



**Centro de Investigaciones y Estudios Avanzados del Instituto Politécnico
Nacional**

Unidad Zacatenco

Departamento de Bioquímica

**“Participación de los elementos 5´ss y BP en el proceso de
circularización intrónica del gen RabX13 de *Entamoeba
histolytica*”**

Tesis que presenta

Lic. Eddy Ezequiel Alfonso Maqueira

En opción al grado de Maestro en Ciencias

Especialidad Bioquímica

Director de la Tesis

Dr. Jesús Valdés Flores

Ciudad de México

Septiembre, 2017

Agradecimientos

Durante este trabajo se contó con la beca 707577 otorgada por CONACYT, por lo cual estoy muy agradecido a esta institución.

Muchas personas han estado involucradas en la obtención de este grado. Desde hace dos años que vine a este país a estudiar, recibí el apoyo de gente que se han convertido en amigos, hermanos y familiares. A todas ellas quisiera expresarles mis más sinceros agradecimientos.

A continuación, quiero hacer una mención aparte para algunas personas muy especiales en este proceso.

A mi familia, porque aunque lejos, siempre me apoyaron mucho para alcanzar este logro que espero sea solo el comienzo de mi vida profesional.

Especialmente, quiero agradecer al Dr. Jesús Valdés Flores por aceptarme en su laboratorio, transmitirme sus conocimientos y permitirme trabajar bajo su tutela.

A José Manuel, por su ayuda en todo momento, en técnicas y conocimientos de Biología Molecular y por las charlas de México. A Martín, por su apoyo en todo este tiempo con los materiales necesarios para realizar este proyecto.

A mis asesores, por las correcciones y aclaraciones hechas a este trabajo que contribuyeron indudablemente a su mejoría.

A todos los profesores del Departamento de Bioquímica del CINVESTAV por todo que me han aportado y por darme las herramientas para crecer como profesional.

A Pris y Angi por todo su apoyo.

A Gretter, Wendy, Mónica, Sandra, Víctor y Noe, compañeros que se convirtieron en familia e hicieron que esta etapa en México fuera excelente.

A todos, muchas gracias.

Índice

LISTA DE ABREVIATURAS	1
RESUMEN	3
ABSTRACT	4
1. INTRODUCCIÓN	5
1.1 EI RNA	5
1.2 Procesamiento del mRNA.....	5
1.3 <i>Splicing</i>	7
1.3.1 Mecanismos del <i>splicing</i> y regulación en <i>cis</i>	8
1.3.2 Elementos reguladores en <i>cis</i> en la transcripción de genes.....	9
1.3.3 Ensamble, estabilidad y actividad del spliceosoma	11
1.3.4 Procesamiento del intrón <i>lariat</i>	12
1.4 RNAs no codificantes	13
1.5 RNAs circulares.....	14
1.5.1 RNAs circulares de eucariotas	15
1.5.2 Biogénesis de los circRNAs eucarióticos	15
1.5.3 Posibles funciones de los circRNAs	19
1.6 <i>Entamoeba histolytica</i> y su mecanismo de <i>splicing</i>	21
1.6.1 <i>Splicing</i> en <i>E. histolytica</i>	22
2. ANTECEDENTES	24
3. JUSTIFICACIÓN.....	27
4. HIPÓTESIS.....	27
5. OBJETIVOS.....	28
5.1 Objetivo general	28

5.2 Objetivos específicos	28
6 MATERIALES Y MÉTODOS	29
6.1 Diseño y síntesis de las clonas	29
6.2 Restricción, construcción de plásmidos y clonación	29
6.3 Transfección de los plásmidos en trofozoitos de <i>E. histolytica</i>	30
6.4 Cultivos de trofozoitos de <i>Entamoeba histolytica</i>	32
6.5 Extracción de RNA.....	32
6.6 RT-PCR.....	34
6.7 PCR en tiempo real (qPCR)	35
6.8 Análisis estadístico.....	35
7. RESULTADOS.....	38
7.1 Construcción de plásmidos y transfección.....	38
7.2 Efecto de los minigenes en la biogénesis de flicRX13 en <i>E. histolytica</i>	39
7.3 Efectos de los minigenes sobre la transcripción del gen parental en <i>E. histolytica</i> ..	41
7.4 Efectos de los minigenes sobre la transcripción del gen RabX13 nativo en <i>E. histolytica</i>	43
8. DISCUSIÓN.....	45
9. CONCLUSIONES.....	53
10. PERSPECTIVAS.....	54
11. BIBLIOGRAFÍA.....	55
APÉNDICES	65
Apéndice 1	65
Apéndice 2	66

Lista de abreviaturas

- RNA: Ácido ribonucleico, del inglés: Ribonucleic acid
- DNA: Ácido desoxirribonucleico, del inglés: Deoxyribonucleic acid
- flicRNAs: RNAs intrónicos circulares de longitud completa, del inglés: full length intronic circular RNA.
- 5´ss: sitio de splice 5´, del inglés: 5´ splice site
- 3´ss: sitio de splice 3´, del inglés: 3´ splice site
- Branch point (BP): sitio de ramificación
- ciRNAs: RNA circular intrónico, del inglés: circular intronic RNA
- PCBP: proteínas de unión a poli citosinas, del inglés: poly C binding proteins
- 3´ UTR: región no traducible 3´, del inglés: 3´ Untranslated region
- 5´ UTR: región no traducible 5´, del inglés: 5´ Untranslated region
- 5´ETS: Espaciador externo 5´ transcrito, del inglés: 5´ External Transcribed Spacer
- bp: pares de bases, del inglés: base pairs
- cDNA: ADN complementario, del inglés: complementary DNA
- CTD: Dominio carboxilo terminal, del inglés: carboxy terminal domain
- hnRNP: Ribonucleoproteína heterogeneonuclear, del inglés: heterogeneous RiboNucleoParticle
- kb: kilobases
- PCR: Reacción en cadena de la polimerasa, del inglés: Polymerase Chain Reaction
- mRNA: RNA mensajero, del inglés: Messenger RNA
- ncRNA: RNA no codificante, del inglés: non coding RNA
- nt: nucleótido
- Pol II: RNA Polimerasa II
- Pre-mRNA: precursor del mRNA, del inglés: precursor Messenger RNA
- PRM: Revoluciones por minuto
- RRN: Motivo de reconocimiento del RNA, del inglés: RNA recognition motif

- rRNA: RNA ribosomal, del inglés: ribosomal RNA
- RT: retro transcripción
- SF1: Factor de *splicing* 1, del inglés: *splicing* factor 1
- siRNA: RNA de interferencia pequeño, del inglés: small interfering RNA
- snRNP: Ribonucleoproteína pequeña nuclear, del inglés: small nuclear RiboNucleoProtein
- tRNA: RNA de transferencia, del inglés: transference RNA
- U2AF: Factor auxiliar de *splicing* de U2, del inglés: U2 Auxiliary Factor
- RT-PCR: Reacción en cadena de la polimerasa acoplada a reverso transcripción, del inglés: reverse-transcription polymerase chain reaction
- qPCR: PCR cuantitativa, del inglés: quantitative PCR

Resumen

Los RNAs circulares no codificantes se encuentran en todos los organismos eucariotas. Estos RNAs son un subproducto del proceso de splicing, recientemente se han asociado a funciones importantes como esponjas de microRNAs, en el *splicing* alternativo y en la regulación de la expresión de sus genes parentales; asociados a esta última función se encuentran los RNAs circulares que contienen regiones intrónicas (ciRNAs). Análisis de estas moléculas en humanos, han permitido determinar que la región 5' ss, rica en GU (5' ss-GU) y 11 bases ricas en citosinas que se encuentran muy conservadas y asociadas al sitio de ramificación (BP-C) en el extremo 3' que participan en la formación de estas estructuras circulares. Recientemente en *Entamoeba histolytica*, agente causal de disentería y de abscesos hepáticos en humanos, se han descrito por primera vez a los RNAs intrónicos circulares de longitud completa (flicRNAs). En este organismo, las regiones 5' ss de los intrones contienen un elemento GU similar al de los ciRNAs humanos, mientras que la región asociada al sitio de ramificación (-11 nt) carece de trectos de citosinas, estando ausente el BP-C característico de los genes RNAs humanos. Mediante la transfección de minigenes en amebas con mutaciones en estas 2 regiones, se analizó la participación de estos elementos en la biogénesis del flicRX13. A su vez se analizó el efecto del flicRX13 sobre la transcripción del gen parental. A diferencia de los ciRNA, la biogénesis de los flicRX13 solo depende del elemento GU y no del tracto rico en citosinas asociado al BP. Por otra parte, nuestros resultados indican que los flicRX13 regulan negativamente la transcripción de su gen parental en *trans*. La adición del tracto de citosinas provocó un silenciamiento del gen parental posiblemente a través de un mecanismo con proteínas del tipo PCBP.

Abstract

Circular non-coding RNAs are present in all eukaryotic organisms. These RNAs are a sub product of splicing process have been associated recently to important functions as microRNAs sponges, alternative splicing and regulation of parental genes expression. Related to this last function, have been described circular RNAs that contain intronic regions (ciRNAs). Genetic studies of these molecules have allowed to determine that in humans the region 5' splice site (5'ss), rich in GU (5'ss-GU) and 11 bases rich in cytosine associated with the branch point (BP-C) in the 3' splice site end, are well conserved and are involved in the formation of these circular structures. In the pathogen *Entamoeba histolytica*, causal agent of dysentery and liver abscess in humans, had been described recently for first time full-length intronic circularized RNAs (flicRNAs) molecules. In this organism, the intronic regions 5'ss has a GU element like in humans, while there is not a cytosine tract as in humans in the region associated (-11 nt) with the branch point. We analyzed the implication of these elements in the flicRX13's biogenesis by minigenes transfections in amoebas with mutations in these regions. Also, we studied the effect of flicRX13 over their parental gene's transcription process. Unlike ciRNAs, the flicRX13's biogenesis only depends of GU element and the cytosine tract associated to the branch point is not essential to this process. In addition, our results indicate that the flicRX13 negatively regulate their parental gene expression in *trans*. The inclusion of a cytosine tract in the intron's sequence led to a silencing of the parental gene's expression, possibly by a mechanism controlled by PCBP like proteins.

1. Introducción

1.1 El RNA

La síntesis de una molécula monocatenaria de ácido ribonucleico (RNA) mediante el proceso de transcripción del DNA, es el primer paso de la expresión génica y puede llevar a la síntesis de proteínas, en todos los tipos celulares. Por sus características moleculares, esta molécula funciona no solamente como un mensajero entre el DNA y las proteínas, además tiene funciones catalíticas y participa en procesos de regulación de la expresión de genes (Morris & Mattick, 2014).

Los tipos de RNA presentes en las células, se pueden clasificar para su estudio en los que tienen función reguladora y los involucrados en la síntesis de proteínas. En el primer grupo se encuentran los RNA reguladores y cumplen esta función en las células, debido a que en muchos casos su secuencia es complementaria a la de regiones de los mRNAs. En este grupo se encuentran: RNAs de interferencia (iRNA) formados por miRNAs (21-25 nucleótidos) y siRNAs (20-25 pb), RNA antisentido y RNA largos no codificantes que en humanos tienen un tamaño mayor a 200 nucleótidos (Esteller, 2011). Recientemente, se han descubierto moléculas de RNA con conformación circular que regulan la expresión génica por su asociación a proteínas y factores involucrados en este proceso (Zhang et al., 2013).

Los RNA asociados a la síntesis de proteínas son: el tRNA cuya función es transportar los aminoácidos correspondientes a cada codón específico para la síntesis de proteínas, los rRNA que unidos a proteínas forman complejos ribonucleoproteicos que participan en la traducción de proteínas y el mRNA, que es la molécula mediadora entre la información conservada en el genoma celular y la síntesis proteica (Lewin, 2008).

1.2 Procesamiento del mRNA

En eucariotas, el proceso de maduración del mRNA requiere varias etapas, *capping*, *splicing* y poliadenilación. Debido a la precisa regulación sobre la transcripción en las células, el pre-mRNA debe ser procesado antes de salir del

núcleo. La secuencia del DNA se transcribe a RNA por la RNA polimerasa II formándose una molécula precursora llamada pre-mRNA. En este pre-mRNA se encuentran secuencias codificantes (exones) y secuencias no codificantes (intrones). La maduración del pre-mRNA es un proceso que ocurre co-transcripcionalmente y donde el dominio carboxilo terminal (CTD) de la Polimerasa II (Pol II) desempeña un rol fundamental (Bentley, 2014; Fong & Bentley, 2001; Schroeder, Schwer, Shuman, & Bentley, 2000).

El *capping* es el proceso mediante el cual al primer nucleótido del transcrito nascente se le adiciona un nucleósido de guanina metilado en la posición N-7, en su extremo 5'. Esta modificación es necesaria para la correcta expresión génica y para la viabilidad celular ya que importantes procesos como la transcripción, el *splicing*, la poliadenilación, la estabilidad y el exporte del núcleo del mRNA y la traducción dependen de esta adición (Cowling & Cole, 2010).

Debido a la presencia de elementos no codificantes en la secuencia del pre-mRNA, estos deben ser removidos antes de la salida del núcleo. La remoción de los intrones ocurre mediante el *splicing*, quedando unidas todas las secuencias que serán traducidas a proteínas. El *splicing* del pre-mRNA es catalizado por el spliceosoma, complejo ribonucleoproteico formado por 5 ribonucleoproteínas nucleares pequeñas (snRNP) U1, U2, U4, U5 y U6. Cada snRNP está formada por varias proteínas y una molécula de RNA rica en uracilos, mediante la cual reconocen a la secuencia intrónica por apareamiento complementario de bases (Matera & Wang, 2014). La activación del spliceosoma es un proceso coordinado que requiere que las snRNP y otras proteínas y factores se ensamblen sobre el sustrato (pre-mRNA), cada vez que se requiera eliminar una de las secuencias no codificantes. Este mecanismo de remoción de secuencias se desarrolla debido a señales moleculares del pre-mRNA, que son reconocidas por los diferentes elementos del spliceosoma (Matera & Wang, 2014; Will & Lührmann, 2011). A este mecanismo de escisión de intrones se le dedicará un análisis más detallado en epígrafes posteriores.

El último paso del procesamiento del pre-mRNA tiene lugar en su extremo 3'. Un corte endonucleolítico en esta región, seguido de la síntesis de una cola de poliadeninas (poli A), es el último evento en la maduración del pre-mRNA. El sitio del extremo 3' donde ocurre el corte, es determinado por la interacción de secuencias específicas entre el pre-mRNA y un complejo multiproteico llamado complejo de procesamiento del extremo 3'. Este ensamblaje macromolecular está compuesto por una polimerasa poli A y 4 factores más: el factor de especificidad de corte de la poliadenilación (CPSF) de 360kDa, el heterodímero de factores de corte I_m y I_{II}m (CFI_m y CFI_{II}m) y el factor de estimulación de corte (CSTF) de 200 kDa (Misra & Green, 2016). La poli A polimerasa se une a este complejo antes de la degradación endonucleolítica, 35 nt río debajo del elemento AAUAAA y de esta manera el sitio 3' generado es rápidamente poliadenilado. El componente central de este complejo es el CPSF que reconoce a la secuencia AAUAAA de manera específica y realiza el corte catalítico (Fong & Bentley, 2001; Misra & Green, 2016).

1.3 Splicing

Los genes eucariotas son transcritos en forma de pre-mRNA, molécula precursora que es convertida a mRNA mediante el *splicing*. Este es un paso esencial en el proceso de expresión génica que permite aumentar de manera eficiente la variedad proteica celular (Bush, Chen, Tovar-Corona, & Urrutia, 2017; Marshall, Montealegre, Jiménez-López, Lorenz, & van Hoof, 2013; Yu, Tian, Yu, & Jiao, 2016). Para esto, las secuencias no codificantes son escindidas y las secuencias codificantes son ligadas para producir RNA maduros funcionales (mRNAs). Este proceso es llevado a cabo por el spliceosoma, el cual cataliza dos reacciones de transesterificación (Fig.1). Este macrocomplejo está formado por 5 ribonucleoproteínas pequeñas nucleares U1, U2, U4, U5 y U6 y más de 100 proteínas auxiliares (Matera & Wang, 2014), las cuales se ensamblan por interacciones coordinadas de RNA-RNA, RNA-proteína y proteína-proteína que conducen a la ligación de los exones y la liberación de los intrones en forma de *lariat* (Patel & Steitz, 2003).

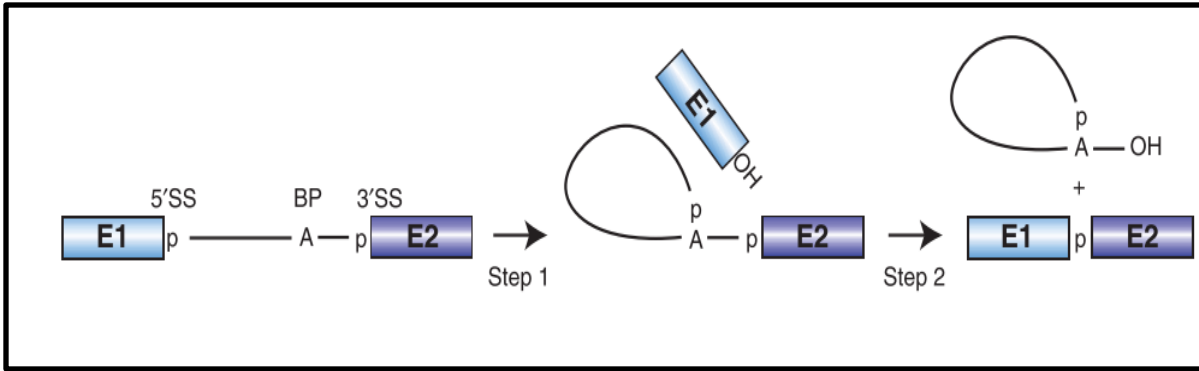


Figura 1. Representación esquemática de las dos reacciones de transesterificación durante el *splicing* del pre-mRNA. Las cajas y la línea negra representan los exones y el intrón, respectivamente. La adenosina del sitio de ramificación se indica con la letra A y los grupos fosfatos (p) en los sitios *splice* 5' y 3'. Ambos ataques nucleofílicos ocurren en las uniones fosfodiéster de cada extremo del intrón (p), el primero es realizado por el 2' OH de la adenosina (A) del sitio de ramificación y posteriormente ocurre un segundo ataque nucleofílico por el 3' OH libre del exón 1. Los productos finales de estas reacciones son los exones ligados y el intrón *lariat*. Figura tomada de (Will & Lührmann, 2011).

1.3.1 Mecanismos del *splicing* y regulación en *cis*

Los intrones son removidos del pre-mRNA mediante dos reacciones consecutivas de transesterificación. Primero, el grupo OH de la posición 2' de la adenina del sitio de ramificación ataca nucleofílicamente al sitio de *splicing* 5'. Esto resulta en un corte en este sitio y la unión del extremo 5' del intrón a la adenina del sitio de ramificación formándose una estructura en forma de *lariat*. En el segundo paso, el sitio de *splicing* 3' es atacado por el grupo OH 3' del exón, provocando el empalme de los exones mediante una unión 5'-3' (formándose el mRNA) y la liberación de la secuencia intrónica en forma de *lariat* (Fig. 1) (H. Chen & Cheng, 2012; Will & Lührmann, 2011).

El reconocimiento de los límites exón-intrón es un paso crítico del proceso de *splicing*, debido a que variaciones de un solo nucleótido podrían generar RNAs no funcionales. La alta conservación de las secuencias en los sitios *splice* 5' y 3' y el sitio de ramificación (*branch point*) garantiza que transcurra sin errores este mecanismo. El sitio de ramificación está típicamente localizado de 18 a 40 nucleótidos río arriba del sitio de *splicing* 3' y en eucariotas superiores (metazoos)

está seguido (río abajo) de un tracto de polipirimidinas (Fig 2) (Patel & Steitz, 2003).

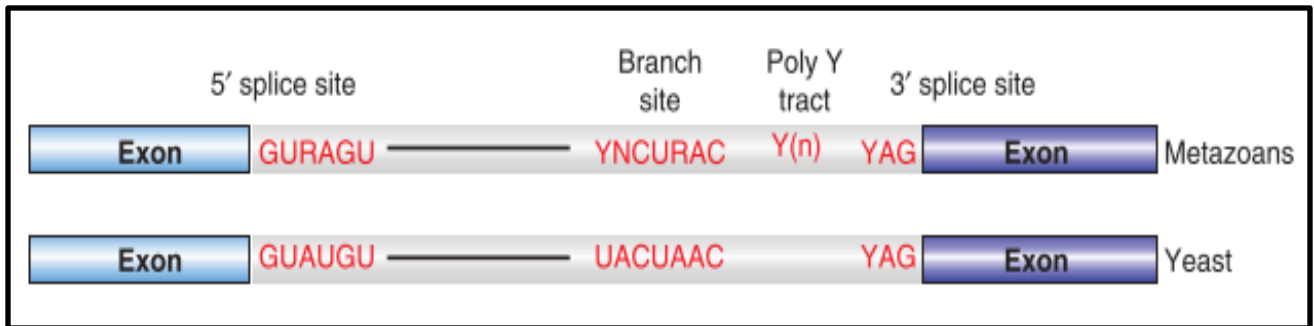


Figura 2. Elementos reguladores en cis del splicing. Secuencias conservadas encontradas en los sitios de splicing 5' and 3' y en el sitio de ramificación de los intrones tipo U2 en metazoos y levaduras (*S. cerevisiae*). Y = pirimidinas y R = purinas. El tracto de pirimidinas se indica con (Yn). Figura tomada de (Will & Lührmann, 2011).

