *et al.*, 2018).



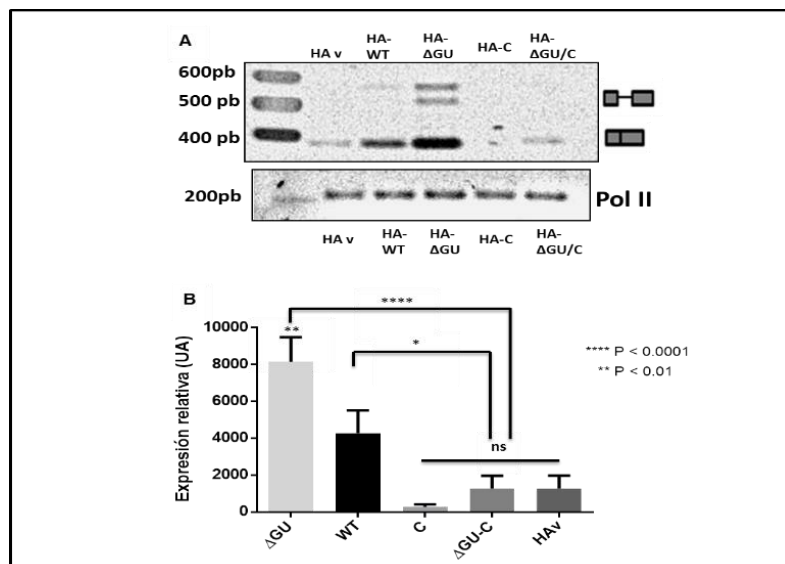
**Figura 7: Expresión relativa de flicRX13 en cada una de las clonas.** (A) Gel nativo de poliacrilamida al 12% con los flicRX13 (79 pb) detectados en cada mutante. Debajo se observa la amplificación del gen de la Polimerasa II (control de carga) en un gel de agarosa al 2%. (B) Comparación estadística (ANOVA de un factor) de la expresión relativa en unidades arbitrarias de los flicRX13 en cada clona detectados mediante PCR en punto final. (C) Comparación estadística (ANOVA de un factor) de la expresión relativa en unidades arbitrarias de los flicRX13 en cada clona detectados mediante qPCR (Tomado de Mendoza-Figueroa *et al.*, 2018 ).



Mediante experimentos donde se sobrepresó el factor U2AF deletado en el carboxilo terminal Mendoza-Figueroa *et al.* (2018), pudieron observar un incremento en la producción tanto de flicRNAs, como de productos de *splicing*, sugiriendo que los flicRNAs se formaban posterior al ensamblaje del complejo E del spliceosoma. Seguidamente, para determinar si la formación de los flicRNAs ocurría durante o después del primer paso de transesterificación del *splicing*, fueron tratadas amebas con ácido bórico, reportado como inhibidor del segundo paso de transesterificación del *splicing in vitro* (Shomron & Ast, 2003). Mediante RT-PCR usando primers marcados radiactivamente, detectaron que en tratamientos de 1.5 horas a concentraciones de 5 mM de ácido bórico, se afecta la amplificación de todos los productos de *splicing* particularmente flicRX13. Con estos resultados se pudo concluir que los flicRNAs se originan por circularización del intrón-*lariat* en eventos post-*splicing* (Mendoza-Figueroa *et al.*, 2018).

Intrones con características similares fueron descritos anteriormente por Taggart *et al.* (2017) en estudios dirigidos al análisis de la configuración del BP en diferentes líneas celulares humanas. Pudieron observar que aproximadamente el 9% de los intrones de tipo U2 con sitios de ramificación localizados a menos de 10 nucleótidos del sitio de *splicing* 3' splice site (AG) producen RNAs intrónicos circulares y el 3% de los casos constituyen intrones circulares completos (Taggart *et al.*, 2017).

Adicionalmente, estudios de la implicación de flicRX13 en la regulación génica, permitieron observar un aumento de la expresión de su gen parental en las cepas de *E. histolytica* transfectantes HA+ΔGU indicando una posible inhibición constitutiva del gen parental RabX13 mediada por flicRX13 (Figura 8) (Mendoza-Figueroa *et al.*, 2018).



**Figura 8: Efecto de cada clona sobre la expresión del gen RabX13 nativo de *E. histolytica*.** (A) Gel de agarosa al 2% con el transcrito y el mensajero de RabX13 correspondiente a cada clona (transcrito primario 520 pb; mRNA de RabX13 376 pb). Debajo se observa la amplificación del gen de la polimerasa II (control de carga) en un gel a agarosa al 2%. (B) Comparación estadística (ANOVA de un factor) de la expresión relativa en unidades arbitrarias del mRNA de RabX13 obtenido por PCR tiempo real en cada una de las clonas (Tomado de Mendoza-Figueroa *et al.*, 2018 ).

Por la novedad de los hallazgos antes descritos y la necesidad de estudios más detallados respecto a la función de los flicRNAs, resulta de interés para el laboratorio vincular aún más los flicRNAs con la regulación de sus genes parentales, para lo cual demostraremos la presencia de estos en el complejo de unión a la RNA Polimerasa II y llegaremos a conclusiones acerca de su función celular .

### **III. Justificación**

Recientemente fueron reportados RNAs intrónicos circulares de longitud completa (flicRNA) en *Entamoeba histolytica*. Se demostró que éstas son secuencias intrónicas estables productos del procesamiento de pre-mRNA y circularizadas por unión covalente 3'-5'. Intrones con características similares han sido descritos en líneas celulares humanas sin embargo, no han sido estudiadas sus posibles funciones. El efecto del flicRX13 sobre la expresión de su gen parental RabX13 es el único estudio reportado que analiza la capacidad funcional de intrones circulares completos. Los resultados permitieron concluir de una inhibición constitutiva del gen RabX13 mediada por flicRX13, resultados similares a los observados para los ciRNAs, círculos intrónicos de unión covalente 2'-5', los cuales se han relacionado a la regulación génica por unión a la RNA polimerasa II fosforilada.

## **IV. Hipótesis**

Los flicRNA se encuentran unidos a la RNA Polimerasa II y otros factores en cepas de *Entamoeba histolytica* y posiblemente participan en la regulación de sus genes parentales.

## V. Objetivos

### Objetivo General:

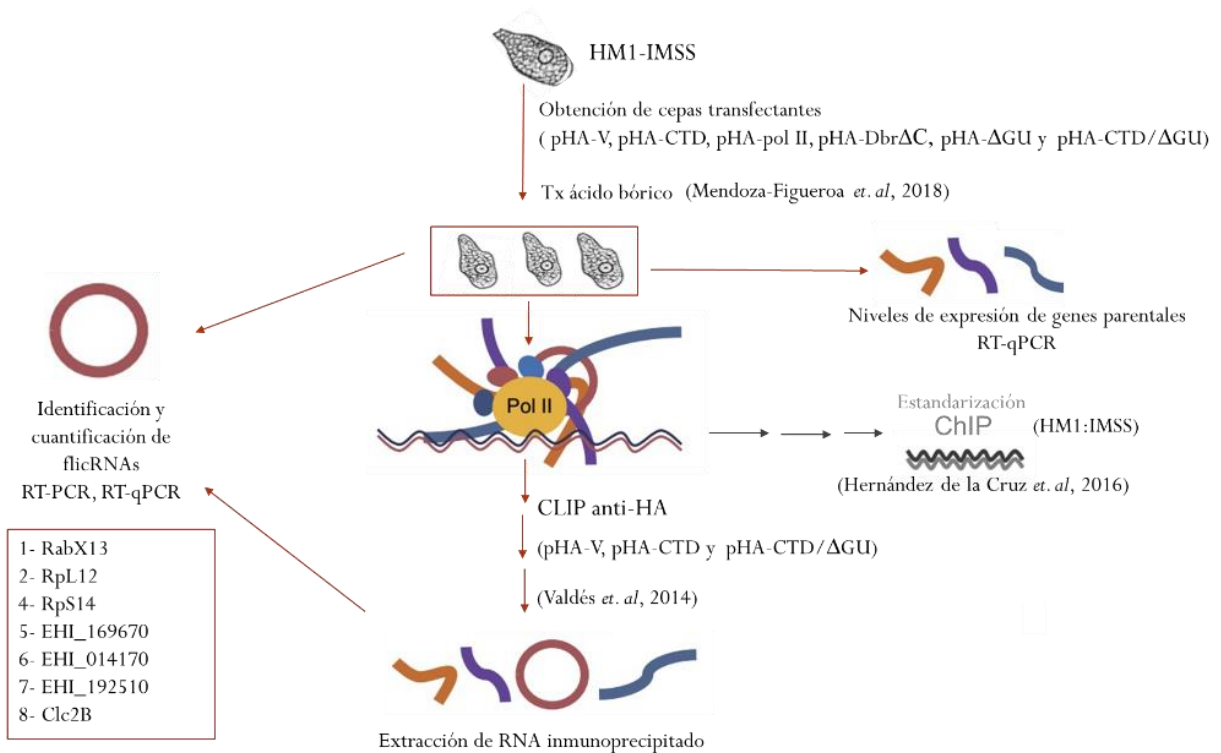
Determinar la presencia y actividad reguladora de los flicRNAs en el complejo transcripcional de sus genes parentales en cepas de *Entamoeba histolytica*.

### Objetivos específicos:

- Determinar la unión de los flicRNA a la RNA polimerasa II mediante ensayos de *Cross-Linking* inmunoprecipitación anti HA (CLIP) en cepas transfectantes pHA-CTD.
- Determinar diferencias en la actividad transcripcional de genes formadores de flicRNAs en diferentes cepas transfectantes y ante diferentes tratamientos.
- Estandarizar ensayo de Inmunoprecipitación de cromatina (ChIP) en cepas HM1-IMSS determinando condiciones para la fragmentación del DNA genómico y eficiencia de *cross-linking*.

## VI. Estrategia experimental

Para dar cumplimiento a los objetivos trazados se obtuvieron cepas amebianas transfectantes con diferentes construcciones para marcar con el péptido de hemaglutinina tanto el extremo carboxilo terminal de la RNA polimerasa II amebiana (pHA-CTD) como la subunidad mayor de la misma (pHA-Pol II). A partir de ellas, se llevaron a cabo ensayos de CLIP seguido de identificación de los flicRNAs de diferentes genes de *E. histolytica* unidos a la Pol II por RT-PCR circular. Para correlacionar la formación de flicRNAs con la expresión de sus genes parentales, las transfectantes pHA-V y pHA-CTD se trataron con ácido bórico, inhibidor de la segunda reacción de transesterificación del *splicing*, resultados que se compararon mediante RT-qPCR con transfectantes pHA-ΔGU para el gen RabX13. Finalmente se estandarizó el ensayo de ChIP en cepas HM1: IMSS para futuros ensayos en transfectantes pHA-Pol II. El siguiente diagrama esquematiza la estrategia seguida:



## VII. Materiales y Métodos

### VII.1 Restricción, construcción de plásmidos y transformación

Para la obtención de la construcción pHA-Pol II, el gen de la RNA polimerasa II de *E. histolytica* se amplificó mediante PCR punto final a partir de DNA genómico extraído mediante reactivo de TRIzol como especifica el fabricante (Invitrogen, No. de catálogo 15596-018). Fueron utilizados oligonucleótidos diseñados con sitios de restricción para las enzimas XmaI y XhoI (Pol II F XmaI, Pol II R XhoI; Anexo 1) (NEB No. de catálogo #R0180L y #R0146L). El resultado de la amplificación se visualizó en gel de agarosa al 1 %. Posteriormente se purificó mediante precipitación con etanol a partir del gel. Para esto se cortaron los fragmentos del gel con las bandas de interés y se colocaron en columnas *Freeze 'N Squeeze DNA Gel Extraction Spin Columns* (BIO-RAD No. de catálogo # 732-6165). Las columnas se incubaron a -20 °C durante 10 minutos y luego se centrifugaron a 10000 rpm por 15 minutos. Al eluido se le añadieron 2 volúmenes de etanol absoluto, 0.1 volúmenes de acetato de potasio al 5 M y 1 µg de glicógeno. Se agitaron y se incubaron a -70 °C durante 1 hora. Pasado el período de incubación se centrifugaron nuevamente las columnas a 15000 rpm por 10 minutos y se desechó el sobrenadante. La pastilla obtenida, se lavó con 1 mL de alcohol al 70 % y se centrifugó 3 minutos a 15000 rpm. Se desechó el sobrenadante y la pastilla se resuspendió en 50 µL de H<sub>2</sub>O mQ estéril. El resultado de la precipitación se verificó en un gel de agarosa al 1 %. La secuencia purificada se insertó en vector pCR2.1 (3.9 kb) para su secuenciación, resultados que arrojaron un 99 % de similitud con la secuencia anotada en la base de datos *AmoebaDB* para el locus EHI\_121760 (Anexo 2). La construcción se sometió a restricción con las enzimas XmaI y XhoI y se procedió a ligarlas con el plásmido de sobreexpresión pEhExHA de *E. histolytica* (vector que confiere resistencia al antibiótico G418) previamente restringido con las enzimas XmaI y XhoI. Para

la reacción de ligación se mantuvo una relación vector e insertos de 1:2 en 4  $\mu\text{L}$  de buffer de ligación 5x con 1  $\mu\text{L}$  de ligasa T4 (Invitrogen No. de catálogo 15224-017) completando a 20  $\mu\text{L}$  con  $\text{H}_2\text{O}$  mQ estéril. La reacción se incubó toda la noche a 16 °C. El resultado de la ligación, se empleó para transformar bacterias *E.coli* cepa TOP 10. Para esto, se mezclaron suavemente 200  $\mu\text{L}$  de bacterias y 10  $\mu\text{L}$  de producto ligado, se incubaron en hielo por 20 minutos, luego se incubaron a 42 °C por 45 segundos para el choque térmico y nuevamente se incubaron las muestras en hielo por 2 minutos. Luego se adicionó a cada tubo de reacción 800  $\mu\text{L}$  de medio súper óptimo (SOB) y se dejaron crecer las bacterias por 1 hora y 20 minutos en agitación a 37 °C. Pasado este período, se tomaron 200  $\mu\text{L}$  de cada transformación y se espatularon en placas con medio sólido Luria-Bertani (LB) con ampicilina (100  $\mu\text{g}/\text{mL}$ ). Las placas fueron incubadas toda la noche a 37 °C. Las colonias obtenidas se crecieron en 3 mL de medio líquido SOB con ampicilina (100  $\mu\text{g}/\text{mL}$ ) durante 18 horas en agitación a 37 °C. Las bacterias se colectaron por centrifugación a 12000 rpm por 1 minuto y se extrajo el DNA plasmídico por miniprep siguiendo el protocolo descrito por (Sal *et al.*, 1988). Las posibles clonas se seleccionaron atendiendo al retraso en la migración respecto al control negativo (pEhExHA vacío) en gel de agarosa al 1 %. Para la comprobación de la eficiencia de ligación se llevaron a cabo ensayos de restricción por 2 horas a 37 °C con las enzimas XmaI-XhoI, EcoRI y Hind III, corroborando la presencia del inserto de interés con base en el patrón de bandas observadas con cada enzima (Hind III: 800 y 10100 pb; EcoR I: 510, 1405, 2776 y 6230 pb; XmaI/Xho I: 4911 y 6000 pb) (Figura 9B).



## VII.2 Transfección de plásmidos en trofozoítos de *E. histolytica* cepa HM1: IMSS

La transfección de la cepa HM1: IMSS de *E. histolytica* se realizó en tubos de 7.5 mL. Se partió de tubos semiconfluentes, se incubaron en hielo durante 35 minutos y las amebas se centrifugaron a 800 rpm por 15 minutos. La pastilla se resuspendió en 6 mL de PBS 1X estéril y se contaron un aproximado de  $8.5 \times 10^5$  amebas. Posteriormente se sembraron  $7.5 \times 10^5$  amebas por tubo, en los cuáles se adicionó 150 mL de medio TYI-S-33 (Diamond *et al.*, 1978) y se incubaron a 37 °C en presencia de CO<sub>2</sub> durante 30 minutos permitiendo la adhesión de las células. Las mezclas de transfección se realizaron en esterilidad, en la primera se agregaron 10 µg del plásmido control pHA-V y en los de sobreexpresión 10 µg de cada uno pHA-CTD, pHA-ΔGU, pHA-Pol II, pHA-CTD/ΔGU y pHA-Dbr-ΔC respectivamente, más 20 µL Superfect (Lipofectamina) (QIAGEN No. de catálogo 301305) y 100 µL de medio M199 (M199 0.476 g, ácido ascórbico 0.005 g, HEPES 0.298 g, L-cisteína 0.05 g, pH 6.8). La mezcla se incubó durante 10 min a temperatura ambiente, luego se agregó el volumen necesario para completar 2000 µL con M199 + SBA 15 %. Posteriormente, se retiró el medio TYI-S-33 de la placa y se colocó la mezcla final de transfección adicionando lentamente sobre la monocapa de amebas de cada pozo, la placa se incubó nuevamente de 4-5 horas. Posteriormente, se completó el volumen de los tubos a 7 mL con medio TYI-S-33 y se incubaron nuevamente en las mismas condiciones durante 24 horas. Posteriormente, se inició la selección con el antibiótico G418 (*Life technologies* No. de catálogo 11811098) a una concentración de 1 µg/mL, la concentración fue aumentando conforme los tubos alcanzaron una confluencia de 80 %, hasta llegar a una concentración de 10 µg/mL de antibiótico en las transfectantes estables. En el caso de la transfectante pHA-Pol II se mantuvo una concentración selectiva de 5 µg/mL de G418. Una vez que se establecieron los cultivos a las concentraciones de antibiótico deseadas, se expandieron en cajas de 25 cm<sup>2</sup>

y se recolectaron las muestras después de 48 horas de incubación para la posterior extracción de RNA.

