



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS
FACULTAD DE ESTUDIOS SUPERIORES IZTACALA
BIOLOGÍA EXPERIMENTAL Y BIOMEDICINA

ANÁLISIS FUNCIONAL DEL PROMOTOR DEL GEN DEL RNA PEQUEÑO NUCLEAR U2 EN
Leishmania major

TESIS

QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS

PRESENTA:

SAÚL ROJAS SÁNCHEZ

TUTOR PRINCIPAL DE TESIS: Dr. Santiago Martínez Calvillo
Facultad de Estudios Superiores Iztacala, UNAM

COMITÉ TUTOR: Dr. Abraham Landa Piedra
Facultad de Medicina, UNAM

COMITÉ TUTOR: Dr. Javier Rolando Ambrosio Hernández
Facultad de Medicina, UNAM

LOS REYES IZTACALA, TLANEPANTLA , EDO. DE MÉXICO
SEPTIEMBRE, 2016.



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

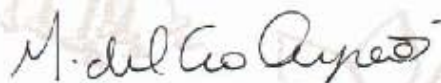
Dr. Isidro Ávila Martínez
Director General de Administración Escolar, UNAM
Presente

Me permito informar a usted que en la reunión ordinaria del Subcomité de Biología Experimental y Biotecnología del Posgrado en Ciencias Biológicas celebrada el día 25 de abril de 2016, se aprobó el siguiente jurado para el examen de grado de **DOCTOR EN CIENCIAS** del alumno **ROJAS SÁNCHEZ SAÚL**, con número de cuenta **303180353**, con la tesis titulada "**Análisis funcional del promotor del gen del RNA pequeño nuclear U2 en *Leishmania major***", realizada bajo la dirección del **DR. SANTIAGO MARTÍNEZ CALVILLO**:

Presidente: DRA. ANA MARÍA CEVALLOS CAOS
Vocal: DR. FELIPE VACA PANIAGUA
Secretario: DR. ABRAHAM LANDA PIEDRA
Suplente: DR. SERGIO VACA PACHECO
Suplente: DR. JORGE EDUARDO CAMPOS CONTRERAS

Sin otro particular, me es grato enviarle un cordial saludo.

ATENTAMENTE
"POR MI RAZA HABLARA EL ESPIRITU"
Cd. Universitaria, Cd. Mx., a 11 de agosto de 2016.



DRA. MARÍA DEL CORO ARIZMENDI ARRIAGA
COORDINADORA DEL PROGRAMA



c.c.p. Expediente del (la) interesado (a)

Agradecimientos

Al Posgrado en Ciencias Biológicas de la Universidad Nacional Autónoma de México por la formación recibida.

Al Consejo Nacional de Ciencia y Tecnología, por la beca otorgada para cursar mis estudios de doctorado (CVU: 384986, becario: 256179).

A la Coordinación de Estudios de Posgrado (CEP) de la Universidad Nacional Autónoma de México por el apoyo otorgado mediante su “programa de becas para estudios de posgrado en la UNAM”, No. de Cta.: 303180353.

Este proyecto fue financiado por los donativos 251831 de CONACyT, IN210712 e IN214715 de PAPIIT (DGAPA, UNAM), y FESI-DIP-PAPCA-2014-24 de PAPCA (FES Iztacala, UNAM) que fueron otorgados al Dr. Santiago Martínez Calvillo.

Al Dr. Santiago Martínez Calvillo, por la dirección del presente trabajo.

Al Dr. Abraham Landa Piedra y a al Dr. Javier Rolando Ambrosio Hernández, por la asesoría brindada durante el transcurso de mis estudios de Doctorado.

Agradecimientos a título personal

A la Universidad Nacional Autónoma de México, institución en donde me he formado académica y personalmente desde mis estudios de nivel medio superior hasta posgrado. Gracias a mi *alma mater*.

Al Dr. Santiago Martínez Calvillo porque a lo largo de todos estos años su tutoría no solo me ha formado en el ámbito académico sino también me ha inspirado personalmente.

A la Dra. Elisa Elvira Figueroa Angulo, los primeros esfuerzos que invirtió en este proyecto han rendido frutos con el presente trabajo. Le agradezco mucho el haberme instruído en mis primeros días de trabajo experimental.

Al Biólogo Luis Enrique Florencio Martínez por la ayuda técnica brindada durante el trabajo experimental.

A los compañeros de laboratorio que me apoyaron cuando surgían problemas de trabajo y que hacían ameno el tiempo laboral.

Por las enriquecedoras críticas y comentarios acerca de esta Tesis, agradezco a los miembros del jurado de mi examen para obtener el grado: Dr. Abraham Landa Piedra, Dra. Ana María Cevallos Gaos, Dr. Felipe Vaca Paniagua, Dr. Jorge Eduardo Campos Contreras y Dr. Sergio Vaca Pacheco.

Quiero agradecer enormemente a mi mamá, María Elena Beatriz Sánchez Vargas, por aguantarme y brindarme su apoyo incondicional durante todo el proceso del Doctorado.

A mi hermana, Dra. Beatriz Elena Rojas Sánchez, por enseñarme a nunca perder la esperanza y por mostrarme cómo ser una buena persona.

A mis sobrinos, Tulio Andrés Padilla Rojas y Ariadna Inés Padilla Rojas, porque aunque al momento en el que escribo estas líneas no lo sepan, llegaron como un rayo de alegría en tiempos oscuros ocurridos durante mis estudios de Doctorado.

A mi compadre, Dr. Daniel Padilla Cicourel, por el apoyo a nuestra familia en los buenos y malos momentos.

A toda mi familia que mostró su apoyo cuando pasé por circunstancias amargas.

A mis amigos, por hacerme olvidar el estrés y ayudarme a sobrellevar con risas todos los problemas personales y laborales que surgían.

A Diana L. Martínez Saucedo, que con su compañía y cariño me hizo fuerte cuando los malos ratos me doblegaban.

Dedicatoria

A todas las personas que entraron, permanecieron y/o salieron de mi vida durante mi Doctorado y que de alguna manera, directa o indirectamente, influyeron en este largo proceso.

Dedico mi Tesis también al presente pues es el tiempo en el que se vive y se ama, se llora y se ríe, se lee y se escribe. El pasado y el futuro, ambos existen sólo en el presente.

Contenido

Lista de Figuras y Tablas.....	IX
Lista de abreviaturas.....	XI
Resumen	XII
<i>Abstract</i>	XIII
1. Introducción.....	1
1.1. Biología Celular de <i>Leishmania</i>.....	1
1.2. Taxonomía de <i>Leishmania</i>	1
1.3. Leishmaniasis.....	3
1.3.1. Tipos de leishmaniasis	3
1.4. Distribución de la leishmaniasis en el mundo	5
1.5. Leishmaniasis en México	6
1.6. Ciclo de vida de <i>Leishmania</i>	7
1.7. <i>Leishmania</i> desde el punto de vista de la biología molecular.....	8
1.8. Transcripción en eucariontes	8
1.8.1. RNA polimerasas	9
1.8.2. Secuencias promotoras.....	10
1.8.3. Factores de transcripción generales de la RNAP III.....	15
1.8.4. Factores de transcripción involucrados en la transcripción de los snRNAs	16
1.8.5. Procesamiento de los snRNAs	18
1.9. Proceso de <i>splicing</i> y el empalmosoma.....	20
1.10. Los snRNAs del empalmosoma y sus modificaciones postranscripcionales.....	22
1.11. Expresión genética en tripanosomátidos	24
1.11.1. Características del genoma nuclear de <i>Leishmania</i>.....	24
1.11.2. Maduración de pre-mRNAs en tripanosomátidos.....	25
2. Antecedentes	27
3. Justificación de la investigación	29
4. Objetivos.....	30
5. Materiales y métodos.....	31
5.1. Análisis <i>in silico</i>.....	31
5.2. Cultivo de <i>L. major</i>	31
5.3. Extracción de RNA total.....	31
5.4. Tratamiento de RNA con DNAsa I	32
5.5. Ensayo 5'-RACE.....	33
5.6. Secuenciación automatizada de plásmidos.....	34
5.7. Mapeo de extremos 3' por RT-PCR	35