Dos únicos tipos de spliceosomas coexisten en la mayoría de las células eucariotas. El spliceosoma U2 dependiente, cataliza la remoción de los intrones tipo U2 y el menos abundante spliceosoma U12 dependiente, este último está presente solo en algunos eucariontes y remueven los intrones de tipo U12 (Sharp & Burge, 1997; Will & Lührmann, 2011). Los intrones de tipo U2 (terminaciones GT-AG) son los más abundantes y tienen un sitio 5' compuesto por 8 nucleótidos altamente conservados así como el 3' y el sitio de ramificación que tiene una secuencia consenso YNCURAC (en negritas la adenina altamente conservada que participa en la primera reacción de transesterificación). Este tipo de intrones en levaduras carece del tracto de polimiridinas (Fig 2) (Patel & Steitz, 2003; Will & Lührmann, 2011). Los intrones de tipo U12 son menos abundantes y presentan terminaciones AT-AC. Principalmente se encuentran en metazoos y están ausentes en levaduras y nematodos. (Patel & Steitz, 2003).

1.3.2 Elementos reguladores en cis en la transcripción de genes

La selección correcta de los sitios de splicing 5' y 3' en el pre-mRNA es dirigida por los elementos reguladores en cis de la secuencia del premRNA (Z. Wang & Burge, 2008). Debido a la alta conservación de estos sitios en la mayoría de los

organismos, este proceso ocurre de manera efectiva por el reconocimiento de los dinucleótidos GU y AG en los sitios 5' y 3' respectivamente.

Los procesos de *splicing* y transcripción se encuentran estrechamente integrados en la célula (Braunschweig, Gueroussov, Plocik, Graveley, & Blencowe, 2013). Tal es el grado de interrelación entre ambos fenómenos que experimentos de inmunoprecipitación de la cromatina en levaduras han demostrado que los factores de *splicing* no se asocian con genes sin intrones y no ocurre la transcripción, pero estos factores si son reclutados por genes con intrones desarrollándose en estos casos la transcripción a partir de los sitios de *splicing* reconocidos (Görnemann, Kotovic, Hujer, & Neugebauer, 2005). Por otra parte, se ha observado que si los sitios de *splicing* son débiles (diferentes a las secuencias consenso), los exones involucrados pueden ser saltados al ocurrir la transcripción. La principal razón asociada a este proceso es la diferencia en afinidad de los factores de splicing de los elementos en *cis* entre estos sitios débiles y las regiones consenso de splicing (Braunschweig et al., 2013).

Desde las 2 décadas finales del siglo pasado, se conoce la importancia de los intrones sobre la transcripción de los genes. Experimentos en modelos animales y cultivos celulares (Callis, Fromm, & Walbot, 1987) mostraron un aumento en la eficiencia (de 10 a 100 veces más) de la transcripción de genes con regiones intrónicas incluidas, comparados con los mismos genes sin intrones (Brinster, Allen, Behringer, Gelinas, & Palmiter, 1988). Algunos de los efectos provocados por los intrones sobre la transcripción son: el potenciamiento en la procesividad de la Polimerasa II, así como en la iniciación de la transcripción: la asociación a través de U1 con factores de iniciación de la transcripción como TFIIH o la interacción de las U snRNPs con factores (TAT-SF1), que a su vez unen a cinasas (pTEFb) que fosforilan la CTD de la Polimerasa II (Le Hir et al., 2003). La fosforilación de la CTD de la Pol II es esencial para la iniciación y la síntesis del transcrito.

1.3.3 Ensamble, estabilidad y actividad del spliceosoma

El spliceosoma es un complejo enzimático que cataliza las reacciones de transesterificación durante el procesamiento del premRNA. Este complejo tiene un gran dinamismo, ya que se va reordenando a medida que ocurre el proceso debido a la entrada y salida de snRNAs y proteínas no snRNPs.

El primer evento en el ensamble del spliceosoma es la unión de U1 al sitio de splicing 5', conformando el complejo temprano o E. Debido a que el sitio de ramificación es degenerado en muchos eucariotas, la unión de U2 requiere la unión de múltiples factores auxiliares, siendo el más importante U2AF, un heterodímero que consiste de dos subunidades de 35kD y 65kD (Jenkins, Agrawal, Gupta, Green, & Kielkopf, 2013). La subunidad de 65kD (U2AF65) que contacta directamente el tracto de polipirimidinas, contiene un dominio rico en arginina/serina en el N-terminal, un dominio de interacción con la subunidad de 35kD (U2AF35) y tres dominios de unión a RNA (RRM). Por su parte, U2AF35 contiene una región con homología a un dominio de unión a RNA que se ha demostrado que interacciona con el dinucleótido AG del sitio de splicing 3' (Cho et al., 2015). Otra proteína que interacciona con el sitio de ramificación es el factor de splicing 1 (SF1), este se une cooperativamente con U2AF65 para terminar de formar el complejo E (Fig 3).

A continuación se forma el complejo A en el que U2 reconoce al sitio de ramificación en una reacción que depende de ATP (Crisci et al., 2015). Este proceso es dependiente de las helicasas Prp5 y Sub2. En una reacción posterior catalizada por Prp28 el tri-snRNP U4-U6-U5 es reclutado para formar el complejo B el cual requiere de las helicasas Brr2, Snu114 y Prp2 para la liberación de los snRNPs U1 y U4. La estabilización de este complejo depende del complejo Prp19. La remoción de los snRNPs U1 y U4 conduce a la formación del complejo B catalítico (Lee & Rio, 2015). Posteriormente, ocurre el primer paso catalítico del *splicing* generando el complejo C el cual contiene un exón libre y el intermediario *lariat*. Durante el complejo C el spliceosoma sufre una serie de re-arreglos permitiendo que se desarrolle la segunda reacción catalítica, resultando en un complejo post-spliceosoma que contiene los exones ligados y el intrón *lariat*. Por

último, los snRNPs U2, U5 y U6 son liberados por la acción catalítica de otras helicasas y son reciclados (Fig. 3) (Matera & Wang, 2014; Papasaikas & Valcárcel, 2016; Will & Lührmann, 2011; Wu & Fu, 2015).

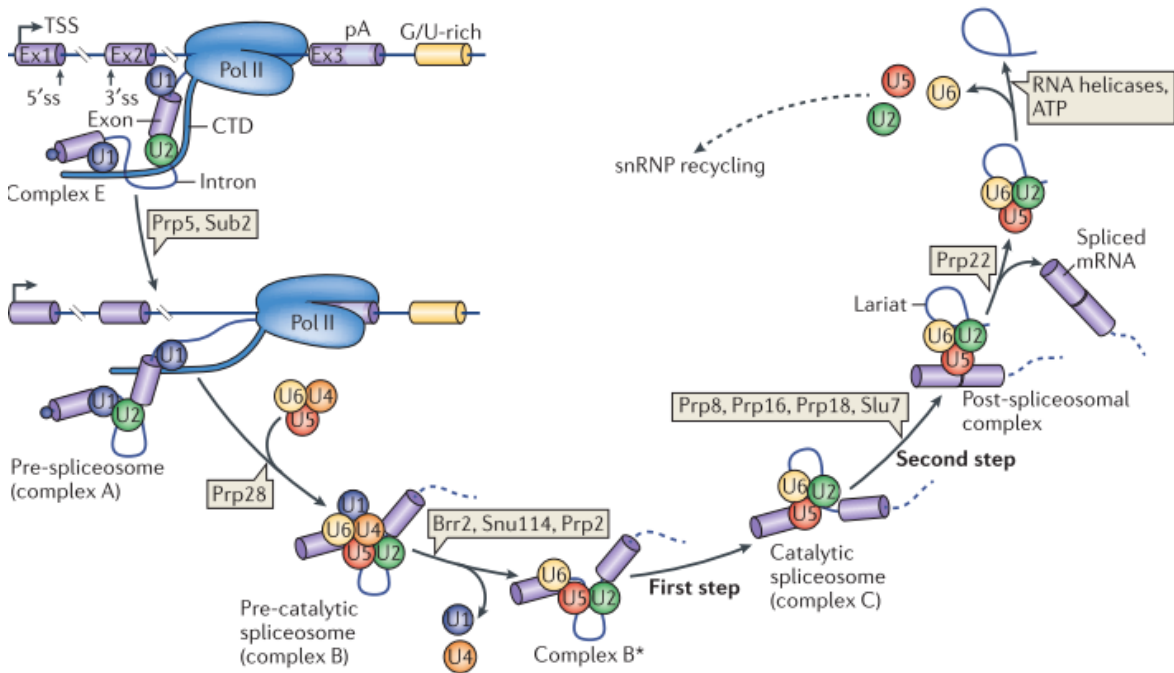


Figura 3. Ensamblaje del spliceosoma y pasos catalíticos del *splicing*. El ensamblaje de este complejo ocurre co-transcripcionalmente. Graficado solo se muestran las interacciones de los snRNPs (círculos). Los complejos del spliceosoma son nombrados de acuerdo a la nomenclatura de metazoos. Los exones e intrones son representados por cilindros y líneas, respectivamente. Se indican los nombres de helicasas que participan en el re-arreglo del spliceosoma.

1.3.4 Procesamiento del intrón *lariat*

El resultado final del *splicing* es la unión de los exones y la liberación de la secuencia intrónica en forma de *lariat*. Este *lariat* presenta un enlace fosfodiéster inusual entre 2'-5' debido al ataque de la adenina del sitio de ramificación a la guanina del sitio de *splice* 5' (Montemayor et al., 2014; Yoshimoto, Kataoka, Okawa, & Ohno, 2009)

Esta unión 2'-5' es hidrolizada por la enzima desramificadora del intrón (Dbr1) (Khalid, Damha, Shuman, & Schwer, 2005; Montemayor et al., 2014), la cual fue

descubierta en 1985 en extractos de células HeLa (Ruskin & Green, 1985). Esta enzima presenta un alto grado de conservación en muchos organismos (Masaki et al., 2015) siendo su función escindir la unión 2'-5' del *lariat* para dar lugar a moléculas intrónicas lineales que pueden ser empleadas en la células para formar microRNAs y RNAs nucleolares pequeños (snoRNAs), o son degradadas por exonucleasas 5'-3' como Xrn1 (Kataoka, Dobashi, Hagiwara, & Ohno, 2013; Montemayor et al., 2014). La actividad de esta proteína es esencial para el correcto funcionamiento celular, la delección del gen DBR1 en *Schizosaccharomyces pombe* y *Saccharomyces cerevisiae* causa defectos morfológicos y un crecimiento lento debido a la acumulación de *lariat* intactos (Katolik et al., 2017). En eucariotas superiores como *Arabidopsis thaliana* su delección es letal debido a la carencia de snoRNAs y microRNAs que se genera, los cuales son necesarios en el proceso de embriogénesis (H. Wang, Hill, & Perry, 2004). En líneas celulares humanas y experimentos en ratas se ha observado que una expresión reducida de la Dbr1 o problemas asociados a la regulación de la actividad de la enzima contribuyen a la formación de tumores (B. Han et al., 2017).

1.4 RNAs no codificantes

Una parte importante de los RNAs que son transcritos en metazoos son no codificantes. Estos RNAs no codificantes producen moléculas de RNA que a pesar de no codificar para proteínas cumplen importantes funciones en la célula (Mattick & Makunin, 2006). Estas moléculas funcionales de RNA se dividen en dos grandes grupos basados en su tamaño, los RNAs no codificantes pequeños (sncRNAs) menores a 200 nt, que incluyen microRNAs, snoRNAs, RNAs nucleares pequeños (snRNAs), RNAs ribosomales (rRNAs), RNAs de transferencia (tRNAs), RNAs heterogéneos nucleares (hnRNAs) los cuales cumplen diversas funciones durante la expresión génica y los RNAs grandes no codificantes (lncRNAs) los cuales tienen tamaños que varían de 200 nt hasta 100 kb y se han implicado en la regulación de la transcripción en *cis* y *trans*, estabilizando o promoviendo la traducción de mRNAs blanco, y participan en la diferenciación, proliferación y control del ciclo celular (Johnsson, Lipovich, Grandér, & Morris, 2014; Shi, Sun, Liu, Yao, & Song, 2013). Debido a los procesos en los que participan los lncRNAs,

las deficiencias en el procesamiento de estos, se ha asociado a muchas enfermedades en humanos (Esteller, 2011; Shi et al., 2013). Recientemente, se ha descrito una nueva clase de RNAs no codificantes clasificados dentro de los lncRNAs, estos son los RNAs circulares. Estas moléculas han cobrado gran interés científico debido a su ubicuidad y a las funciones que realizan en el contexto celular (Qu et al., 2015; Salzman, 2014; Zhang et al., 2013).

1.5 RNAs circulares

Actualmente, los RNAs circulares se han detectado en todos los sistemas biológicos (Salzman, 2014), estos pueden sintetizarse a partir de distintas reacciones. Debido a su conformación, esta clase de moléculas de RNA tiene propiedades únicas, entre las que se encuentran: servir como molde para amplificación por círculo rodante, esponjas de miRNAs o de proteínas de unión a RNA, re-arreglar el orden de la información genética, resistencia a exonucleasas, característica que les confiere mayor estabilidad que las moléculas lineales, regulación de la transcripción génica y además su estructura circular les permite restringir la formación de estructuras secundarias (Ebbesen, Kjems, & Hansen, 2016; Lasda & Parker, 2014).

Se han descrito varios tipos de RNAs circulares: los genomas circulares de RNA en viroides encontrados en parásitos, los cuales se forman por la ligación de los extremos 3' y 5' y que permiten la replicación mediante el mecanismo de círculo rodante; los RNAs de secuencias intrónicas encontrados en algunas bacterias, arqueas y virus los cuales son un producto secundario del procesamiento del RNA y pertenecen a los intrones grupo I e intrones del grupo II y que posiblemente propician la movilidad de elementos genéticos. En esta misma categoría se encuentran círculos procedente de intrones *lariat* encontrados en los eucariotas, los cuales se forman por el ataque nucleofílico 2'-5' de la adenosina del sitio de ramificación, y se les ha asociado con la regulación de la transcripción (Zhang et al., 2013) ; los RNAs circulares intermediarios encontrados en arqueas, los cuales se forman a partir de precursores de rRNAs y tRNAs, estas moléculas circulares también son originadas a partir de la unión de los extremos 3' y 5'; los RNAs circulares no codificantes encontrados en arqueas, como los snoRNAs y la

RNasa P los cuales se originan a partir de la unión de los extremos 3' y 5' lo que les proporciona estabilidad; y los RNAs circulares constituidos por secuencias exónicas o exónicas e intrónicas encontrados en eucariotas, productos secundarios generados por *backsplicing* y que funcionan como reguladores de la traducción y esponjas de miRNAs (Lasda & Parker, 2014).

1.5.1 RNAs circulares de eucariotas

Debido a que se han encontrado isoformas de circRNAs en varios tipos de eucariontes desde hongos (*S. pombe* y *S. cerevisiae*) y protozoos (*P. falciparum* y *D. discoideum*) hasta plantas (*A. thaliana*) y humanos, incluso cuando los primeros tienen intrones pequeños (100 nucleótidos o menos) y debido a que su ancestro común más reciente existió hace más de un billón de años, se ha propuesto que los RNAs circulares pueden ser una característica funcional antigua y conservada en los sistemas de expresión génica (Salzman, 2014).

Los RNAs circulares (circRNAs) difieren conformacionalmente de otros lncRNAs encontrados en eucariotas debido a que sus extremos 3' y 5' están unidos covalentemente, además de ser expresados en todos los tejidos y organismos eucariotas investigados hasta la fecha (Ebbesen et al., 2016). Recientemente se ha encontrado que estas moléculas están involucradas en un gran número de procesos biológicos (Qu et al., 2015). En eucariotas superiores estos circRNAs son generados por diferentes mecanismos, algunos de estos son: formación a partir de secuencias exónicas o intrónicas, regiones intergénicas, UTR's y secuencias antisentido de transcritos conocidos (L.-L. Chen, 2016; Vicens & Westhof, 2014).

1.5.2 Biogénesis de los circRNAs eucarióticos

Independientemente del mecanismo de formación, los circRNAs se forman a partir de pre-mRNAs que son transcritos por la RNA polimerasa II. La regulación de la biogénesis depende, en principio, de elementos en *cis* y *trans* que regulan el proceso de *splicing* (L.-L. Chen, 2016).

Biogénesis de circRNAs dependiente del spliceosoma

Este tipo de moléculas en eucariotas provienen del *splicing* del premRNA mediado por el spliceosoma (Fig. 4A) donde el 5'ss río abajo se une a un 3'ss río arriba generando así una molécula circular (Lasda & Parker, 2014) con secuencias exónicas. A pesar de que este mecanismo de formación de moléculas circulares de RNA depende del spliceosoma, el mecanismo por el cual este participa en el proceso de *backsplicing* aún no se ha dilucidado (L.-L. Chen, 2016). Se ha demostrado mediante experimentos de mutagénesis con minigenes que la eficiencia en la formación de circRNAs depende de elementos de *splicing* canónicos (Starke et al., 2015). Debido a que los niveles de estos circRNAs son menores que los de los productos lineales, se piensa que la eficiencia del *backsplicing* es menor a la del *splicing* canónico, lo cual pudiera estar explicado por el hecho de que posiblemente el spliceosoma es ensamblado incorrectamente durante este proceso.

Biogénesis por apareamiento de elementos en cis

Muchas circRNAs en eucariontes superiores se forman debido al apareamiento de secuencias intrónicas complementarias. La unión de estas secuencias provoca en muchas ocasiones la circularización del RNA con secuencias exónicas involucradas. Muchos circRNAs en mamíferos, en *C. elegans* y algunos en *D. melanogaster* presentan secuencias de exones flanqueados por intrones de gran tamaño que usualmente contienen secuencias complementarias que pueden aparearse y forman dúplex de RNA que incrementan el proceso de *backsplicing* (Fig. 4B) (L.-L. Chen, 2016; Qu et al., 2015). Estos elementos en *cis* facilitan la formación de circRNAs. Secuencias complementarias de 30 a 40 nucleótidos son suficientes para propiciar la biogénesis de algunas moléculas circulares.