### **VII.3 Cultivos de trofozoítos de *Entamoeba histolytica* y tratamientos**

Los trofozoítos de la cepa HM1:IMSS de *E. histolytica* sin transfectar y transfectadas con las construcciones (pHA-V, pHA-CTD, pHA- $\Delta$ GU, pHA-Pol II y pHA-CTD/ $\Delta$ GU y pHA-Dbr- $\Delta$ C), se crecieron de forma axénica en medio TYI-S-33 suplementado con Suero Bovino Adulto Inactivado (SBA) al 10 %, con 3% de una mezcla de vitaminas Diamond (North American Biologicals) y penicilina-estreptomicina al 1X en tubos de cultivo (Gibco, No. De catálogo 430720) a 37 °C. Cuando alcanzaron un 80 % de confluencia (se determinó por conteo de células en cámara de Neubauer) se pasaron a frascos de cultivo de 25 cm<sup>2</sup> donde se incubaron toda la noche en las mismas condiciones de incubación para la posterior extracción de RNA. Para los estudios *in vivo*, se trataron las transfectantes (pHA-V y pHA-CTD) a una concentración de 5 mM de ácido bórico (pH 7.9) por 1.5, 2 y 5 horas (Mendoza-Figueroa *et al.*, 2018) y posteriormente se extrajo RNA para posteriores estudios.

### **VII.4 Extracción de RNA**

La extracción de RNA se realizó a partir de amebas crecidas en frascos de cultivo con una confluencia del 95 %. Las células se colectaron sumergiendo los frascos en un baño de hielo-agua por 20 min. La suspensión celular se centrifugó a 1000 rpm durante 5 min a 4 °C, se eliminó el medio de cultivo y se resuspendieron los en el reactivo TRIzol (Invitrogen, No. de catálogo 15596-018). Se transfirió la suspensión a un tubo eppendorf de 1.5 mL y se incubó a 4 °C durante 5 minutos. Posteriormente se le adicionó una mezcla de cloroformo: isoamílico (24:1), las muestras se agitaron vigorosamente y se centrifugaron a 12000 g durante 10 minutos a 4 °C. La fase acuosa se transfirió a un tubo eppendorf nuevo y se agregaron 200  $\mu$ L de cloroformo, se agitó

nuevamente durante 15 segundos y se incubaron a temperatura ambiente durante 3 minutos. Se centrifugaron a 10000 rpm durante 15 minutos a 4 °C, se eliminó el sobrenadante y el precipitado se resuspendió en 225 µL de buffer de extracción SDS 1X [SDS 0.5 %, EDTA 1 mM (pH 8), Tris-HCl 20 mM (pH 7.5)]. Luego se agregó acetato de sodio pH 5.3 a una concentración final de 0.3 M, dos volúmenes de fenol-cloroformo-isoamílico (25:24:1) y se mezcló en vortex durante un minuto, posteriormente se centrifugaron a 10000 rpm por 3 minutos a 4 °C y se recuperó la fase acuosa, a esta se le agregó nuevamente acetato de sodio a una concentración final de 0.3 M, 2.5 volúmenes de etanol absoluto, se mezcló y el RNA se dejó precipitando a -70 °C durante 16 horas. Posteriormente se centrifugaron las muestras a 10000 rpm durante 15 minutos a 4 °C, se eliminó el sobrenadante y el precipitado se lavó con 300 µL de etanol 70 % y nuevamente se centrifugó a 8000 rpm durante 15 minutos a 4 °C. El precipitado se dejó secar a temperatura ambiente y se re-suspendió en 50 µL de agua miliQ estéril. Para una disolución completa del RNA, las muestras se incubaron a 55 °C durante 5 minutos. La integridad del RNA se verificó en un gel de agarosa 2 % y se cuantificó espectrofotométricamente.

Las muestras de RNA se trataron con DNasa I (Sigma-Aldrich No. de catálogo 4716728001), durante 5 minutos a 37 °C. Luego se llevaron a un volumen final de 200 µL con agua miliQ estéril y se agregó un volumen de fenol-cloroformo-isoamílico (25:24:1), se agitaron en vortex durante 30 segundos y se centrifugaron a 13000 rpm durante 3 minutos, posteriormente se recuperó la fase acuosa y a ésta se le agregó un volumen de cloroformo-isoamílico (24:1), se repitió la agitación en vortex y la centrifugación. Se recuperó la fase acuosa y se agregó a cada muestra 1 µL de glicógeno (5 mg/mL) y un volumen de isopropanol frío. El RNA se precipitó a -70 °C durante 16 horas. Posteriormente, las muestras se centrifugaron a 13000 rpm durante 15 minutos a 4 °C, se eliminó el sobrenadante y el precipitado se lavó con 300 µL de etanol al 70 %. Se centrifugó nuevamente bajo las mismas condiciones, se

eliminó el sobrenadante y el precipitado se dejó secar a temperatura ambiente hasta que se evaporó el etanol. A continuación, las muestras fueron resuspendidas en 50  $\mu\text{L}$  de agua miliQ y se cuantificaron espectrofotométricamente.

## **VII.5 RT-PCR**

Para las reacciones de retro-transcripción se empleó el kit RevertAid First Strand cDNA Synthesis siguiendo el protocolo recomendado por el fabricante. La concentración de RNA empleada como molde fue 1.5  $\mu\text{g}$  en todos los casos. Se adicionó 1  $\mu\text{L}$  de hexámeros y se incubó a 65  $^{\circ}\text{C}$  durante 5 minutos. Posteriormente, se agregaron 4  $\mu\text{L}$  de buffer de reacción, 1  $\mu\text{L}$  de dNTPs, 1  $\mu\text{L}$  de RivoLock y 1  $\mu\text{L}$  de M-MLV, completando un volumen final de 20  $\mu\text{L}$ . Se agitó suavemente y se incubó 10 minutos a temperatura ambiente. Posteriormente para el desarrollo de la reacción, la mezcla se incubó a 42  $^{\circ}\text{C}$  por una hora. La inactivación se realizó a 75  $^{\circ}\text{C}$  por 10 minutos. El cDNA obtenido se usó inmediatamente o se almacenó a -20  $^{\circ}\text{C}$  hasta su uso. En todos los casos se empleó un volumen de 1  $\mu\text{L}$  de cDNA ( $\approx$  10 ng de cDNA) para las PCR, completándose con los cebadores específicos para cada gen (Anexo 1), 1  $\mu\text{L}$  de dNTPs, 1X de buffer de PCR 10X, (1.5, 2.5 y 3)  $\mu\text{L}$  de  $\text{MgCl}_2$  (25 mM) y 0.15  $\mu\text{L}$  de Taq DNA Polimerasa a completar con  $\text{H}_2\text{O}$  en un volumen final de 25  $\mu\text{L}$ .

La detección de los flircRNAs, y del transcrito de RabX13 y del gen parental de las clonas se realizó por PCR empleando oligos específicos para cada caso y las condiciones mostradas en la Tabla 1. Como control de carga, se realizó bajo las mismas condiciones la RT-PCR para el gen constitutivo de Actina. Como control negativo, se corrió en cada uno de los experimentos reacciones de PCR utilizando como molde RNA sin tratar con la enzima reverso-transcriptasa. Los productos de PCR se analizaron en geles de agarosa al 2.5%, teñidos con bromuro de etidio para visualizar las muestras. Las

imágenes fueron procesadas con el software ImageJ, que permitió obtener los valores de IntDen de cada una de las bandas y normalizar estos a partir de los valores de IntDen de la actina, del fondo y de los controles en cada experimento.

## **VII.6 PCR en tiempo real (qPCR)**

Para los ensayos de qPCR se empleó el kit KAPA SYBR FAST Universal One-Step qRT-PCR REF: U46510000284 y el equipo CFX96 Real-Time PCR Detection System—IVD siguiendo las indicaciones del fabricante. Los transcritos del gen de la polimerasa II se tomaron como normalizador (gen constitutivo) y contra su expresión en cada transfectante, tratada o no tratada, se calcularon los valores de expresión relativa. (ver Apéndice 1). En cada caso se agregó un volumen de 1  $\mu$ L de cDNA ( $\approx$  10 ng de cDNA) para las qPCR de la polimerasa II, así como en el caso de flicRX13 y mRabX13; se adicionaron los cebadores específicos (Pol II: RNAPIIs, RNAPIIas; mRabX13: Rab2 F, Rab2 R; flicRX13: Rab2BSs, NBRX13) y se completó cada mezcla con 2  $\mu$ L del Master Mix del kit (ver Tabla 2). El análisis de las expresiones relativas de los genes se realizó mediante el método de Livak o  $\Delta\Delta$ Ct (Livak & Schmittgen, 2001).

## **VII.7 *Cross-Linking* inmunoprecipitación anti-HA (CLIP)**

Se partieron de  $8 \times 10^7$  trofozoítos de las transfectantes (pHA-V, pHA-CTD, pHA- $\Delta$ GU y pHA-CTD/ $\Delta$ GU) crecidos en frascos de 50 mL con medio TYI-S-33 hasta una confluencia del 95 %. Se decantó el medio y fueron lavados los frascos con 5 mL de PBS 1X estéril a 37 °C servido por el lado contrario donde está la monocapa de amebas, enjuagando muy suavemente y decantando de nuevo. Posteriormente se agregaron 5 mL de PBS tibio nuevamente y se llevó a cabo el *cross-linking* con UV por 30 minutos. Seguidamente, se agitó el recipiente por 10-15 minutos para despegar las amebas y se cosechan en tubos de 50 mL en frío, centrifugando a 1000 rpm por 5 minutos. Las pastillas

se resuspendieron en 1.5 mL de PBS frío y se trasvasaron a tubos eppendorf y se centrifugaron a 1200 rpm por 5 minutos. Se eliminó cuidadosamente el sobrenadante por aspiración con pipeta y la pastilla se pesó en una balanza granataria (debe pesar entre 100 y 250 mg para 5-6 frascos). Para pastillas de 0.2 gramos se resuspendió en 400  $\mu$ L de HEPES ++ (1 pastilla Complete Mini, 7  $\mu$ L E64 (10  $\mu$ M), 30 U/mL inhibidor de nucleasas, 250  $\mu$ L HEPES 0.4 M, 15  $\mu$ L  $MgCl_2$  1 M, 100  $\mu$ L KCl 1 M, mQ Cbp 10 mL) con el fin de obtener un volumen de 0.6 mL. Se incubó por rotación a 4 °C por 20 min. La suspensión celular cambia de apariencia más opaca a tersa/aperlada. Se añadió 0.15 mL de 10 % NP-40 en HEPES ++ (concentración final de NP-40 2 %) y se mezcló por inversión. La suspensión se clarificó. Se incubó la solución 3 minutos en hielo y posteriormente se centrifugó a 3560 rpm por 10 minutos a 4 °C para empastillar los núcleos. El sobrenadante se apartó como fracción citosólica y se lavaron los núcleos dos veces con 0.3 mL de HEPES ++ centrifugando a 1200 rcf 5 minutos a 4 °C. A continuación, se resuspendieron los núcleos en 0.3 mL de buffer de *splicing* (20  $\mu$ L ATP/CP 25x, 20  $\mu$ L 80 mM  $MgCl_2$ , 25  $\mu$ L HEPES ++, 435  $\mu$ L  $H_2O$  mQ) con 30  $\mu$ L de E64. Los núcleos se rompieron con 5 rondas de congelación (nitrógeno líquido) / descongelación (37 °C). La preparación se clarificó mediante centrifugación a 20000 rcf, 20 minutos a 4 °C. El sobrenadante se guardó como extracto nuclear y 1/10 se apartó como input nuclear control.

Posteriormente, se añadieron 25  $\mu$ L de tRNA a 10 mg/mL como bloqueador/competidor y se preabsorbió en 20  $\mu$ L de proteína G-Sepharosa. La mezcla se incubó en agitación por media hora a temperatura ambiente y luego se centrifugó a 3000 rpm por 3 minutos a 4 °C. El sobrenadante se colectó en 90  $\mu$ L de *slurry* de perlas de HA-agarosa y se incubó por rotación 3 horas a 4°C. Seguidamente se centrifugó a 3000 rpm tres minutos a 4 °C y el sobrenadante se colecta como (fracción no unida). La muestra se lavó 3 veces con buffer de lavado (2.5 mL HEPES 0.4 M, 2.5/25 mL KCl 3 M, 0.125 mL 20

% tritonx-100, completar a 50 mL de H<sub>2</sub>O mQ) por 1 minuto a 10000 rpm. La muestra se dividió en dos para llevar a cabo proteómica resuspendiendo en 600 de buffer de lavado para proteínas y 600 de buffer TES (500 mM Tris-HCl pH 7.5, 10 mM EDTA y 1 % SDS) para extracción de RNA. Se agregó péptido HA (Sigma) a una concentración final de 0.2 mg/mL y se incubó a 65 °C por 10 minutos. Seguidamente se incubó en agitación toda la noche a 4 °C. Pasada la incubación se centrifugó para la precipitación de las perlas y se recuperó el sobrenadante. Se centrifugó a 3000 rpm por 3 minutos a 4 °C y el sobrenadante se colectó y se trató con proteinasa K, DNasa RQ1 y se sometió a posterior extracción de RNA con fenol.

### **VII.8 Estandarización de ensayo de Inmunoprecipitación de cromatina (ChIP)**

Se partió de  $2 \times 10^7$  trofozoítos de la cepa HM1: IMSS los cuales fueron crecidos en medio TYI-S-33. Se cosecharon en la fase exponencial. Las células fueron incubadas en hielo durante 30 minutos. Las amebas se despegaron y se transvasaron a tubos estériles de 50 mL. Se centrifugó a 1500 rpm durante 5 minutos. Se decantó el medio si se lavó una vez con PBS 1X frío. La muestra se centrifugó nuevamente a 1500 rpm durante 15 minutos, se decantó el sobrenadante y se resuspendieron en 1 mL de formaldehído al 1 % durante 10 minutos a temperatura ambiente. El formaldehído es un aldehído pequeño que reacciona con diferentes macromoléculas en varios pasos. Inicialmente, un grupo nucleofílico en una molécula de DNA o en una proteína forma un enlace covalente con el formaldehído dando lugar a una base de *Schiff*. Las bases de *Schiff* son muy inestables, pero se pueden estabilizar ante una segunda interacción molecular para la formación de puentes metileno por los cuales ocurre el *cross-linking* (Hoffman *et al.*, 2015).

Las células fueron colectadas por centrifugación a 1500 rpm por 5 minutos y lavadas 3 veces una vez con PBS 1X frío, con amortiguador I (HEPES 10 mM,

pH 6.5; EDTA 10 mM; EGTA 0.5 mM; Triton X-100 0.25 %) y finalmente con amortiguador II (HEPES 10 mM, pH 6.5; EDTA 1 mM; EGTA 0.5 mM; NaCl 200 mM). Seguidamente las células fueron resuspendidas en 500  $\mu$ L de amortiguador de lisis (Tris 50 mM, pH 8.1; EDTA 10 mM; SDS 1 %) con inhibidores de proteasas 1X (Hernández de la Cruz *et al.*, 2016). Se sonicó la muestra 16 veces en sonicador modelo GEX 130 en pulsos de 30 segundos y descansos de 30 segundos con una amplitud de 65 % y 75 % de un voltaje de 115V a una frecuencia de 20 kHz, potencia: 130W y corriente 3.0 A para producir fragmentos de un tamaño aproximado de 300-500 pares de bases. Los resultados de la fragmentación en cada pulso se monitorearon en geles de agarosa al 1.5%. Una alícuota de los pulsos 15 y 16 se sometieron a *reverse-crosslinking* incubando a 65 °C durante dos horas en 200 mM de NaCl. Posteriormente fueron tratados con RNasa A, 15 minutos a temperatura ambiente y purificados en columnas *Freeze 'N Squeeze DNA Gel Extraction Spin Columns* (BIO-RAD No. de catálogo # 732-6165) centrifugando a 10000 rpm por 15 minutos. El eluido se precipitó con 2.5 volúmenes de etanol absoluto, 0.2 volúmenes de acetato de potasio 5 M y 1  $\mu$ g de glicógeno. Se agitaron y se incubaron a -70 °C durante 1 hora. Pasado el período de incubación, las columnas se centrifugaron nuevamente las a 13000 rpm por 15 minutos y se desechó el sobrenadante. El precipitado se lavó con 1 mL de etanol al 70 % y se volvió a centrifugar durante 5 minutos a 13000 rpm. Se desechó el sobrenadante y la pastilla se resuspendió en H<sub>2</sub>O mQ estéril. La concentración de DNA se midió en Nanodrop lite (*Thermo Scientific*) a una absorbancia de 260/280 y se llevó a cabo PCR punto final utilizando cebadores para amplificación de los genes constitutivos Hsp70 (HSP70s, HSP70as) y Enolasa (EnolsF, EnolsR) (Anexo 1).

## VII.9 Análisis estadístico

Se utilizó el software GraphPad Prism versión v5.01 para el análisis de los datos (Radushev, 2007). Todos los experimentos se realizaron por duplicado,



se realizaron análisis de varianzas (ANOVA) de clasificación simple con pruebas de Tukey *a posteriori* para el análisis de los datos. En todos los casos se asumieron diferencias significativas para  $p \leq 0.05$  y  $p \leq 0.01$ .

## VIII. Resultados

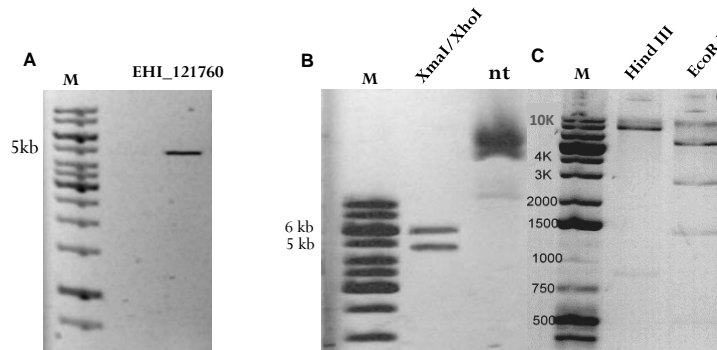
### VIII.1 Transfección en cepas HM1-IMSS

Para el desarrollo de los ensayos subsecuentes, procedimos a la obtención de cepas HM1-IMSS transfectantes con diferentes construcciones que garantizaban la expresión exógena del extremo carboxilo terminal de la Pol II (pHA-CTD), así como la subunidad mayor de la RNA polimerasa II (pHA-Pol II) fusionados a un péptido de hemaglutinina en su extremo amino, como bandera. Como controles de experimentación fueron transfectadas cepas con el vector de sobreexpresión vacío (pHA-V), con la construcción que contiene el minigen mutado en la secuencia 5' ss del intrón del gen RabX13 (pHA- $\Delta$ GU) y una doble transfección (pHA-CTD/ $\Delta$ GU).