5.8. Marcaje radioactivo de oligonucleótidos	36
5.9. Ensayo de <i>primer extension</i>	36
5.10. Mapeo de pseudouridinas	37
5.11. Secuenciación manual de DNA	38
5.12. Electroforesis en geles desnaturalizantes	39
5.13. Ensayos tipo Northern blot	40
5.14. Extracción de DNA genómico	41
5.15. Amplificación por PCR	41
5.16. Electroforesis en geles de agarosa	43
5.17. Purificación de fragmentos de DNA	43
5.18. Ligaciones y transformaciones	44
5.19. Extracción de DNA plasmídico	44
5.20. Ensayo de mutagénesis dirigida	45
5.21. Digestión de plásmidos con enzimas de restricción	47
5.22. Transfecciones transitorias	47
5.23. RT-qPCR	48
6. Resultados	50
6.1. Contexto genómico del snRNA U2 en <i>L. major</i> y otros tripanosomátidos	50
6.2. Mapeo de sitios de inicio y término de la transcripción del snRNA U2 de <i>L. major</i>	52
6.2.1. Mapeo del sitio de inicio de la transcripción del snRNA U2 en <i>L. major</i>	53
6.2.2. Mapeo del extremo 3' maduro del snRNA U2 en <i>L. major</i>	56
6.3. Localización de residuos de pseudouridina en el snRNA U2 de <i>L. major</i>	58
6.4. Localización de residuos de pseudouridina en dominios importantes del snRNA U2 de <i>L. major</i> ...	60
6.5. Análisis comparativo de las secuencias de snRNAs U2 de tripanosomátidos	61
6.6. Identificación de un tRNA-like asociado al snRNA U2 de <i>L. major</i>	65
6.7. Caracterización de elementos promotores del gen del snRNA U2 de <i>L. major</i>	66
6.7.1. Amplificación y clonación del gen del snRNA U2	68
6.7.2. Inserción de una secuencia <i>tag</i> en la región codificante de U2	69
6.7.3. Adición de la señal de término de la transcripción del gen del snRNA U2	71
6.7.4. Obtención de vectores con deleciones de la región flanqueante 5' de U2	72
6.7.5. Obtención de vectores con mutaciones	75
6.7.6. Actividad transcripcional <i>in vivo</i> del promotor del snRNA U2 en <i>L. major</i>	78
6.8. Análisis transcripcional del tRNA-like asociado a U2 en <i>L. major</i>	83
6.9. Análisis de las secuencias de las cajas A y B de tRNAs-like asociados a snRNAs U2 de tripanosomátidos	85
7. Discusión	89

7.1. Número de copias del snRNA U2	89
7.2. Término de la transcripción del snRNA U2	90
7.3. Pseudouridinas en el snRNA U2.....	91
7.4. Promotor del snRNA U2 en <i>L. major</i> y otros organismos.....	92
7.5. Transcripción de tRNAs-like asociados a snRNAs en tripanosomátidos	94
7.6. Modelo de la transcripción del snRNA U2 en <i>L. major</i>.....	95
8. Conclusiones.....	99
9. Bibliografía.....	101
10. Apéndice	115
10.1. Artículo de investigación derivado de los estudios de Doctorado	115

Lista de Figuras y Tablas

Figura 1. Taxonomía del género <i>Leishmania</i>	2
Figura 2. Tipos principales de leishmaniasis.....	4
Figura 3. Distribución de la leishmaniasis cutánea y visceral en el mundo.	6
Figura 4. Distribución de la leishmaniasis en México.....	7
Figura 5. Tipos de RNAs que transcriben las RNAPs de eucariontes.....	9
Figura 6. Promotores reconocidos por la RNAP III.	11
Figura 7. Estructura de los promotores de snRNAs de algunos eucariontes.	13
Figura 8. Factores de transcripción TFIIB y TFIIC.....	15
Figura 9. Factores involucrados en la transcripción de genes de snRNAs llevada a cabo por la RNAP II (panel A) o la RNAP III (panel B).	16
Figura 10. Formación del extremo 3' de snRNAs transcritos por la RNAP II.....	19
Figura 11. Ciclo del <i>splicing</i> y el ensamblaje del empalmosoma.	22
Figura 12. Secuencia, estructura secundaria y pseudouridinas de los snRNAs del empalmosoma.	23
Figura 13. Producción de mRNAs maduros en tripanosomátidos.....	26
Figura 14. Trans- <i>splicing</i>	27
Figura 15. Los snRNAs son transcritos por diferentes RNAPs en eucariontes superiores y tripanosomátidos.	28
Figura 16. Los promotores de snRNAs de <i>T. brucei</i> difieren de los de eucariontes superiores.....	28
Figura 17. Contexto genómico del snRNA U2 en <i>Leishmania</i> spp.....	50
Figura 18. Secuencia codificante y 3' flanqueante del gen del snRNA U2 en <i>L. major</i>	52
Figura 19. Electroforesis de RNA total extraído de <i>L. major</i>	53
Figura 20. Sitio de inicio de la transcripción del snRNA U2 de <i>L. major</i> determinado mediante <i>primer extension</i>	54
Figura 21. Productos de PCR del ensayo 5'-RACE.	55
Figura 22. Digestión con las enzimas <i>Nco</i> I y <i>Sac</i> I de las clonas obtenidas del ensayo 5'-RACE.	55
Figura 23. Productos de RT-PCR del ensayo para mapear el extremo 3' del snRNA U2.	56
Figura 24. Digestión con <i>Eco</i> RI de las clonas obtenidas en el ensayo para mapear el extremo 3' del snRNA U2.	57
Figura 25. Mapeo de pseudouridinas en el snRNA de <i>L. major</i>	59
Figura 26. Mapeo de la Ψ 14 en el snRNA U2 de <i>L. major</i>	60
Figura 27. Posición de las Ψ en la estructura secundaria del snRNA U2 de <i>L. major</i>	60
Figura 28. Alineamiento múltiple de los genes del snRNA U2 de tripanosomátidos.	64
Figura 29. Secuencia consenso del gen del snRNA U2 en tripanosomátidos.	65
Figura 30. Esquema del snRNA U2 de <i>L. major</i> y sus tRNAs asociados.	66
Figura 31. Localización y secuencia de los oligonucleótidos usados para el proceso de caracterización del promotor del snRNA U2 en <i>L. major</i>	67
Figura 32. Obtención de ptRNAU2.....	68
Figura 33. Análisis de secuencia de ptRNAU2.	69
Figura 34. Digestión de siete clonas de ptRNAU2M con <i>Nco</i> I y <i>Sac</i> I.....	70
Figura 35. Comparación de las secuencias de los vectores ptRNAU2M y ptRNAU2.....	70
Figura 36. Obtención de pComp.	71
Figura 37. Comparación de las secuencias de los vectores pComp, ptRNAU2M y ptRNAU2.	72
Figura 38. Esquema de los insertos de los plásmidos con deleciones de la región flanqueante 5' del snRNA U2.	72
Figura 39. Deleción por PCR de secuencias flanqueantes 5' del snRNA U2 de <i>L. major</i>	73
Figura 40. Análisis por restricción de los vectores con deleciones de la región flanqueante 5' del snRNA U2...74	74