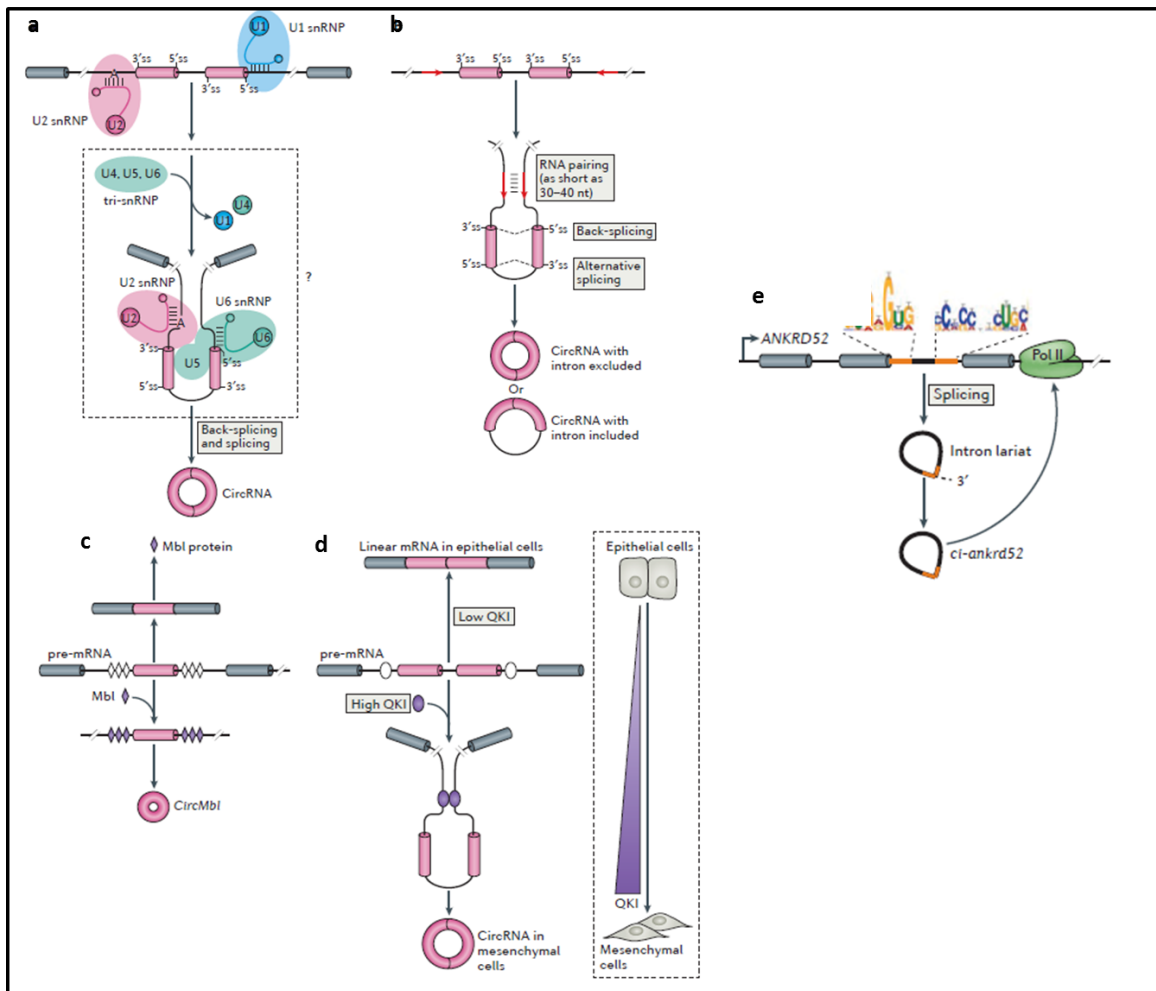


Figura 4. Biogénesis de circRNAs y su regulación. A) *Backsplicing*; ligación de un sitio donador 5' río abajo del sitio aceptor 3' mediado por la maquinaria canónica del *splicing*. B) Apareamiento entre intrones; secuencias complementarias propician mediante su apareamiento la formación de estructuras circulares conteniendo regiones exónicas e intrónicas. C) Proteínas de unión a RNA; la proteína Mbl incrementa la formación de RNAs circulares de su propio gen uniéndose a secuencias de los intrones que flanquean al exón circularizado. Los diamantes blancos representan los sitios de unión de Mbl, D) QK1 es sobreexpresada en células mesenquimales lo cual facilita la circularización de exones. Los círculos blancos representan los sitios de unión de QK1. E) El intrón *lariat* es generado durante el *splicing*. La presencia de secuencias ricas en GU asociadas al sitio de *splice* 5' y ricas en C en la proximidad del sitio de ramificación permiten que el *lariat* escape de la acción de la DBR1. Figura tomada y modificada de (L.-L. Chen, 2016).

Biogénesis regulada por proteínas de unión a RNA

En adición a los elementos en *cis*, se ha reportado también que proteínas de unión a RNA (RBPs) pueden participar en la biogénesis de circRNAs (Conn et al., 2015;

Kramer et al., 2015). Entre estas proteínas se ha descrito el factor de splicing Mbl (Muscleblind) en *D. melanogaster* que regula la producción de circRNAs debido a que los intrones que se encuentran flanqueando al exón que es circularizado tienen múltiples sitios de unión para Mbl (Fig. 4C). Además una sobreexpresión de esta proteína o la adición de sitios de unión extra para la proteína estimularon la circularización del exón, sugiriendo que Mbl estimula la producción de circRNAs (Ashwal-Fluss et al., 2014).

Otra RBP que aumenta la formación de circRNAs debido a la unión al premRNA es *quaking* (QKI) (Fig. 4D). Esta proteína promueve la biogénesis de estas moléculas durante la transición epitelial- mesenquimal humana debido a que más de un tercio de los circRNAs que se forman en este tipo celular son regulados por el factor de *splicing* QKI. La adición sintética de sitios de unión de esta proteína en los intrones de genes que no circularizan es suficiente para promover la formación de circRNAs (Conn et al., 2015).

Los resultados de ambos estudios (Ashwal-Fluss et al., 2014; Conn et al., 2015) sugieren que estos factores de *splicing* modifican el balance entre la biogénesis de los circRNAs y el *splicing* canónico debido a que se observó una competencia en la formación de ambos productos.

Biogénesis dependiente de splicing

Uno de los productos del *splicing* es el intrón *lariat*. El primer paso en el procesamiento de este es su linearización. La acción de la enzima desramificadora permite la ruptura del enlace 2'-5' formándose una molécula lineal que puede ser degradada o empleada para generar microRNAs y snoRNAs. En investigaciones recientes (Zhang et al., 2013) se detectó un tipo de RNA circular en células humanas formado solamente por secuencias intrónicas, el cual se nombró ciRNA (circular intronic RNA).

Experimentos de mutagénesis y construcción de minigenes permitieron determinar que la biogénesis de estos nuevos círculos depende de dos elementos conservados en algunos intrones humanos. Un elemento de 7 nucleótidos rico en

GU asociado al sitio de *splice* 5' y otro elemento de 11 bases rico en C cerca del sitio de ramificación. Estos dos elementos son suficientes para que el intrón por un mecanismo aún no dilucidado escape a la acción de la enzima DBR1, por lo que la unión 2'-5' se mantiene. El extremo 3' que se encuentra río abajo del sitio de ramificación es degradado por nucleasas generándose de esta manera un RNA circular estable (Fig. 4E).

1.5.3 Posibles funciones de los circRNAs

La relativa baja expresión de los circRNAs comparados con su contraparte lineal pudiera sugerir que estos productos de la transcripción en eucariontes carecen de función. No obstante, estudios recientes han mostrado que algunos circRNAs tienen importantes funciones biológicas y pueden regular la expresión génica en diferentes niveles.

Los circRNAs pueden actuar como esponjas de miRNAs

La gran mayoría de los RNAs circulares formados por *backsplicing* se localizan en el citoplasma. A pesar de encontrarse en este sitio, aún no se ha reportado que estos se asocien con ribosomas y puedan ser traducidos (Conn et al., 2015; Y. Wang & Wang, 2015). Se sugiere que debido a la mayor estabilidad de los circRNAs comparados con el mRNA, estos compitan en el citoplasma con el mRNA por la unión a miRNAs y a través de este mecanismo regulen la expresión genética (L.-L. Chen, 2016; Jeck et al., 2013).

De los circRNAs estudiados, el que más confirma este modelo es ciRS-7 el cual actúa como esponja de miR-7. Este RNA circular se produce a partir del transcrito antisentido del gen CDR1 y se ha detectado principalmente en el cerebro de humanos y ratones. Su localización citoplasmática y la presencia en su estructura de más de 60 sitios blanco conservados para miR-7, le confieren a este circRNA la capacidad de unir hasta a 20,000 moléculas (miR-7) por célula, lo cual tiene un marcado efecto en la traducción ya que no permite que miR-7 se una a sus sitios blanco en el mRNA (Hansen et al., 2013; Memczak et al., 2013).

Aunque los sitios de unión a miRNA en circRNAs necesitan estar conservados para funcionar como esponjas de miRNAs, la mayoría de los circRNAs analizados hasta ahora en humanos y ratones no presentan de manera frecuente estos sitios conservados, indicando que la mayoría de los círculos de RNA no funcionan como esponjas de miRNAs (L.-L. Chen, 2016).

Los RNAs circulares regulan la transcripción en cis

A pesar de que la mayoría de los RNAs circulares se localizan en el citoplasma, los circRNAs que contienen intrones se localizan en el núcleo. Se cree que la retención de estos circRNAs en el núcleo es un mecanismo similar al que ocurre con los RNAs lineales que contienen intrones (L.-L. Chen, 2016). Debido a que la ubicación de estos es nuclear y tienen pocos sitios de unión a miRNAs se sugiere que tengan diferente función que los citoplasmáticos. Algunos ciRNAs que se expresan abundantemente como ci-ankrd52 y ci-sirt7 interactúan con el complejo de elongación de la RNA polimerasa II. De manera interesante la depleción de estos ciRNAs provoca una disminución en los niveles de transcripción de sus genes correspondientes (ANKRD52 y SIRT7), lo que sugiere que estos RNAs circulares promueven la transcripción de sus genes parentales (Zhang et al., 2013).

En ensayos de inmunoprecipitación contra la polimerasa II (CLIP) seguidos de secuenciación de RNA, se han detectado RNAs circulares con secuencias intrónicas asociados a esta enzima (Li et al., 2015). Estos circRNAs predominantemente localizados en el núcleo, interactúan con U1 snRNP, con la Pol II y con la secuencia promotora de su gen parental. El complejo circRNA – U1snRNP se une a la Pol II promoviendo la transcripción de su gen parental. Estos análisis revelan un nuevo rol de los circRNAs; la regulación de la expresión génica, al provocar un potenciamiento en la expresión del gen parental en *cis*. Por otra parte, estos estudios muestran una nueva vía de regulación del control transcripcional a través de la interacción RNA-RNA entre U1 snRNA y este tipo de circRNA (Li et al., 2015).

Por otra parte, el “*knockdown*” de estas moléculas circulares redujo la transcripción de sus genes parentales. Cuando se bloquean las interacciones RNA-RNA entre U1 snRNP y el circRNA poniendo en el sistema U1 antisentido morpholino (AMO) disminuyen la interacción de estos circRNAs con la polimerasa II y disminuye los niveles del mRNA del gen parental (Li et al., 2015). Estos resultados sugieren que algunos RNAs circulares nucleares con secuencias intrónicas participan en la regulación de la transcripción. No obstante, los mecanismos mediante los cuales esta regulación tiene lugar aún no se han dilucidado completamente (L.-L. Chen, 2016).

1.6 *Entamoeba histolytica* y su mecanismo de *splicing*

Entamoeba histolytica es el patógeno intestinal protozoario con mayor prevalencia a nivel mundial. Presenta dos estadios en su ciclo de vida, el trofozoito que es la forma patogénica móvil con capacidad replicativa (diámetro 10-60 μm) y que pueden invadir múltiples órganos y el quiste (diámetro 10-20 μm) que se desarrolla bajo condiciones desfavorables para la vida del trofozoito. La infección se materializa al ingerir los quistes. *E. histolytica* es el agente causal de la disentería amebiana y/o abscesos hepáticos, padecimientos que provocan alrededor de 100 000 muertes anuales, principalmente en países tropicales y en vías de desarrollo (Begum, Quach, & Chadee, 2015; Stauffer, Abd-Alla, & Ravdin, 2006; Ximénez, Morán, Rojas, Valadez, & Gómez, 2009)

La virulencia de este protozoo se atribuye principalmente a su capacidad de adhesión a los tejidos y de secreción de factores que conducen a la muerte celular. Moléculas en la superficie de la ameba controlan la adherencia, la señalización y la modulación de la respuesta inmune en la interface hospedero-parásito. La destrucción de la matriz extracelular, inducción de apoptosis y la trogocitosis (proceso mediante el cual el trofozoito ingiere pedazos de las células vivas) son algunos de los mecanismos principales de patogenicidad de amebas (Marie & Petri Jr, 2014). A pesar de estos mecanismos, las infecciones asintomáticas por *E. histolytica* son muy comunes ya que solo entre el 10 y 20% de los individuos infectados presentan síntomas de amebiasis invasiva (Ximénez et al., 2009).

La cepa HM1-IMSS de *E.histolytica* (ATCC 30459) fue aislada a partir de una ulcera rectal de un adulto masculino humano con disentería amebiana en la Ciudad de México en 1967. A pesar de ser morfológicamente idéntica a otras cepas de *Entamoeba*, esta cepa es bioquímica y genéticamente diferente a otras siendo capaz de provocar amebiasis (Solaymani-Mohammadi et al., 2006).

Debido a la importancia clínica de este patógeno el estudio de sus mecanismos de virulencia así como de la regulación de su transcripción resulta de gran interés (Valdés et al., 2014).

1.6.1 Splicing en *E. histolytica*

Los eucariontes de ramificación temprana presentan pocos o carecen de intrones en sus genomas. De manera interesante, en *E. histolytica* se han identificado alrededor de 4328 intrones dentro de los casi 8201 genes reportados (Lorenzi et al., 2010; Valdés et al., 2014).

La secuenciación del genoma de *E. histolytica* (Loftus, 2005) permitió determinar que este se caracteriza por su alto contenido de adeninas y timinas (75.9%) y por la abundancia de secuencias repetitivas de DNA (Willhoeft, Campos-Góngora, Touzni, Bruchhaus, & Tannich, 2001). El genoma de 20.8 Mb contiene 8201 genes con un tamaño promedio de 1260 pb y el de los intrones de 74.1 bp (Loftus, 2005; Lorenzi et al., 2010; Willhoeft et al., 2001).

Este organismo presenta un alto grado de conservación en las secuencias asociadas a los sitios de *splicing* 5' (GUUUGUUU) y 3' (AUG) (Valdés et al., 2014; Willhoeft et al., 2001). A diferencia de estas regiones, el sitio de ramificación no se encuentra conservado y solo ha podido identificarse *in silico*, siendo la secuencia consenso propuesta YNYYRAY (Willhoeft et al., 2001).

Análisis de secuenciación de RNAs poliadenilados permitieron conocer que el 48% de los genes de *E. histolytica* originan transcritos con uniones exónicas alternativas, sugiriendo el uso de sitios de *splicing* 5' y 3' alternativos. Se encontró una correlación positiva entre la abundancia de mRNA y los procesos de *splicing* alternativo en el 17% de los genes modelos sugiriendo que estos eventos

alternativos son más probables que ocurran en los transcritos más abundantes (Hon et al., 2013).

La identificación de los snRNPs U2, U4, U5 y U6 (Davis, Brown, & Singh, 2007; Miranda, Salgado, Sánchez-López, Alagón, & Lizardi, 1996) han permitido comprender mejor el proceso de splicing en este organismo, mientras que estudios más recientes permitieron identificar 32 factores de splicing (Valdés et al., 2014).

2. Antecedentes

Los RNAs circulares no codificantes se encuentran en todos los organismos eucariotas y son un subproducto del proceso de *splicing*. Aunque estos RNAs son generalmente expresados en bajos niveles, recientemente se han asociado a importantes funciones como son: esponjas de microRNAs, *splicing* alternativo y regulación de la expresión de sus genes parentales (L.-L. Chen, 2016). El proceso de biogénesis, regulación y las posibles funciones de estos RNAs circulares aún no está bien dilucidado, por lo que el estudio de los circRNAs se ha convertido en uno de las áreas más atractivas dentro del campo de la biología molecular (Qu et al., 2015).

Recientemente, se encontraron RNAs no codificantes (microRNAs) en *E. histolytica* (Mar-Aguilar et al., 2013) y RNAs circulares del espaciador transcrito externo 5' (5'ETS) de los genes de rRNA. Estos RNAs circulares provocan la acumulación de moléculas no procesadas de pre-rRNA, por lo que se cree que los etsRNAs participan en la inhibición del procesamiento del pre-rRNA y no son un producto destinado a degradación (Bhattacharya, 2015).

Dentro de los circRNAs, aquellos con regiones intrónicas parecen estar particularmente asociados a la regulación de la transcripción de los genes parentales (Li et al., 2015). Estudios en intrones humanos que no forman círculos de RNA, a los que se les adicionó secuencias de intrones que forman circRNAs (ci-ankrd52) (Zhang et al., 2013) permitieron determinar que las secuencias ricas en GU en el sitio de *splice* 5' y un tracto de 11 bases rico en C asociado al sitio de ramificación son suficientes para promover la circularización intrónica a partir del intrón *ariat*. En este mismo estudio se comparó la secuencia de estos sitios en 20 genes humanos que presentan intrones que forman círculos y se demostró en todos un alto grado de conservación en estas regiones, lo que sugiere que estos elementos son necesarios para la formación de circRNAs de naturaleza intrónica.

En *E. histolytica*, agente causal de disentería y de abscesos hepáticos han sido descritos recientemente por primera vez los RNAs intrónicos circulares (Mendoza-Figueroa, 2016). A diferencia de los detectados por (Zhang et al., 2013) estos

RNAs circulares están constituidos por la secuencia intrónica completa, por lo que se les dio el nombre de flicRNAs (RNAs circulares intrónicos de secuencia completa).

Debido a que lo descrito por (Zhang et al., 2013) fueron moléculas circulares de RNA de naturaleza intrónica pero con unión 2'-5' por su origen de intrón *ariat*, el hallazgo en amebas de los flicRNAs resultó de gran interés ya que sugiere un mecanismo de formación diferente al propuesto por (Zhang et al., 2013). Los flicRNAs presentan la característica de estar unidos por los sitios de *splicing* 3' y 5' y se originan en un evento co-transcripcional, ya que al inhibir la transcripción con Actinomicina D las variantes de *splicing* de su gen parental dejaron de detectarse, no así el RNA circular, el cual permaneció detectable por mayor duración de tiempo, aunque finalmente dejó de detectarse (Mendoza-Figueroa, 2016). Esto indica que estos flicRNAs tiene una vida media mayor que su contraparte lineal, característica común en los circRNAs.

Experimentos de sobreexpresión de la enzima Dbr1 y una mutante sin el carboxilo terminal (Dbr1 Δ C), región que ha sido reportada que incapacita a la enzima para realizar la reacción de desramificación (Montemayor et al., 2014) demostraron que en *E. histolytica* esta enzima participa de alguna manera en la formación del flicRNA del gen RabX13 (Palacio-Morales, 2016).

Según el modelo propuesto por (Zhang et al., 2013) las secuencias consenso cerca del sitio de *splicing* 5' y el sitio de ramificación del mRNA de *ci-ankrd52* provocan que la enzima Dbr1 no puede escindir el enlace 2'-5' que se genera. El intrón del gen Rabx13, modelo de este trabajo (que codifica para una GTPasa pequeña), presenta una secuencia rica en GU en el sitio de *splicing* 5', de manera similar al gen analizado por (Zhang et al., 2013). La región cercana al sitio de ramificación (-11 nt) es rica en adeninas, característica de los intrones amebianos (Fig 5).

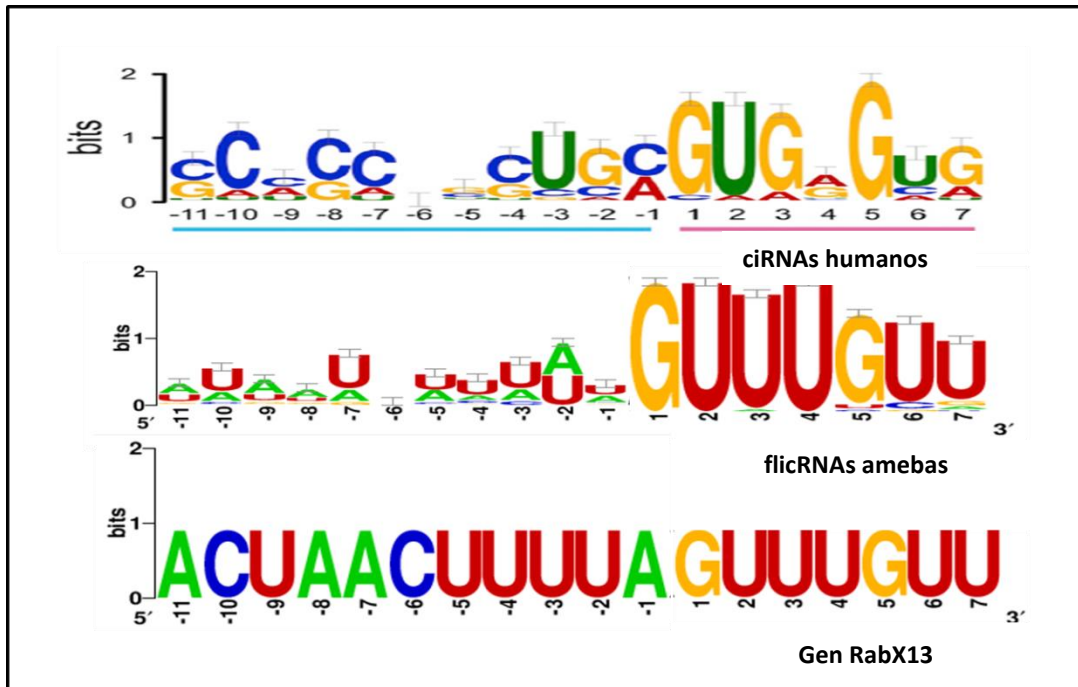


Figura 5. Elementos consensos del RNA en los sitios de *splice* 5' y el sitio de ramificación (WebLogo). Se muestran la conservación de estos sitios en los ciRNAs (Zhang et al., 2013) y en los flicRNAs de amebas detectados hasta la fecha y en el gen RabX13 (modelo de trabajo). Figura modificada de (Zhang et al., 2013).

Debido a que no se conocen los mecanismos ni los factores involucrados en la biogénesis y la regulación de la síntesis de flicRNAs, el análisis de la implicación de las secuencias de estos elementos en la formación de estas moléculas es de gran interés. El estudio del efecto sobre la formación de los flicRNAs de diferentes mutantes en estos sitios, permitirá ampliar el conocimiento acerca de este mecanismo en *E. histolytica*. Por su naturaleza intrónica, es muy probable que estas moléculas participen en procesos de regulación génica, por lo que también se analizará el efecto sobre la expresión del gen parental (RabX13) de cada uno de los mutantes desarrollados.