Para obtener la construcción pHA-Pol II se procedió a la amplificación del gen de la RNA polimerasa II de *E. histolytica* (Figura 9A). Posteriormente se purificó e insertó en vector pCR2.1 para su secuenciación usando los cebadores del propio vector M13F y M13R (Anexo 1), resultados que arrojaron un 99 % de similitud con la secuencia anotada en la base de datos *AmoebaDB* para el locus EHI\_121760 (Anexo 2). La construcción se sometió a restricción con las enzimas XmaI y XhoI y el amplicón se insertó en vector de sobreexpresión pEhExHA. Para la comprobación de la eficiencia de clonación, se llevaron a cabo ensayos de restricción usando las enzimas XmaI-XhoI, EcoRI y Hind III, corroborando la presencia del inserto de interés en base al patrón de bandas observadas con cada enzima (Hind III: 800 y 10100 pb; EcoRI: 510, 1405, 2776 y 6230 pb; XmaI/XhoI: 4911y 6000 pb) (Figuras 9A-C).

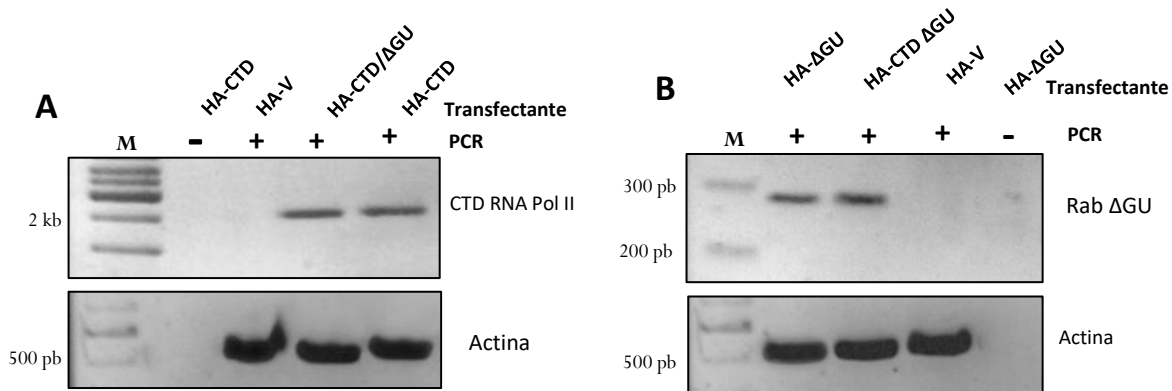
Con las construcciones pHA-V, pHA-CTD, pHA- $\Delta$ GU, pHA-Dbr1 $\Delta$ C y pHA-Pol II ya disponibles, se procedió a la transfección y selección con geneticina-418 hasta alcanzar una concentración de 10  $\mu$ g de antibiótico por mL de medio

mientras que las clonas pHA-Pol II se mantuvieron a menor presión selectiva con valores de 5 µg/mL.



**Figura 9: Obtención de la construcción pHA-Pol II . (A)** Amplificación por PCR del gen de la RNA Polimerasa II (4911 pb) con el uso de los cebadores Pol II F Xma I y Pol II R Xho I (Anexo 1). Carril 2: locus EHI\_121760. **(B)** Ensayos de restricción del DNA de la construcción pHA-Pol II con las enzimas XmaI-XhoI, sin tratar (nt), Hind III y EcoR I. Geles de agarosa al 1%. M: marcadores de peso molecular.

Para determinar si la disminución de los flicRX13 por la mutación al 5' ss del intrón del gen RabX13 provocaba cambios en la unión de flicRX13 a la CTD de la Pol II, realizamos la doble transfección de los plásmidos pHA-CTD y pHA-ΔGU, haciéndose necesario el control de que ambos plásmidos estuvieran presentes en la cepa doble transfectante. Así, se procedió a la identificación de los insertos CTD y ΔGU marcados por la bandera de hemaglutinina, mediante PCR a partir del DNA genómico extraído de las cepas en estudio. Fueron utilizadas combinaciones de cebadores que garantizaran la obtención del fragmento de interés (CS+116 y Rab2Rex 2 para ΔGU; CS+116 y Pol II R Xho I para CTD (Figuras 10A y B). Como se puede observar, se logró identificar específicamente ambos amplicones marcados a los tamaños estimados.

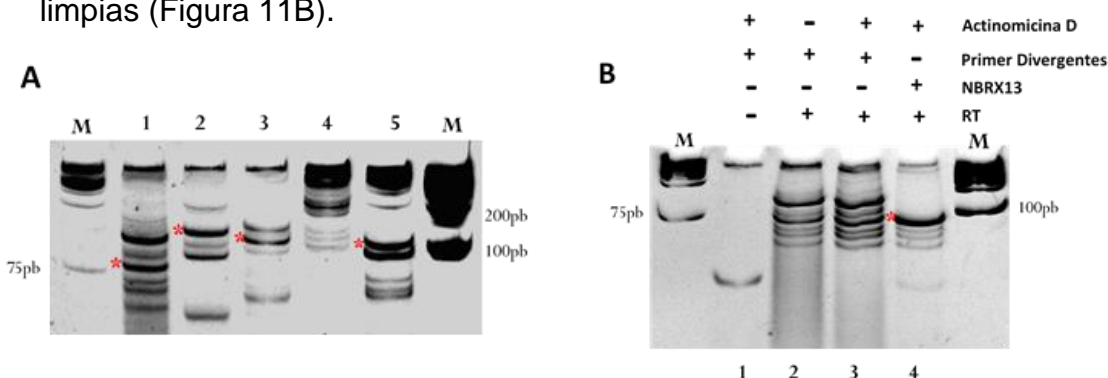


**Figura 10: Presencia de plásmidos en las transfectantes pHA-CTD, pHA-ΔGU y pHA-CTD/ΔGU.** (A) Amplificación por PCR del carboxilo terminal de la RNA Polimerasa II marcado con cola de hemaglutinina (2240 pb) con el uso de los cebadores CS+116 y Pol II R Xho I, gel de agarosa al 1%. (B) Amplificación por PCR del minigen ΔGU marcado con cola de hemaglutinina (287 pb) con el uso de los cebadores CS+116 y Rab2Rex 2, gel a agarosa al 2%. La tranfectante pHA-V se utilizó como control negativo. Debajo se observa la amplificación del gen de actina (control de carga) en gel a agarosa al 2%. Los experimentos se realizaron en presencia (+) y ausencia (-) de reacción de PCR. M, marcadores de peso molecular.

## VIII.2 Identificación de flicRNAs en cepas transfectantes

A modo de control se identificó la presencia de los flicRNAs correspondientes al gen de referencia RabX13, de los genes codificantes de proteínas ribosomales RpL12 y RpS14 y los loci asociados a virulencia EHI\_169670, EHI\_014170 y EHI\_192510 en las transfectantes obtenidas mediante RT-PCR divergente utilizando además sondas de reconocimiento de la unión covalente 3'-5' presente en los flicRNAs. Con asteriscos en rojo se muestran los flicRNAs identificados (Figura 11A). No se logró identificar el flicRNA proveniente del gen que codifica para el canal de cloro Cl2B en ninguna cepa ni en los estudios posteriores. A pesar de la baja tasa de error de la taq polimerasa y las retrotranscriptasas, durante las reacciones de RT y PCR, tienden sintetizarse secuencias espurias mayormente en el caso de la retrotranscripción. Así, consideramos que la amplificación inespecífica observada durante la identificación de los flicRNAs es producto de la formación de artefactos (Figura 11A).

Con el fin de eliminar el DNA contaminante, llevamos a cabo las RT-PCR en presencia de 200  $\mu$ M de actinomicina D (ActD), antibiótico inhibidor de la transcripción y la replicación del DNA, que usado en reacciones de PCR previene la síntesis de copias de DNA inespecíficas (Goldberg *et al.*, 1962; Perocchi *et al.*, 2007). De este modo, se obtuvieron las bandas de interés más limpias (Figura 11B).



**Figura 11: Identificación de FlicRNAs.** (A): Carril 1: flicRX13, carril 2: flic014170, carril 3: flic169670 Intrón I, carril 4: flic192510, carril 5: flicRpL12..(B): RT-PCR con (+) y sin (-) actinomicina D para la eliminación de artefactos durante la reacción de amplificación de flicRX13 con primers divergentes Rab2BSs y Rab2BSas y usando la sonda NBRX13 de reconocimiento a la unión 5'-3'. Carril 1: Control negativo en ausencia de RT. Carril 2: Reacción con primers divergentes en ausencia de ActD. Carril 3: Reacción con primers divergentes en presencia de ActD. Carril 4: Reacción con sonda NBRX13 en presencia de ActD. FlicRNAs marcados con asteriscos en rojo. M, marcadores de peso molecular. Los productos fueron resueltos en geles de acrilamida al 12%.

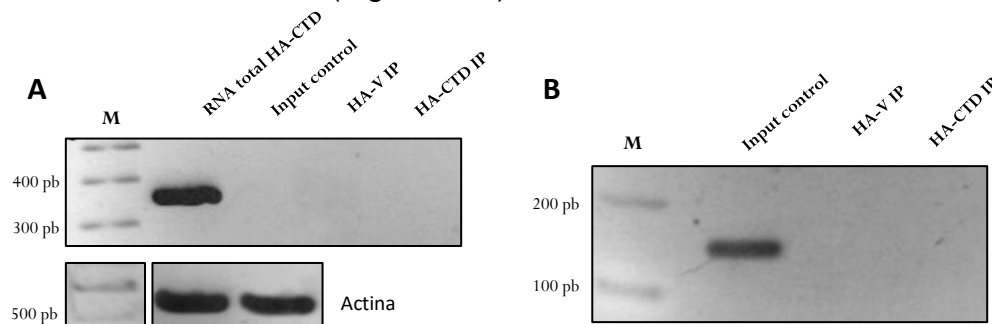
### VIII.3 *Cross-linking* Inmunoprecipitación anti-HA (CLIP) en cepas pHA-CTD

En este estudio nos propusimos identificar la unión de los flicRNAs al carboxilo terminal de la RNA polimerasa II mediante estudios de CLIP a partir de las transfectantes pHA-CTD y pHA-CTD/ $\Delta$ GU, las cuales presentan marcado el CTD con péptido de hemaglutinina. A partir de  $8 \times 10^7$  trofozoitos en botellas semiconfluentes de 50 mL, sometidos a *cross-linking* con UV por treinta

minutos, se procedió a la inmunoprecipitación con anticuerpos anti-HA y al aislamiento de RNA e identificación de flicRNAs por RT-PCR circular.

### VIII.3.1 Eficiencia del ensayo de inmunoprecipitación

Uno de los principales problemas de los protocolos que implican el fraccionamiento celular radica en la contaminación de la fracción nuclear con la citoplasmática o viceversa, por lo que los controles de experimentación son esenciales para garantizar resultados adecuados (Valdés *et al.*, 2014; Talhouarne & Gall, 2018). Consideramos que, como se ha realizado para proteínas, es necesario analizar la presencia de RNAs de localización citoplasmática en los pasos posteriores a la obtención de la fracción nuclear como indicativo de contaminación. Es por ello, que identificamos al circRNA exónico (circEh169670 exón II), como control de la eficiencia de fraccionamiento celular (Figura 12A). Este circRNA fue recientemente identificado en nuestro laboratorio mediante análisis *in silico* en cepas de *E. histolytica* y validado *in vivo* en este trabajo. Como fue mencionado anteriormente, los RNA circulares exónicos son los más abundantes de las especies circulares y presentan localización citoplasmática (Holdt *et al.*, 2018), por lo que resulta útil para este propósito. El exón fue identificado a partir de RNA total de cepas pHA-CTD y apareció ausente en los pasos subsiguientes a la lisis de la fracción nuclear, lo cual indica alta eficiencia durante el fraccionamiento celular (Figura 12A).



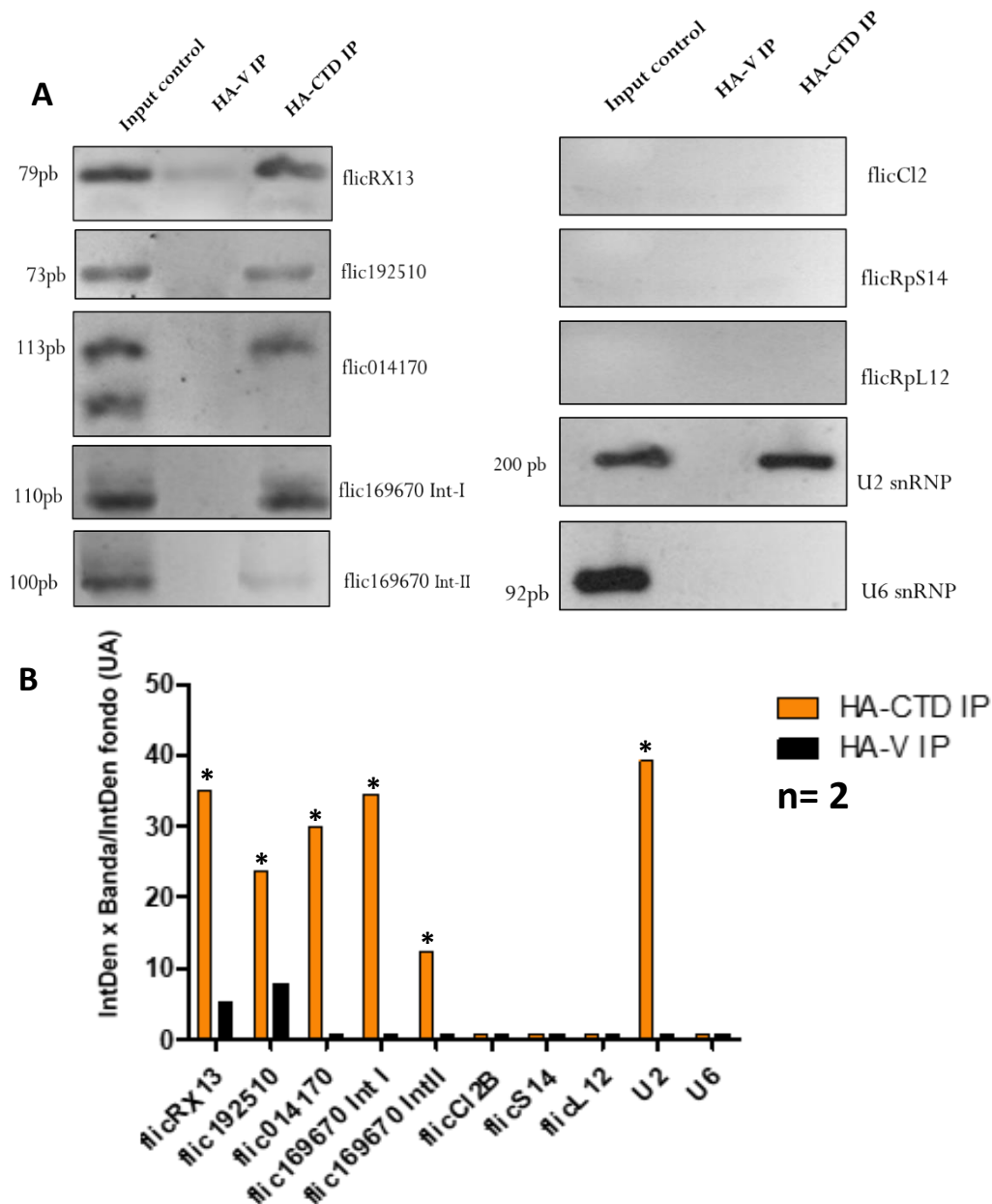
**Figura 12: Eficiencia de fraccionamiento celular y especificidad de inmunoprecipitación anti-HA.** (A): Identificación de circEh169670 exón II a partir de RNA total, en el input de la fracción nuclear y en las inmunoprecipitaciones anti-HA de las transfectantes pHA-V y pHA-

CTD usando los cebadores (EHI169670\_E2s y EHI169670\_E2as). Debajo se observa la amplificación del gen de actina como control de carga (EhActF y EhActR). **(B)**: Identificación del transcrito de Hsp70 a partir del input de la fracción nuclear e inmunoprecipitación anti-HA en las transfectantes pHA-V y pHA-CTD usando los cebadores (HSP70s y HSP70as). Los productos fueron resueltos en geles de agarosa al 2.5 % o al 2 % para actina. M, marcadores de peso molecular.

Para el análisis de la especificidad de inmunoprecipitación procedimos a la identificación del transcrito del gen constitutivo Hsp70 en la fracción nuclear y las inmunoprecipitaciones específicas (Figura 12B). Se logró identificar a Hsp70 en el extracto nuclear pero no en la inmunoprecipitación problema, lo cual indica que este transcrito no se encuentra interaccionando con el complejo de asociación a la Pol II y permite concluir que el protocolo de inmunoprecipitación anti-HA presenta alta especificidad.

### VIII.3.2 CLIP

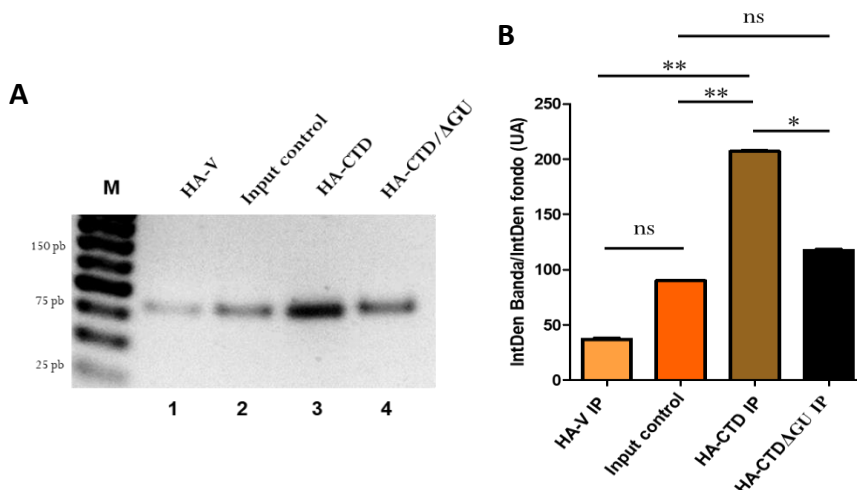
Posterior a la comprobación de los parámetros de control de experimentación, se procedió a identificar la asociación de los flicRNAs con la CTD de la Pol II. Mediante RT-PCR divergente a partir del RNA obtenido de la inmunoprecipitación anti-HA, identificamos los flicRNAs del gen RabX13 y los loci asociados a virulencia EHI\_169670, EHI\_014170 y EHI\_192510 en unión a la CTD de la Pol II en dos ensayos CLIP diferentes (n=2), sugiriendo la participación en la regulación en *cis* (Figura 13). Los resultados muestran diferencias estadísticamente significativas en todos los flicRNAs identificados en comparación con sus respectivos controles negativos (Figura 13B). Resultados similares fueron encontrados por Zhang *et al.* (2013) al estudiar la asociación de los ciRNAs a la RNA Pol II, observando que ciRNAs abundantes como ci-*ankrd52*, se encontraban acumulados en el sitio de inicio de la transcripción de su gen parental y asociados al complejo de elongación de la Polimerasa.