Figura 41. Comparación de la secuencia de los vectores pComp, pΔtRNA-Ala, pΔBox-B, pΔBox-A y pΔ-293/-1.....	74
Figura 42. Esquema de los insertos de los plásmidos con mutaciones por sustitución de bases.....	75
Figura 43. Análisis por restricción de los vectores que contienen mutaciones por sustitución de bases.	76
Figura 44. Comparación de las secuencias de los vectores que contienen mutaciones.....	77
Figura 45. Identificación de elementos de secuencia requeridos para la expresión de U2 por medio de <i>primer extension</i>	79
Figura 46. Identificación de elementos de secuencia requeridos para la expresión de U2 por medio de RT-qPCR.	80
Figura 47. Análisis tipo Northern blot del snRNA U2 exógeno.	82
Figura 48. Transcripción del gen del tRNA-like asociado al snRNA U2.	84
Figura 49. Posible estructura secundaria del tRNA-like.	85
Figura 50. Análisis de secuencias de las cajas A y B de genes de tRNA-like asociados a genes del snRNA U2 en tripanosomátidos.....	86
Figura 51. Contexto genómico de los genes del snRNA U2 en tripanosomátidos.....	88
Figura 52. Comparaciones de las secuencias de los genes del snRNA U2 y sus secuencias flanqueantes presentes en <i>T. cruzi</i> (CL Brener Non-Esmeraldo-like).	89
Figura 53. Modelo de la transcripción de snRNAs en tripanosomátidos.	96
Figura 54. El elemento de control interno en los snRNAs de tripanosomátidos se asemeja a elementos promotores de snRNAs de eucariontes superiores.	97
Tabla 1. Oligonucleótidos usados para generar los plásmidos con deleciones de la región flanqueante 5' del snRNA U2 de <i>L. major</i>	42
Tabla 2. Oligonucleótidos usados para generar los plásmidos con mutaciones en la región codificante y flanqueante 5' del snRNA U2 de <i>L. major</i>	47
Tabla 3. Porcentajes de identidad de la secuencia del snRNA U2 entre diferentes tripanosomátidos.	63

Lista de abreviaturas

2'-Ome: metilación del oxígeno 2' de la ribosa

3'SS: sitio de empalme 3'

5'SS: sitio de empalme 5'

Ψ: residuo de pseudouridina

A: residuo de Adenina

BP: punto de ramificación

BPRR: región de reconocimiento del punto de ramificación

Brf: factor relacionado a TFIIB

C: residuo de Citosina

cDNA: DNA copia

CP: promotor núcleo

CTD: dominio carboxilo terminal de la RNAP II

dNTPs: desoxinucleósidos trifosfato

DPE: elemento promotor río abajo

dsDNA: DNA de doble cadena

DSE: elemento de secuencia distal

EI: elemento intermedio

G: residuo de Guanina

gDNA: DNA genómico

Inr: elemento iniciador

Kb: kilobases

m3G: cap tipo 2,2,7-trimetil-guanosina

m7G: cap tipo 7-metil-guanosina (m7GpppG)

Mb: megabases

miRNAs: micro RNAs

MLEE: electroforesis de enzimas de varios locus

mRNA: RNA mensajero

nt: nucleótidos

OH: tratamiento alcalino

pb: pares de bases

PE: *primer extension*

pre-mRNA: precursor de mRNA

PSE: elemento de secuencia proximal

RACE: amplificación rápida de extremos de cDNA

RB: proteína supresora de retinoblastomas

rInr: iniciador ribosomal

RNAP: RNA polimerasa

rRNA: RNA ribosomal

RT-qPCR: transcripción reversa-PCR tiempo real

SIT: Sitio de inicio de la transcripción (+1)

SL RNA: RNA del *spliced leader*

Sm: sitio que reconocen las proteínas Sm

SNAPc: complejo proteico activador de la transcripción de snRNAs

snoRNAs: RNAs pequeños nucleolares

snRNAs: RNAs pequeños nucleares

snRNP: ribonucleoproteína pequeña nuclear

snTAFc: complejo de TAFs específico de snRNAs

T: residuo de Timina

TBP: proteína de unión a caja TATA

TF: Factor de transcripción

Tritryps: *T. brucei*, *T. cruzi* y *L. major*

tRNA: RNA de transferencia

tRNA-Ala: RNA de transferencia de alanina

tRNA-like: gen parecido a un RNA de transferencia

U: residuo de Uracilo

UPE: elemento promotor río arriba

USE: elemento de secuencia río arriba

WT: células silvestres

Resumen

Leishmania, como otros tripanosomátidos parásitos, presenta mecanismos de expresión genética atípicos que incluyen a la maduración de mRNAs por *trans-splicing* y la transcripción de todos los snRNAs por la RNA polimerasa III. Debido a que los snRNAs son esenciales para el proceso de *trans-splicing*, estamos interesados en el estudio de las regiones de DNA que regulan su síntesis. En este trabajo se reporta la caracterización de la región promotora del gen del snRNA U2 de *L. major*. Un análisis de secuencia mostró que todas las especies de *Leishmania* contienen un solo gen del snRNA U2, y que río arriba de este gen se encuentra un gen de tRNA-Ala ubicado en la hebra de DNA opuesta. Entre estos dos genes, se encontró un gen parecido a un tRNA (tRNA-like) que presenta los elementos de secuencia típicos de un tRNA: las cajas A y B. Por medio de ensayos de *primer extension* y RT-qPCR en los que se usó RNA total de células transfectadas transitoriamente, se demostró que la transcripción del gen del snRNA U2 de *L. major* es abatida cuando las cajas A y B del tRNA-like son deletadas o mutadas. Los niveles del snRNA U2 también se vieron afectados cuando se mutó la caja B del tRNA-Ala y a los primeros nucleótidos del gen del snRNA U2. Así, nuestros resultados demostraron que el promotor del snRNA U2 de *L. major* es diferente a cualquier otro promotor descrito para snRNAs en otros organismos. También se demostró que el gen del tRNA-like es transcrito *in vivo*, generándose un RNA de alrededor de 109 nucleótidos. Como las pseudouridinas de los snRNAs son requeridas para el proceso de *splicing* y la formación de snRNPs en otros organismos, se realizó una búsqueda de este nucleótido modificado en el snRNA U2 de *L. major* usando carbodiimida soluble (CMCT). Los resultados de estos ensayos muestran la presencia de ocho pseudouridinas en el snRNA U2, incluyendo una pseudouridina ubicada en el sitio Sm, la cual no ha sido reportada en ningún otro organismo.

Abstract

Leishmania and other trypanosomatid parasites possess atypical mechanisms of gene expression, including the maturation of mRNAs by trans-splicing and the involvement of RNA Polymerase III in the transcription of all snRNA molecules. Since snRNAs are essential for trans-splicing, we are interested in the functional analysis of the sequences that direct their expression. In this work, the characterization of the promoter region of the U2 snRNA from *L. major* is reported. A sequence analysis showed that all the *Leishmania* species have a single U2 snRNA gene that contains a divergently-oriented tRNA-Ala gene in the upstream region. Between these two genes we found a tRNA-like gene that contains conserved boxes A and B. Primer extension and RT-qPCR analyses performed with RNA from transiently-transfected cells showed that transcription of the *L. major* U2 snRNA is almost abolished when boxes A and B from the tRNA-like gene are deleted or mutated. The levels of the U2 snRNA were also highly affected when base substitutions were introduced into box B from the tRNA-Ala gene and the first nucleotides of the U2 snRNA gene itself. Thus, the promoter region of the *L. major* U2 snRNA is different from any other promoter reported for snRNAs. The *in vivo* transcription of tRNA-like gene was demonstrated by Northern blot and primer extension assays, identifying a main transcript of around 109 nucleotides. As pseudouridines in snRNAs are required for splicing and snRNP formation in other organisms, a search for this modified nucleotide was carried out in the *L. major* U2 snRNA using soluble carbodiimide (CMCT). Results from these assays showed the presence of eight pseudouridines in the U2 snRNA, including one in the Sm site that has not been reported in other organisms.

1. Introducción

1.1. Biología Celular de *Leishmania*

Leishmania es un protozooario parásito que pertenece a la familia Trypanosomatidae, la cual incluye a otros parásitos como *Trypanosoma cruzi* y *Trypanosoma brucei*. Estos organismos causan graves enfermedades en el humano y otros mamíferos. Algunas de las características de la biología celular de *Leishmania* son: presencia de un flagelo con una invaginación de la membrana, llamada bolsa flagelar, en donde se lleva a cabo la endocitosis y exocitosis; contiene un cinetoplasto, el cual es una masa de DNA localizada cerca de la bolsa flagelar, embebida en la única mitocondria presente en la célula; el citoesqueleto está compuesto por una matriz sub-pelicular de microtúbulos; las enzimas necesarias para la glicólisis se concentran en compartimentos altamente específicos llamados glicosomas [1]; y se reproduce asexualmente por fisión binaria, proceso durante el cual sus cromosomas no se condensan ni la envoltura nuclear se disuelve. Se ha propuesto que los ancestros del género *Leishmania* surgieron de organismos flagelados de vida libre, los cuales eran ingeridos por mosquitos que se alimentaban de la sangre de vertebrados hace al menos 100 millones de años en el Paleoceno [2], por lo que la diversificación de especies de *Leishmania* ha sido dirigida desde hace ya varios millones de años por los complejos microambientes de su vida parasitaria, así como por su ubicación geográfica.