3. Justificación

Recientemente se describieron en *Entamoeba histolytica*, RNAs intrónicos circulares de longitud completa (flicRNA). Se demostró que éstos, son productos del procesamiento de premRNA y que son más estables que las variantes de *splicing* relacionadas. Los flicRNAs amebianos están formados por la secuencia completa del intrón y son circularizados por una unión 3'-5', a diferencia de los ciRNAs que provienen del intrón *lariat* y presentan unión 2'-5'. En el caso de los ciRNAs, los elementos asociados a los sitios de *splice* 5' y al sitio de ramificación presentan un alto grado de conservación en secuencias ricas en GU y C respectivamente, lo que permite que escapen de la acción de la enzima Dbr1. En los flicRNAs se ha comprobado la presencia del elemento rico en GU en el sitio de *splice* 5', no así la presencia de C en las bases relacionadas al sitio de ramificación. Los datos anteriores indican que estos elementos (5' ss y BP) están relacionados con la circularización del RNA, a pesar de sugerir que la biogénesis de los flicRNAs es distinta a la de los ciRNAs de mamíferos. La incidencia de estos elementos en la formación de los flicRNAs se comprobará removiendo o adicionando los mismos a la secuencia intrónica de RabX13.

4. Hipótesis

Las regiones 5'ss y BP participan en la circularización intrónica de RabX13, su modificación en la secuencia del intrón de este gen en *E. histolytica* afecta este proceso.

5. Objetivos

5.1 Objetivo general

Analizar cómo se afecta la circularización intrónica y la expresión del gen parental (RabX13) en *Entamoeba histolytica* HM-1:IMSS transfectada con minigenes con intrones mutados en los elementos 5' ss y BP.

5.2 Objetivos específicos

- Obtener construcciones del intrón del gen RabX13 (mutadas en los sitios 5' ss y BP) y el vector pEhExHA.
- Transfectar amebas con las construcciones pEhExHA + intrón RabX13.
- Comprobar la formación de flicRNAs de RabX13 en cada uno de los casos.
- Evaluar la expresión génica de RabX13 en las amebas transfectadas con cada construcción.

6 Materiales y Métodos

6.1 Diseño y síntesis de las clonas

A partir del análisis de la secuencia del gen RabX13 (EHI_065790), se diseñaron secuencias intrónicas mutadas en sitios asociados al 5' SS y BP. A cada secuencia se le colocó un sitio de corte para la enzima Hind III (AAGCTT), una secuencia complementaria para el oligo RabRex 2 (ver Tabla 2) y se flanquearon con los sitios de corte de XmaI (CCCGGG) y XhoI (CTCGAG). La información referente a las secuencias sintetizadas se encuentra resumida en la Tabla 1. La síntesis de las secuencias fue realizada por la compañía GenScript.

6.2 Restricción, construcción de plásmidos y clonación

Las secuencias ordenadas fueron enviadas en el vector pUC57 (2710 pb). Estas construcciones se restringieron con las enzimas XmaI y XhoI (NEB No. de catálogo #R0180L y #R0146L) por dos horas a 37°C para liberar cada una de las secuencias. El resultado de la restricción se visualizó en un gel de agarosa al 1%.

Una vez liberados los fragmentos de interés, se procedió a purificarlos 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 (BIO-RAD Freeze 'N Squeeze DNA Gel Extraction Spin Columns No. de catálogo # 732-6165). Las columnas se incubaron a -20°C durante 10 minutos y luego se centrifugaron a 10000 revoluciones por minuto (RPM) por 15 minutos. Al precipitado se le añadió 2.5 volúmenes de etanol absoluto, 0.2 volúmenes de acetato de sodio al 3M y 1µg de glicógeno. Se agitaron y se incubaron a -70°C durante 1 hora. Pasado el período de incubación se centrifugaron nuevamente las columnas a 10000 RPM por 15 minutos y se desechó el sobrenadante. La pastilla que se formó, se lavó con 700 µL de alcohol al 70% y se volvió a centrifugar durante 10 minutos a 10000 RPM. Se desechó el sobrenadante y la pastilla se resuspendió en 20 µL de H₂O mQ estéril. El resultado de la precipitación se verificó en un gel de agarosa al 1%.

Después de purificadas las secuencias de interés se procedió a ligarlas con el plásmido de sobreexpresión pEhExHA de *E. hystolitica* (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 1:3

en 4 μL de Buffer de Ligación 5x con 1 μL de Ligasa T4 (Invitrogen No. de catálogo 15224-017) completando a 20 μL con H_2O 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 μL de bacterias y 10 μL de producto ligado, se incubaron en hielo por 20 minutos, luego se pusieron a 42 °C por 45 segundos y nuevamente se incubaron las muestras en hielo por 2 minutos. Luego de este tiempo se adicionó a cada tubo de reacción 800 μL de medio súper óptimo (SOB) y se dejaron crecer las bacterias por 1 hora y 20 minutos en agitación a 37°C. Pasado este periodo, se tomaron 200 μL de cada transformación y se espatularon en placas con medio sólido Luria-Bertani 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, Manfioletti, & Schneider, 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 comprobar la inserción de las secuencias de interés en el vector de expresión, se realizó una restricción por 2 horas a 37 °C con las enzimas XmaI y XhoI de de las clonas con retraso, para comprobar la liberación de los fragmentos.

6.3 Transfección de los plásmidos en trofozoitos de *E. histolytica*

La transfección de la cepa HM1: IMSS de *E. histolytica* se realizó en una placa de 12 pozos. Se sembraron 750,000 amebas por pozo, en los cuáles se adicionó 150 mL de medio TYI-S-33 y se incubaron a 37°C en presencia de CO_2 durante 30 minutos permitiendo la adhesión a la superficie de los pozos.

Tabla 1. Nombre y secuencia de las moléculas sintetizadas. Las mutaciones se muestran entre []. En letras mayúsculas se muestran regiones exónicas y en minúsculas la secuencia de los intrones. La secuencia complementaria al oligo RabRex 2 se muestra en mayúsculas y en formato de negrita. El sitio de corte de Hind III está en letras cursivas y los sitios de corte de XmaI y XhoI se encuentran subrayados.

Nombre de las secuencias	Secuencia
Rab WT con sitio de HindIII (<i>ttcga</i>)	<p><u>ccggg</u>AAAAGATATGAAGTAGATGGAACGACATATAAA GTTCAAATATGGGATACTGgtttgttttaaataacttta aaagataaaaaacaaaaatgattttcaagacagtaatt cggttttataagtagaaaaagaagatgagatgagataa gattttatctttcgaaaataaatataaactaactttta ttttagCTGGACAAGAAAAATTTAGAACTATTATTACT AGTTATTATCGTGGTGTTCAA<u>tgactcgag</u></p>
Rab Box ΔGU con sitio de HindIII (<i>ttcga</i>)	<p><u>ccggg</u>AAAAGATATGAAGTAGATGGAACGACATATAAA GTTCAAATATGGGATACTGgt [aa]g [aa] ttaaatac tttaaaagataaaaaacaaaaatgattttcaagacagt aattcggttttataagtagaaaaagaagatgagatgag ataagattttatctttcgaaaataaatataaactaac ttttatttagCTGGACAAGAAAAATTTAGAACTATTAT TACTAGTTATTATCGTGGTGTTCAA<u>TGActcgag</u></p>
Rab Box C con sitio de HindIII (<i>ttcga</i>)	<p><u>ccggg</u>AAAAGATATGAAGTAGATGGAACGACATATAAA GTTCAAATATGGGATACTGgtttgttttaaataacttta aaagataaaaaacaaaaatgattttcaagacagtaatt cggttttataagtagaaaaagaagatgagatgagataa gattttatctttcgaaaataaatataa [c] ct [cc] ct [c] ttatttagCTGGACAAGAAAAATTTAGAACTATTA TTACTAGTTATTATCGTGGTGTTCAA<u>tgActcgag</u></p>
Rab Box ΔGU/Box C con sitio de HindIII (<i>ttcga</i>)	<p><u>ccggg</u>AAAAGATATGAAGTAGATGGAACGACATATAAA GTTCAAATATGGGATACTGgt [aa]g [aa] ttaaatac tttaaaagataaaaaacaaaaatgattttcaagacagt aattcggttttataagtagaaaaagaagatgagatgag ataagattttatctttcgaaaataaatataa [c] ct [cc] ct [c] ttatttagCTGGACAAGAAAAATTTAGAAC TATTATTACTAGTTATTATCGTGGTGTTCAA<u>tga</u> <u>ctcgag</u></p>

Se prepararon las mezclas de transfección en esterilidad, agregando 10 μ L del plásmido correspondiente (pEhExHA vacío, pEhExHA-WT, pEhExHA- Δ GU, pEhExHA-C y pEhExHA- Δ GU/C), 100 μ L de medio M199 y 20 μ L de SuperFect.

Transcurrido el periodo de incubación, se retiró el medio TYI-S-33 y se agregó lentamente la mezcla final de transfección sobre la monocapa de amebas en cada pozo. La placa se incubó nuevamente a 37°C durante 5 horas. Posteriormente, la placa se colocó en un baño de hielo-agua durante 20 minutos y se colectó el medio con las amebas, las cuales se sembraron en tubos de 7 mL con medio TYI-S-33 y se incubaron nuevamente en las mismas condiciones durante 24 horas.

Después de esta incubación, se inició el proceso de selección con el antibiótico G418 (Life technologies No. de catálogo 11811098) a una concentración de 1 μ g/mL. La concentración del antibiótico se fue aumentando atendiendo a la confluencia de los cultivos de amebas (80%), hasta llegar a una concentración de 3 y 5 μ g/mL de antibiótico. Una vez alcanzada esta concentración de antibiótico, se pasaron las amebas a cajas de 25 cm² y se recolectaron las muestras después de 24 horas de incubación para la posterior extracción de RNA.

6.4 Cultivos de trofozoitos de *Entamoeba histolytica*

Los trofozoitos de la cepa HM1:IMSS de *E. histolytica* sin transfectar y transfectados con las diferentes mutaciones, se crecieron de forma axénica en medio TYI-S-33 (Diamond, L. S., Harlow, D. R., & Cunnick, 1978) suplementado con Suero Bovino Adulto Inactivado al 10% y Penicilina-Estreptomicina al 1X en tubos de cultivo 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 cm² donde se dejaron toda la noche en las mismas condiciones de incubación para la posterior extracción de RNA.

6.5 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 conservaron los trofozoitos. Se adicionó TRIzol (Invitrogen, No. de catálogo 15596-018) y cloroformo: isoamílico (24:1), las muestras se agitaron vigorosamente y se centrifugaron a 12000 g durante 10 minutos a 4°C. La fase más clara se transfirió a un tubo eppendorf nuevo y se agregaron 200 µL 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 re-suspendió 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 75 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.

Posteriormente, las muestras de RNA se trataron con DNasa I (Sigma-Aldrich No. de catálogo 4716728001), durante 20 minutos a 25°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 esta se le agregó un volumen de cloroformo-isoamílico (25:1), 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 re-suspendidas en agua miliQ y se cuantificaron espectrofotométricamente.

6.6 RT-PCR

Para las reacciones de retro-transcripción se empleó el kit RevertAid First Strand cDNA Synthesis #1622 siguiendo el protocolo recomendado por el fabricante. La concentración de RNA empleada como molde fue 1.5 µg en todos los casos. Se adicionó 1 µL de hexámeros y se incubó a 65°C durante 5 minutos. Posteriormente, se agregaron 4 µL de Buffer de reacción, 1 µL de dNTPs, 1 µL de RivolLock y 1 µL de M-MLV, completando un volumen final de 20 µL. Se agitó suavemente y se incubó la mezcla a 42°C por una hora. La inactivación se realizó a 75°C por 10 minutos.

El cDNA obtenido se usó inmediatamente o se almacenó a -20°C hasta su uso. En todos los casos se empleó un volumen de 1 µL de cDNA (≈ 10ng de cDNA) para las PCR, completándose con los cebadores específicos, 1 µL de dNTPs, 1X de buffer de PCR 10X, 1.5 µL de MgCl₂ (50mM) y 0.15 µL de Taq DNA Polimerasa a completar con H₂OmQ estéril un volumen de 25 µL (ver Tabla 2).

La detección de los flircRX13, de los transcritos 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 2. Como control de carga, se realizó bajo las mismas condiciones la RT-PCR para la Polimerasa II. 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 (2%) o acrilamida al 12% (flircRX13), los cuales se tiñeron 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 Pol II y de los controles en cada experimento.

6.7 PCR en tiempo real (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 para los ensayos de qPCR, siguiendo en todos los casos las indicaciones del fabricante. El gen de la polimerasa II se tomó como normalizador (*house-keeping*) y contra su expresión se calculó la de los otros genes analizados (transcritos de pEhExHA y las construcciones y los flicRX13s.) (ver Apéndice 1).

En cada caso se agregó un volumen de 1 μ L de cDNA (\approx 10ng de cDNA) para las qPCR de la polimerasa II, transcritos de pEhExHA y las construcciones, mientras que para la amplificación de los flicRX13s se agregaron como templado concentraciones en el orden de picogramos; se adicionaron los cebadores específicos y se completó cada mezcla con 5 μ L del Master Mix de los kit (ver Tabla 3).

El análisis de las expresiones relativas de los genes se realizó mediante el método de Livak o $\Delta\Delta$ Ct (Livak & Schmittgen, 2001).

6.8 Análisis estadístico

Se utilizó el software GraphPad Prism versión 6.0 para el análisis de los datos. Todos los experimentos se realizaron por triplicado en cada caso y empleando cDNA de tres diferentes transfecciones para cada clona. Una vez obtenidos los valores numéricos de cada experimento, se realizaron pruebas de ANOVA de una o dos direcciones para determinar diferencias significativas entre los distintos tratamientos.

Tabla 2. Oligonucleótidos y condiciones de amplificación. Se indica el nombre y número de acceso en Amoeba DB de los genes analizados y los oligonucleótidos sentido (indicados como s) y antisentido (as) y sus secuencias. Se muestran las condiciones de amplificación para cada gen utilizando el par de oligonucleótidos respectivo y las concentraciones de MgCl₂ empleadas. Se indica el tamaño de los respectivos amplicones. US: *unspliced*, S: *spliced*, flicRX13: Círculo de *RabX13*.

Gen	Oligonucleotidos	Secuencia 5'-3'	MgCl ₂ (mM)	Condiciones de amplificación	Tamaño del amplicón (pb)
EhRabX13 (EHI_065790)	Rab2F	cgttggtggagactcttc agttgg	3	35 ciclos (94°C/45'',	US= 510pb S= 374 pb
	Rab2R	gaccatttcagttgaa acagttc	3	62°C/45'', 72 °C/1')	
	Rab2BSs	ccgaattactgtcttgaa aatc	3	30 ciclos (94°C/45'',	flicRX13 = 79 pb
	NBRX13	cttttaaagtatttaaaa caaacctaaataaaagt tagtttatattatt	3	55°C/45'', 72 °C/30'')	
EhRNA Pol II (EHI7A_158420)	RNAPII-s	gatccaacatatcctaa aacaaca	1	35 ciclos (94°C/45'',	210 pb
	RNAPII-as	tcaattattttctgaccc gtcttc	1	60°C/4', 72 °C/1')	
Transcrito de pEhExHA	CS+116	tagaatagactctcatc aac	3	35 ciclos (94°C/45'',	469 pb
	Rab-REx2	ttgaacaccacgataat aac	3	55°C/45'', 72 °C/45'')	

Tabla 3. Oligonucleótidos y condiciones de amplificación para las qPCR Se indica el nombre y número de acceso en Amoeba DB de los genes analizados y los oligonucleótidos sentido (indicados como s) y antisentido (as), sus secuencias y el tamaño del amplicón. En todos los casos se adicionaron 5 µL de Master Mix a cada reacción. flicRX13: Círculo de *RabX13*.

Gen	Oligonucleotidos	Secuencia 5'-3'	Condiciones de amplificación	Tamaño del amplicón (pb)
EhRabX13 (EHI_065790)	Rab2BSs	ccgaattactgtcttgaa aatc	39 ciclos (94°C/15'', 55°C/15'', 72 °C/15'')	flicRX13 = 79 pb
	NBRX13	cttttaaagattttaaaa caaacctaaataaaaagt tagtttatatttatt		
EhRNA Pol II (EHI7A_158420)	RNAPII-s	gatccaacatatcctaa aacaaca	39 ciclos (94°C/15'', 60°C/30'', 72 °C/30'')	210 pb
	RNAPII-as	tcaattattttctgaccc gtcttc		
Transcrito de pEhExHA	CS+116	tagaatagactctcatc aac	39 ciclos (94°C/15'', 55°C/30'', 72 °C/30'')	469 pb
	Rab-REx2	ttgaacaccacgataat aac		

7. Resultados

7.1 Construcción de plásmidos y transfección

La restricción de los vectores pUC57 que contienen los minigenes sintetizados con las enzimas XmaI y XhoI, por dos horas a 37°C, permitió la liberación de los fragmentos intrónicos, por sobre expresar (Fig. 6A). En todos los casos, se observó un fragmento con un peso correspondiente a 260 pb, tamaño esperado para cada construcción.

Cada fragmento se cortó del gel y se purificó mediante precipitación etanólica agregando acetato de sodio a la mezcla y glicógeno como acarreador de DNA. Una vez monitoreado el resultado de la precipitación, se ligó cada minigen con el vector de expresión en amebas pEhExHA y se transformaron bacterias *E. coli* cepa TOP 10. El análisis por restricción con XmaI y XhoI de candidatos por retraso en su migración en gel de agarosa al 1%, con respecto a un control de pEhExHA vacío, permitió comprobar la presencia de cada minigen en este vector de expresión (Fig. 6B). En todas las restricciones se liberó un fragmento con un peso aproximado de 260 pb, correspondiente al tamaño esperado.

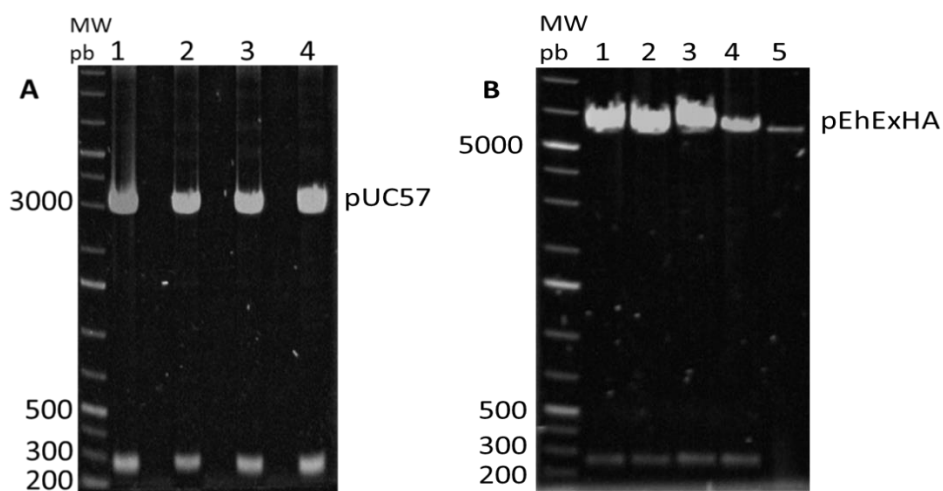


Figura 6. Restricción con XmaI y XhoI de las construcciones correspondientes. (A) Vector pUC57 + los minigenes correspondientes. (B) Vector pEhExHA con los minigenes después del proceso de clonación. En ambos casos se muestra el marcador de peso molecular y los carriles 1, 2, 3 y 4 indicando WT, Δ GU, C y Δ GU/C respectivamente; en B el carril 5 muestra el vector pEhExHA vacío restringido con las mismas enzimas

Una vez comprobada la presencia de cada clona en el vector pEhExHA, se procedió a transfectar a la cepa HM1-IMSS de *E.histolytica* con estas construcciones, para analizar el efecto sobre la biogénesis del flicRX13, así como la transcripción del gen parental plasmídico y del gen nativo (RabX13) en estas células.

7.2 Efecto de los minigenes en la biogénesis de flicRX13 en *E. histolytica*

Debido a que los minigenes presentan la misma secuencia que el gen nativo de amebas (RabX13), con la excepción de algunas bases mutadas, se colocó un sitio de corte de la enzima HindIII en cada una de las construcciones, para diferenciar los productos nativos de aquellos transcritos a partir del vector de expresión. Aunque el sitio de corte de esta enzima no se modificó en ningún momento del proceso de clonación y transfección y la actividad de la enzima empleada no estaba afectada, no se detectaron diferencias estadísticamente significativas entre los tratamientos en presencia y ausencia de la enzima HindIII (Fig. 7 A y B). No obstante, si se detectaron diferencias estadísticamente significativas en los niveles de expresión de flicRX13 entre las clonas (Fig. 8 A y B). En las amebas transfectadas con las construcciones HA-WT, HA-C y HA- Δ GU/C la detección de flicRX13 fue mayor que en las transfectadas con las construcciones pEhExHA vacío y HA- Δ GU.