**Figura 13. Asociación de los flicRNAs a la CTD de la RNA Pol II.** (A) Identificación de flicRNAs por RT-PCR divergente a partir del RNA nuclear (Carril 1), de RNA aislado de la inmunoprecipitación anti-HA en transfectantes pHA-V (Carril 2) y pHA-CTD (Carril 3) en gel de agarosa al 2.5 %. Los pesos moleculares correspondiente a cada flicRNAs se muestra a la izquierda de cada gel. (B) Estudio densitométrico comparativo entre las inmunoprecipitaciones control (pHA-V) y problema (pHA-CTD) para una n=2. Los resultados se analizaron mediante anova de clasificación simple con prueba de Tukey *a posteriori* para la comparación de medias a un nivel de significancia \*:  $p \leq 0.01$ .



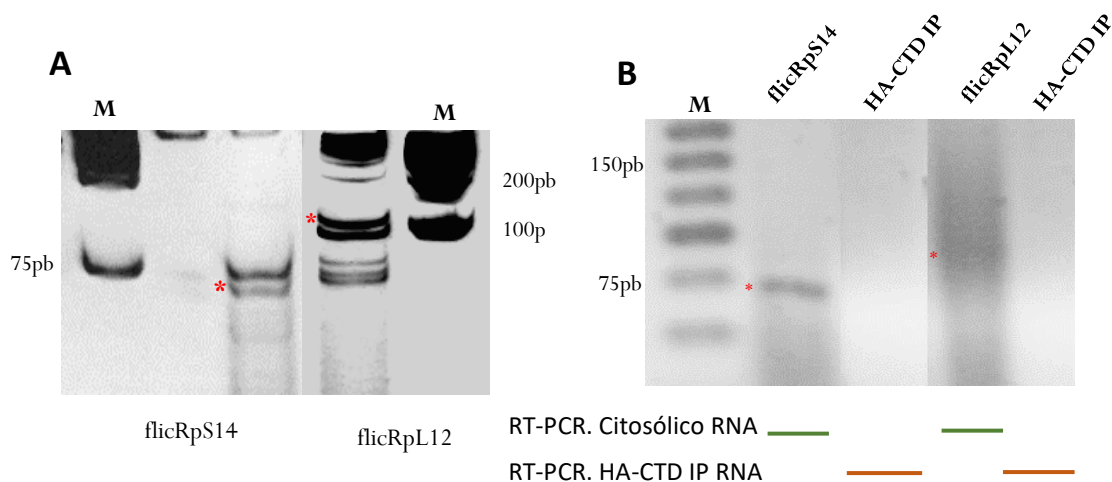
También llevamos a cabo ensayos de CLIP en transfectantes pHA-CTD/ $\Delta$ GU, cepa que presenta mutaciones en la secuencia 5' ss del intrón del gen RabX13 y pudimos observar una disminución estadísticamente significativa de la unión de flicRX13 a la CTD de la Pol II en comparación con los ensayos de CLIP en transfectantes pHA-CTD (Figura 14A y B).



**Figura 14. Comparación de la presencia de flicRX13 en complejo con la CTD de la Pol II en cepas pHA-CTD y pHA-CTD/ $\Delta$ GU.** (A) Identificación de flicRX13 por RT-PCR divergente a partir del RNA aislado de la inmunoprecipitación anti-HA en transfectantes pHA-V (Carril 1), de RNA nuclear de cepas pHA-CTD/ $\Delta$ GU (Carril 2) y de RNA aislado de la inmunoprecipitación anti-HA en transfectantes pHA-CTD (Carril 3) y pHA-CTD/ $\Delta$ GU (Carril 4) en gel de agarosa al 2.5%. (B) Densitométris de los resultados en A comparativo entre las inmunoprecipitaciones pHA-V, pHA-CTD, pHA-CTD/ $\Delta$ GU para flicRX13. Los resultados se analizaron mediante anova de clasificación simple con prueba de Tukey *a posteriori* para la comparación de medias a un nivel de significancia  $p \leq 0.01$ .

Sorprendentemente no fueron identificados los flicRNAs de los genes codificadores de proteínas ribosomales RpS14 y RpL12 en unión a la CTD de la Pol II y tampoco en el input nuclear (Figura 13A), sin embargo, fueron identificados a partir de RNA total tanto en cepas silvestres (wt), pHA-CTD como en pHA-V (Figura 15A).

Estos resultados indicaban una posible localización citoplasmática de flicRps14 y flicRpL12 por lo que procedimos a la extracción de RNA citoplasmático y la identificación de estos. En los primeros ensayos de RT-PCR no se obtuvieron resultados satisfactorios y supusimos que podría deberse a la inhibición de la reacción de amplificación por algún componente de la muestra. Por lo que fue tratado el RNA citosólico con Proteinasa K y DNasa RQ1 para la limpieza de contaminantes en la muestra. En nuevos ensayos de RT-PCR, se logró identificar la presencia de flicRpS14 en el citoplasma. En el caso de flicRpL12, se observa una amplificación inespecífica aproximadamente a la altura del amplicón esperado de 100 pb, por lo que podría estar ocurriendo aún la inhibición de la reacción por la presencia de carbohidratos o sales en la muestra (Figura 15B).



**Figura 15. Identificación de flicRNAs de los genes codificantes de proteínas ribosomales Rps14 y RpL12.** (A): Identificación de flicRNAs por RT-PCR divergente a partir del RNA nuclear (Carril 1) y del RNA aislado de la inmunoprecipitación anti-HA en transfectantes pHA-V (Carril 2) y pHA-CTD(Carril 3) en gel de agarosa al 2.5%. (B): Identificación de flicRpS14 y flicRpL12 mediante RT-PCR divergente a partir de RNA total en gel de agarosa al 2.5 %. (C): Identificación de flicRpS14 y flicRpL12 mediante RT-PCR divergente a partir de RNA citosólico en gel de agarosa al 2.5 %. FlicRNAs marcados con asteriscos en rojo. M, marcadores de peso molecular.

### VIII.3.3 Localización de los snRNAs U2 y U6

Múltiples evidencias sugieren que la síntesis del pre-mRNA y el mecanismo de *splicing* son eventos coordinados e interdependientes. Estudios de inmunoprecipitación contra la RNA polimerasa II mostraron la unión de los elementos del spliceosoma catalítico U2 y U5 snRNA de manera funcional a la CTD fosforilada de la Pol II durante la elongación de la transcripción (McKay & Johnson, 2011; Nojima *et al.*, 2018). Con base en estas evidencias, decidimos utilizar U2 snRNA como control positivo de unión a la CTD de la Pol II en nuestros ensayos de inmunoprecipitación.

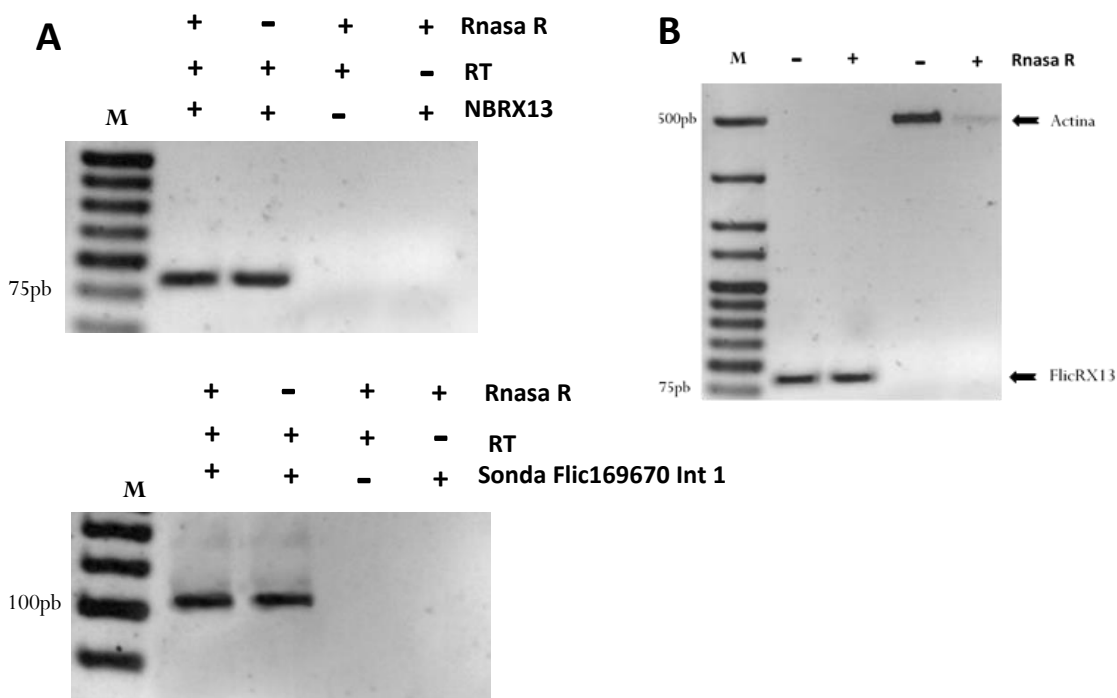
Como se puede observar en la figura 13A, hubo un reconocimiento marcado de U2 snRNA en unión a la CTD, lo cual resulta muy informativo, pues nos permite hipotetizar acerca de la implicación de los flicRNAs en la regulación de la elongación de la transcripción, etapa en la que se encuentran unidos U2 y U5 snRNA a la CTD fosforilada en la serina 2 (Nojima *et al.*, 2018).

El snRNA U6 es un transcrito de la RNA Polimerasa III esencial para la formación del spliceosoma catalítico. Sin embargo, no hay evidencia de su unión directa a la RNA Pol II (Nojima *et al.*, 2018), lo cual podría estar dado por la propia organización de complejo spliceosomal en unión a la Pol II que impide la interacción de este factor con la CTD fosforilada. De este modo, se utilizó como control negativo de unión en los ensayos de inmunoprecipitación con resultados esperados y en concordancia a los reportado por Nojima *et al.*, (2018) (Figura 13A).

### VIII.3.4 Validación de CLIP, tratamiento con RNasa R

Para verificar la circularidad de los flicRNAs en unión a la CTD y descartar identificaciones inespecíficas, una fracción del RNA inmunoprecipitado se trató con RNasa R con el fin de degradar RNAs lineales previo a la reacción de RT-PCR divergente con sonda de reconocimiento a la unión covalente 5´-3´ para cada caso. Los resultados nos permitieron concluir que los flicRNAs

identificados son RNA intrónicos circulares *bona fide* al no observarse diferencias en la amplificación en presencia o ausencia de tratamientos con RNasa R (Figura 16A). Para probar la eficiencia de esta exorribonucleasa, adicionalmente se analizó el efecto sobre el transcrito lineal de actina observandose una marcada disminución en el número de copias del mismo (Figura 16B).



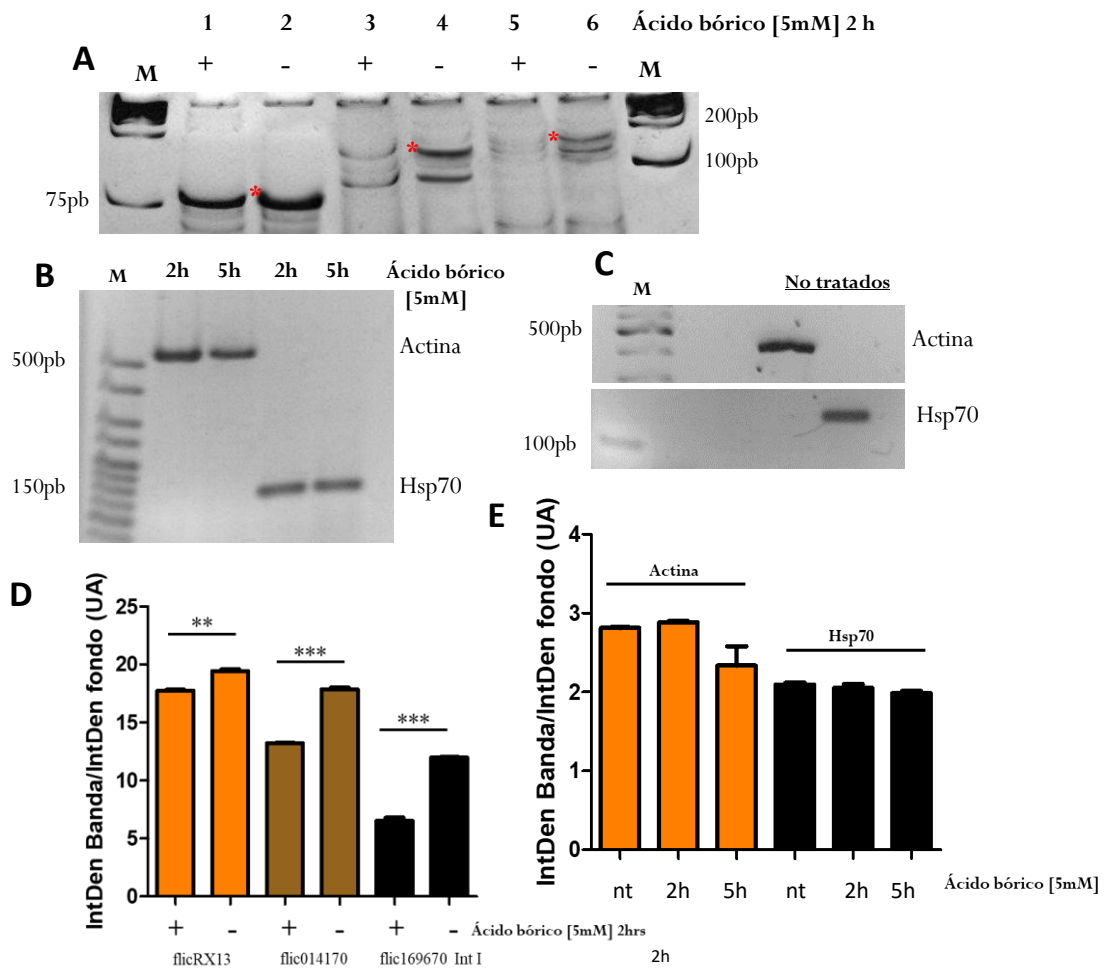
**Figura 16. Tratamientos con RNasa R y validación de flicRNAs** (A): Identificación de flicRNAs de RabX13 y del intron 1 del locus EHI\_169679 por RT-PCR divergente a partir del RNA tratado y no tratado con RNasa R. en gel de agarosa al 2.5 %. (B): Análisis de la eficiencia de degradación de la exorribonucleasa mediante la obtención de cDNA del transcrito lineal de actina a partir de input de RNA tratado y no tratado con RNasa R. M, marcadores de peso molecular.

#### VIII.4 Tratamiento con ácido bórico en cepas transfectantes

Con el objetivo de correlacionar la actividad de los flicRNAs con la expresión de sus genes parentales, establecimos un control donde garantizamos la disminución en la producción de círculos para así observar su efecto sobre la transcripción génica. Para ello, se trataron las cepas transfectantes pHA-CTD y pHA-V con ácido bórico ( $H_3BO_3$ ) 5 mM por 1.5, 2 y 5 horas (Figura 17A-E). El ácido bórico es un agente bacteriostático, fungistático y protozoostático que ha sido reportado como inhibidor reversible de la segunda reacción de transesterificación del *splicing*, tanto *in vitro* como *in vivo* (Shomron & Ast, 2003; Brittingham & Wilson, 2014; Mendoza-Figueroa *et al.*, 2018).

A partir del análisis densitométrico del RT-PCR entre cepas tratadas y no tratadas por un tiempo de 2 horas observamos una disminución estadísticamente significativa ( $p \leq 0.05$ ) en la producción de cada uno de los flicRNAs en estudio, en concordancia a lo reportado *in vivo* en *E. histolytica* por Mendoza-Figueroa *et al.* (2018), donde observaron una disminución de los productos post-*splicing* para flicRX13 y flicRpS14 ante el uso de ácido bórico (Figura 17A y D).

Para descartar efectos sobre la transcripción global durante los tratamientos, se analizó la expresión de los genes constitutivos que codifican para actina y la proteína chaperona Hsp70 a las 2 y 5 horas de incubación con  $H_3BO_3$  (Figura 17B y C). Los estudios densitométricos no arrojaron diferencias significativas entre los tiempos de incubación en comparación con cepas no tratadas, lo cual nos permite asegurar que la acción del  $H_3BO_3$  no afecta eventos de transcripción global en los tiempos de incubación y a las concentraciones utilizadas en este estudio (Figura 17E), mientras que se observa una disminución significativa de los flicRNAs, evidencia de una acción específica sobre el mecanismo de *splicing*.