1.2. Taxonomía de *Leishmania*

Dependiendo del autor, se han reportado de 31 a 38 especies del género *Leishmania* (Fig. 1). La taxonomía actual de *Leishmania* se ha construido con base en los patrones obtenidos con la técnica de “electroforesis de enzimas de varios locus” (*Multilocus Enzyme Electrophoresis*, MLEE), así como en el análisis de su DNA [3-5]. *Leishmania* junto con los géneros *Crithidia*, *Leptomonas*, *Herpetomonas*, *Blastocrithidia*, *Trypanosoma*, *Phytomonas* y *Endotrypanum* componen a los miembros de la familia Trypanosomatidae. El género *Leishmania* está subdividido en dos secciones, *Paraleishmania* y *Euleishmania*, identificadas gracias a una variedad de técnicas moleculares [6]. *Paraleishmania* está

compuesta por las especies *L. colombiensis*, *L. equatorensis*, *L. hertigi*, *L. herreri* y *L. deanei*. Euleishmania por su parte está compuesta por los subgéneros *Sauroleishmania*, *Leishmania* y *Viannia*, y la especie *L. siamensis* [7-9]. *Sauroleishmania* se caracteriza porque sus miembros son parásitos de lagartijas del Viejo Mundo (Fig. 1). Los otros dos subgéneros de Euleishmania (*Leishmania* y *Viannia*) incluyen a diversos miembros parásitos de humanos (Fig. 1).

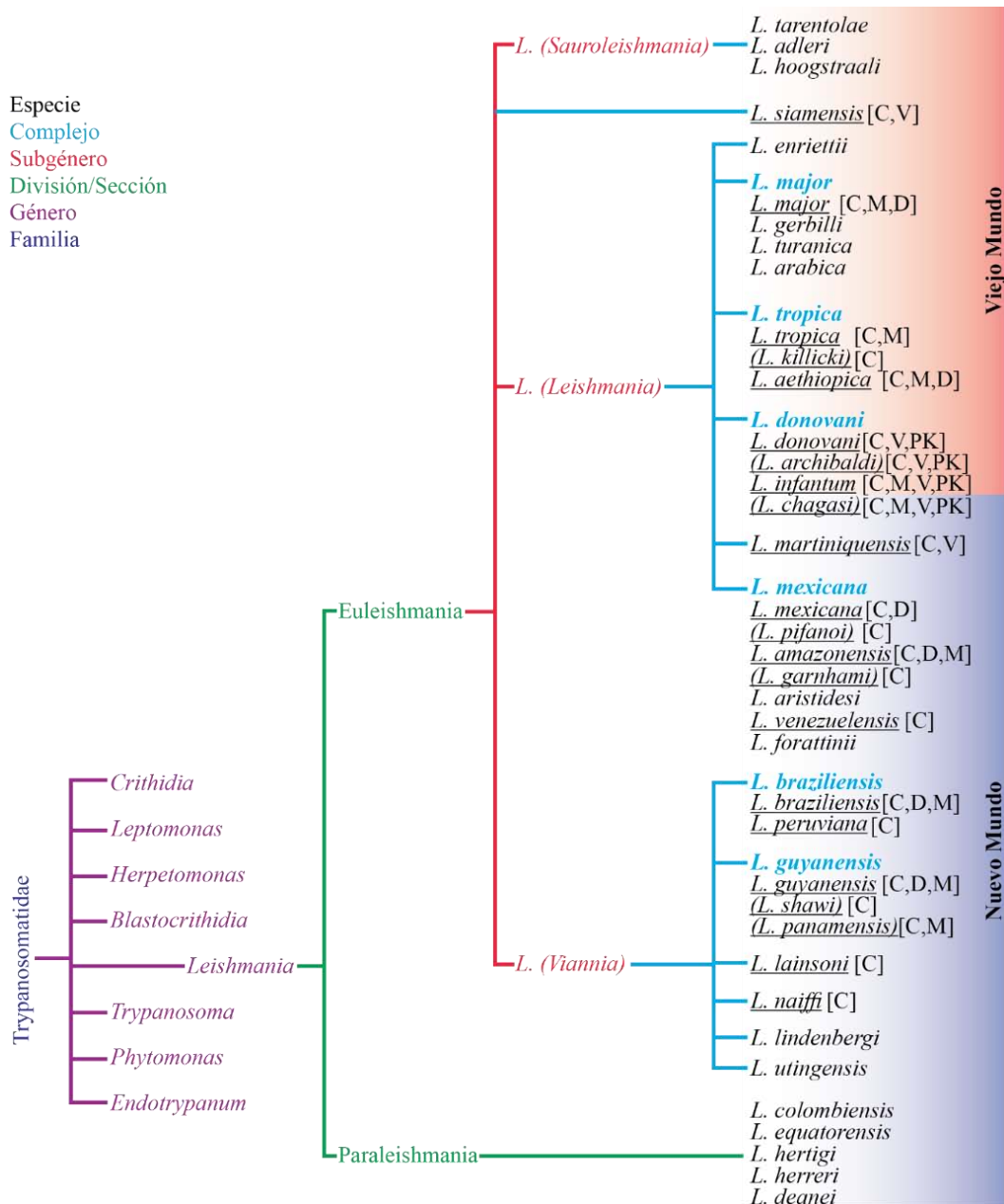


Figura 1. Taxonomía del género *Leishmania*. Los diferentes niveles taxonómicos están indicados con líneas y nombres en colores: Especies en negro, Complejos en azul, Subgéneros en rojo, Secciones en verde, Géneros en morado y Familia en azul marino. Las especies subrayadas son causantes de algún tipo de leishmaniasis: [C], cutánea localizada; [D], cutánea difusa; [M], mucocutánea; [V], visceral o “kala azar”; [PK], Post-kala azar. Las especies entre paréntesis no son reconocidas como especies verdaderas por algunos autores y se posicionan abajo de su especie más relacionada. Modificado de [4] con información basada en [4,10-12].

El subgénero *Leishmania*, único que tiene especies tanto del Viejo como del Nuevo Mundo, está formado por cuatro complejos: *L. major*, *L. tropica*, *L. donovani* y *L. mexicana*, y dos especies: *L. enrietti* y *L. martiniquensis* (Fig. 1). El subgénero *Viannia* [13] sólo tiene miembros en el Nuevo Mundo y está formado por los complejos *L. braziliensis* y *L. guyanensis*, además de las especies *L. lainsoni*, *L. naiffi*, *L. lindenbergi* y *L. utingensis*.

1.3. Leishmaniasis

Se ha determinado que al menos 22 de las ~38 especies de *Leishmania* causan la enfermedad llamada leishmaniasis. El agente causal de dicha enfermedad se identificó por vez primera en 1903, de manera simultánea pero independiente, por William Boog Leishman y Charles Donovan [14]. Sin embargo, la historia de la leishmaniasis se puede rastrear hasta el primer siglo de nuestra era como evidencian vasijas antropomórficas y restos humanos momificados [15,16].

1.3.1. Tipos de leishmaniasis

Existen distintos tipos de leishmaniasis, siendo los principales: leishmaniasis cutánea, visceral y post-kala azar (Fig. 2). Como regla general, la especie de *Leishmania* que causa la infección determina el tipo de leishmaniasis que se va a desarrollar (Fig. 1). Sin embargo, esto depende de diversos factores que incluyen: la cepa del parásito, las características de la saliva del vector, la posible infección del parásito con el virus de RNA de *Leishmania*, la genética del hospedero y la inmunosupresión cuando se presenta co-infección con VIH, entre otras [17]. A continuación se detallan algunos aspectos de los tipos principales de leishmaniasis.