El análisis estadístico de los resultados obtenidos por PCR a punto final, no detectó diferencias estadísticamente significativas en la formación de flicRX13 entre las clonas HA-WT, HA-C y HA- Δ GU/C, pero estas si mostraron diferencias significativas con respecto a las construcciones pEhExHA vacío y HA- Δ GU, las cuales no presentaron entre ellas diferencias estadísticamente significativas ($\alpha = 0.05$) (Fig. 8B). Esto sugiere que algunas clonas favorecieron la formación de los flicRX13, mientras que otras afectaron la biogénesis.

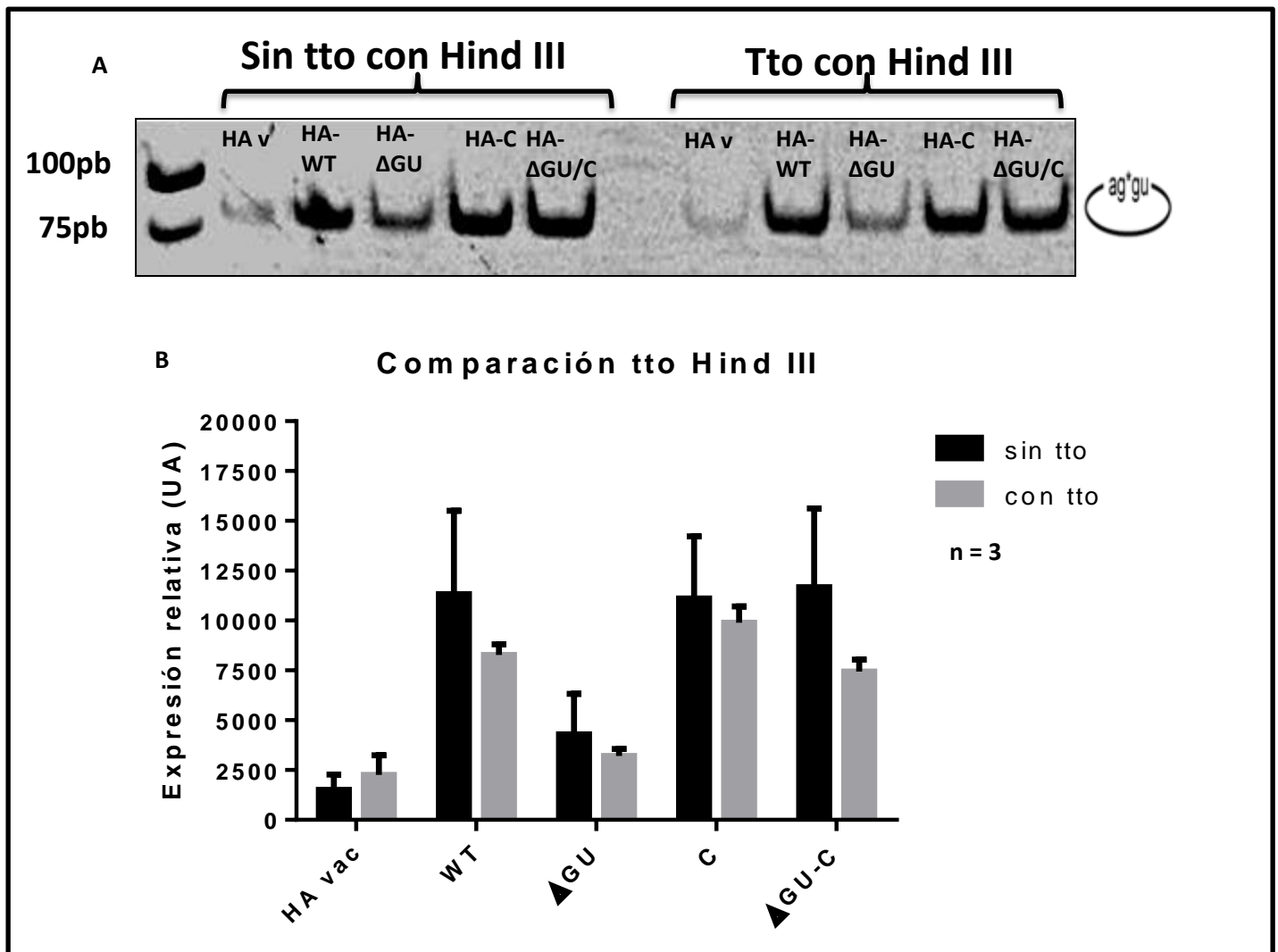


Figura 7. Análisis de la expresión de flicRX13 en cada una de las clonas antes y después del tratamiento (tto) con la enzima HindIII. (A) Gel nativo de poliacrilamida al 12% con los flicRX13 (79 pb) detectados en cada mutante antes y después de tratar con la enzima HindIII. (B) Comparación estadística (ANOVA de dos factores) de la expresión relativa en unidades arbitrarias de cada construcción antes y después del tratamiento con HindIII (n = 3).

Mediante los análisis de qPCR se pudo determinar que la formación de los flicRX13 se favoreció en aquellas clonas en las que no se removió el elemento GU (HA+WT y HA+C). Los niveles de flicRX13 detectados en el RNA extraído de amebas transfectadas con estas dos construcciones mostraron diferencias estadísticamente significativas con los flicRX13 detectados en pEhExHA vacío, HA-ΔGU y HA-ΔGU/C (Fig. 8C). En estos casos los niveles de moléculas

circulares de naturaleza intrónica detectados fueron superiores al doble de lo detectado en amebas transfectadas con secuencias sin el elemento GU. La remoción del elemento GU provocó una disminución significativa en la biogénesis del flicRX13 (ver Figura 8C), incluso en presencia del tracto de citosinas asociado al sitio de ramificación (HA- Δ GU/C). Estos resultados indican la importancia del elemento GU en la biogénesis del flicRX13.

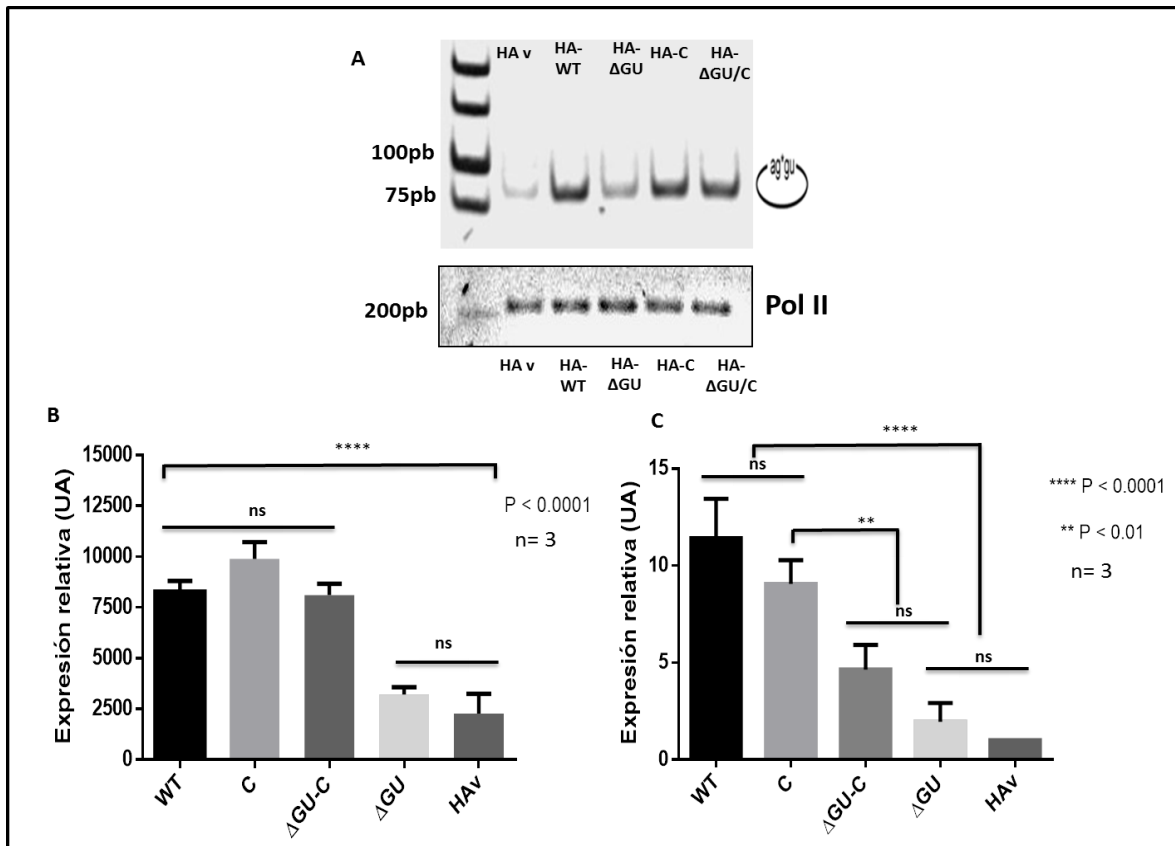


Figura 8. Expresión relativa de flicRX13 en cada una de las clonas. (A) Gel nativo de poliacrilamida al 12% con los flicRX13 (79 pb) detectados en cada mutante. Debajo se observa la amplificación del gen de la polimerasa II (control de carga) en un gel a agarosa al 2%. (B) Comparación estadística (ANOVA de un factor) de la expresión relativa en unidades arbitrarias de los flicRX13 en cada clona detectados mediante PCR en punto final (n = 3). (C) Comparación estadística (ANOVA de un factor) de la expresión relativa en unidades arbitrarias de los flicRX13 en cada clona detectados mediante qPCR (n = 3).

7.3 Efectos de los minigenes sobre la transcripción del gen parental en *E. histolytica*

Las amebas transfectadas con el vector vacío no arrojaron señal debido a la ausencia de la región complementaria al cebador Rab-REx 2 (Fig. 9A). Mediante análisis estadísticos

de los resultados obtenidos por PCR a punto final se detectaron diferencias estadísticamente significativas ($\alpha = 0.05$) en la transcripción a partir del promotor del gen de cada clona (HA-WT, HA-C, HA- Δ GU/C y HA- Δ GU) (Fig. 9B). Este resultado sugiere que el elemento GU asociado al 5' ss es esencial para la transcripción de estos genes, ya que al removerse (Fig. 9A) los niveles de expresión del gen disminuyen drásticamente. En este sentido, la adición de citosinas en regiones asociadas al sitio de ramificación provocó una recuperación en los niveles de transcripción de estos genes aún en ausencia del elemento GU (Fig. 9A).

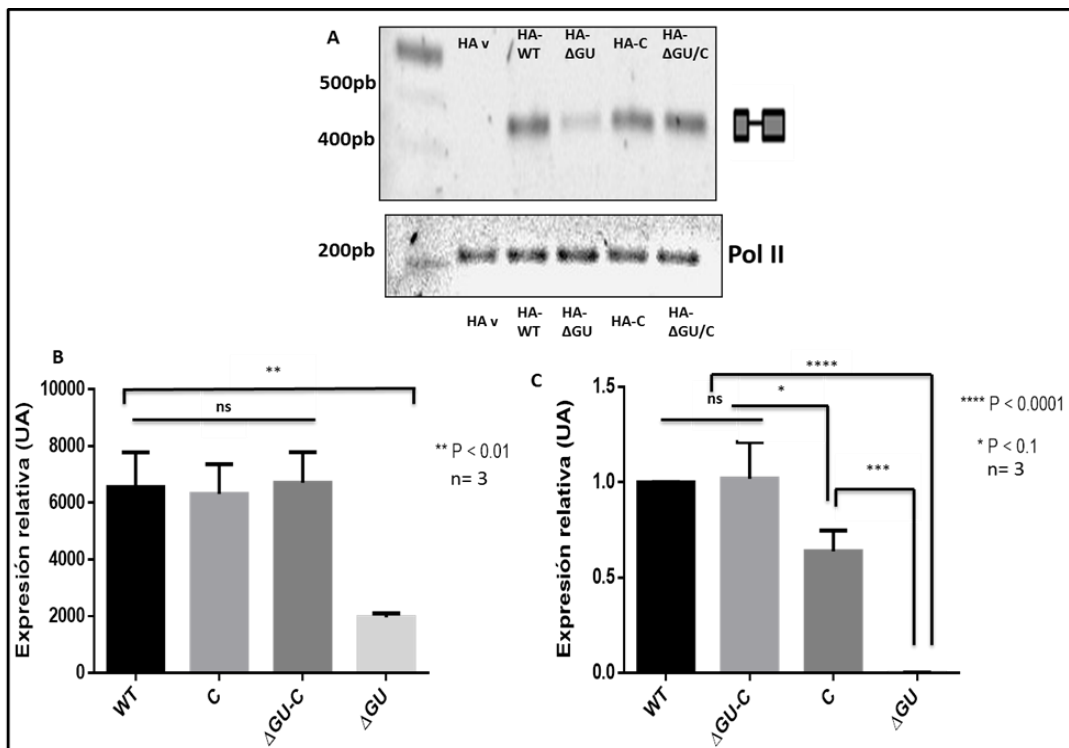


Figura 9. Expresión del gen de cada clona. (A) Gel de agarosa al 2% con el transcrito parental plasmídico correspondiente a cada clona (469 pb). Debajo se observa la amplificación del gen de la polimerasa II (control de carga) en un gel a agarosa al 2%. (B) Comparación estadística (ANOVA de un factor) de la expresión relativa en unidades arbitrarias de los transcritos parentales plasmídicos en cada clona detectados mediante PCR en punto final ($n = 3$). (C) Comparación estadística (ANOVA de un factor) de la expresión relativa en unidades arbitrarias de los transcritos parentales plasmídicos en cada clona detectados mediante qPCR ($n = 3$).

La amplificación de este fragmento mediante qPCR permitió detectar diferencias estadísticamente significativas entre la expresión de las construcciones HA-WT y HA- Δ GU/C con las construcciones HA-C y HA- Δ GU (ver Fig. 9C). Mediante ambos

procedimientos se constató que los niveles de transcripción de la construcción HA- Δ GU fueron mucho menores que el de las otras construcciones con señales intrónicas conservadas.

7.4 Efectos de los minigenes sobre la transcripción del gen RabX13 nativo en *E. histolytica*

Para comprobar si la presencia de los minigenes tenía efecto sobre la transcripción del gen RabX13 nativo en las amebas transfectadas, se realizaron experimentos de PCR a punto final con oligos específicos al gen RabX13; mostrando estos análisis resultados no esperados.

El análisis de densitometría del transcrito primario y del mensajero de RabX13 obtenido mediante amplificación por PCR a punto final permitió determinar diferencias significativas estadísticamente ($\alpha = 0.05$) entre las construcciones (Fig. 10A y B). Las amebas transfectadas con las clonas HA- Δ GU y HA-WT presentaron mayores niveles de transcripción y procesamiento del transcrito de RabX13 que las clonas pEhExHA vacío, HA-C y HA- Δ GU/C (Fig. 10B). Aunque los mayores niveles de transcripción se observaron en las amebas transfectadas con las construcciones HA- Δ GU y HA-WT, entre estas se detectaron diferencias estadísticamente significativas, observándose que los niveles del mRNA de RabX13 en HA- Δ GU son el doble de lo detectado en HA-WT (Fig. 10B). En ambos casos los niveles detectados son mayores que el control.

Por su parte, la adición de un tracto de citosinas provocó una disminución en el procesamiento del gen RabX13 nativo. Esta reducción en los niveles detectados se presentó tanto en presencia como en ausencia del elemento GU, lo que sugiere un mecanismo de regulación de la expresión génica diferente, donde el tracto de citosinas asociado al sitio de ramificación esté participando. De manera interesante el aumento del procesamiento de RabX13 nativo en la transfectante HA- Δ GU podría estar involucrado con la disminución de la expresión de este gen, así como de los niveles de flicRX13 asociados a esta construcción (ver Fig. 8 y 9).

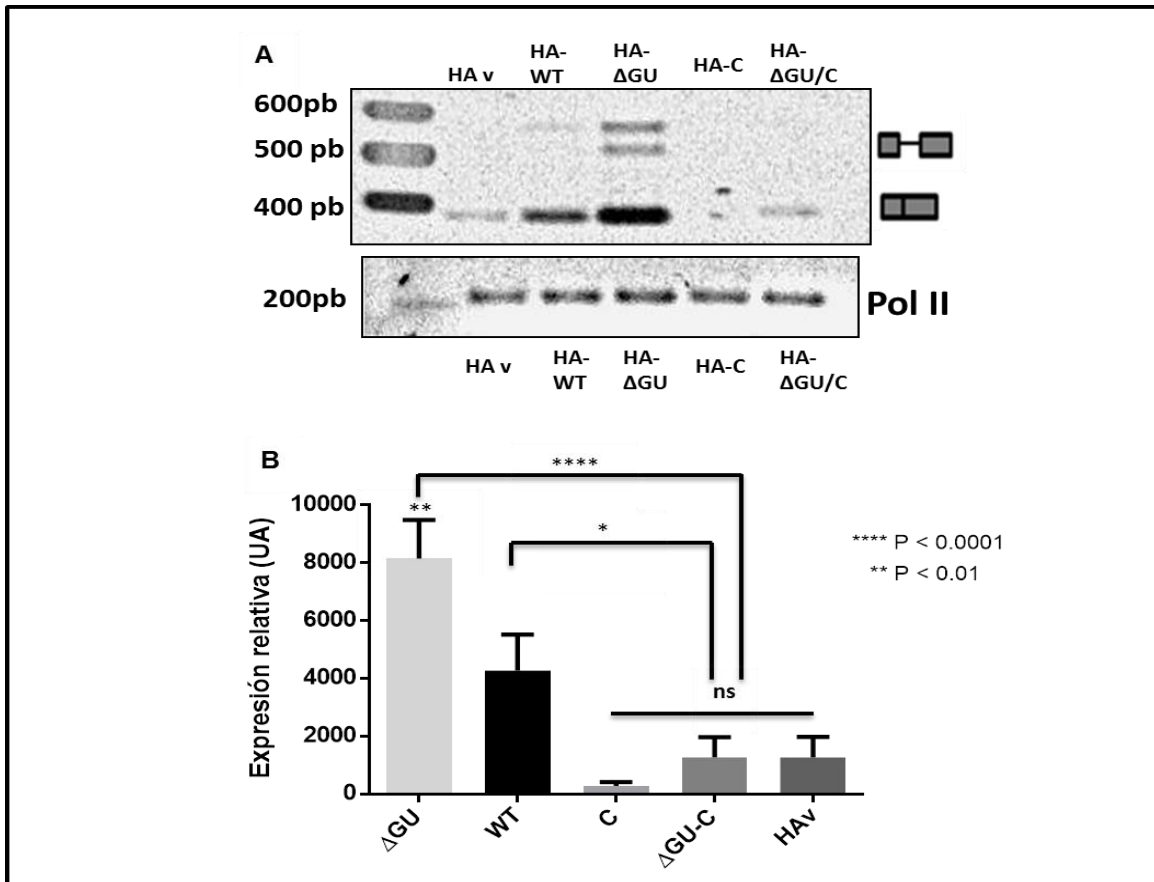


Figura 10. Efecto de cada clona sobre la expresión del gen RabX13 nativo de *E. histolytica*. (A) Gel de agarosa al 2% con el transcrito y el mensajero de RabX13 correspondiente a cada clona (transcrito primario 520 pb; mRNA de RabX13 376 pb). Debajo se observa la amplificación del gen de la polimerasa 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 a punto final en cada una de las clonas (n=3).

8. Discusión

Los estudios del proceso de circularización del intrón del gen ANKRD52 (Zhang et al., 2013) permitieron conocer la participación de elementos presentes en las regiones 5' ss y del sitio de ramificación en el correcto ensamblaje de los ciRNAs humanos. Por otra parte, las características moleculares de los flicRNAs (Mendoza-Figueroa, 2016) sugieren un mecanismo de ligación de los mismos diferente al detectado en los ciRNAs.

Aunque es muy probable que se formen por mecanismos diferentes, los flicRNAs y los ciRNAs comparten características, siendo algunas de estas: vida media mayor que su correspondiente lineal, moléculas circulares de RNA no codificantes, son resultado del proceso de *splicing* y ambos tipos escapan a la acción de la enzima desramificadora.

No obstante a la posibilidad de formarse por mecanismos diferentes, debido a que ambos tipos de moléculas son resultado del proceso de *splicing*, el reconocimiento en la secuencia intrónica de señales o elementos específicos por parte de factores involucrados en su formación parece ser un requerimiento esencial en la formación de ambas moléculas de RNA circulares.

Con el empleo de herramientas y estrategias moleculares como la construcción y posterior transfección en amebas de minigenes con secuencias intrónicas circularizables, pero con mutaciones en los elementos asociados a su ligación en el 5' ss y el sitio de ramificación, según (Zhang et al., 2013), se analizó la participación de los mismos en el proceso de formación del flicRX13. La adición o remoción de estos elementos a la secuencia intrónica permitió conocer su efecto sobre la biogénesis del flicRX13, así como de la transcripción a partir del promotor de los genes parentales de cada construcción. Por otro lado, se evaluó la influencia de estas moléculas mutadas sobre la regulación de la transcripción en *trans* y el procesamiento del gen RabX13 nativo.