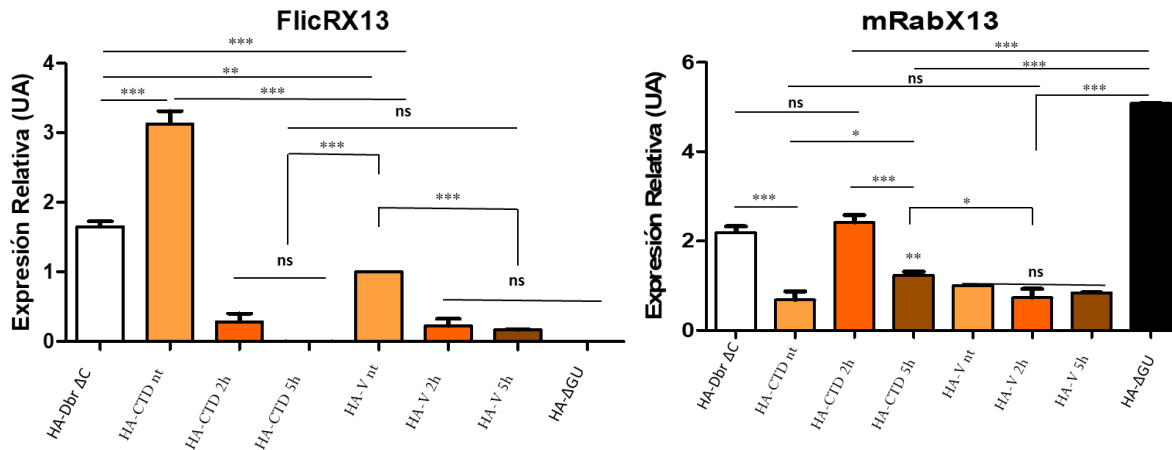
**Figura 17: Efectos del  $H_3BO_3$  en la transcripción global y la producción de flicRNAs .**

(A): RT-PCR divergente para la identificación de flicRNAs en cepas tratadas y no tratadas con  $H_3BO_3$  en gel de acrilamida al 12%. Carril 1-2: flicRX13, carril 3-4: flic014170, carril 5-6: flic169670 Intrón I. flicRNAs marcados con asteriscos en rojo. A ambos extremos, corren marcadores de peso molecular. (B): RT-PCR para identificación de los transcritos actina y Hsp70 en cepas tratadas con  $H_3BO_3$  por 2 y 5 horas en gel de agarosa al 2.5%. (C): RT-PCR control de actina y Hsp70 en cepas no tratadas. (D): Análisis densitométrico de los niveles de producción de círculos en cepas tratadas y no tratadas con  $H_3BO_3$ . (E): Análisis densitométrico de los niveles de producción de transcritos de lo genes constitutivos actina y Hsp70 en cepas tratadas y no tratadas con  $H_3BO_3$ . Los resultados se analizaron mediante anova de clasificación simple con prueba de Tukey *a posteriori* para la comparación de medias a un nivel de significancia de ( $p \leq 0.05$ ). M, marcadores de peso molecular.

En estudios previos, donde fueron obtenidas cepas con mutación en la secuencia rica en GU al 5' del intrón del gen RabX13 (pHA- $\Delta$ GU), se observó una disminución significativa de la producción de flicRX13 y a su vez, un aumento de los niveles de transcrito parental, lo cual permitió concluir acerca de la inhibición constitutiva ejercida por flicRX13 sobre su gen parental. Por otro lado, se observó un incremento en la producción de flicRX13 en transfectadas transitorias (pHA-Dbr $\Delta$ C), con la enzima Dbr1 catalíticamente deficiente a baja presión selectiva (1.5  $\mu$ g/mL G-418), lo cual corroboró el origen post-*splicing* de flicRX13 (Mendoza-Figueroa *et al.*, 2018). Con base en estos resultados, diseñamos un estudio comparativo de los efectos del ácido bórico sobre la producción de flicRX13 y la expresión de su gen parental mRabX13 mediante RT-qPCR en cepas pHA-CTD, pHA-V tratadas durante 2 y 5 horas con ácido bórico en comparación con cepas no tratadas y con las transfectantes antes mencionadas pHA- $\Delta$ GU y pHA-Dbr $\Delta$ C (Figura 18).

Los resultados permitieron observar que ante la disminución en la producción de flicRX13 aumentan los niveles de transcrito mRabX13 en cepas pHA-CTD tratadas durante dos horas. A las cinco horas de tratamiento, disminuye aún más la producción de flicRNAs, sin embargo, comienzan a restablecerse los niveles de transcrito basales (Figura 18).

En las transfectantes pHA- $\Delta$ GU se eliminó la producción de flicRX13 totalmente, lo cual ocasionó un incremento sustancial de los niveles de transcrito mRabX13. Por su parte las transfectantes pHA-Dbr $\Delta$ C mantuvieron elevados los niveles de flicRX13 con la expresión de transcrito relativamente alta, resultados en concordancia a lo reportado por Mendoza-Figueroa *et al.* (2018). En el caso de las transfectantes pHA-V tanto los niveles de flicRX13 y transcrito se mantuvieron bajos. La producción de flicRX13 disminuyó significativamente en cepas tratadas con ácido bórico, sin embargo, los niveles de transcrito no presentan diferencias significativas.



**Figura 18: Efectos del  $H_3BO_3$  en la producción de flicRX13 y los niveles de transcrito mRabX13.** RT-qPCR comparativa de los niveles de flicRX13 (A): y de transcrito mRabX13 (B) en transfectantes pHA-V y pHA-CTD tratadas y no tratadas con  $H_3BO_3$  por 2 y 5 h. Como controles se usaron los niveles de producción de flicRX13 en las transfectantes pHA- $\Delta$ GU y pHA-Dbr $\Delta$ C sin tratar. Los resultados se analizaron mediante anova de clasificación simple con prueba de Tukey *a posteriori* para la comparación de medias a un nivel de significancia ( $p \leq 0.01$ ).

### VIII.5 Estandarización de ensayo de Inmunoprecipitación de cromatina (ChIP) en cepas HM1-IMSS

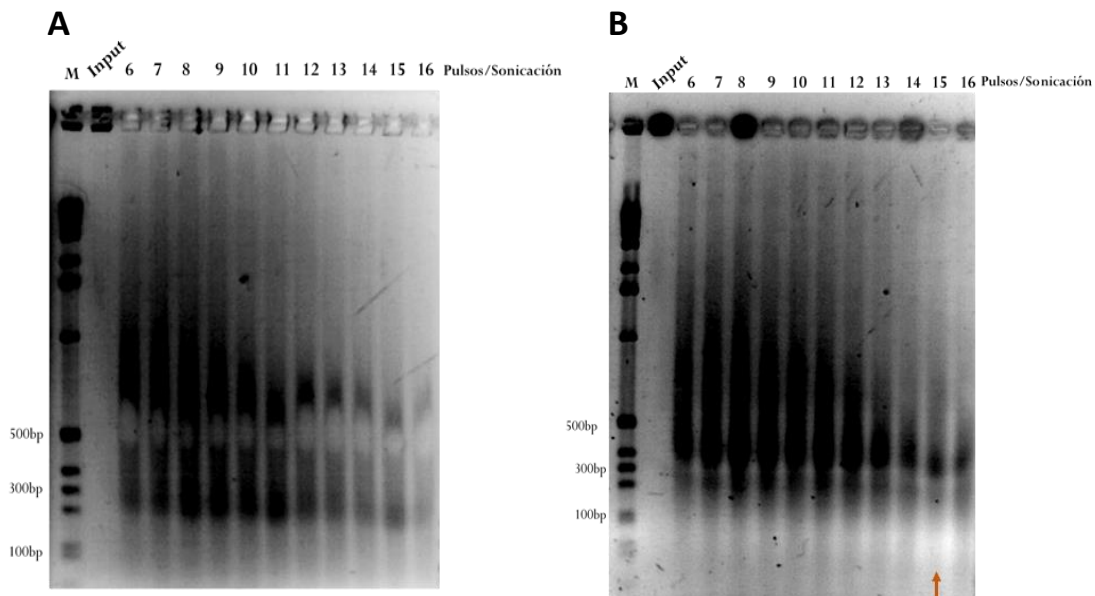
La inmunoprecipitación de cromatina (ChIP) es una herramienta comúnmente utilizada para estudiar interacción entre DNA y proteínas. De manera general el ensayo consiste en la fijación de DNA y proteínas con formaldehído seguido de una serie de pasos de lisis con el fin de separar los componentes celulares y retener la cromatina nuclear. Posteriormente, se lleva a cabo la ruptura del genoma y la purificación de los fragmentos de DNA de interés que deben tener entre 300-600 pb, basados en el tamaño promedio de los promotores eucarióticos (Bortz & Wamhoff, 2011).

Con el fin de establecer los parámetros de fragmentación del DNA y analizar las condiciones adecuadas de *cross-linking* con formaldehído en cepas HM1-



IMSS, se cultivaron un total de  $2 \times 10^7$  trofozoítos los cuales, posterior a su colecta, se resuspendieron en 1 mL de formaldehído al 1 % durante 10 minutos a temperatura ambiente para realizar el *cross-linking*.

Posterior a varios lavados y pasos de lisis para la extracción del genoma, la muestra se sonicó 16 veces en pulsos de 30 segundos con 30 segundos de descanso, probando dos amplitudes de sonicación (65 % y 75 %) con el fin de obtener fragmentos de DNA entre 300-500 pares de bases. Como se puede observar en la figura 19B, el uso de amplitudes de sonicación del 75 % aumenta la eficiencia de fragmentación de la cromatina, permitiendo obtener un tamaño de DNA deseado entre 300-400 pb. En contraste, trofozoítos sonicados a una amplitud de 65 %, mantuvieron fragmentos mayoritarios de DNA de alrededor de 700-800 pb en todos los pulsos (Figura 19A).

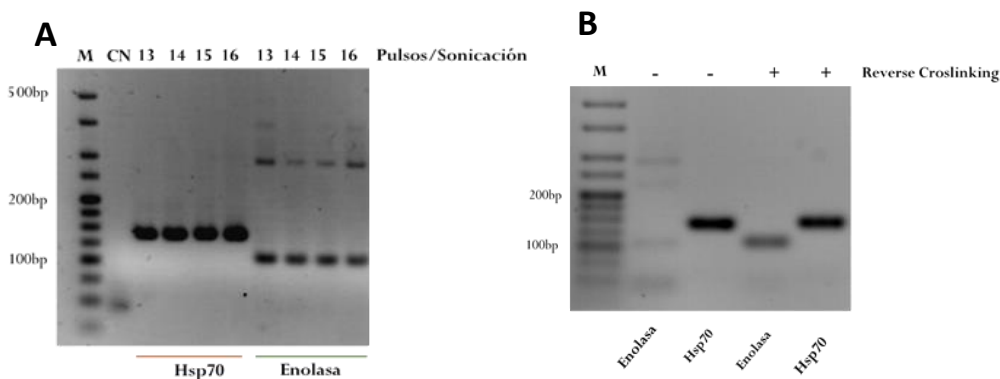


**Figura 19: Efectos de los tiempos de sonicación en la fragmentación de la cromatina en suspensiones celulares frescas.** Partiendo de  $2 \times 10^7$  trofozoítos, se sonicaron en intervalos de 30 segundos durante un total de 11 minutos y se apartó un volumen de 20  $\mu$ L en cada pulso de sonicación para verificar los niveles de fragmentación en gel de agarosa 1.5 %. Fragmentación a amplitud de 65 % (A) y 75 % (B). Carril 2: input no sonicado, carriles 3-13: pulsos de sonicación del 6 al 16. La flecha naranja indica el número de pulsos óptimo para la

obtención de fragmentos de DNA mayoritarios de 300-400 pb. Geles de agarosa al 1.5 %. M, marcadores de peso molecular.

Para analizar la integridad del DNA en los fragmentos obtenidos de los pulsos 13-16 de sonicación a 75 % de amplitud se trataron con RNasa A y se purificó de la presencia de sales y carbohidratos que podrían inhibir la reacción de PCR con las columnas (BIO-RAD *Freeze 'N Squeeze DNA*). Posteriormente, se midió la concentración de DNA purificado y se llevó a cabo PCR punto final para la identificación de fragmentos de los genes constitutivos Hsp70 (150 pb) y enolasa (100 pb) (Figura 20A). Se logró amplificar los fragmentos de los genes controles pero se observó la presencia de bandas inespecíficas, principalmente en la amplificación de la enolasa.

Para la eliminación de las amplificaciones inespecíficas y el análisis de la eficiencia de *cross-linking* bajo las condiciones y cantidades de formaldehído utilizadas, una alícuota del pulso 15 se incubó durante 2 horas a 65 °C en 200 mM de NaCl para revertir el *cross-linking* y se amplificaron nuevamente los genes constitutivos Hsp70 y enolasa con los cebadores HSP70s; HSP70as y EnolsF; EnolsR respectivamente (Figura 20B).



**Figura 20: Análisis de la integridad del DNA en la fragmentación de la cromatina y la eficiencia de *cross-linking*.** (A): PCR punto final para identificación de los genes constitutivos enolasa y Hsp70 en los pulsos de sonicación 13, 14, 15 y 16. (B): PCR punto final del pulso 15 de sonicación antes y después de revertir el *cross-linking* para la eliminación de artefactos producto de la asociación de proteínas a la cadena molde. Geles de agarosa al 2.5%. M, marcadores de peso molecular.

## IX. Discusión

Los RNA circulares fueron clasificados inicialmente productos de errores en el *splicing* por lo que se descartaron como elementos funcionales en procesos biológicos. Sin embargo, debido a su abundancia, conservación evolutiva y capacidades funcionales observadas en las últimas décadas, se han considerado uno de los principales mecanismos para la generación del transcriptoma celular (Hocine *et al.*, 2010; Chen & Yang, 2015; Bangru & Kalsotra, 2016). Solamente en metazoos, se han identificados alrededor de 10000 circRNAs, pero en la mayoría de los casos aún no han sido estudiadas sus funciones en la célula (Zhang *et al.*, 2016).

Los RNA circulares con secuencias intrónicas se han asociado preferentemente a la regulación en *cis* de sus genes parentales por unión a la RNA polimerasa II. Dentro de estos fueron descritos un pequeño grupo que se forma a partir de la circularización intrónica por unión covalente de los sitios de *splicing* 5´-3´ de intrones humanos, murinos y recientemente descritos en *E. histolytica* en nuestro laboratorio denominados flicRNAs (Taggart *et al.*, 2017; Mendoza-Figueroa *et al.*, 2018). Estudios dirigidos a analizar la implicación de los flicRNAs en la regulación génica, permitió observar un aumento de los niveles de transcrito en cepas de *E. histolytica* transfectadas con minigenes donde fue mutada la secuencia 5´ss del intrón del gen RabX13, lo cual permitió inferir acerca de la posible inhibición de la transcripción que ejerce flicRX13 sobre su gen parental (Figura 8) (Mendoza-Figueroa *et al.*, 2018).

Los ensayos de CLIP en transfectantes pHA-CTD nos permitieron identificar la unión de los flicRNAs del gen RabX13 y de los loci asociados a virulencia EHI\_169670, EHI\_014170 y EHI\_192510 con la CTD de Pol II (Figura 13), los cuales se validaron como círculos *bona fide* mediante ensayos con RNAsa R (Figura 16). Por otro lado, pudimos observar una disminución de la unión de

flicRX13 a la Pol II en ensayos de CLIP llevados a cabo en transfectantes pHA-CTD/ $\Delta$ GU (Figura 14). De este modo, relacionamos directamente el aumento de los niveles de transcrito parental de RabX13, observado en transfectantes con mutaciones en la secuencia 5' ss del intrón con la unión de flicRX13 a la CTD de la Pol II. Estos resultados nos permiten sugerir la participación de los flicRNAs en la regulación en *cis* de sus genes parentales por medio de la unión a la RNA polimerasa II y la importancia de la secuencia GUUGU al 5' del intrón para la formación del complejo de unión a la CTD.

Ensayos similares llevados a cabo en células HeLa para el estudio de ElciRNAs permitió identificar al snRNA U1 como elemento esencial para la actividad regulatoria de circEIF3J y circPAIP2. Ellos pudieron observar que el bloqueo de U1, con oligos morfolino antisentido específicos, eliminó la actividad de los ElciRNAs y estimaron que a unos 300 pb del sitio de inicio de la transcripción dentro de la secuencia del promotor interactuaban en complejo snRNA U1/ElciRNAs con la RNA Pol II, para la regulación de la transcripción (Li *et al.*, 2015). Hasta la fecha no existe evidencia de homólogos para el snRNA U1 en *E. histolytica*, por lo que estimamos la posible participación del snRNA U6 snRNA como mediador de la interacción de flicRNAs con base en su homología de secuencia con el 5' ss y lo reportado para otros sistemas (Valdés *et al.*, 2014). Contrario a lo esperado y en concordancia a lo reportado por Nojima *et al.* (2018), el snRNA U6 no fue identificado como parte del complejo inmunoprecipitado en ninguno de los ensayos (Figura 13). Hasta la fecha solo existen reportes de la unión de U2 y U5 snRNA, elementos del spliceosoma catalítico, a la CTD de la pol II. Podríamos pensar que los factores del snRNP U1, U1C en particular (Du & Rosbash, 2002), pudieran estar asociados al 5' ss del flicRNA y ambos acomplejados a Pol II. Sin embargo, los resultados de U6 snRNA se analizaron como control negativo de unión en los ensayos de inmunoprecipitación ya que el snRNA U6 es un transcrito de la RNA polimerasa III (Paule & White, 2000).

Adicionalmente, identificamos al snRNA U2 como parte del inmunoprecipitado (Figura 13). Los componentes de U2 snRNA, SF3B3, así como SNRP116 y SNRP200 del snRNA U5, fueron identificados en cepas de *S. cerevisiae* en unión a la Polimerasa fosforilada en la serina 2 del repetido de heptapéptido de la CTD de la Pol II, fosforilación que indica eventos de elongación de la transcripción (McKay & Johnon, 2011; Nojima *et al.*, 2018), lo cual nos permite inferir acerca de una posible regulación de la elongación de la transcripción por parte de los flicRNAs. Por otro lado, Li *et al.* (2015) estudiaron a U2 snRNA como posible factor implicado en la unión de los ElciRNAs a la CTD de la polimerasa y reportaron que podría existir una asociación indirecta entre ambos, debido a que no observaron variación significativa sobre la acción de los ElciRNAs ante el bloqueo de esta ribonucleoproteína. Sin embargo, no descartamos la posible implicación de los snRNAs U2 o U5 como mediadores de la unión de los flicRNAs.