1.3.1.1. Leishmaniasis cutánea

La leishmaniasis cutánea se ha sub-dividido en tres tipos: leishmaniasis cutánea localizada, cutánea difusa y mucocutánea. La leishmaniasis cutánea localizada se manifiesta con una lesión ubicada en el

sitio de inoculación que puede variar de una pequeña ulceración a heridas grandes, eritematosas y sangrantes (Fig. 2) [11,19-21].



Cutánea localizada por *L. mexicana*.
Kevric, et al 2015.



Mucocutánea por *L. braziliensis*.
Cochabamba, Bolivia. Reithinger et al 2007.



Cutánea difusa.
Brasil. Hartley, et al 2014.



Post-kala azar por *L. infantum*.
Sudán. Hartley, et al 2014.



Visceral.
Libo Kemkem, Etiopía. OMS, 2007.

Figura 2. Tipos principales de leishmaniasis. Debajo de las fotografías se indica el tipo de leishmaniasis correspondiente. Tomadas de [10,18,19].

La leishmaniasis cutánea difusa se caracteriza por lesiones nodulares no ulceradas diseminadas en todo el cuerpo. Las lesiones aparentemente localizadas pueden persistir por meses o años y después diseminarse a otras áreas del cuerpo (Fig. 2) [4,10-12,22-24]. Los síntomas de la leishmaniasis mucocutánea incluyen la congestión nasal persistente, sangrado y ronquera, seguidos de eritema y edema que eventualmente progresan a ulceración con un exudado mucopurulento. El desarrollo de la enfermedad conlleva a la destrucción del septo nasal y la mutilación de las membranas de mucosas de la nariz, boca, garganta y los tejidos circundantes (Fig. 2) [4,10-12,18].

1.3.1.2. Leishmaniasis visceral

La leishmaniasis visceral es el resultado de una infección crónica cuyas características incluyen fiebre, pérdida de peso, hepatoesplenomegalia y pancitopenia (Fig. 2). Pueden presentarse hemorragias por trombocitopenia e infecciones por leucopenia. El oscurecimiento de la piel le ha dado el nombre de

kala azar (fiebre negra) a la leishmaniasis visceral. Este tipo de leishmaniasis es el único que puede causar la muerte si no se trata apropiadamente [4,10-12,18,25].

1.3.1.3. Post-kala azar

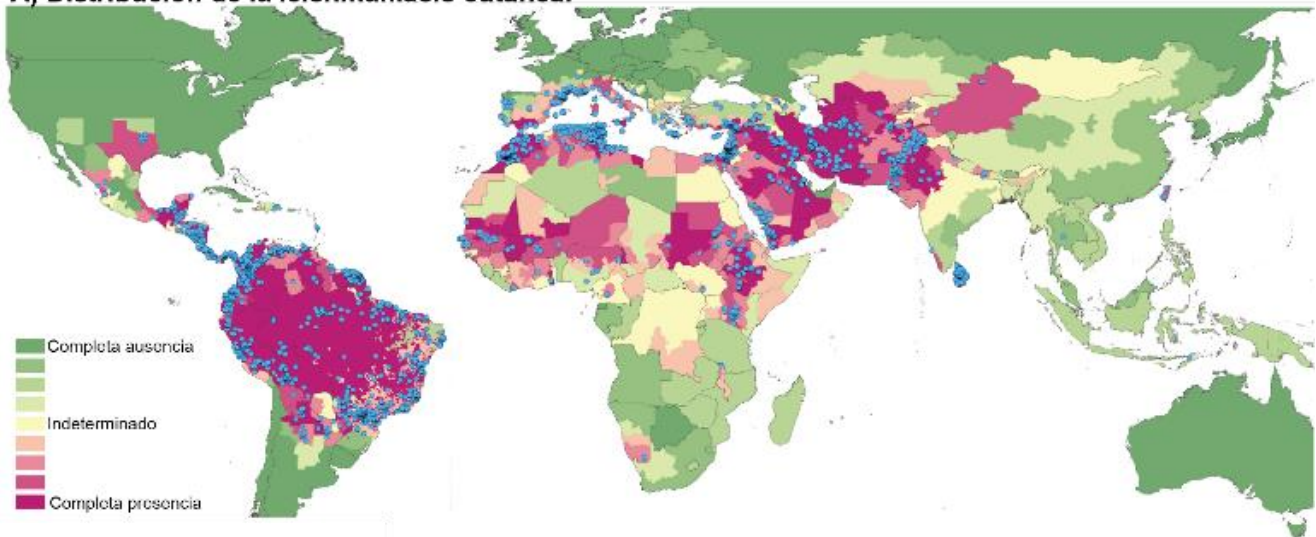
El post-kala azar se desarrolla de meses a años después de que los enfermos de leishmaniasis visceral se han curado aparentemente. Los síntomas varían desde manchas hipopigmentadas en todo el cuerpo, hasta pápulas eritematosas o nódulos ubicados predominantemente sobre la cara (Fig. 2) [4,10-12,18,25-28].

1.4. Distribución de la leishmaniasis en el mundo

La leishmaniasis está ampliamente distribuida alrededor del mundo (Fig. 3). Las estimaciones más recientes indican que cada año ocurren en el mundo de 0.2 a 0.4 millones de nuevos casos de leishmaniasis tipo visceral y de 0.7 a 1.2 millones de nuevos casos de leishmaniasis cutánea. Además, se calcula que hay entre 20 y 40 mil muertes por año en todo el mundo [29].

La leishmaniasis cutánea se extiende en el Nuevo Mundo desde el sur de Estados Unidos hasta el norte de Argentina, en el Viejo Mundo por toda la cuenca del Mediterráneo, en el norte y centro de África y en el oeste de Asia (Fig. 3, panel A) [17,29]. La leishmaniasis visceral está distribuida en América desde México hasta Argentina; en el Viejo Mundo está presente en la cuenca del mediterráneo, en el este de África, y en una gran parte de Asia que se extiende desde medio oriente hasta el norte de China (Fig. 3, panel B) [17,29].

A) Distribución de la leishmaniasis cutánea.



B) Distribución de la leishmaniasis visceral.

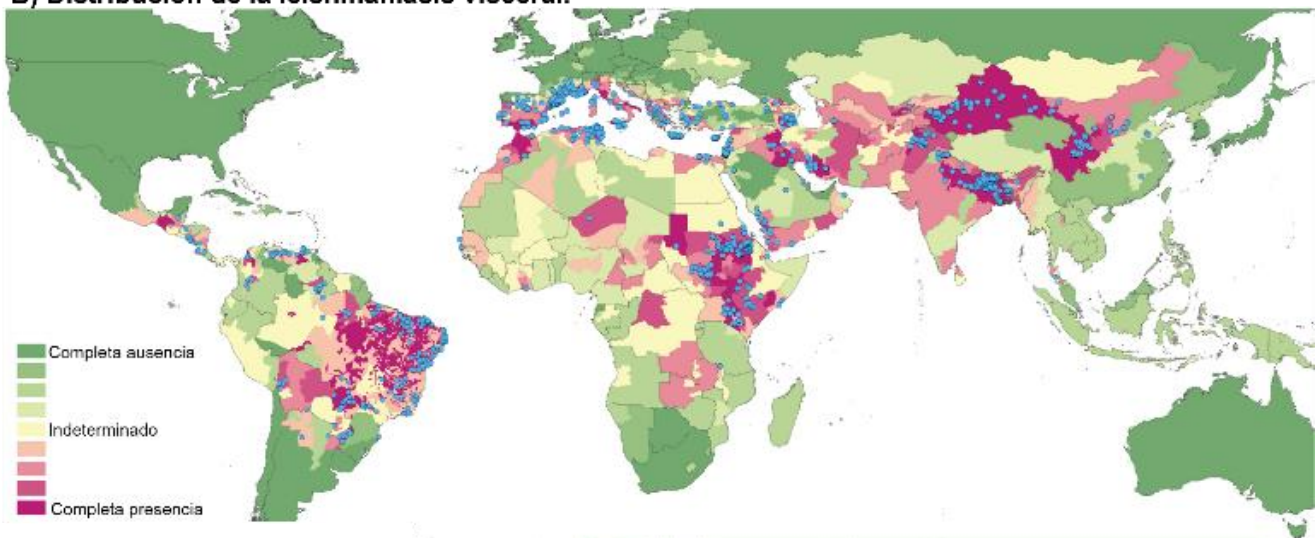


Figura 3. Distribución de la leishmaniasis cutánea y visceral en el mundo. Se representa la presencia de la leishmaniasis en diferentes regiones con un intervalo de colores que va del verde (completa ausencia) al morado (completa presencia). Los puntos azules representan los lugares donde se han reportado los casos de leishmaniasis. Panel A): distribución de la leishmaniasis cutánea. Panel B): distribución de la leishmaniasis visceral. Modificado de [17].