A diferencia de los ciRNAs que requieren de dos elementos esenciales para su formación y escape de la acción de la enzima DBR1 (GU asociado al 5' ss y varias C asociadas al sitio de ramificación), los flicRNAs descritos hasta el momento en

amebas solo presentan el elemento GU del 5' ss. Con una baja presencia de citosinas en el genoma de *E. histolytica* (Loftus, 2005), los trectos de esta base no son frecuentes en este parásito. De manera que este, sea otro hecho que indique que es un mecanismo de biogénesis diferente en los flicRNAs.

El análisis de detección del flicRX13 mediante qPCR se realizó normalizando la expresión relativa a los niveles de flicRX13 detectados en las amebas transfectadas con pEhExHA vacío. La introducción en amebas de minigenes de regiones intrónicas con trectos de citosinas (HA-C y HA- Δ GU/C) provocó un aumento aparente en la biogénesis del flicRX13 (ver figura 8A y B). En contraste, la ausencia de uno de los elementos analizados (la clona HA-WT carece del tracto de citosinas asociado al sitio de ramificación, mientras que en la HA- Δ GU/C falta el elemento rico en GU asociado al 5' ss) no permitió observar diferencias en la formación de flicRNAs entre estos minigenes. Sin embargo, el análisis cuantitativo con PCR en tiempo real debido a su mayor sensibilidad permitió detectar diferencias estadísticamente significativas entre estas muestras (ver figura 8C) lo que sugiere que en *E. histolytica* la biogénesis del flicRX13 depende de un solo elemento (GU), a diferencia de lo señalado por (Zhang et al., 2013) para los ciRNAs.

En ambos casos en los que se removió el elemento GU (HA- Δ GU y HA- Δ GU/C) hubo una disminución en los niveles de flicRX13. Los resultados de qPCR para estas clonas no muestran diferencias significativas entre los niveles de flicRX13 detectados. Este dato, reafirma la importancia del elemento GU en la formación de los flicRX13, ya que incluso en presencia del tracto de citosinas (HA- Δ GU/C) los niveles de flicRX13 no son elevados si no hay presencia del factor GU. Por otro lado, esto indica que el tracto de citosinas asociado al sitio de ramificación por sí solo no provoca un aumento en la biogénesis de flicRX13.

Debido a lo mencionado anteriormente y a que no se encontraron diferencias estadísticamente significativas entre HA-C (clona con los dos elementos) y HA-WT (ver Figura 8C) se puede afirmar que a diferencia de lo que ocurre en células eucariotas donde es necesaria la presencia de los dos elementos para asegurar la

formación de los ciRNAs, en amebas el proceso de circularización intrónica no depende de estos dos elementos. Los datos obtenidos en este estudio sugieren que el elemento GU es el necesario para la formación de este tipo de moléculas. Esto indica un mecanismo de ligación diferente al de los ciRNAs y adaptado a un genoma con una baja prevalencia de citosinas, donde la presencia del elemento GU es imprescindible para la formación de moléculas circulares de naturaleza intrónica de RabX13 en *E. histolytica*. Además, debido a que no todos los flicRNAs observados contienen elementos GUUUGUUU completos, es necesario identificar el tamaño mínimo de dicho elemento.

Debido a que los flicRNAs son moléculas resultantes del proceso de *splicing*, el análisis de la transcripción del gen parental plasmídico resulta esencial para comprender su proceso de biogénesis. Además, una de las funciones atribuidas a las moléculas circulares de RNA no codificante es la regulación de la transcripción de su gen parental. Empleando oligos específicos se determinaron los niveles de transcripción del plásmido pEhExHA que contenía cada una de las clonas. Como se observa en las figura 9 A y B, en todos los casos ocurrió la transcripción de las construcciones, desde el promotor de la Cisteín sintasa hasta las secuencias mutadas. Los niveles de transcripción mostrados por las diferentes clonas varió de una construcción a otra. Cuando se removió el elemento GU (HA+ Δ GU) la transcripción disminuyó drásticamente. Mientras que cuando se eliminó este elemento pero se adicionó el tracto de citosinas (HA+ Δ GU/C), la transcripción mostró un nivel similar al de la clona nativa (HA+WT) por ambos métodos de detección (PCR a punto final y qPCR) (Fig.9). Esto indica que la presencia de al menos uno de los elementos reconocibles por la maquinaria de *splicing* es necesaria para la correcta transcripción de los genes en *E. histolytica*. Este aspecto es conocido desde la década del 80 del pasado siglo, cuando diferentes investigadores notaron que al clonar en diferentes organismos era necesario colocar secuencias intrónicas en los vectores de expresión para que el proceso de transcripción se desarrollara efectivamente (Callis et al., 1987; Le Hir et al., 2003). La remoción del único elemento de reconocimiento canónico del intrón en *E. histolytica* por la maquinaria de *splicing* (HA+ Δ GU), posiblemente provocó que

este minigen no se transcribiera en el mismo grado que el resto de las construcciones analizadas. De esta manera no se observa casi producto resultante de este plásmido.

La cuantificación relativa de la expresión de estos genes mediante ensayos de qPCR, se normalizó a la expresión de la construcción HA+WT. Existen diferencias significativas entre las expresiones de HA+WT, HA+ Δ GU/C y HA+C con la expresión de HA+ Δ GU, así como entre HA+WT y HA+ Δ GU/C con HA+C. Este resultado, en conjunto con los datos de expresión relativa de los flicRX13, es de gran importancia ya que demuestran que incluso aunque la expresión del gen resultante de la construcción HA+ Δ GU/C fue mayor que la del HA+C, en el proceso de formación de flicRX13 ocurrió lo contrario.

Después de analizados los resultados de la transcripción del vector de expresión con los minigenes correspondientes integrados, se procedió a analizar la relación entre este proceso y la biogénesis del flicRX13. Los menores niveles de flicRX13 detectados en HA+ Δ GU, pueden estar asociados a la baja transcripción de esta clona. En este caso en particular, es probable que la baja detección de círculos de RNA se deba a que este vector no se transcribió de manera eficiente, por lo que no se pudieron formar moléculas circulares. A pesar de que la clona HA+ Δ GU/C se transcribió de manera correcta y con niveles mayores que la clona HA+C, la formación de flicRX13 en la primera fue menor que en HA+C. Esto indica que la formación de los flicRNAs es un proceso que depende de la presencia del elemento GU. En el caso de HA+C se detectó el doble de expresión relativa de flicRX13 que en la clona HA+ Δ GU/C, independientemente de que esta última se transcribió más que HA+C. Este resultado refuerza la teoría de que la presencia del elemento GU es determinante para la formación del flicRX13.

Cuando se comparan los niveles de flicRX13 detectados en HA+WT con los de HA+C, se observa que son similares, a pesar de las diferencias en la transcripción de estos genes (Fig. 8). La diferencia en la secuencia de estas clonas es la adición del tracto de citosinas a la segunda. Este resultado indica que el elemento C no participa directamente en la formación del flicRX13, por lo que no es requerido

para la formación de esta molécula. Si bien hay una relación entre la transcripción del vector más los minigenes con la formación de los flicRNAs, realmente la biogénesis de estas moléculas no depende de los niveles de transcripción sino de la secuencia intrónica involucrada. Cuando se altera esta secuencia (remoción del elemento GU), la formación del flicRX13 desciende aunque la transcripción sea elevada.

La adición del tracto de citosinas mientras se mantuvo el elemento GU asociado al 5' ss (HA+ C) no provocó una transcripción mayor del gen con esta mutación al compararse con las clonas HA+WT y HA+ Δ GU/C. Esto indica que el tracto de pirimidinas no es reconocido *per se* como un señalizador intrónico en *E. histolytica*. Este resultado concuerda con lo señalado por (Loftus, 2005) cuando se determinó el genoma de *E. histolytica* y se comprobó la baja presencia de citosinas en el mismo. Si el mecanismo de transcripción en *E. histolytica* dependiera de trectos de citosinas este no pudiera desarrollarse eficientemente debido al bajo porcentaje de esta base en el genoma del parásito.

El efecto del flicRX13 sobre la transcripción del gen parental nativo en *E. histolytica* es una novedad de este trabajo. Hasta el momento, solo se han descrito mecanismos de regulación en *cis* para los ciRNAs. En esta investigación se describe por primera vez el efecto del flicRX13 en *trans* actuando sobre el gen nativo de la ameba. Si se tiene en cuenta que la transfección de las clonas provocó alteraciones en la biogénesis del flicRX13 y que estos se vieron reflejados en la expresión del RabX13 nativo se puede afirmar que, posiblemente los flicRX13 están participando en la regulación de la transcripción y procesamiento del gen nativo. El análisis de la expresión del gen RabX13 por PCR a punto final mostró que la presencia del elemento C en las clonas provocó que cesara la actividad transcripcional y procesiva del transcrito de este gen. En los casos en los que se mutaron las bases y se colocó el tracto de citosinas, se dejó de detectar el transcrito primario y el mensajero de este gen. También se detectaron diferencias estadísticamente significativas entre la expresión relativa del mRNA de RabX13 en las clonas HA+ Δ GU y HA+WT.

La cepa HA+WT fue de las que presentó mayores niveles de flicRX13 y la detección del transcrito primario y del mensajero de RabX13 nativo fue menor que lo detectado para HA+ Δ GU. Por su parte, la clona HA+ Δ GU que fue la que mostró menores niveles de círculos presentó mejor actividad procesiva del gen RabX13 nativo. Este resultado, sugiere que los flicRX13 pueden estar regulando negativamente la expresión del gen RabX13.

Al insertar estos minigenes en las amebas, se provocó en algunos casos una sobreexpresión de flicRX13 (HA+WT y HA+C). Debido a que se ha propuesto que los ciRNAs regulan positivamente la transcripción de sus genes parentales, esto explicaría la disminución del producto de RabX13 en la clona HA+WT comparado con HA+ Δ GU, ya que los flicRX13 formados estarían favoreciendo la transcripción del gen parental plasmídico y no del nativo. En este mismo orden, al formarse menos círculos de RNA en la clona HA+ Δ GU, no se afectaría la síntesis del producto nativo de RabX13 debido a que no habría competencia entre productos.

Atendiendo a este análisis, se podrían esperar unos niveles de procesamiento de RabX13 en la clona HA+C similares a los de HA+WT debido a que no se encontraron diferencias significativas entre los niveles de flicRX13. No obstante, no se detectó ni transcrito primario ni el mensajero del gen RabX13 en las amebas transfectadas con HA+C. Por su parte, se observó el mismo resultado cuando se comparó entre HA+ Δ GU y HA+ Δ GU/C (ambas clonas con el elemento GU removido). En ambos casos, el factor común diferente entre las clonas es la presencia del tracto de citosinas asociado al sitio de ramificación. Este resultado indica un mecanismo de regulación que involucra trectos de citosinas y proteínas o factores que reconozcan estos sitios. Debido al complejo mecanismo de regulación de la expresión génica que posiblemente esté involucrado en este caso, estudios posteriores al respecto son requeridos.

La presencia del tracto de citosinas en estas dos clonas puede estar provocando la activación de una vía de regulación alternativa dirigida por proteínas hnRNPs del tipo PCBP (*poly C binding proteins*) (Chaudhury, Chander, & Howe, 2010; S. P. Han, Tang, & Smith, 2010). Se conoce que estas proteínas al unirse a un

transcrito favorecen su expresión así como su transcripción por sobre otros genes (Kim et al., 2005; Meng et al., 2007). En el 2007, Meng y colaboradores describieron como la proteína PCBP1 podía actuar simultáneamente como factor de splicing y activador de la transcripción de un mismo gen, estando ambos procesos regulados por 3 dominios KH presentes en esta proteína que le permiten unirse a otras proteínas y al RNA. Los dominios KH se han encontrado en proteínas asociadas con la regulación de la transcripción y la traducción. La superficie de unión de estos dominios está compuesta por un surco en el que caben solamente 4 bases no apareadas. Interacciones hidrofóbicas, de Van der Waals y en menor medida de tipo electrostáticas contribuyen a la afinidad por la unión de ácidos nucleicos (Leffers, Dejgaard, & Celis, 1995; Valverde, Edwards, & Regan, 2008).

La proteína PCBP1 presenta tres de estos dominios (Leffers et al., 1995) y se ha visto en cultivos celulares que puede estar participando en eventos de splicing vía la proteína CAPER α y además estar regulando de manera positiva la transcripción de genes (Meng et al., 2007). Aunque en amebas no están descritas de manera específica proteínas del tipo PCBP, un alineamiento de la secuencia de esta proteína en humanos en AmoebaDB permitió detectar al menos 2 proteínas en *E. histolytica* con un porcentaje de identidad del 26 al 41% (ver Apéndice 2). Estas proteínas están descritas en la base de datos (AmoebaDB) como proteínas putativas de unión al RNA. Los porcentaje de similitud entre estas proteínas y la humana inducen a pensar que sean estas, proteínas amebianas del tipo PCBP.

La búsqueda en AmoebaDB de proteínas con secuencias similares a la del factor de *splicing* CAPER α resultó en la detección de varias proteínas en *E. histolytica* con un porcentaje de identidad del 22 al 47 % y regiones altamente conservadas (ver Apéndice 2).

Por otra parte, la comparación de secuencias entre las bases esenciales para el reconocimiento de la proteína PCBP1 (Kim et al., 2005) y las secuencias generadas con las mutaciones en las bases asociadas al sitio de ramificación

(HA+C y HA+ Δ GU/C) por adiciones de citosinas mostraron un 100% de identidad en la región central reconocida por PCBP1 (ver Apéndice 2).

Estos análisis sugieren que un mecanismo similar al estudiado por Meng y colaboradores en el 2007, donde mediante el reconocimiento de la proteína CAPER α y PCBP1 a secuencias ricas en C, se favorece la transcripción y *splicing* de estos genes, pudiera estarse desarrollando en las amebas transfectadas con las clonas HA+C y HA+ Δ GU/C. En este sentido, se requieren estudios posteriores.

De estar sucediendo un mecanismo de regulación génica similar a lo observado por Meng et al., 2007, la disminución del procesamiento de RabX13 en las amebas transfectadas con genes con trectos de citosinas pudiera deberse al procesamiento de estas moléculas mediado por la proteína PCBP 1 y por tanto que estuviera favorecida esta ruta y no así la vía de regulación nativa. Esto explicaría la disminución en los productos de RabX13 nativo detectados en estas clonas. Si se probara esta hipótesis, se explicaría que los flicRX13 participan en la regulación de la transcripción de su gen parental y que al parecer en los casos de que las mutaciones realizadas en las secuencias intrónicas consistan en trectos de citosinas pueden provocar la activación de otras vías de regulación de la expresión génica.

Sin menoscabo de la posible regulación negativa mediada por trectos ricos en C introducidos en forma de episoma, más importante, revelador e informativo resultó el incremento en la expresión de RabX13 endógeno en *trans* provocado por la introducción de HA+ Δ GU. Una posible interpretación funcional en *E. histolytica* es que los flicRNAs silvestres pudieran inhibir la transcripción de sus genes parentales mediante un mecanismo en *cis*. Una regulación de tipo negativa sobre la expresión del gen parental explicaría los resultados obtenidos en esta investigación. De ser así, el mecanismo de regulación de los flicRX13 sería opuesto a lo observado por (Zhang et al., 2013) para los ciRNAs, ya que el silenciamiento de estas moléculas se refleja en una disminución de la expresión de su gen parental. En nuestro caso los resultados obtenidos indican lo contrario.

9. Conclusiones

En esta investigación se analizó la participación de los elementos 5´ss y bases asociadas al sitio de ramificación en el proceso de formación del flicRX13 con herramientas y estrategias moleculares. La adición o remoción de estos elementos a la secuencia intrónica, permitió conocer el efecto de estos sobre la biogénesis del flicRX13, así como de las variaciones a nivel de transcrito de los genes parentales de cada construcción. También se evaluó la influencia de estas moléculas mutadas sobre la regulación en *trans* de de las variaciones a nivel de transcrito y procesamiento del gen RabX13 nativo. Los resultados obtenidos permitieron arribar a las siguientes conclusiones:

- A diferencias de los observado en ciRNAs humanos, en *E. histolytica*, solo se requiere la secuencia GU, para la formación del flicRX13.
- La formación de los flicRX13 no depende de los niveles de transcripción de los genes sino de la presencia del elemento GU en el 5´ss.
- El tracto de citosinas asociado al BP no aumenta *per se* los niveles de transcripción de los genes ni la biogénesis del flicRX13.
- La adición del tracto de citosinas provocó, al parecer mediante una vía alternativa de regulación de la expresión génica una disminución en los niveles de procesividad del gen parental nativo (RabX13).
- Los flicRX13 regulan negativamente la transcripción del gen parental nativo (RabX13) mediante un mecanismo que involucra la presencia del elemento GU.

10. Perspectivas

Debido a que son necesarios análisis adicionales para comprender mejor los procesos de biogénesis y de regulación de los flicRNAs, se proponen algunos estudios posteriores:

- Ensayos de inmunoprecipitación de la RNA polimerasa II, seguidos de secuenciación de RNA para identificar si el flicRX13 está interactuando con el complejo de transcripción.
- Ensayos de silenciamiento de los flicRNAs (siRNAs o DNazimas) y determinación del efecto sobre el splicing del gen parental.
- Transfección de amebas con las moléculas circulares obtenidas *in vitro* a partir de las construcciones HA+WT y HA+C para comprobar el efecto sobre la transcripción y el splicing del gen RabX13 nativo.
- Silenciar la proteína amebiana similar a la PCBP1 humana para comprobar si hay afectaciones en la transcripción del gen RabX13 nativo en amebas transfectadas con HA+C o HA+ Δ GU/C.

11. Bibliografía

- 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. <https://doi.org/10.1016/j.molcel.2014.08.019>
- Begum, S., Quach, J., & Chadee, K. (2015). Immune evasion mechanisms of *Entamoeba histolytica*: Progression to disease. *Frontiers in Microbiology*, 6(DEC), 1–8. <https://doi.org/10.3389/fmicb.2015.01394>
- Bentley, D. L. (2014). Coupling mRNA processing with transcription in time and space. *Nature Reviews Genetics*, 15(3), 163–175. <https://doi.org/10.1038/nrg3662>
- Bhattacharya, S. (2015). Circular noncoding ribosomal transcripts and recombination of retrotransposons: Two charming secrets revealed by a less-explored human parasite. *Proceedings of the Indian National Science Academy*, 81(2), 431–446. <https://doi.org/10.16943/ptinsa/2015/v81i2/48097>
- Braunschweig, U., Gueroussov, S., Plocik, A. M., Graveley, B. R., & Blencowe, B. J. (2013). Dynamic integration of splicing within gene regulatory pathways. *Cell*, 152(6), 1252–1269. <https://doi.org/10.1016/j.cell.2013.02.034>
- Brinster, R. L., Allen, J. M., Behringer, R. R., Gelinas, R. E., & Palmiter, R. D. (1988). Intróns increase transcriptional efficiency in transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America*, 85(3), 836–40. <https://doi.org/10.1073/pnas.85.3.836>
- Bush, S. J., Chen, L., Tovar-Corona, J. M., & Urrutia, A. O. (2017). Alternative splicing and the evolution of phenotypic novelty. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 372(1713), 1–7. <https://doi.org/10.1098/rstb.2015.0474>
- Callis, J., Fromm, M., & Walbot, V. (1987). Intróns increase gene expression in cultured maize cells. *Genes & Development*, 1(10), 1183–1200. <https://doi.org/10.1101/gad.1.10.1183>
- Chaudhury, A., Chander, P., & Howe, P. H. (2010). Heterogeneous nuclear

ribonucleoproteins (hnRNPs) in cellular processes: Focus on hnRNP E1's multifunctional regulatory roles. *RNA (New York, N.Y.)*, 16(8), 1449–1462. <https://doi.org/10.1261/rna.2254110>

- Chen, H., & Cheng, S. (2012). Functional roles of protein splicing factors. *Bioscience Reports*, 32(4), 345–359. <https://doi.org/10.1042/BSR20120007>
- Chen, L.-L. (2016). The biogenesis and emerging roles of circular RNAs. *Nature Reviews Molecular Cell Biology*, 17(4), 205–211. <https://doi.org/10.1038/nrm.2015.32>
- Cho, S., Moon, H., Loh, T. J., Jang, H. N., Liu, Y., Zhou, J., ... Shen, H. (2015). Splicing inhibition of U2AF⁶⁵ leads to alternative exon skipping. *Proceedings of the National Academy of Sciences*, 112(32), 9926–9931. <https://doi.org/10.1073/pnas.1500639112>
- Conn, S. J., Pillman, K. A., Toubia, J., Conn, V. M., Salmanidis, M., Phillips, C. A., ... Goodall, G. J. (2015). The RNA binding protein quaking regulates formation of circRNAs. *Cell*, 160(6), 1125–1134. <https://doi.org/10.1016/j.cell.2015.02.014>
- Cowling, V. H., & Cole, M. D. (2010). Myc Regulation of mRNA Cap Methylation. *Genes & Cancer*, 1(6), 576–579. <https://doi.org/10.1177/1947601910378025>
- Crisci, A., Raleff, F., Bagdiul, I., Raabe, M., Urlaub, H., Rain, J. C., & Kr??mer, A. (2015). Mammalian splicing factor SF1 interacts with SURP domains of U2 snRNP-associated proteins. *Nucleic Acids Research*, 43(21), 10456–10473. <https://doi.org/10.1093/nar/gkv952>
- Davis, C. A., Brown, M. P. S., & Singh, U. (2007). Functional characterization of spliceosomal introns and identification of U2, U4, and U5 snRNAs in the deep-branching eukaryote *Entamoeba histolytica*. *Eukaryotic Cell*, 6(6), 940–948. <https://doi.org/10.1128/EC.00059-07>
- 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.