La localización citoplasmática de flicRpS14 del gen codificante de proteínas ribosomales, no fue un resultado esperado (Figura 15B). Los círculos intrónicos tanto ciRNAs como flicRNAs han sido previamente reportados en grandes cantidades en citoplasma de ovocitos de ranas del género *Xenopus*, así como en diversas líneas celulares isogénicas en humanos, ratón, pez cebra y rata, incluyendo eritrocitos, células que perdieron el núcleo durante su maduración a partir de precursores eritroblastos. A pesar de ello, no hay reportados estudios de sus funciones en el citoplasma. Independientemente, se infiere que se asocian a la regulación de la traducción o la estabilidad del mensajero maduro debido a la gran conservación interespecie de ciertos loci. En base a nuestros hallazgos, se podría agregar que los círculos intrónicos encabezan mecanismos regulatorios primitivos y altamente conservados en la evolución debido a su presencia en organismo que se diferenciaron antes de los eventos endosimbióticos aproximadamente un millón de años atrás como es el caso de *Entamoeba histolytica* (Morf & Singh, 2012; Taggart *et al.*, 2017;

Mendoza-Figueroa *et al.*, 2018; Talhouarne & Gall, 2018). De este modo, proponemos que flicRpS14 podría cumplir funciones en la regulación postranscripcional de elementos del complejo ribosomal, aunque son requeridos estudios adicionales para la determinación de dichas funciones.

El empleo del ácido bórico como inhibidor del segundo paso de transesterificación del *splicing*, nos permitió establecer un control para analizar el efecto de los flicRNAs sobre la expresión de sus genes parentales. Pudimos observar una disminución en la producción de flicRNAs ante tratamientos de dos horas con ácido bórico 5 mM en todos los flicRNAs en estudios (Figura 17A y D) y sin afectar la transcripción global de la célula significativamente (Figura 17B,C y E).

Mediante estudios cuantitativos por RT-qPCR pudimos observar que tanto en cepas tratadas con ácido bórico, como en transfectantes pHA- $\Delta$ GU, hubo una disminución significativa de la producción de flicRX13 que conllevó a un aumento de los niveles de transcrito de RabX13 en comparación con cepas no tratadas (Figura 18). En cepas de *E. histolytica* transfectadas con pHA-V, el efecto sobre los niveles de expresión tanto de flicRX13 como su gen parental disminuyeron, esto podría estar dado por la sobreexpresión de péptidos de hemaglutinina libres que, al no estar asociados a proteínas del hospedero, podrían ser reconocidas por mecanismo de respuesta celular que causen un cambio total de la expresión del transcriptoma amebiano sin que la disminución de la producción de flicRNAs ocasione efectos significativos en la regulación de la transcripción. Estos hallazgos validan el uso del ácido bórico como tratamiento control para la disminución de la producción de flicRNAs de acuerdo a lo reportado por Mendoza-Figueroa *et al.*, (2018).

Los mecanismos subyacentes a estas observaciones aún no están totalmente descritos. Se plantea que el boro podría formar complejos reversibles a través de enlaces de hidrógeno intramoleculares en la posición *cis*-2'-3' diol de los nucleósidos, interfiriendo con el desarrollo del ataque nucleofílico para la

liberación del intrón *lariat*, y causando el aumento del intermediario *lariat*-exon 2 en experimentos *in vitro* (Shomron & Ast, 2003). El efecto inhibitorio específico se observó solamente a partir de boros con 3 átomos de oxígeno asociados en tanto que ácidos borínicos, como los ácidos metil y butil borínicos, no afectan el desarrollo del *splicing* (Shomron & Ast, 2003). De igual modo, teniendo en cuenta que se ha propuesto que el mecanismo catalítico del *splicing* es mediado por iones metálicos divalentes donde los snRNAs llevan a cabo la coordinación de la catálisis, podría ocurrir que el ácido bórico interaccione con estos iones del sitio catalítico durante la segunda reacción del *splicing*. Teniendo en cuenta que ambas reacciones de *splicing* transcurren mediante reacciones de transesterificación, se esperaría un efecto similar del ácido bórico en la primera reacción, sin embargo no se reportan tales cambios en ensayos *in vitro* o *in vivo* (Shomron & Ast, 2003; Mendoza-Figueroa *et al.*, 2018).

En cada paso del *splicing*, los componentes y la conformación del spliceosoma es distinto. Durante la formación de los complejos catalíticamente activos del spliceosoma son necesarios una serie de re-arreglos esturcturales que permiten el desarrollo de las reacciones de transesterificación (Ritchie *et al.*, 2009; Ohrt *et al.*, 2012; Fica & Nagai, 2017). Los factores Yju2 y Cwc25 se unen al complejo B\* activado y forman el complejo C que lleva a cabo la primera reacción de *splicing*. Seguidamente, la incorporación de la helicasa Prp16 ocasiona nuevos cambios conformacionales y forma el complejo catalíticamente activado C\* para el desarrollo de la segunda reacción de transesterificación y la remoción del intrón *lariat* con la subsecuente unión de los exones (Will & Lührmann, 2011; Matera & Wang, 2014). Estos cambios conformacionales entre ambas reacciones varían también la configuración del sitio activo del spliceosoma. De este modo, la coordinación y orientación de los iones divalentes del core, así como la composición de residuos aminoacídicos varía; lo cual podría explicar la inhibición específica de la

segunda reacción de transesterificación del *splicing* ante tratamientos con ácido bórico. Ya sea por mayor accesibilidad o por orientación idónea del sitio catalítico para la formación de estados de transición estables con el ácido bórico, que impiden el ataque nucleofílico del exón libre sobre la unión *ariat*-exón 2. También se descarta la acción del ácido bórico sobre factores implicados en el desarrollo de la segunda reacción de transesterificación como el factor hSlu7, proteína que presenta dominios en dedos de zinc, los cuales en principio podrían verse afectados ante la presencia de ácido bórico, sin embargo la adición de hSlu7 recombinante activa a ensayos de *splicing in vitro*, no elimina el efecto inhibitorio del H<sub>3</sub>BO<sub>3</sub> (Shomron & Ast, 2003).

Por otro lado, las transfectantes pHA-Dbr $\Delta$ C mantuvieron niveles de flicRX13 elevados con la expresión de transcrito relativamente alta, similar a lo reportado por Mendoza-Figueroa *et al.* (2018) (Figura 18). Si estudiamos las características de los intrones formadores de flicRNAs en *E. histolytica*, podríamos analizar los resultados antes mencionados. Observamos que estos presentan una secuencia conservada al 5' ss del intrón esencial para su biogénesis, homóloga a la encontrada para ciRNAs en organismos superiores. Además, los flicRNAs tienen como característica común distancias menores a 8 nucleótidos entre el BP y el 3' ss del intrón según estudios de secuenciación. Estos factores, sumados a la baja conservación del BP en el caso de *E. histolytica* y a las secuencias degeneradas de BP reportadas en organismos superiores, propicia todos los elementos para la circularización y se plantea el desarrollo de un posible tercer ataque nucleofílico producto de la propia cercanía que conlleva a la circularización completa de este grupo de intrones (Taggart *et al.*, 2017; Mendoza-Figueroa *et al.*, 2018).

Así, los resultados observados ante la Dbr1 catalíticamente deficiente, donde se mantiene relativamente alta la producción de flicRNAs y bajo el fundamento del tercer ataque nucleofílico podría significar que la Dbr1 no participa en la circularización de los flicRNAs y en su lugar, impide su formación mediante el



procesamiento de la unión 2'-5' del *lariat* que ocasiona un cambio en la configuración y cercanía necesaria para el desarrollo del tercer ataque nucleofílico por linealización del intrón, lo cual estaría en concordancia con los resultados obtenidos en este estudio sin descartar el probado origen post-*splicing* de los flicRNAs (Mendoza-Figueroa *et al.*, 2018).

Hay que tener en cuenta que todas estas inferencias de los mecanismos de biogénesis son producto de estudios *in vitro*, por lo que son necesarios estudios adicionales *in vivo*. De igual modo, el ácido bórico podría afectar directamente la actividad de la Dbr1 de *E. histolytica* al interactuar en su sitio activo, el cual fue recientemente determinado por cristalografía de rayos X y presenta iones  $Mn^{2+}$  y  $Zn^{2+}$  coordinados por residuos de histidina y un grupo hidroxilo puente (Ransey & Heroux, 2017). Por último, habría que analizar la posible competencia en el procesamiento de la unión 2'-5' entre el spliceosoma y la Dbr1 reportado por Tseng & Cheng, (2013) como un mecanismo alternativo en la formación de los flicRNAs.

Finalmente, con el fin de responder a muchas de las interrogantes para la identificación de la unión de los flicRNAs al promotor de sus genes parentales en futuros estudios, se estandarizó la fragmentación de la cromatina utilizando diferentes intensidades de sonicación. Observamos que los trofozoítos sonicados a una amplitud de 65 %, mantuvieron fragmentos mayoritarios de DNA de 700-800 pb (Figuras 19A), lo cual podría comprometer la calidad de la cromatina en ensayos de ChIP (Arrigoni *et al.*, 2015), además de abarcar una secuencia extensa que introduce inespecificidades en la posterior identificación de la unión de factores a la región promotora que en *E. histolytica* tienen un tamaño promedio de 200 pb. En contraste, con pulsos de sonicación con amplitud del 75%, logramos la obtención de fragmentos mayoritarios de 400 pb a los 15 pulsos (Figuras 19B). A partir de este pulso, se logró amplificar los fragmentos de los genes controles Hsp70 y enolasa, pero se observó inicialmente la presencia de bandas inespecíficas, principalmente en la

amplificación de la enolasa (Figuras 20A), lo cual podría estar dado por la acción de proteínas unidas al DNA posterior al *cross-linking* que afectan la procesatividad de la enzima y su unión al DNA molde durante la amplificación (Perocchi *et al.*, 2007). Para la eliminación de estas bandas se revirtió el cross-linking y efectivamente, se obtuvieron amplificaciones específicas para nuestros genes controles. De este modo, se logró estandarizar exitosamente la técnica de CHIP con la obtención eficiente de fragmentos de DNA de 300-400 pb no degradados a los 15 pulsos de sonicación de treinta segundos con una amplitud del 75 % e incubación previa de 10 minutos con formaldehído al 1%.

## X. Conclusiones

En este trabajo se demostró la presencia de los flicRNAs del gen RabX13 y de los loci asociados a virulencia EHI\_192510, EHI\_014170 y EHI\_169670 en complejo con la RNA polimerasa II y la participación de flicRX13 en la regulación de sus genes parentales.

- En el mismo complejo de unión a la pol II, se identificó la presencia de snRNA U2, lo cual nos permite inferir una posible participación de los flicRNAs en la iniciación y elongación de la transcripción.
- Se descarta la presencia del snRNA U6 como mediador de la interacción de flicRNAs con la RNA Polimerasa II.
- Se demostró que la secuencia GUUUGU al 5' ss de intrón del gen RabX13 es esencial para la formación del complejo de unión a la RNA polimerasa II observándose una disminución de la unión de flicRX13 a la CTD en transfectantes con mutación en esta secuencia.
- Sorprendentemente los flicRNAs de los genes ribosomales RpL12 y RpS14 no fueron encontrados asociados a la RNA polimerasa II. Estudios adicionales nos permitieron identificar la localización citosólica de flicRpS14, por lo que proponemos una posible participación en la regulación postranscripcional de elementos del complejo ribosomal.
- Se corroboró que la inhibición del segundo paso de *splicing* con ácido bórico disminuyó la producción de flicRNAs (flicRX13, flic014170 y flic169670 int1), indicando su origen *post-splicing*.
- Demostramos que la producción constitutiva del mRNA de RabX13 es inhibida en respuesta a la disminución de flicRX13 ante el tratamiento con ácido bórico.
- Se estandarizaron los parámetros de fragmentación de la cromatina y condiciones de *cross-linking* del ensayo de CHIP.

## **XI. Perspectivas**

Para continuar con el estudio de las funciones de los flicRNAs, nos planteamos los siguientes experimentos:

- Determinar la presencia del complejo flicRNA-RNA Pol II unido a los promotores de sus genes parentales mediante ChIP para determinar la etapa de la transcripción regulada por los flicRNAs.
- Identificar los factores involucrados en la unión de los flicRNAs a la RNA Polimerasa II mediante estudios de proteómica.
- Determinar la funcionalidad de los flicRNAs de localización citoplasmática mediante perfiles ribosomales.

## XII. Bibliografía

1. Arrigoni L., Andreas S. Richter, Emily Betancourt, Kerstin Bruder, Sarah Diehl, Thomas Manke & Ulrike Bonisch (2015). Standardizing chromatin research: a simple and universal method for ChIP-seq. *Nucleic Acids Research*, 2016, Vol. 44, No. 7 e67.
2. Ashwal-Fluss, R., Meyer, M., Pamudurti, N. R., Ivanov, A., Bartok, O., Hanan, M. Kadener, S. (2014). CircRNA Biogenesis competes with Pre-mRNA splicing. *Molecular Cell*, 56(1), 55–66.
3. Bachellerie, J.P., Cavaille, J. and Huttenhofer, A. (2002). The expanding snoRNA world. *Biochimie*, 84, 775–790.
4. Bangru S. & Kalsotra A. (2016). Advances in analyzing RNA diversity in eukaryotic transcriptomes: peering through the Omics lens. *F1000Research*, 2668.
5. Begum, S., Quach, J., & Chadee, K. (2015). Immune evasion mechanisms of *Entamoeba histolytica*: Progression to disease. *Frontiers in Microbiology*, 6(DEC), 1–8.
6. Bentley, D. L. (2014). Coupling mRNA processing with transcription in time and space. *Nature Reviews Genetics*, 15(3), 163–175.
7. Bernstein, E. & Allis, C.D. (2005). RNA meets chromatin. *Genes Dev.*19, 1635–1655.
8. Boeckel JN et al (2015). Identification and characterization of hypoxia-regulated endothelial circular RNA. *Circ Res* 117:884–890.
9. Bortz P. D. S.& Wamhoff R. B.(2011). Chromatin Immunoprecipitation (ChIP): Revisiting the Efficacy of Sample Preparation, Sonication, Quantification of Sheared DNA, and Analysis via PCR. *PLoS ONE* 6(10): e26015.
10. Braunschweig U., Nuno L. Barbosa-Morais, Qun Pan, Emil N. Nachman, Babak Alipanahi, Thomas Gonatopoulos-Pournatzis, Brendan Frey, Manuel Irimia & Benjamin J. Blencowe (2014). Widespread intron retention in mammals functionally tunes transcriptomes. *Genome Res* (11): 1774–1786.
11. Brittingham A. & Wilson W. A., (2014). The Antimicrobial Effect of Boric Acid on *Trichomonas vaginalis*. *Sexually Transmitted Diseases*, vol 41, Number 12.

12. Burd CE, Jeck WR, Liu Y, Sanof HK, Wang Z, Sharpless NE (2010). Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. *PLoS Genet* 6:e1001233.
13. Chedin F (2016). Nascent connections: R-loops and chromatin patterning. *Trends Genet* 32:828–838.
14. Chen B. & Huang S.(2018). Circular RNA: an emerging non-coding RNA as a regulator and biomarker in cancer. *Cancer Letters*.
15. Chen, L. & Yang, L. (2015). Gear Up in Circles. *Molecular Cell*. 58.
16. Chen, L.-L. (2016). The biogenesis and emerging roles of circular RNAs. *Nature Reviews Molecular Cell Biology*, 17(4), 205–211.
17. Conn, S. J., Pillman, K. A., Toubia, J., Conn, V. M., Salmandis, M., Phillips, C. A., Goodall, G. J. (2015). The RNA binding protein quaking regulates formation of circRNAs. *Cell*, 160(6), 1125–1134.
18. Cornick S. & Chadee K.. (2017). *Entamoeba histolytica*: Host parasite interactions at the colonic epithelium, *Tissue Barriers*, 5:1, e1283386.
19. Cowling, V. H., & Cole, M. D. (2010). Myc Regulation of mRNA Cap Methylation. *Genes & Cancer*, 1(6), 576–579.
20. Crick F. (1970). Central Dogma of Molecular Biology. *Nature*, vol. 227.
21. Danan M., Schraga Schwartz, Sarit Edelheit and Rotem Sorek.(2011). Transcriptome-wide discovery of circular RNAs in Archaea. *Nucleic Acids Research*, vol. 40, No. 7 3131–3142.
22. Diamond, L. S., Harlow, D. R., & Cunnick, C. C. (1978). A new medium for the axenic cultivation of *Entamoeba histolytica* and other *Entamoeba*. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 72(4), 431–432.
23. Du H. & Rosbash M. (2002). The U1 snRNP protein U1C recognizes the 5' splice site in the absence of base pairing. *Nature* 419: 86–90.
24. Eddy, S. R. (2001). Non-coding RNA genes and the modern RNA world. *Nature Reviews Genetics*, 2(12), 919–929.
25. Elnekave K., Siman-Tov R, Ankri, S. (2003). Composition of L-arginine mediated by *Entamoeba histolytica* L-arginase (EhArg) inhibits amoebicidal activity and nitric oxide production by activated macrophages. *Parasite Immunol.* 25.597-608.

26. Errichelli L et al (2017). FUS affects circular RNA expression in murine embryonic stem cell-derived motor neurons. *Nat Commun* 8:14741.
27. Fica S. M. & Nagai K (2017). Cryo-EM snapshots of the spliceosome: structural insights into a dynamic ribonucleoprotein machine. *Nat Struct Mol Biol* 24 (10): 791-799.
28. Fourmann J-B, Schmitzova J, Christian H, Urlaub H, Ficner R, Boon K-L, Fabrizio P, Luhrmann R (2013). Dissection of the factor requirements for spliceosome disassembly and the elucidation of its dissociation products using a purified splicing system. *Genes Dev*, 27:413–428.
29. Goldberg I. H., M. Rabinowitz, & E. Reich (1962) Asis of actinomycin action. DNA binding and inhibition of RNA-Polymerase synthetic reactions by actinomycin. *Biochemistry*, vol 48.
30. Hansen, T. B., Jensen, T. I., Clausen, B. H., Bramsen, J. B., Finsen, B., Damgaard, C. K., & Kjems, J. (2013). Natural RNA circles function as efficient microRNA sponges. *Nature*, 495(7441), 384–388.
31. Hernández de la Cruz O., Marchat L.A., Guillén N., Weber C., López R. I., Díaz-Chávez J., Herrera L., Rojo Domínguez A., Orozco E., López-Camarillo C. (2016). Multinucleation and Polykaryon Formation is Promoted by the EhPC4 Transcription Factor in *Entamoeba histolytica*. *Nature. Scientific reports*, 6:19611.
32. Hesselberth J. R. (2013). *Lives that introns lead after splicing*. John Wiley & Sons, Ltd, vol4.
33. Hocine S., Singer R. H., Grunwald D.(2010). RNA Processing and Export. *Cold Spring Harb Perspect Biol* 2010;2:a000752.
34. Hoffman E. A., Brian L. Frey, Lloyd M. Smith & David T. Auble (2015). Formaldehyde Crosslinking: A Tool for the Study of Chromatin Complexes. *THE JOURNAL OF BIOLOGICAL CHEMISTRY* VOL. 290, NO. 44, pp. 26404–26411.
35. Holdt LM et al (2016). Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans. *Nat Commun* 7:12429.
36. Holdt L. M., Alexander Kohlmaier, Daniel Teupser (2018). Molecular roles and function of circular RNAs in eukaryotic cells. *Cell. Mol. Life Sci.*75:1071–1098.