1.5. Leishmaniasis en México

En México existen reportes de los tipos principales de leishmaniasis (Fig. 4). La población en riesgo de enfermarse de leishmaniasis cutánea es de 7,613,221 personas, mientras que 921,273 personas se encuentran en riesgo de contraer leishmaniasis visceral. La leishmaniasis cutánea localizada es la más frecuente en México y se conoce como “úlceras del chiclero” por presentarse en trabajadores que extraían

la goma del *Manilkara zapota*. La mayoría de reportes de leishmaniasis han ocurrido en los estados de Yucatán, Quintana Roo, Campeche, Tabasco, Chiapas, Oaxaca, Veracruz, Guerrero y Puebla; aunque también se ha reportado en Coahuila, Michoacán, Morelos, Nayarit, Jalisco, Sinaloa, Durango y Ciudad de México (Fig. 4) [29].

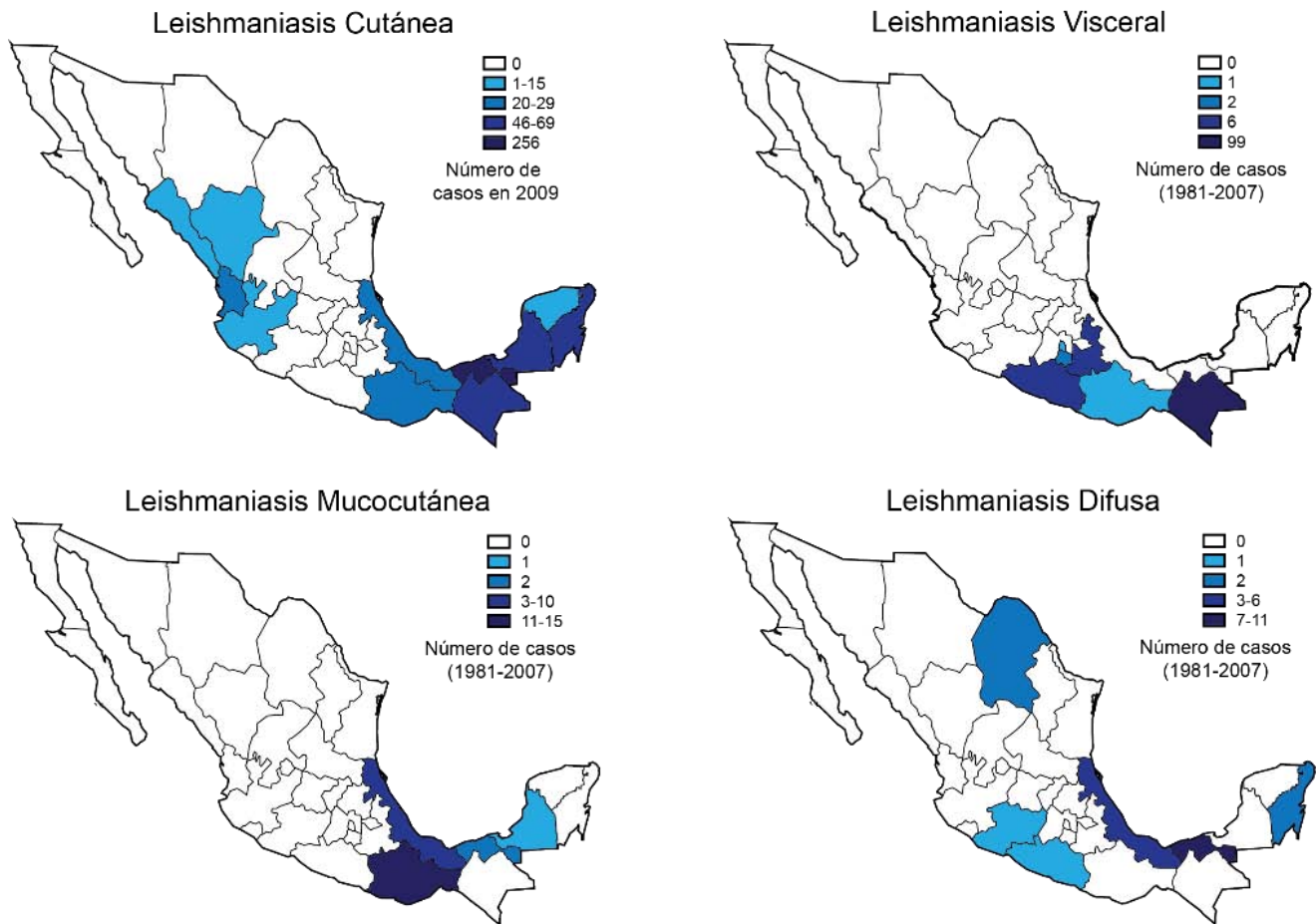


Figura 4. Distribución de la leishmaniasis en México. Se representa el número de casos reportados en un determinado lapso de tiempo con un gradiente de colores en azul. Modificado de [29].

1.6. Ciclo de vida de *Leishmania*

El ciclo de vida de *Leishmania* spp. comienza cuando el insecto vector, la mosca de la arena de los géneros *Phlebotomus* (en el Viejo Mundo) o *Lutzomyia* (en América), se alimenta de la sangre de algún vertebrado mamífero, ingiriendo macrófagos infectados con amastigotes. Los parásitos posteriormente se diferencian a promastigotes procíclicos y viajan al intestino medio del insecto para que en ese lugar se repliquen activamente. Subsecuentemente, los promastigotes procíclicos migran al tracto

digestivo anterior del mosquito y ahí se diferencian a promastigotes metacíclicos infectivos. Posteriormente cuando el mosquito vuelve a alimentarse, inyecta los promastigotes metacíclicos al hospedero definitivo a través de la piel. Es en el torrente sanguíneo del hospedero donde los macrófagos fagocitan al parásito formando una vacuola parasitófora la cual contiene un microambiente que es modificado por el propio parásito durante el desarrollo de la infección [30]. Dentro de las 12 a 24 h siguientes, los parásitos se transforman a amastigotes en el interior del macrófago infectado. Después, los amastigotes se multiplican por fisión binaria dentro de la vacuola parasitófora, resultando en la ruptura del macrófago y la expulsión de los amastigotes, los cuales infectan nuevos macrófagos y otras células del sistema inmune. Esta última parte del ciclo es crónica ya que puede continuar por varios meses o años, incluso toda la vida del hospedero.

1.7. *Leishmania* desde el punto de vista de la biología molecular

La leishmaniasis continúa siendo un gran problema de salud a nivel mundial. Además, *Leishmania* constituye un modelo de estudio complejo desde el punto de vista de la biología molecular pues presenta diversos mecanismos de expresión genética diferentes a los que se encuentran en eucariontes superiores, incluyendo la transcripción policistrónica y la maduración de sus RNAs mensajeros por *trans-splicing*. En nuestro laboratorio estamos interesados en estudiar la expresión genética en *Leishmania major* Friedlin, cepa usada en el “Proyecto del Genoma de *Leishmania*”. Particularmente, nuestro interés se enfoca en el análisis de la transcripción de la RNA polimerasa III (RNAP III), la cual sintetiza moléculas de RNA pequeñas esenciales para el crecimiento celular. Antes de hablar de la expresión genética en *Leishmania*, se tratarán algunas generalidades de la expresión genética en eucariontes superiores.