- Ebbesen, K. K., Kjems, J., & Hansen, T. B. (2016). Circular RNAs: Identification, biogenesis and function. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1859(1), 163–168. <https://doi.org/10.1016/j.bbagr.2015.07.007>
- Eddy, S. R. (2001). Non-coding RNA genes and the modern RNA world. *Nature Reviews Genetics*, 2(12), 919–929. <https://doi.org/10.1038/35103511>
- Esteller, M. (2011). Non-coding RNAs in human disease. *Nature Reviews Genetics*, 12(12), 861–74. <https://doi.org/10.1038/nrg3074>
- Fong, N., & Bentley, D. L. (2001). Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: Different functions for different segments of the CTD. *Genes and Development*, 15(14), 1783–1795. <https://doi.org/10.1101/gad.889101>
- Görnemann, J., Kotovic, K. M., Hujer, K., & Neugebauer, K. M. (2005). Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. *Molecular Cell*, 19(1), 53–63. <https://doi.org/10.1016/j.molcel.2005.05.007>
- Han, B., Park, H. K., Ching, T., Panneerselvam, J., Wang, H., Shen, Y., ... Fei, P. (2017). Human DBR1 modulates the recycling of snRNPs to affect alternative RNA splicing and contributes to the suppression of cancer development. *Oncogene*, (April), 1–10. <https://doi.org/10.1038/onc.2017.150>
- Han, S. P., Tang, Y. H., & Smith, R. (2010). Functional diversity of the hnRNPs: past, present and perspectives. *Biochemical Journal*, 430(3), 379–392. <https://doi.org/10.1042/BJ20100396>
- 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. <https://doi.org/10.1038/nature11993>
- Hon, C. C., Weber, C., Sismeiro, O., Proux, C., Koutero, M., Deloger, M., ... Guillen, N. (2013). Quantification of stochastic noise of splicing and polyadenylation in *Entamoeba histolytica*. *Nucleic Acids Research*, 41(3),

1936–1952. <https://doi.org/10.1093/nar/gks1271>

- Jeck, W. R., Sorrentino, J. A., Wang, K., Slevin, M. K., Burd, C. E., Liu, J., ... Sharpless, N. E. (2013). Circular RNAs are abundant, conserved, and associated with ALU repeats. *Rna*, 19(2), 141–157. <https://doi.org/10.1261/rna.035667.112>
- 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. <https://doi.org/10.1093/nar/gkt046>
- Johnsson, P., Lipovich, L., Grandér, D., & Morris, K. V. (2014). Evolutionary conservation of long non-coding RNAs; sequence, structure, function. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1840(3), 1063–1071. <https://doi.org/10.1016/j.bbagen.2013.10.035>
- Kataoka, N., Dobashi, I., Hagiwara, M., & Ohno, M. (2013). hDbr1 is a nucleocytoplasmic shuttling protein with a protein phosphatase-like motif essential for debranching activity. *Scientific Reports*, 3. <https://doi.org/10.1038/srep01090>
- 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. <https://doi.org/10.1021/acscchembio.6b00971>
- Khalid, M. F., Damha, M. J., Shuman, S., & Schwer, B. (2005). Structure-function analysis of yeast RNA debranching enzyme (Dbr1), a manganese-dependent phosphodiesterase. *Nucleic Acids Research*, 33(19), 6349–6360. <https://doi.org/10.1093/nar/gki934>
- Kim, S., Pandey, K. K., Choi, H. S., Kim, S., Law, P., Wei, L., & Loh, H. H. (2005). Poly (C) Binding Protein Family Is a Transcription Factor in \square - Opioid Receptor Gene Expression. *Molecular Pharmacology*, 68(3), 729–736. <https://doi.org/10.1124/mol.105.012245.laboratories>
- Kramer, M. C., Liang, D., Tatomer, D. C., Gold, B., March, Z. M., Cherry, S.,

- & Wilusz, J. E. (2015). Combinatorial control of *Drosophila* circular RNA expression by intronic repeats, hnRNPs, and SR proteins. *Genes and Development*, 29(20), 2168–2182. <https://doi.org/10.1101/gad.270421.115>
- Lasda, E., & Parker, R. (2014). Circular RNAs: diversity of form and function. *Rna*, 20(12), 1829–1842. <https://doi.org/10.1261/rna.047126.114>
 - Le Hir, H., Nott, A., & Moore, M. J. (2003). How introns influence and enhance eukaryotic gene expression. *Trends in Biochemical Sciences*, 28(4), 215–220. [https://doi.org/10.1016/S0968-0004\(03\)00052-5](https://doi.org/10.1016/S0968-0004(03)00052-5)
 - Lee, Y., & Rio, D. C. (2015). Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annual Review of Biochemistry*, 84(1), 291–323. <https://doi.org/10.1146/annurev-biochem-060614-034316>
 - Leffers, H., Dejgaard, K., & Celis, J. E. (1995). Characterisation of Two Major Cellular Poly(rC) Binding Human Proteins, Each Containing Three K homologous (KH) Domains. *European Journal of Biochemistry*, 230(2), 447–453. <https://doi.org/10.1111/j.1432-1033.1995.0447h.x>
 - Lewin, B. (2008). *Genes 9*. Jones & Bartlett Learning.
 - Li, Z., Huang, C., Bao, C., Chen, L., Lin, M., Wang, X., ... Shan, G. (2015). Exon-intron circular RNAs regulate transcription in the nucleus. *Nature Structural & Molecular Biology*, 22(3), 256–264. <https://doi.org/10.1038/nsmb.2959>
 - 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. <https://doi.org/10.1006/meth.2001.1262>
 - Loftus, B. (2005). The genome of the protist parasite *Entamoeba histolytica*. *Nature*, 433, 865–868. Retrieved from <http://dx.doi.org/10.1038/nature03291>
 - 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). <https://doi.org/10.1371/journal.pntd.0000716>
 - Mar-Aguilar, F., Trevino, V., Salinas-Hernández, J. E., Taméz-Guerrero, M.

- M., Barrón-González, M. P., Morales-Rubio, E., ... Reséndez-Pérez, D. (2013). Identification and Characterization of microRNAs from *Entamoeba histolytica* HM1-IMSS. *PLoS ONE*, *8*(7). <https://doi.org/10.1371/journal.pone.0068202>
- Marie, C., & Petri Jr, W. a. (2014). Regulation of Virulence of *Entamoeba histolytica*. *Annual Review of Microbiology*, (June), 493–520. <https://doi.org/10.1146/annurev-micro-091313-103550>
 - Marshall, A. N., Montealegre, M. C., Jiménez-López, C., Lorenz, M. C., & van Hoof, A. (2013). Alternative Splicing and Subfunctionalization Generates Functional Diversity in Fungal Proteomes. *PLoS Genetics*, *9*(3). <https://doi.org/10.1371/journal.pgen.1003376>
 - Masaki, S., Yoshimoto, R., Kaida, D., Hata, A., Satoh, T., Ohno, M., & Kataoka, N. (2015). Identification of the specific interactors of the human lariat RNA debranching enzyme 1 protein. *International Journal of Molecular Sciences*, *16*(2), 3705–3721. <https://doi.org/10.3390/ijms16023705>
 - Matera, A. G., & Wang, Z. (2014). A day in the life of the spliceosome. *Nature Reviews Molecular Cell Biology*, *15*(2), 108–121. <https://doi.org/10.1038/nrm3742>
 - Mattick, J. S., & Makunin, I. V. (2006). Non-coding RNA. *Human Molecular Genetics*, *15 Spec No(1)*, 17–29. <https://doi.org/10.1093/hmg/ddl046>
 - Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., ... Rajewsky, N. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature*, *495*(7441), 333–338. <https://doi.org/10.1038/nature11928>
 - Mendoza-Figueroa, S. (2016). *Tesis de doctorado*.
 - Meng, Q., Rayala, S. K., Gururaj, A. E., Talukder, A. H., O'Malley, B. W., & Kumar, R. (2007). Signaling-dependent and coordinated regulation of transcription, splicing, and translation resides in a single coregulator, PCBP1. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(14), 5866–5871.

<https://doi.org/10.1073/pnas.0701065104>

- Miranda, R., Salgado, L. M., Sánchez-López, R., Alagón, A., & Lizardi, P. M. (1996). Identification and analysis of the u6 small nuclear RNA gene from *Entamoeba histolytica*. *Gene*, 180(1–2), 37–42. [https://doi.org/https://doi.org/10.1016/S0378-1119\(96\)00397-6](https://doi.org/https://doi.org/10.1016/S0378-1119(96)00397-6)
- Misra, A., & Green, M. R. (2016). From polyadenylation to splicing: Dual role for mRNA 3' end formation factors. *RNA Biology*, 13(3), 259–264. <https://doi.org/10.1080/15476286.2015.1112490>
- Montemayor, E. J., Katolik, A., Clark, N. E., Taylor, A. B., Schuermann, J. P., Combs, D. J., ... Hart, P. J. (2014). Structural basis of lariat RNA recognition by the intrón debranching enzyme Dbr1. *Nucleic Acids Research*, 42(16), 10845–10855. <https://doi.org/10.1093/nar/gku725>
- Morris, K. V., & Mattick, J. S. (2014). The rise of regulatory RNA. *Nature Reviews Genetics*, 15(6), 423–437. <https://doi.org/10.1038/nrg3722>
- Palacio-Morales. (2016). *Tesis de Maestría "Identificación de algunos componentes de la biogénesis de flicRNAs en Entamoeba histolytica."*
- Papasaikas, P., & Valcárcel, J. (2016). The Spliceosome: The Ultimate RNA Chaperone and Sculptor. *Trends in Biochemical Sciences*, 41(1), 33–45. <https://doi.org/10.1016/j.tibs.2015.11.003>
- Patel, A. A., & Steitz, J. A. (2003). Splicing double: insights from the second spliceosome. *Nature Reviews Molecular Cell Biology*, 4(12), 960–970. <https://doi.org/10.1038/nrm1259>
- 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. <https://doi.org/10.1016/j.canlet.2015.06.003>
- Ruskin, B., & Green, M. R. (1985). An RNA processing activity that debranches RNA lariats. *Science*, 229(4709), 135–140. <https://doi.org/10.1126/science.2990042>
- Sal, G. Del, Manfioletti, G., & Schneider, C. (1988). *Acids Research*, 16(20), 34012.
- Salzman, J. (2014). Circular RNA is expressed across the eukaryotic tree of

- life. *PloS One*, 9(6), e90859. <https://doi.org/10.1371/journal.pone.0090859>
- Schroeder, S. C., Schwer, B., Shuman, S., & Bentley, D. (2000). Dynamic association of capping enzymes with transcribing RNA polymerase II service Dynamic association of capping enzymes with transcribing RNA polymerase II, (303), 2435–2440. <https://doi.org/10.1101/gad.836300>
 - Sharp, P. A., & Burge, C. B. (1997). Classification of intróns: U2-type or U12-type. *Cell*, 91(7), 875–879. [https://doi.org/10.1016/S0092-8674\(00\)80479-1](https://doi.org/10.1016/S0092-8674(00)80479-1)
 - Shi, X., Sun, M., Liu, H., Yao, Y., & Song, Y. (2013). Long non-coding RNAs: A new frontier in the study of human diseases. *Cancer Letters*, 339(2), 159–166. <https://doi.org/10.1016/j.canlet.2013.06.013>
 - 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. <https://doi.org/10.1128/JCM.00530-06>
 - 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. <https://doi.org/10.1016/j.celrep.2014.12.002>
 - Stauffer, W., Abd-Alla, M., & Ravdin, J. I. (2006). Prevalence and incidence of *Entamoeba histolytica* infection in South Africa and Egypt. *Archives of Medical Research*, 37(2), 266–269. <https://doi.org/10.1016/j.arcmed.2005.10.006>
 - Valdés, 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. <https://doi.org/10.1016/j.jprot.2014.07.027>
 - Valverde, R., Edwards, L., & Regan, L. (2008). Structure and function of KH domains. *FEBS Journal*, 275(11), 2712–2726.

<https://doi.org/10.1111/j.1742-4658.2008.06411.x>

- Vicens, Q., & Westhof, E. (2014). Biogenesis of circular RNAs. *Cell*, 159(1), 13–14. <https://doi.org/10.1016/j.cell.2014.09.005>
- 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. <https://doi.org/10.1074/jbc.M309106200>
- Wang, Y., & Wang, Z. (2015). Efficient backsplicing produces translatable circular mRNAs. *Rna*, 21(2), 172–179. <https://doi.org/10.1261/rna.048272.114>
- Wang, Z., & Burge, C. B. (2008). Splicing regulation : From a parts list of regulatory elements to an integrated splicing code Splicing regulation : From a parts list of regulatory elements to an integrated splicing code. *Rna*, 14(617), 802–813. <https://doi.org/10.1261/rna.876308.802>
- Will, C. L., & Lührmann, R. (2011). Spliceosome structure and function. TL - 3. *Cold Spring Harbor Perspectives in Biology*, 3 VN-re(7), 1–23. <https://doi.org/10.1101/cshperspect.a003707>
- 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. <https://doi.org/http://dx.doi.org/10.1078/1434-4610-00053>
- Wu, T., & Fu, X.-D. (2015). Genomic functions of U2AF in constitutive and regulated splicing. *RNA Biology*, 12(5), 479–485. <https://doi.org/10.1080/15476286.2015.1020272>
- Ximénez, C., Morán, P., Rojas, L., Valadez, A., & Gómez, A. (2009). Reassessment of the epidemiology of amebiasis: State of the art. *Infection, Genetics and Evolution*, 9(6), 1023–1032. <https://doi.org/10.1016/j.meegid.2009.06.008>
- Yoshimoto, R., Kataoka, N., Okawa, K., & Ohno, M. (2009). Isolation and characterization of post-splicing lariat -intrón complexes. *Nucleic Acids Research*, 37(3), 891–902. <https://doi.org/10.1093/nar/gkn1002>
- Yu, H., Tian, C., Yu, Y., & Jiao, Y. (2016). Transcriptome Survey of the

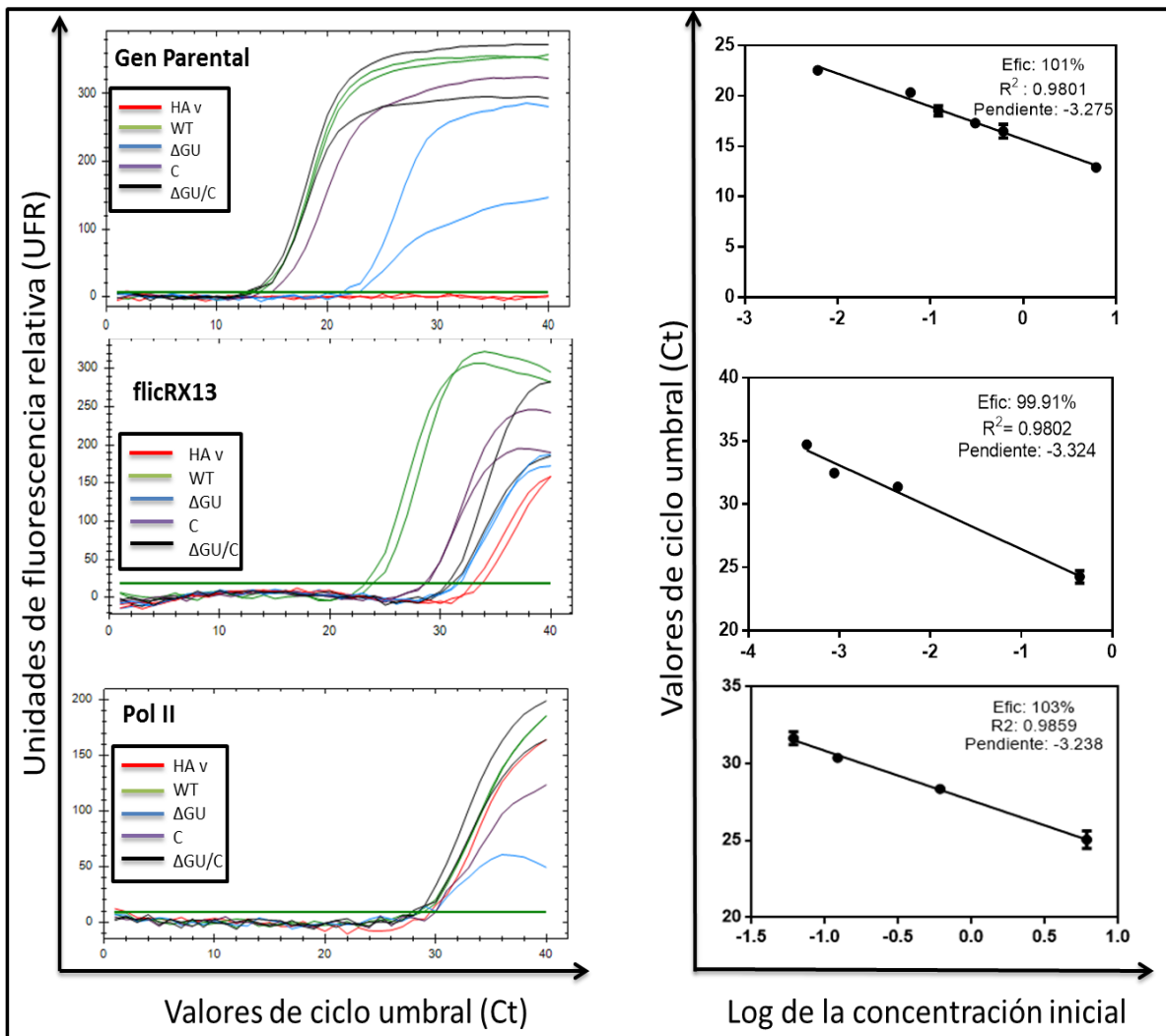
Contribution of Alternative Splicing to Proteome Diversity in *Arabidopsis thaliana*. *Molecular Plant*, 9(5), 749–752.
<https://doi.org/10.1016/j.molp.2015.12.018>

- Zhang, Y., Zhang, X. O., Chen, T., Xiang, J. F., Yin, Q. F., Xing, Y. H., ... Chen, L. L. (2013). Circular Intronic Long Noncoding RNAs. *Molecular Cell*, 51(6), 792–806. <https://doi.org/10.1016/j.molcel.2013.08.017>

Apéndices

Apéndice 1

Curvas representativas de la amplificación por qPCR de los genes analizados y curvas de concentraciones con los valores de eficiencia respectivos a cada proceso.