37. Holley R. W., Apgar J., Everett G. A., Madison J. T., Marquisee M., Merrill S. H., Periswick J. R., Zamir A. (1965). Structure of a Ribonucleic Acid, 147 (3664), 1462-1465.
38. Hon, C., Weber, C., Sismeiro, O., Proux, C., Koutero, M., Deloger, M., Das. S., Agrahari, M., Dillies, M., Jagla, B., Coppee, J., Bhattacharya, A. y Gullen, N. (2013). Quantification of stochastic noise of splicing and polyadenylation in *Entamoeba histolytica*. *Nucleic Acids Research*. 41: 1936-52.
39. Hopper K. A. & Huang H-Y (2015). Quality Control Pathways for Nuclear-encoded Eukaryotic tRNA Biosynthesis and Subcellular Trafficking. *Mol. Cell. Biol*.
40. Hsu M-T. & Coca-Padros M.(1979). Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. *Nature*, Vol. 280.
41. Jenkins, J. L., Agrawal, A. A., Gupta, A., Green, M. R., & Kielkopf, C. L. (2013). U2AF65 adapts to diverse pre-mRNA splice sites through conformational selection of specific and promiscuous RNA recognition motifs. *Nucleic Acids Research*, 41(6), 3859–3873.
42. Kantor M. ,Anarella Abrantes ,Andrea Estevez ,Alan Schiller,Jose Torrent, Jose Gascon , Robert Hernandez,Christopher Ochner (2018). *Entamoeba Histolytica: Updates in Clinical Manifestation, Pathogenesis, and Vaccine Development*. *Canadian Journal of Gastroenterology and Hepatology* vol. 2018, Article ID 4601420, 6 pag.
43. Katolik, A., Clark, N. E., Tago, N., Montemayor, E. J., Hart, P. J., & Damha, M. J. (2017). Fluorescent Branched RNAs for High-Throughput Analysis of Dbr1 Enzyme Kinetics and Inhibition. *ACS Chemical Biology*, 12(3), 622–627.
44. Keating, K. S., Toor, N., Perlman, P. S. & Pyle, A. M. (2010). A structural analysis of the group II intron active site and implications for the spliceosome. *RNA* 16, 1–9.
45. Kelemen O, Convertini P, Zhang Z, Wen Y, Shen M, et al. (2013). Function of alternative splicing. *Gene* 514:1–30.
46. Kishore, S. & Stamm, S. (2006). The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. *Science*, 311, 230–232.
47. Lasda, E., & Parker, R. (2014). Circular RNAs: diversity of form and function. *Rna*, 20(12), 1829–1842.



48. Le Hir, H., Nott, A., & Moore, M. J. (2003). How intróns influence and enhance eukaryotic gene expression. *Trends in Biochemical Sciences*, 28(4), 215–220.
49. Lee, Y., & Rio, D. C. (2015). Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annual Review of Biochemistry*, 84(1), 291–323.
50. Lewin, B. (2008). *Genes 9*. Jones & Bartlett Learning.
51. Li, Z., Huang, C., Bao, C., Chen, L., Lin, M., Wang, X., ... Shan, G. (2015). Exon-intrón circular RNAs regulate transcription in the nucleus. *Nature Structural & Molecular Biology*, 22(3), 256–264.
52. Liang D & Wilusz JE (2014). Short intronic repeat sequences facilitate circular RNA production. *Genes Dev* 28:2233–2247.
53. Lidell, M. E., Moncada, D. M., Chadee, K., and Hansson, G. C. (2006). *Entamoeba histolytica* cysteine proteases cleave the MUC2 mucin in its C-terminal domain and dissolve the protective colonic mucus gel. *Proc. Natl. Acad. Sci. U.S.A.* 103, 9298–9303. doi: 10.1073/pnas.0600623103.
54. Lin J.-Y., Seguin R., Keller K., Chadee K. (1994). Tumor necrosis factor alpha augments nitric oxide-dependent macrophage cytotoxicity against *Entamoeba histolytica* by enhanced expression of the nitric oxide synthase gene. *Infect. Immun.* 62, 1534–1541.
55. Livak, K. J., & Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the  $2^{-\Delta\Delta CT}$  Method. *Methods*, 25(4), 402–408.
56. Loftus, B. (2005). The genome of the protist parasite *Entamoeba histolytica*. *Nature*, 433, 865–868.
57. Lorenzi, H. A., Puiu, D., Miller, J. R., Brinkac, L. M., Amedeo, P., Hall, N., & Caler, E. V. (2010). New assembly, reannotation and analysis of the *Entamoeba histolytica* genome reveal new genomic features and protein content information. *PLoS Neglected Tropical Diseases*, 4(6).
58. Lewis, B.P., Burge, C.B. and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, 120, 15–20.
59. Mandal SS, Chu C, Wada T, Handa H, Shatkin AJ, Reinberg D. (2004). Functional interactions of RNA-capping enzyme with factors that positively and

- negatively regulate promoter escape by RNA polymerase II. *Proc Natl Acad Sci* 101: 7572–7577.
60. Matera, A. G., & Wang, Z. (2014). A day in the life of the spliceosome. *Nature Reviews Molecular Cell Biology*, 15(2), 108–121.
  61. Mattick, J. S., & Makunin, I. V. (2006). Non-coding RNA. *Human Molecular Genetics*, 15 Spec No(1), 17–29.
  62. McKay S. L. & Johnson T. L. (2011). An Investigation of a Role for U2 snRNP Spliceosomal Components in Regulating Transcription. *Plos One*, vol. 6, e16077.
  63. Meier, U.T. (2005). The many facets of H/ACA ribonucleoproteins. *Chromosoma*, 114, 1–14.
  64. Mendoza-Figueroa M. S., Alfonso-Maqueira E. E., Vélez Cristina, Azuara-Liceaga E. I., Zárate S., Villegas-Sepúlveda N., Saucedo-Cárdenas O., Valdés F.J.(2018). Postsplicing-Derived Full-Length Intron Circles in the Protozoan Parasite *Entamoeba histolytica*. *Frontiers in Cellular and Infection Microbiology*, Vol. 9, Art. 255.
  65. Meyer, M., Fehling, H., Matthiesen, J., Lorenzen, S., Schuldt, K., Bernin, H., Zaruba, M., Lender, C., Ernst, T., Ittrich, H., Roeder, T., Tannich, E., Lotter, H., Bruchhaus, I. (2016). Overexpression of Differentially Expressed Genes Identified in Non-pathogenic and Pathogenic *Entamoeba histolytica* Clones Allow Identification of New Pathogenicity Factors Involved in Amoebic Liver Abscess Formation. *PLoS Pathog* 12, e1005853.
  66. Miao Z. & Westhof E. (2017). RNA Structure: Advances and Assessment of 3D Structure Prediction. *Annu. Rev. Biophys.*46.
  67. Misteli T. & Spector DL. (1999). RNA polymerase II targets premRNA splicing factors to transcription sites in vivo. *Mol Cell* 3: 697–705.
  68. Moonah S. N., Nona M. Jiang, William A. Petri, Jr. (2013). Host Immune Response to Intestinal Amebiasis. *PLOS Pathogens*, vol 9, Issue 8, e1003489.
  69. Morris, K. V., & Mattick, J. S. (2014). The rise of regulatory RNA. *Nature Reviews Genetics*, 15(6), 423–437.
  70. Morf Laura & Singh Upinder. (2012). *Entamoeba histolytica*: a snapshot of current research and methods for genetic analysis. *Curr Opin Microbiol.* 15(4): 469–475.

71. Nojima T., Rebelo K., Gomes T., Grosso A. R., Proudfoot N. J., Carmo-Fonseca M. (2018). RNA Polymerase II Phosphorylated on CTD Serine 5 Interacts with the Spliceosome during Co-transcriptional Splicing. *Molecular Cell* 72, 369–379.
72. O’Gorman, W., Thomas, B., Kwek, K.Y., Furger, A. and Akoulitchev, A. (2005). Analysis of U1 snRNA interaction with cyclin H. *J. Biol. Chem.*, 280, 36920–36925.
73. Ohrt T. *et al.*, (2012). Prp2-mediated protein rearrangements at the catalytic core of the spliceosome as revealed by dcFCCS. *RNA*, 18:1244-1256.
74. Olivos-García Alfonso a, Emma Saavedra b, Mario Nequiz Avendaño a y Ruy PérezTama. (2011). amibiasis: mecanismos moleculares de la patogenicidad de *Entamoeba histolytica*. *Res Microbiol*; 162(6-2): 637–645.
75. Pamudurti NR, Osnat Bartok, Marvin Jens, Reut Ashwal-Fluss, Christin Stottmeister, Larissa Ruhe, Mor Hanan, Emanuel Wyler, Daniel Perez-Hernandez, Evelyn Ramberger, Shlomo Shenzis, Moshe Samson, Gunnar Dittmar, Markus Landthaler, Marina Chekulaeva, Nikolaus Rajewsky, Sebastian Kadener (2017). Translation of CircRNAs. *Mol Cell* 66(1):9-21.
76. Pasman Z, Been MD, Garcia-Blanco MA (1996). Exon circularization in mammalian nuclear extracts. *RNA* 2:603–610.
77. Patel, A. A., & Steitz, J. A. (2003). Splicing double: insights from the second spliceosome. *Nature Reviews Molecular Cell Biology*, 4(12), 960–970.
78. Paule M. R &, White R. J.(2000). SURVEY AND SUMMARY mRNA capping: biological functions and applications. *Nucleic Acid Research*, vol. 28, no. 6 1283-1298.
79. Pearson Richard J. & Singh Upinder (2010).Approaches to characterizing *Entamoeba histolytica* transcriptional regulation *Cellular Microbiology*, 12(12), 1681–1690.
80. Perocchi F., Xu Z., Clauder-Munster S., Steinmetz L. M. (2007). Antisense artifacts in transcriptome microarray experiments are resolved by actinomycin D. *Nucleic Acids Research*, 2007, Vol. 35, No. 19.
81. Qin M., Gang Wei c, Xiaomeng Sun (2018). Circ-UBR5: An exonic circular RNA and novel small nuclear RNA involved in RNA splicing. *Biochemical and Biophysical Research Communications* 503 ,1027-1034.

82. Qu, S., Yang, X., Li, X., Wang, J., Gao, Y., Shang, R., ... Li, H. (2015). Circular RNA: A new star of noncoding RNAs. *Cancer Letters*, 365(2), 141–148.
83. Radushev D. (2007). GraphPad Prism. 5.01 ed. [www.graphpad.com](http://www.graphpad.com).
84. Ramanathan A., Robb G. B. and Chan S.H.(2016). SURVEY AND SUMMARY mRNA capping: biological functions and applications *Nucleic Acids Research*, Vol. 44, No. 16 7511–7526.
85. Ransey E. & Heroux A. (2017). Crystal structure of the *Entamoeba histolytica* RNA lariat debranching enzyme EhDbr1 reveals a catalytic Zn<sup>2+</sup>/Mn<sup>2+</sup> heterobinucleation. Brookhaven National Laboratory, BNL-203614-2018-JAAM.
86. Ritchie *et al.*, (2009). Spliceosome structure: Piece by piece. *Biochimica et Biophysica Acta*, 1789,624-633.
87. Roy B., Larisa M. Haupt and Lyn R. Griffiths (2013). Alternative Splicing (AS) of Genes As An Approach for Generating Protein Complexity *Current Genomics*, 14, 182-194.
88. Ruskin, B., & Green, M. R. (1985). An RNA processing activity that debranches RNA lariat s. *Science*, 229(4709), 135–140.
89. Sal, G. Del, Manfioletti, G., & Schneider, C. (1988). *Acids Research*, 16(20), 34012.
90. Salzman, J. (2014). Circular RNA is expressed across the eukaryotic tree of life. *PloS One*, 9(6), e90859.
91. Sambrook J. (1977). Adenovirus amazes at Cold Spring Harbor. *Nature*.
92. Sanger H. I., klotzt G., Riesnert D., Gross H. J., Kleinschmidt A. K.(1976). Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proc. Natl. Acad. Sci. USA*, vol. 73, no. 11, pp. 3852-3856.
93. Shi Yigong (2017). Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *NATURE REVIEWS | MOLECULAR CELL BIOLOGY*, vol 18.
94. Shomron N. & Ast G.(2003). Boric acid reversibly inhibits the second step of pre-mRNA splicing. *FEBS Letters* 552, 219-224.
95. Sim S., Yong T.-S., Park S.-J., Im K.-I., Kong Y., Ryu J.-S., et al. . (2005). NADPH oxidase-derived reactive oxygen species-mediated activation of

- ERK1/2 is required for apoptosis of human neutrophils induced by *Entamoeba histolytica*. *J. Immunol.* 174, 4279–4288.
96. Small EC, Leggett SR, Winans AA, Staley JP (2006). The EF-G-like GTPase Snu114p Regulates Spliceosome Dynamics Mediated by Brr2p, a DExD/H Box ATPase. *Mol Cell*, 23:389–399.
97. Solaymani-Mohammadi, S., Rezaian, M., Babaei, Z., Rajabpour, A., Meamar, A. R., Pourbabai, A. A., & Petri, W. A. (2006). Comparison of a stool antigen detection kit and PCR for diagnosis of *Entamoeba histolytica* and *Entamoeba dispar* infections in asymptomatic cyst passers in Iran. *Journal of Clinical Microbiology*, 44(6), 2258–2261.
98. Starke, S., Jost, I., Rossbach, O., Schneider, T., Schreiner, S., Hung, L. H., & Bindereif, A. (2015). Exon circularization requires canonical splice signals. *Cell Reports*, 10(1), 103–111.
99. Taggart A. J., Chien-Ling Lin, Barsha Shrestha, Claire Heintzelman, Seongwon Kim, & William G. Fairbrother (2017). Large-scale analysis of branchpoint usage across species and cell lines. Cold Spring Harbor Laboratory Press; 27:639–649.
100. Talhouarne G. J. S. & Gall J. G. (2018). Lariat intronic RNAs in the cytoplasm of vertebrate cells. *PNAS*, vol. 115, no. 34.
101. Tseng C-K & Cheng S-C (2013). The spliceosome catalyzes debranching in competition with reverse of the first chemical reaction. *RNA*, 320:1782–1784.
102. Uribarren B. T.. 2017. ENTAMOEBOISIS o AMIBIASIS o AMEBIASIS. Departamento de Microbiología y Parasitología, Facultad de Medicina, UNAM.
103. Valdés F. J., Nozaki, T., Sato, E., Chiba, Y., Nakada-Tsukui, K., Villegas-Sepúlveda, N., ... Galindo-Rosales, J. M. (2014). Proteomic analysis of *Entamoeba histolytica* in vivo assembled pre-mRNA splicing complexes. *Journal of Proteomics*, 111, 30–45.
104. Valdés F. J, Itzel López-Rosas, César López-Camarillo, Esther Ramírez-Moreno, Juan D. Ospina-Villa, Laurence A. Marchat. (2018). Life and Death of mRNA Molecules in *Entamoeba histolytica*. *Frontiers in Cellular and Infection Microbiology*, vol.8, Article 199.

105. Wang, H., Hill, K., & Perry, S. E. (2004). An Arabidopsis RNA Lariat Debranching Enzyme Is Essential for Embryogenesis. *Journal of Biological Chemistry*, 279(2), 1468–1473.
106. Wang Y., Jing Liu, Bo Huang, Yan-Mei Xu, Jing Li, Lin-Feng Huang<sup>1</sup>, Jin Lin, Jing Zhang, Qing-hua Min, Wei-ming Yang & Xiao-zhong Wang (2014). Mechanism of alternative splicing and its regulation. *BIOMEDICAL REPORTS* 3: 152-158.
107. Weber, C., Koutero, M., Dillies, M. A., Varet, H., Lopez-Camarillo, C., Coppee, J.Y., Hon, C. C., and Guillen, N. (2016). Extensive transcriptome analysis correlates the plasticity of *Entamoeba histolytica* pathogenesis to rapid phenotype changes depending on the environment. *Scientific reports* 6, 35852.
108. Will, C. L., & Lührmann, R. (2011). Spliceosome structure and function. *TL-3. Cold Spring Harbor Perspectives in Biology*, 3 VN-re(7), 1–23.
109. Willhoeft, U. t. e., Campos-Góngora, E., Touzni, S., Bruchhaus, I., & Tannich, E. (2001). Intróns of *Entamoeba histolytica* and *Entamoeba dispar*. *Protist*, 152(2), 149–156.
110. Wilusz J. E., Sunwoo H. & Spector D. L.(2009). Long noncoding RNAs: functional surprises from the RNA world. *GENES & DEVELOPMENT* 23:1494–1504.
111. Wu, J., & Fu, Xiao J, Zhang Z., Wang X., Hu S., Yu J. (2014). Ribogenomics: the Science and Knowledge of RNA, *Genomics Proteomics Bioinformatics* 12 ,57–63.
112. Wu, T., & Fu, X.-D. (2015). Genomic functions of U2AF in constitutive and regulated splicing. *RNA Biology*, 12(5), 479–485.
113. Yean, S. L., Wuenschell, G., Termini, J. & Lin, R. J. (2000). Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature* 408, 881–884.
114. Zhang, Y., Zhang, X. O., Chen, T., Xiang, J. F., Yin, Q. F., Xing, Y. H., ... Chen, L. L. (2013). Circular Intrónic Long Noncoding RNAs. *Molecular Cell*, 51(6), 792–806.
115. Zhang Y, Xue W, Li X, Zhang J, Chen S, Zhang JL, Yang L, Chen LL (2016). The biogenesis of nascent circular RNAs. *Cell. Rep* 15:611–624.