1.8. Transcripción en eucariontes

La transcripción es un proceso celular fundamental en el que se sintetizan hebras de RNA usando secuencias de DNA como molde. La hebra de RNA resultante es idéntica en secuencia a la cadena de

DNA codificante y complementaria a la cadena molde. Esta reacción bioquímica es llevada a cabo por una enzima conocida como RNA polimerasa (RNAP) que adiciona unidades ribonucleótidas a los extremos 3'-hidroxilo de la cadena de RNA en elongación. Como consecuencia, la nueva cadena de RNA, llamada transcrito primario, crece en dirección 5'→3' y de forma antiparalela a la cadena de DNA molde. La transcripción puede ser dividida en tres etapas: inicio, elongación y término. Al inicio de la transcripción ciertos factores proteicos, llamados factores de transcripción, reconocen una secuencia dentro del DNA denominada promotor. Posteriormente la RNAP, a través de interacciones con los factores de transcripción, reconoce a su vez la secuencia promotora. El complejo formado por la RNAP unida a los factores de transcripción y al promotor, determina el sitio de inicio de la transcripción (SIT, nucleótido +1) de un gen determinado.

1.8.1. RNA polimerasas

Mientras que la transcripción en procariontes es llevada a cabo por sólo una RNAP, en eucariontes se han identificado tres RNAPs nucleares llamadas I, II y III, las cuales difieren en el número de subunidades que las componen, su localización nuclear y, de manera más relevante, en los RNAs que transcriben [31]. La RNAP I transcribe al RNA ribosomal (rRNA) 45S, que se procesa para generar a los rRNAs 28S, 18S y 5.8S. La RNAP II sintetiza RNAs mensajeros (mRNAs), y a los RNAs pequeños nucleares (snRNAs) U1, U2, U4 y U5, los cuales están involucrados en la maduración de los precursores de mRNAs (pre-mRNAs). La RNAP III transcribe al rRNA 5S, los RNAs de transferencia (tRNAs), el RNA 7SL y el snRNA U6, así como otros RNAs pequeños esenciales para la viabilidad celular (Fig. 5).

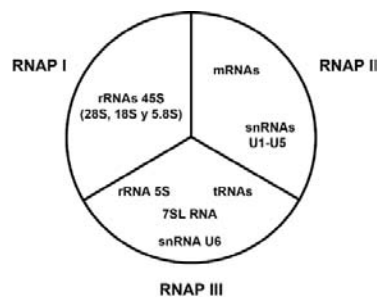


Figura 5. Tipos de RNAs que transcriben las RNAPs de eucariontes. La RNAP I transcribe exclusivamente al rRNA 45S, que se procesa para formar a los rRNAs 28S, 18S y 5.8S; la RNAP II a los mRNAs y snRNAs U1-U5; la RNAP III a los tRNAs, 7SL RNA, rRNA 5S y snRNA U6.

Las RNAPs I, II y III están formadas por 14, 12 y 17 subunidades y sus pesos moleculares son de ~589, ~514 y ~693 kDa, respectivamente [32,33]. La composición de las tres RNAPs es muy similar, pues del total de subunidades que las forman, cinco (ABC27, ABC23, ABC14.5, ABC10 α y ABC10 β) son compartidas por las tres RNAPs; dos (AC40 y AC19) están presentes en la RNAP I y III, aunque son homólogas de las subunidades B44 y B12.5 de la RNAP II; y cinco subunidades (A135/B220/C160, A135/B150/C128, A43/B16/C25, A14/B32/C17 y A12/B12.6/C11) de las tres RNAPs presentan un alto grado de homología. Existen también subunidades específicas, pues las subunidades A49 y A34.5 sólo se encuentran en la RNAP I y las subunidades C82, C53, C37, C34 y C31 sólo están presentes en la RNAP III [34]. Resulta interesante que cuatro de las subunidades específicas de la RNAP III (C82, C34, C37 y C53) son homólogas de los factores de transcripción TFIIE y TFIIIF de la RNAP II [35]. De manera similar, las dos subunidades específicas de la RNAP I son homólogas del factor de transcripción TFIIIF [32]. Las RNAPs de eucariontes también presentan cierta similitud con la RNAP de procariontes, pues las dos subunidades más grandes de las RNAPs I, II y III son homólogas de los factores β y β' de la RNAP de eubacterias [36]. La similitud que existe entre las RNAPs de todos los seres vivos sugiere un origen evolutivo común.

1.8.2. Secuencias promotoras

Los promotores, como se ha mencionado, son secuencias de DNA que son reconocidas por los factores de transcripción y las RNAPs para poder iniciar la transcripción de un gen. Cada RNAP se une a un tipo diferente de promotor, como se resume a continuación.

1.8.2.1. Promotores de la RNAP I

La RNAP I transcribe únicamente a los genes del rRNA 45S. El promotor de dichos genes consta de dos elementos: el primero es conocido como CP o promotor núcleo y se extiende de la posición -45 a la posición +20, en relación al SIT (+1); el segundo elemento se conoce como UPE o elemento promotor

río arriba y se ubica de -180 a -107. El CP contiene el único elemento del promotor conservado en secuencia, el iniciador ribosomal o rInr, el cual es una secuencia rica en AT localizada alrededor de +1 [37,38].

1.8.2.2. Promotores de la RNAP II

La RNAP II transcribe a todos los mRNAs y a los snRNAs U1, U2, U4 y U5. Los promotores de genes de mRNAs muestran una gran variación en el arreglo de elementos que los conforman. Algunos de estos elementos son: un elemento iniciador (Inr) ubicado entre las posiciones -3 y +5, una caja TATA ubicada generalmente ~30 pares de bases (pb) río arriba del SIT y un elemento promotor río abajo (DPE) localizado entre las posiciones +28 y +32. Típicamente un promotor de la RNAP II presenta al menos dos elementos, la caja TATA y el Inr, o bien el Inr y el DPE [39,40]. Los promotores de los snRNAs U1-U5 se describirán más adelante.

1.8.2.3. Promotores de la RNAP III

La RNAP III transcribe al rRNA 5S, los tRNAs, al RNA 7SL y al snRNA U6, así como otros RNAs pequeños esenciales para el crecimiento celular. La característica más notable e inusual de los promotores de la RNAP III es que la mayoría están constituidos por elementos de secuencia ubicados dentro de la región transcrita. Los promotores de genes transcritos por la RNAP III son principalmente de tres tipos (Fig. 6).

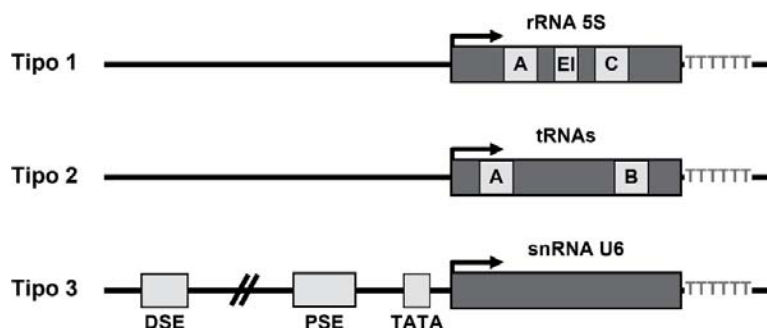


Figura 6. Promotores reconocidos por la RNAP III. Se muestran los promotores tipo 1 (del rRNA 5S), tipo 2 (típico de tRNAs) y tipo 3 (del snRNA U6). Se señala la posición de las cajas A, B y C, así como del elemento intermedio (EI), el elemento de secuencia proximal (PSE) y el elemento de secuencia distal (DSE). Un trazo de Ts al final de los genes representa la señal de término de la transcripción de la RNAP III. El esquema no se muestra a escala.

1.8.2.3.1. Promotor de la RNAP III tipo 1

El promotor tipo 1 está representado por el promotor del gen del rRNA 5S. Lo componen tres elementos intragénicos: una caja A localizada entre +50 y +64, un elemento intermedio (EI) de +67 a +72 y una caja C de +80 a +97 (Fig. 6).