Apéndice 2

- **Alineamiento de la secuencia de aminoácidos de la proteína CAPER α humana con proteínas putativas de unión al RNA de *E. histolytica***

Query= MySeq1

Length=373

	Score	E
EHI_056660A-p1 transcript=EHI_056660A gene=EHI_056660 org...	87.4	4e-19
EHI_010060A-p1 transcript=EHI_010060A gene=EHI_010060 org...	79.3	3e-17
EHI_183100A-p1 transcript=EHI_183100A gene=EHI_183100 org...	78.2	2e-15
EHI_121780A-p1 transcript=EHI_121780A gene=EHI_121780 org...	64.7	4e-11
EHI_030430A-p1 transcript=EHI_030430A gene=EHI_030430 org...	61.6	4e-10
EHI_079880A-p1 transcript=EHI_079880A gene=EHI_079880 org...	61.6	4e-10
EHI_048210A-p1 transcript=EHI_048210A gene=EHI_048210 org...	60.8	2e-10

[EHI_056660A-p1](#) | transcript=EHI_056660A | gene=EHI_056660 |
organism=Entamoeba_histolytica_HM-1:IMSS
| gene_product=RNA-binding
protein TIA-1, putative | transcript_product=RNA-binding
protein TIA-1, putative | location=DS571423:12408-13328 (+)
| protein_length=306 | sequence_SO=supercontig | SO=protein_coding
Length=306

Score = 87.4 bits (215), Expect = 4e-19, Method: Compositional matrix adjust.
Identities = 54/177 (31%), Positives = 92/177 (52%), Gaps = 13/177 (7%)

```
Query 12  LEEFFSTVGKVRDVRMISDRNSRRSKGIAYVEFVDVSSVPLAI-GLTGQRVLGVPIIVQA 70
          L  FS VG V  +++ D++  + G  +VEFVD ++  A  + G+ V G  + V
Sbjct 59  LGRIFSIVGHVVSCKIMRDKSGVHA-GYGFVEFVDSTTARFAKDNMDGRVVYGRELKVNW 117

Query 71  SQAEKNRAAAMANNLQKGSAGPMRLYVGLHFNITEDMLRGIFEPFGRIESIQLMMDSET 130
          S              Q+ + G  +++VG L  + +D+L  F+ FGR+  +++ ++T
Sbjct 118 SYT-----AQQENQGNKIFVGGGLQPEVNDLLYKTFQKFGRVTDARVLKFTQT 166

Query 131 GRSKGYGFITFSDSECAKKALEQLNGFELAGRPMKVGHVTERTDASSASSFLDSDEL 187
          G+SKGYGF+TF  E A+ A++ +NG +L GR +KV  VT  + +  DE+
Sbjct 167  GKSKGYGFVTFIRKEDAETAMQMMNGEKLEGRNIKVNWVTSNIASKTEQPKRSYDEI 223
```

Score = 59.7 bits (143), Expect = 1e-09, Method: Compositional matrix adjust.
Identities = 46/162 (28%), Positives = 78/162 (48%), Gaps = 16/162 (10%)

```
Query 12 LEEFFSTVGKVRDVRMISDRNSRRSKGIAYVEFVDVSSVPLAIGL-TGQRVLGVPIIVQ- 69
          L + F  G+V D R++  + +SKG +V F+  A+ +  G+++ G I V
Sbjct 145 LYKTFQKFGRVTDARVLKFTQTGKSKGYGFVTFIRKEDAETAMQMMNGEKLEGRNIKVNW 204

Query 70 -----ASQAEK-NRAAAMANNLQKGSAGPMRLYVGSLSLHFNITDMLRGIFEPFGRIESIQ 123
          AS+ E+ R+  NN + S+  +Y+G++ N+ D L+ +  +G IE ++
Sbjct 205 VTSNIASKTEQPKRSYDEINN--ETSSQNCTVYIGNIPKNVESDDLKQLLAEYGSIEEVR 262

Query 124 LMMDSETGRSKGYGFITFSDSECAKKALEQLNGFELAGRPMK 165
          L D  KGY FI FS E A A+  NG + G ++
Sbjct 263 LNKD-----KGYAFIKFSKHESATSAILMCNGKIINGSTLR 298
```

```
> EHI\_010060A-p1 | transcript=EHI_010060A | gene=EHI_010060 |
organism=Entamoeba_histolytica_HM-1:IMSS
| gene_product=RNA recognition
motif domain containing protein | transcript_product=RNA
recognition motif domain containing protein | location=DS571200:51551-52012(-)
| protein_length=136 | sequence_SO=supercontig
| SO=protein_coding
Length=136
```

Score = 79.3 bits (194), Expect = 3e-17, Method: Compositional matrix adjust.
Identities = 34/73 (47%), Positives = 53/73 (73%), Gaps = 0/73 (0%)

```
Query 94 RLYVGSLSLHFNITDMLRGIFEPFGRIESIQLMMDSETGRSKGYGFITFSDSECAKKALEQ 153
          RLY+GSL +++T++ L+ FE FG +  +++ D ++ RSKG+GF+TF  E AKKA+E+
Sbjct 3  RLYIGSLAYSVTDESLKAAFEKFGTVDCKVVTDRDSQRSKGFGFVTFEKDEDAKKAIEE 62

Query 154 LNGFELAGRPMKV 166
          +N EL GR +KV
Sbjct 63 MNEQELEGRRRIKV 75
```

Score = 35.8 bits (81), Expect = 0.034, Method: Compositional matrix adjust.
Identities = 20/58 (34%), Positives = 30/58 (52%), Gaps = 1/58 (2%)

```

Query 16 FSTVGKVRDVRMISDRNSRRSKGIAYVEFVDVSSVPLAI-GLTGQRVLGVPIIVQASQ 72
          F G V D +++DR+S+RSKG +V F AI + Q + G I V S+
Sbjct 22 FEKFGTVDCKVVTDRDSQRSKGFVTFEKDEDAKKAIEEMNEQELEGRRRIKVDVSR 79

```

```

> EHI\_183100A-p1 | transcript=EHI_183100A | gene=EHI_183100 |
organism=Entamoeba_histolytica_HM-1:IMSS
| gene_product=RNA recognition
motif domain containing protein | transcript_product=RNA
recognition motif domain containing protein | location=DS571156:101729-102933(+)
| protein_length=381 | sequence_SO=supercontig
| SO=protein_coding
Length=381

```

```

Score = 78.2 bits (191), Expect = 2e-15, Method: Compositional matrix adjust.
Identities = 78/349 (22%), Positives = 144/349 (41%), Gaps = 65/349 (19%)

```

```

Query 11 DLEEFFSTVGKVRDVRMISDRNSRRSKGIAYVEFVDVSSVPLAIGLTGQRVL-GVPIIVQ 69
          +++EFF+ G+V V + D ++RS+G YVEF + AI ++G PI V
Sbjct 92 EIKEFFNQAGEVVHVEQVIDTTTTKRSRGGFYVEFKTIEGAMKAIEMSGMFFKESPIYVS 151

```

```

Query 70 ASQAEKNRAAAMANNLQKGSAGPMRLYVGLHFNITDMLRGIFEPFGRIESIQLMMDSE 129
          S A++NR +N Q + R+ V +++ +++++ + +F G+I + + + E
Sbjct 152 ESNAQQNRNTITISNTQLQTN---RIKVKNMNKQLSKEDIEKVFGIAGKIIELSIQEEEE 208

```

```

Query 130 TGRSKGYGFITFSDSECAKKALEQLNGFELAGRPMKVGHVTERTDASSASSFLDSDELER 189
          + I + + AK+A+E +G +V + E+ S + D LE
Sbjct 209 SNNV----IIEYDTIKMAKRAIELYDGRSFGNMKWEVFSICEK--GSDVITNEDEQLLES 262

```

```

Query 190 TGIDLGTTGRLQLMARLAEGTGLQIPPAQQALQMSGSLAFGAVAEFSFVIDLQTRLSQQ 249
          ++ ++ R+ G G FG
Sbjct 263 KSKEM-----MLKRIQGGVG-----TLFGQ----- 282

```

```

Query 250 TEASALAAAASVQPLATQCFQLSNMFPQTEEEVGDTEIKDDVIEECNKHGGVIHIYVD 309
          E A VQ + TQ +E +G++ E+++D++EE ++ V + VD
Sbjct 283 -ENINHYKALFVQNVFTQG-----KEPIGFEKELRNDILEELKQYCQVKDVIDV 330

```

```

Query 310 KNSAQGNVYVKCPSIAAAIAAVNALHGRWFAGKMITAAYVP---LPTYH 355
          +G V+V C + A A + +H RWF ++ Y P +P H
Sbjct 331 LIHPKGVVFLCETEVDAKKAFSVMHLRWFNMHLRVEYYPESKVPIVH 379

```



```

> EHI\_121780A-p1 | transcript=EHI_121780A | gene=EHI_121780 |
organism=Entamoeba_histolytica_HM-1:IMSS
| gene_product=enhancer
binding protein-1 | transcript_product=enhancer binding
protein-1 | location=DS571276:14324-15244(-) | protein_length=306
| sequence_SO=supercontig | SO=protein_coding
Length=306

```

```

Score = 64.7 bits (156), Expect = 4e-11, Method: Compositional matrix adjust.
Identities = 49/182 (27%), Positives = 87/182 (48%), Gaps = 9/182 (5%)

```

```

Query 7 IRPRDLEEFFSTVGKVRDVRMISDR---NSRRSKGIAYVEFVDVSSVPLAIG-LTGQRVL 62
      I L+E FS G V + + R N ++ GIA++ AI + V
Sbjct 124 ITEEKLKEMFSKFGVVEITLIKTHRKKGNVTKNNGIAFITVKTAEQKKAIAEMNNFEVE 183

Query 63 GVPIIVQASQAE-KNRAAAMANNLQKGSAGPMRLYVGLHFNITEDMLRGIFEPFGRIES 121
      G I V A+ + +N+ N K S ++V +L F +T++ + +FE + +E+
Sbjct 184 GRKITVAAAYKKVENKQTTKKTNEPK-SLSETNVFVKNLPFTLTDEGFKKLFEKYDVVEA 242

Query 122 IQLMMDSE---TGRSKGYGFITFSDSECAKKALEQLNGFELAGRPMKVGHVTDASSA 178
      + ++ RSKGYGF+TF +E KKA+ +++ FE+ GR + V +R +
Sbjct 243 TIVKRHNKKNLVDRSKGYGFVTFKTAEQKKAIAEMDNFEVEGRKITVTSAYQRAEKKEE 302

Query 179 SS 180
      ++
Sbjct 303 TT 304

```

```

Score = 52.0 bits (123), Expect = 5e-07, Method: Compositional matrix adjust.
Identities = 39/151 (26%), Positives = 66/151 (44%), Gaps = 19/151 (13%)

```

```

Query 36 SKGIAYVEFVDVSSVPLAIGLTGQRVLGVPIIVQAS--QAEKNRAAAMANNLQKGSAGPM 93
      S+G A++EF A+ + + + I VQ + + E M KG
Sbjct 40 SRGFATFIEFATPQEAKEKALEMDRHTIGKMEIKVQKALPKEETTERKMNRRGFKGRNTR 99

Query 94 R-----LYVGLHFNITEDMLRGIFEPFGRIESIQLMMDSETG---RSKGY 136
      + L++ +L F ITE+ L+ +F FG +E + + G ++ G
Sbjct 100 KRHLVKRTNEVTDNMLFIKKNLPFAITEEKLKEMFSKFGVVEITLIKTHRKKGNVTKNNGI 159

Query 137 GFITFSDSECAKKALEQLNGFELAGRPMKVG 167
      FIT +E KKA+ ++N FE+ GR + V
Sbjct 160 AFITVKTAEQKKAIAEMNNFEVEGRKITVA 190

```

```
> EHI\_030430A-p1 | transcript=EHI_030430A | gene=EHI_030430 |
organism=Entamoeba_histolytica_HM-1:IMSS
| gene_product=RNA recognition
motif domain containing protein | transcript_product=RNA
recognition motif domain containing protein | location=DS571253:34612-35562(-)
| protein_length=316 | sequence_SO=supercontig
| SO=protein_coding
Length=316
```

```
Score = 61.6 bits (148), Expect = 4e-10, Method: Compositional matrix adjust.
Identities = 41/159 (26%), Positives = 78/159 (49%), Gaps = 16/159 (10%)
```

```
Query 7 IRPRDLEEFFSTVGKVRDVRMISDRNSRRSKGIAYVEFVDVSSVPLAIGLTGQRVLGVPI 66
      I+ +L+++ G V+D+ M+ D ++R KG +V+F+ L + +R P
Sbjct 26 IKKEELKQYCEQFGPVKDISMMVDTETQRPKGCQFVKFMTHEDAKLCVEEAPKRRNDP- 84

Query 67 IVQASQAEKNRAA--AMANNLQKGSAGPMRLYVGS LH-FNITEDMLRGIFEPFGRIESIQ 123
      ++N A ++ +++G M +Y+ +L+ N EDM+R F +G IE +
Sbjct 85 -----HKRNWVIEWAKSSQIKEGDLDKMTIYISNLNQNDNEDMIRQKFSVYGNIEKVT 138

Query 124 LMMDSETGRSKG-YGFITFSDSECAKKALEQLNGFELAG 161
      + + SKG + F+ + E A +A+ + NG E G
Sbjct 139 TIEN-----SKGMFAFVKYEQMESAI EAINKENGKEWCG 172
```

```
Lambda K H a alpha
0.318 0.132 0.374 0.792 4.96
```

Gapped

```
Lambda K H a alpha sigma
0.267 0.0410 0.140 1.90 42.6 43.6
```

```
> EHI\_079880A-p1 | transcript=EHI_079880A | gene=EHI_079880 |
organism=Entamoeba_histolytica_HM-1:IMSS
| gene_product=RNA recognition
motif domain containing protein | transcript_product=RNA
recognition motif domain containing protein | location=DS571226:25094-26044(-)
| protein_length=316 | sequence_SO=supercontig
```

| SO=protein_coding
Length=316

Score = 61.6 bits (148), Expect = 4e-10, Method: Compositional matrix adjust.
Identities = 41/159 (26%), Positives = 78/159 (49%), Gaps = 16/159 (10%)

Query 7 IRPRDLEEFFSTVGKVRDVRMISDRNSRRSKGIAYVEFVDVSSVPLAIGLTGQRVLGVPI 66
I+ +L+++ G V+D+ M+ D ++R KG +V+F+ L + +R P
Sbjct 26 IKKEELKQYCEQFGPVKDISMMVDTEETQRPKGCQFVKFMTHEDAKLCVEEAPKRRNDP- 84

Query 67 IVQASQAEKNRAA--AMANNLQKGSAGPMRLYVGS LH-FNITEDMLRGIFEPFGRIESIQ 123
++N A ++ +++G M +Y+ +L+ N EDM+R F +G IE +
Sbjct 85 -----HKRNWVIEWAKSSQIKEGDLDKMTIYISNLNQNDNEDMIRQKFSVYGNIEKVT 138

Query 124 LMMDSETGRSKG-YGFITFSDSECAKKALEQLNGFELAG 161
+ + SKG + F+ + E A +A+ + NG E G
Sbjct 139 TIEN-----SKGMFAFVKYEQMESAIEAINKENGKEWCG 172

> [EHI_048210A-p1](#) | transcript=EHI_048210A | gene=EHI_048210 |
organism=Entamoeba_histolytica_HM-1:IMSS
| gene_product=splicing
factor, arginine/serine-rich, putative | transcript_product=splicing
factor, arginine/serine-rich, putative | location=DS571256:21739-22408(-)
| protein_length=199 | sequence_SO=supercontig
| SO=protein_coding
Length=199

Score = 60.8 bits (146), Expect = 2e-10, Method: Compositional matrix adjust.
Identities = 27/69 (39%), Positives = 46/69 (67%), Gaps = 1/69 (1%)

Query 93 MRLYVGS LHFNITEDMLRGIFEPFGRIESIQ LMMDSETGRSKGYGFITFSDSECAKKALE 152
+ +YVG+L N +D L+ +F G I++ ++ + G SKG+GF+TF+D E A+KA+
Sbjct 4 LEVYVGNLSLNTDKDKLKEVFASVGEIKNSRVATHKD-GASKGFGFVTFNDEETAEKAVN 62

Query 153 QLNGFELAG 161
++NG E+ G
Sbjct 63 EMNGKEIDG 71

- **Alineamiento de la secuencia de aminoácidos de la proteína PCBP1 humana con dos proteínas putativas de unión al RNA de E. histolytica y porcentaje de similitud según el software BLAST del NIH.**

Consensus key (see documentation for details)

* - single, fully conserved residue

: - conservation of strong groups

. - conservation of weak groups

- no consensus

```

PCBP1      MDAGVTE SGLNMTLTI RLLMHGKEV GSIIGKKGESVKRIR EESGARINI SEGNCP E---R 57
EHI_128160 -----MSCDEYD-IVEVD-NRKYRVK KSIQPHLSVSNELSEKTSYD 39
              * . * . * : . : * : : . : : : * : . :

PCBP1      IITLTGPTNAIFKAFAMIIDKLEEDINSSMTNSTAASRPPVTLRLVVPATQCGSLIGKGG 117
EHI_128160 MKEYCDEPDCIYVPSLSSKEQVEEIK---KAEITPRTTDCTIKLLIPNKLHGQIIGKGG 96
              : . . : : * : . : : : * : : . : . : * : : * . * : * * * *

PCBP1      CKIKEIRESTGAQVQVAGDMLPNSTERAITIAGVPQ---SVTECVKQICLVMLETLSQS 173
EHI_128160 INISPIKKEECQCKIAFASD---DTPDLVRIEGVNEEDKVR EAMERLKNVCINVFCSLKS- 152
              : * . * : . : : . * . . * : * * : . . * : * : * : : : * .

PCBP1      PQGRVMTIPYQMPAS-----SPVICAGGQDR C 201
EHI_128160 --THTIQIPFKSSDLLKQIEIFQEQLVKIDGLNPNYLQKQALHLTLTTL SICTEQQVEV 210
              : : * : : : : : : : : : : : : : : : : : : : : * : * .

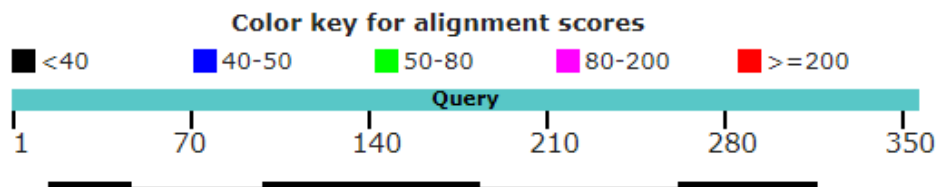
PCBP1      SDAAGYPHATHDLEGPPLDAYSIQGQHTISPLDLAKLNQVARQOS-HFAMMHGGTGFAGI 260
EHI_128160 -----ASKLLEDMAPRIYSILNNKPL-LFELCGISAMGNEEATRVYIKVKTNDEQI 261
              * : : * . * * * : : : : * . : . : : : : : : : : : : * .

PCBP1      DSSSPEVKGYWA-SLDASTQTTHELTIPNNLIGCIIGRQGANINEIRQMSGAIKIANPV 319
EHI_128160 NKIIEEIKQTFNKYLDKQNS--EVI LHM TLF-----NTNKLKNGKSFVIDASDIV 310
              : . * : * : * * * : * : : . * : * * : : : . . * . : : *

PCBP1      EGSSGRQVTITGSAASISLAQYLINARL---SSEKGMGCS----- 356
EHI_128160 RKYTGK-----SFGKYKAESLVFSSMINKIMGTKYALINSIPLP 349
              . : * : * : : * : * : : : : * * * .

```

Max score	Total score	Query cover	E value	Ident
32.7	91.2	48%	8e-06	26%



```

PCBP1      MDAGVTESGLNVTLTIRLLMHGKEVGSIIIGKKGESVKRIREEESGARINISEGMCPERIIT 60
EHI_193510 ----- 0

PCBP1      LTGPTNAIFKAFAMIIDKLEEDINSSMTNSTAASRPPVTLRLVVPATQCGSLIGKGGCKI 120
EHI_193510 -----MEEVKVEDYKERFTHPGGLPI 21
                    :. * . : ** *

PCBP1      KEIRESTGAQVQVAGDMLPNSTERAITIAGVPQSVTECVKQICLVMLETLSQSPQGRVMT 180
EHI_193510 HEKRVSRW--DRKIRDVLPNIPV-----HIPPSLDEKY-----WDT-----LAIR 59
:* * * : * :*** : * * : * : * : * : *

PCBP1      IPYQMPASSPVICAGGQDRCSDAAGYPHATHDLEGPPLDAYSIQGQHTISP-----LD 234
EHI_193510 IRYEELQY-----ALGTHR----LGLNTERDRTPSPK-KQYNNQOEITREMRREELK 108
* * : : * * * * . . ** . * : * . *

PCBP1      LAKLNQVARQQSHFAMMHGGTGFAGIDSSSPEVKGWASLDASTQTTHLTIIPNNLIGCI 294
EHI_193510 NERLYVVDRAIEIYPSFRIPAELAK-PSGKRKKIYFP-----KDRPDTNFIGLI 157
: * * * . : : : : * * . * * : : : * : * * *

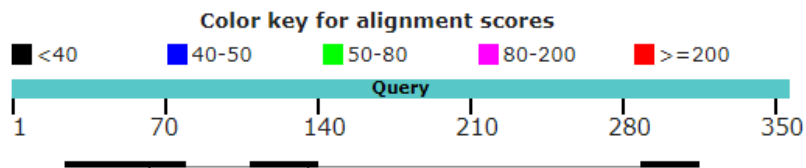
PCBP1      IGRQGANINEIRQMSGAQIKIANPVEGSSGRQVTITGSAASISLAQYLINARLSSEKGMG 354
EHI_193510 IGRGDNQKRLKDSGAKISIRGKDPKLGKLSGYGDKD-N-EDSHVFITADTQEAQLDLA 215
** : * * : : : : *** : * . . * : . . . . : : * * . . . .

PCBP1      CS----- 356
EHI_193510 CEEITKIISAPSEEIIVLKHNLRLALWNGTFREDRVYVEQYESGVKCGFCGSSSHAT 275
*

PCBP1      ----- 356
EHI_193510 CDCPLKKQKMSEHQLELERAFDEFMEIKIQLQLIN 309

```

Max score	Total score	Query cover	E value	Ident
32.7	110	32%	7e-06	41%



- Alineamiento de la secuencia de reconocimiento de la proteína PCBP1 humana y las secuencias intrónicas generadas con la adición del tracto de citosinas.

Range 1: 6 to 13 [Graphics](#) ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
16.4 bits(8)	0.004	8/8(100%)	0/8(0%)	Plus/Plus

```

Query 6 CCTCCCTC 13
      |||||
Sbjct 6 CCTCCCTC 13

```