### XIII. Anexos

**Anexo 1:** Genes, oligonucleótidos y condiciones usadas en los experimentos descritos. El número de acceso de cada gen está entre paréntesis. Todas las amplificaciones se hicieron en 35 ciclos. US: transcrito no procesado por *splicing*, S: transcrito procesado por *splicing*.

Gen	Oligonucleótidos		Condiciones de PCR	Tamaño del Amplicón
	Nombre	Secuencia (5' → 3')		
<i>RabX13</i> (EHI_065790)	Rab2 F	CGTTGTTGGAGACTCTTCAGTTGG	94°C/45", 60°C/45", 72°C/1'	US = 510 pb S = 374 pb
	Rab2 R	GACCCATTTTCAGTTGAAACAGTTC		
	Rab2BSs	GATGAGATGAGATAAGATTTTATC	94°,45"; 55°,45"; 72°C,30"	<i>Lariat</i> ≈ 105 pb flicRX13 = 116 pb
	Rab2BSas	CCGAATTACTGTCTTGAAAATC		
	NBRX13	CTTTTAAAGTATTTAAAACAAACCTA AATAAAAGTTAGTTTATATTTATT	94°C,45"; 55°C,45"; 72°C,30"	79 pb (Al usar Rab2BSs)
	Rab2Rex 2	TTGAACACCACGATAATAACTAGT		
<i>rpL12</i> (EHI_191750)	L12BSs	GATACTTTATAACAATATTTTGGTGG	94°C,45"; 56°C,45"; 72°C,30"	<i>Lariat</i> ≈ 87 pb flicRpL12 = 100 pb
	L12BSas	GAATACTTTAATAACGTATGGATG		
<i>rpS14</i> (EHI_074090)	S14BSs	TTAACAAATGAAGTTGTTTTAAC	94°C,45"; 51°C,45"; 72°C,30"	<i>Lariat</i> ≈ 57 pb flicRpS14 ≈ 69 pb
	S14BSas	AATGAAATGTGTAGTGAACAAAC		
	Cdc2BSas	ATTTATTTTTATTTTCATTTTATATT		
<i>ClcB</i> (EHI_186860)	ClcBSs	GAAAACTATGTTTATAAAAAC	94°C,45"; 47°C,45"; 72°C,30"	<i>Lariat</i> ≈ 50 pb
	ClcBSas	CCAACAAAAATAAATAACAAAC		
EHI_169670	circRev169670	GGCAAGAGAATTGATTGATTAAG	94°C,45"; 55°C,45"; 72°C,30"	Flic169670 int-I = 110 pb
	circF169670	CTATTATCTACACACTCATTTTC		
	circSonda Intron 1	GGAGTATATTCAAAATTAACAAACC TATAAATGATTAGTG		
	circINT2Rev169670	CAAATATTCTCCCCTCAACAACTC	94°C,45"; 55°C,45"; 72°C,30"	Flic169670 int-II = 100 pb
	circINT2F169670	GAGTAAGTTAGTTCTTTTAG		
	circSonda Intron 2	CATATATAATAAAAAACAAACCT ACAAATAAG		
	EHI169670_E2s	AAGAAAGTTAATGATTCTGAGAAAG AG	94°C,45"; 58°C,45"; 72°C,45"	Circ169670 exon II= 354 pb
	EHI169670_E2as	CTTCTTTTTCTTTTTCTAATTCCTCAC CC		
EHI_192510	circRev192510	CATATGCTTCTTCATTTCTTC	94°C,45"; 55°C,45"; 72°C,30"	Flic192510= 73 pb
	circF192510	GAGAAAGGAATGAGCCAAC		
	circSonda 192510	CCATTCTACTCTCATCTCCCCTATT ATTTCTACCTTGAATATGC		
EHI_014170	circRev014170	GAGAGGTAGAAAGGAGAG	94°C,45"; 55°C,45"; 72°C,30"	Flic014170= 113 pb
	circF014170	GTTGGAAATACTCATTTTTATTAG		
	circSonda 014170	GAAGTGGTATGGAGAAATAAAAAC AAACCTTATTTAATTC		
	M13R	CAGGAAACAGCTATGACC		

Eh actina (EHI_107290)	EhActF	GAGCTGTATTCCCATCCATTGTTG	94°C/45", 65°C/45", 72°C,45"	535 pb
	EhActR	CTTTCAGCAGTAGTGGTGAAAGC		
EhRNA polimerasa II (EHI_121760)	Pol II F Xma I	ATCCCCCGGGATGACTGAACTCAAC ATGGAAACGAAAC	94°C/45", 58.5 °C/30", 72 °C/5'	4911 pb
	Pol II R Xho I	GCAACTCGAGTCATTTATCTTGCGG TTTAGGGAATACTAGG		
EhRNA polimerasa II (EHI_056690)	RNAPIIs	GATCCAACATATCCTAAAACAACA	94°C/45", 60 °C/45", 72 °C/45"	204 pb
	RNAPIIas	TCAATTATTTTCTGACCCGTCTTC		
enolasa (EHI_130700)	EnolsF	GCTGAAGATGATTGGGCAGCA	94°C/45", 58 °C/45", 72 °C/45"	100 pb
	EnolsR	TGGACACGAGCTGGGTTTGT		
HSP70 (EHI_052860)	HSP70s	GCAGAAGCTGAAAAGTTCAAGGC	94°C/45", 55 °C/45", 72 °C/45"	150 pb
	HSP70as	CTTCTGCAGCAATCTTATCAGCAA		
U6 snRNA (EHU43841)	EhU6 F	GGATCCACTTCGGTGAAAT	94°C/45", 62°C/45", 72°C/30"	92 bp
	EhU6 R	CTTCTCGTATGAGCGTGCATC		
U2 snRNA (BK006130)	T7U2	GCATGCTAATACGACTTCTCGGCCT TTATG	94°C/45", 60°C/45", 72°C/45"	200 bp
	hU2siR	TAACAGATCTGTTTCCATGCACATC CTCG		
pEhExHA (Cistein sintasa)	CS+116	AGAATAGACTCTCATCAA	94°C/45", 55°C/45", 72°C/45"	-
	CS-103	GCTTTGTAACATCTTCTTGC		
pCR2.1	M13F	GTA AACGACGGCCAGT	94°C/45", 53°C/45", 72°C/45"	-
	M13R	CAGGAAACAGCTATGACC		

**Anexo 2:** Secuenciación del gen correspondiente a la RNA polimerasa II de *Entamoeba histolytica* (EHI.121760) en plásmido pCR2.1 utilizando los cebadores M13F y M13R del plásmido.

**Primer M13 F:**

Pol: secuencia de locus EHI\_121760 de RNA polimerasa II  
Sec: Secuenciación en plásmido pCR2.1

```
>>pol 4911 bp ou (4911 nt)
Waterman-Eggert score: 4720; 331.6 bits; E(1) < 7.5e-94
99.3% identity (99.3% similar) in 961 nt overlap (1039-79:3955-4911)

1040      1030      1020      1010      1000      990
sec      GCAATTTGTAGTTGANCCAGAAACAAAAAATNTTAAGAGATCAATTTACAATGGGTAGT
          ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
pol      GCAATT-GTAGTTGATCCAGAAACAAA--ATGTTAAGAGATCAATT-ACAATGGGTAGT
          3960      3970      3980      3990      4000      4010

980      970      960      950      940      930
sec      ATTTACTGAAGGATCTAATATGGCAGGTTTAGCTGAATGTGATAATGTTGATTTTATAA
          ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
pol      ATTTACTGAAGGATCTAATATGGCAGGTTTAGCTGAATGTGATAATGTTGATTTTATAA
          4020      4030      4040      4050      4060      4070

920      910      900      890      880      870
sec      AACTGTTACAAATAATGTTTATGAACTTTCCTTTATTTAGGTATTGAAGCAGGTAGAAC
          ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
pol      AACTGTTACAAATAATGTTTATGAACTTTCCTTTATTTAGGTATTGAAGCAGGTAGAAC
          4080      4090      4100      4110      4120      4130
```



```

860      850      840      830      820      810
sec      TCAACTTGAAGCAGAACTTATTTAAACATTATCTGGGTCATATGTTAATTTTAGACATTT
          .....
pol      TCAACTTGAAGCAGAACTTATTTAAACATTATCTGGGTCATATGTTAATTTTAGACATTT
          4140      4150      4160      4170      4180      4190

800      790      780      770      760      750
sec      AGCTCTTATGGCTGATGTTATGACATTTACTGGTAGAATTCAACCTTTTAATAGAATGGG
          .....
pol      AGCTCTTATGGCTGATGTTATGACATTTACTGGTAGAATTCAACCTTTTAATAGAATGGG
          4200      4210      4220      4230      4240      4250

740      730      720      710      700      690
sec      TTTATCACGTGGAAGAGCAGGAGTTATTACTAGAGCTTCATTTGAACAAACATTAGAACA
          .....
pol      TTTATCACGTGGAAGAGCAGGAGTTATTACTAGAGCTTCATTTGAACAAACATTAGAACA
          4260      4270      4280      4290      4300      4310

680      670      660      650      640      630
sec      ATTTAGAAGAGCAGCTGCATTTAGTGAATCAGATATTTTAAATGGTATTTCTCAAAATAT
          .....
pol      ATTTAGAAGAGCAGCTGCATTTAGTGAATCAGATATTTTAAATGGTATTTCTCAAAATAT
          4320      4330      4340      4350      4360      4370

620      610      600      590      580      570
sec      TTTAATGGGACAAAGGACTGTTGCAGGAAGTGGTGCATTTACTGTATTACTTGATATAGA
          .....
pol      TTTAATGGGACAAAGGACTGTTGCAGGAAGTGGTGCATTTACTGTATTACTTGATATAGA
          4380      4390      4400      4410      4420      4430

560      550      540      530      520      510
sec      TTCATTTAAAAGCATCAAATATTGATGAAAACTCATTTAGTGAAGAAGAAGCTGATAAAAA
          .....
pol      TTCATTTAAAAGCATCAAATATTGATGAAAACTCATTTAGTGAAGAAGAAGCTGATAAAAA
          4440      4450      4460      4470      4480      4490

500      490      480      470      460      450
sec      AGCATTTATTGATGGATTTGATGTGTCTCCAATTAGAACTCATTGGTTATGTCTCCAAC
          .....
pol      AGCATTTATTGATGGATTTGATGTGTCTCCAATTAGAACTCATTGGTTATGTCTCCAAC
          4500      4510      4520      4530      4540      4550

440      430      420      410      400      390
sec      AAATACTAATAAACCAAAGCCAATGGGATTCTCTCCTTCCCCAGTTTCTCCAAGAAAAAC
          .....
pol      AAATACTAATAAACCAAAGCCAATGGGATTCTCTCCTTCCCCAGTTTCTCCAAGAAAAAC
          4560      4570      4580      4590      4600      4610

380      370      360      350      340      330
sec      ACATCCACAATTTGCTACTCCATCACCATTTATAACCTCTTCTTCTCCAAGATCTTCACC
          .....
pol      ACATCCACAATTTGCTACTCCATCACCATTTATAACCTCTTCTTCTCCAAGATCTTCACC
          4620      4630      4640      4650      4660      4670

320      310      300      290      280      270
sec      TATGTTAATGCAATTTACCCCTCGGTTTAGTTTCAGGTCCTTCACCATATAATCCAGCTAT
          .....
pol      TATGTTAATGCAATTTACCCCTCGGTTTAGTTTCAGGTCCTTCACCATATAATCCAGCTAC
          4680      4690      4700      4710      4720      4730

```

```

260      250      240      230      220      210
sec      TGGTGCATTATCTAAACCTGCATTTTTATCATCACCAATGTTTGCTGGTAAACAATCTTC
          .....
pol      TGGTGCATTATCTAAACCTGCATTTTTATCATCACCAATGTTTGCTGGTAAACAATCTTC
          4740      4750      4760      4770      4780      4790

200      190      180      170      160      150
sec      ACCTTATAATAACCCTAATCCATCACCATTCTTAGGTTCAACACCTGCGTTCCAAGGTAT
          .....
pol      ACCTTATAATAACCCTAATCCATCACCATTCTTAGGTTCAACACCTGCGTTCCAAGGTAT
          4800      4810      4820      4830      4840      4850

140      130      120      110      100      90
sec      TTCATCACCATTTATTAATTACTCATCACCTAGTATATTCCTAAACCGCAAGATAAATG
          .....
pol      TTCATCACCATTTATTAATTACTCATCACCTAGTATATTCCTAAACCGCAAGATAAATG
          4860      4870      4880      4890      4900      4910

80
sec      A
          :
pol

```

**Primer M13 R:**

Pol: secuencia de locus EHI\_121760 de RNA polimerasa II  
 Sec: Secuenciación en plásmido pCR2.1

>>pol 4911 bp (4911 nt)  
 Waterman-Eggert score: 4759; 306.2 bits; E(1) < 3.3e-86  
 99.0% identity (**99.0% similar**) in 981 nt overlap (982-2:1-973)

```

980      970      960      950      940      930
sec      ATGACTGAACTCAACATGGAAACGAAACGAGTCAAACAATAAAATTTACTTTACTTTCT
          .....
pol      ATGACTGAACTCAACATGGAAACGAAACGAGTCAAACAATAAAATTTACTTTACTTTCT
          10      20      30      40      50      60

920      910      900      890      880      870
sec      TCTGAACAAATCAGAAAAATGGGTGTTTGTAAGATTACTGAAAGTACTTCAGTTAGTAGC
          .....
pol      TCTGAACAAATCAGAAAAATGGGTGTTTGTAAGATTACTGAAAGTACTTCAGTTAGTAGC
          70      80      90      100      110      120

860      850      840      830      820      810
sec      GATGGTAAACCAGTTTTAGGTGGACTTGCCGATCCTCGTATGGGACCATTTGAAAAAGGA
          .....
pol      GATGGTAAACCAGTTTTAGGTGGACTTGCCGATCCTCGTATGGGACCATTTGAAAAAGGA
          130      140      150      160      170      180

800      790      780      770      760      750
sec      GAAGAATGTCAAACATGTAGAGGAACAAGGACAGAATGTCCTGGTCATTTTGGATATATA
          .....
pol      GAAGAATGTCAAACATGTAGAGGAACAAGGACAGAATGTCCTGGTCATTTTGGATATATA
          190      200      210      220      230      240

740      730      720      710      700      690
sec      GAATTAACACAACCAGTTTATAATGCTATATATATGAAATATATTTTACATGTATTAAG
          .....

```

```

pol GAATTAACACAACCAGTTTATAATGCTATATATATGAAATATATTTTACATGTATTAAG
      250      260      270      280      290      300

      680      670      660      650      640      630
sec TGTGTTTGTCCAATTTGTAACGAATATATTGTCATGGAGATGAAAGAGTTAAAAATATT
      .....
pol TGTGTTTGTCCAATTTGTAACGAATATATTGTCATGGAGATGAAAGAGTTAAAAATATT
      310      320      330      340      350      360

      620      610      600      590      580      570
sec AAAAGTACTGGGGAAGAACGATTTAAAAGAAATTTATGAGTTTATTCATAAAAAATTATAAA
      .....
pol AAAAGTACTGGGGAAGAACGATTTAAAAGAAATTTATGAGTTTATTCATAAAAAATTATAAA
      370      380      390      400      410      420

      560      550      540      530      520      510
sec GCTTGTGGAGTTTATCTGAGAAGAAAAAGATCATGATAAACACCAAATAATGAAGAA
      .....
pol GCTTGTGGAGTTTATCTGAGAAGAAAAAGATCATGATAAACACCAAATAATGAAGAA
      430      440      450      460      470      480

      500      490      480      470      460      450
sec GAAGAAGAACAAAATCCATGTAATGGACCAATTGTAGATTATGAACTTCCTAAATTAATA
      .....
pol GAAGAAGAACAAAATCCATGTAATGGACCAATTGTAGATTATGAACTTCCTAAATTAATA
      490      500      510      520      530      540

      440      430      420      410      400      390
sec TCAGAAAATAAATTAGAAATTAGAAGAAAAAATGGAATAAAAACTAATGAAGAAAATGAT
      .....
pol TCAGAAAATAAATTAGAAATTAGAAGAAAAAATGGAATAAAAACTAATGAAGAAAATGAT
      550      560      570      580      590      600

      380      370      360      350      340      330
sec TTTAAACCAAGAGATGCACTTGCAATATTTAAAGATATGACTGATGAAGATGTTAGAGTA
      .....
pol TTTAAACCAAGAGATGCACTTGCAATATTTAAAGATATGACTGATGAAGATGTTAGAGTA
      610      620      630      640      650      660

      320      310      300      290      280      270
sec CTTGGATTTGATCCAATTGATTGTCATCCTAAATTTATGATTTTACATTAATTCCAGTA
      .....
pol CTTGGATTTGATCCAATTGATTGTCATCCTAAATTTATGATTTTACATTAATTCCAGTA
      670      680      690      700      710      720

      260      250      240      230      220      210
sec CCACCACCATGTGTTAGACCATCTATTGTATCAGAAGGAGCAAAAAGAATCTAAAGATGAT
      .....
pol CCACCACCATGTGTTAGACCATCTATTGTATCAGAAGGAGCAAAAAGAATCTAAAGATGAT
      730      740      750      760      770      780

      200      190      180      170      160      150
sec TTAAGTGTATTATATGAAAATGTAATAACTGCTAATGAAAGACTAAAAGAAGAAAATCAA
      .....
pol TTAAGTGTATTATATGAAAATGTAATAACTGCTAATGAAAGACTAAAAGAAGAAAATCAA
      790      800      810      820      830      840

      140      130      120      110      100      90
sec AAAATGTGAAGGGTATCAGAATATTATAATAAATTAATGGAATTACAAACTATTTGTGC
      :: .....

```

```

pol   AAA-TGTGAAGGACTATCAGAATATTATAATAAATTAATGGAATTACAAACTATTTGTGC
      850      860      870      880      890

      80      70      60      50      40      30
sec   AAAAAATATCAATGAAAAAATGTTAACTTCCAGAAACCCATGTTGAGAAAGAAAAATT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
pol   AAAAAATATCAATGAAAAA-TGTTAACTTCCAGAAACC-ATGTT-GAGAAAGAAAAATT
      900      910      920      930      940      950

      20      10
sec   TTTAAAAATCAATTGAAAGAAA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
pol   T---AAATCAATTGAA-GAAA
      960      970

```