1.8.2.3.2. Promotor de la RNAP III tipo 2

El promotor de tipo 2 es característico de los genes de tRNAs. Consiste de dos bloques intragénicos muy conservados: la caja A y la caja B (Fig. 6). La longitud de la caja A varía de entre 11 y 12 pb mientras que la caja B siempre es de 11 pb. Las cajas A se encuentran muy cercanas al SIT, entre las posiciones +10 y +20; por otro lado, la locación de la caja B es muy variable, debido principalmente a que algunos genes de tRNAs tienen intrones pequeños entre la caja A y B. La distancia óptima entre las cajas A y B es de 30 a 60 pb, aunque una distancia de 365 pb puede ser tolerada.

1.8.2.3.3. Promotor de la RNAP III tipo 3

El promotor de tipo 3 es característico del gen del snRNA U6. A diferencia de los promotores tipo 1 y 2, el de tipo 3 se encuentra fuera de la región codificante (Figs. 6 y 7). En humanos este promotor contiene tres elementos: un elemento de secuencia proximal (PSE) ubicado de -50 a -65, un elemento de secuencia distal (DSE) localizado de -200 a -220, y una caja TATA encontrada alrededor de -25.

1.8.2.4. Promotores de snRNAs

Es necesario hacer notar que mientras el snRNA U6 es transcrito por la RNAP III, los snRNAs U1, U2, U4 y U5 son transcritos por la RNAP II. Es de gran interés el hecho de que a pesar de ser transcritos por dos RNAPs, los promotores de snRNAs son muy similares dentro de una especie determinada (Fig. 7). Como regla general, los promotores de snRNAs están compuestos por al menos dos elementos ubicados río arriba del SIT (Fig. 7) [41,42]. En humanos, estos elementos son el PSE y el DSE mencionados anteriormente; sin embargo, mientras que el promotor reconocido por la RNAP II está

compuesto únicamente por el PSE y el DSE, el promotor reconocido por la RNAP III, además de estos dos elementos, contiene una caja TATA localizada alrededor de la posición -25 (Figs. 6 y 7, Hs) [43]. La caja TATA y/o el PSE funcionan como el promotor núcleo, mientras que el DSE se encarga de activar la transcripción de los snRNAs. Tanto el PSE como el DSE pueden ser intercambiados entre los promotores de las RNAP II y III sin afectar la especificidad de las polimerasas. Sin embargo, cuando se remueve la caja TATA del promotor del snRNA U6, el gen comienza a ser transcrito por la RNAP II (Fig. 7, Hs) [44].

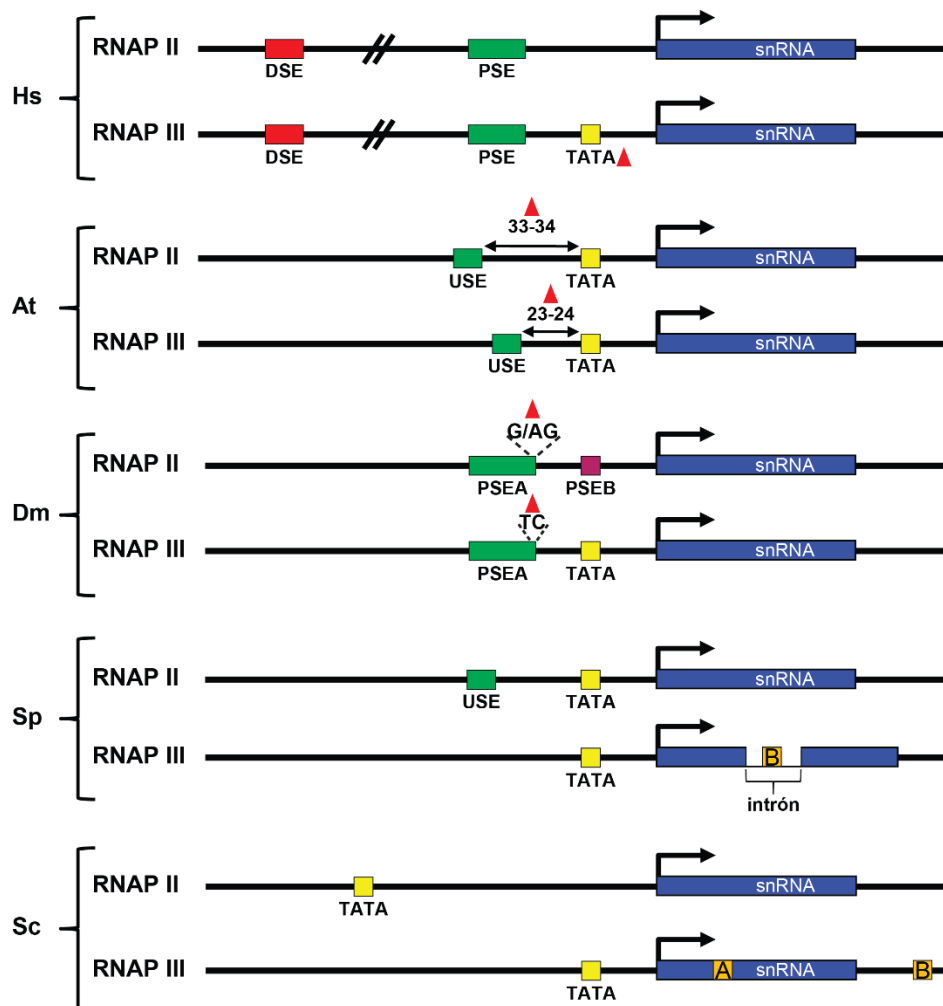


Figura 7. Estructura de los promotores de snRNAs de algunos eucariotes. Se muestran los promotores de snRNAs transcritos por la RNAP II o RNAP III en *Homo sapiens* (Hs), *Arabidopsis thaliana* (At), *Drosophila melanogaster* (Dm), *Schizosaccharomyces pombe* (Sp) y *Saccharomyces cerevisiae* (Sc). Los genes de snRNAs se muestran en azul. Con cajas de diferentes colores se muestran los elementos promotores. Un triángulo rojo indica la característica que marca la especificidad de las RNAPs. En los casos donde no se muestran triángulos rojos, el elemento que dicta la especificidad de las RNAPs no se ha determinado experimentalmente (ver texto). La figura no está a escala. Modificado de [42].

En *A. thaliana* los promotores de snRNAs transcritos por las dos RNAPs están compuestos por un elemento de secuencia río arriba (USE) y una caja TATA. En este caso, la distancia entre los dos elementos promotores determina la especificidad de las RNAPs, pues en el promotor reconocido por la RNAP II dicha distancia es de 33 a 34 pb, mientras que en el promotor de la RNAP III la distancia es de 23 a 24 pb (Fig.7, At) [45]. En *D. melanogaster* los promotores de snRNAs transcritos por la RNAP II consisten de un PSEA ubicado entre -61 y -40 y un PSEB localizado entre -32 y -24; por otro lado, los promotores reconocidos por la RNAP III están compuestos por un PSEA ubicado entre -64 y -43 y una caja TATA localizada entre -31 y -23. La secuencia del elemento PSEA mide 21 nt y está muy conservada. Sin embargo, las posiciones 19 y 20 siempre son GG o AG en el promotor reconocido por la RNAP II y TC en el reconocido por la RNAP III (Fig. 7, Dm). Así, estas diferencias en el elemento PSEA determinan cuál RNAP transcribirá a los genes de snRNAs [46,47]. El promotor de snRNAs transcritos por la RNAP II en *Schizosaccharomyces pombe* está formado por una caja TATA ubicada alrededor de la posición -30 y un USE localizado entre -63 y -44 [48-50]. Las secuencias promotoras de snRNAs transcritos por la RNAP III en este organismo no se han caracterizado funcionalmente; sin embargo, la caja TATA presente en los promotores de la RNAP II se encuentra también río arriba del snRNA U6. Además se ha reportado la presencia de una caja B localizada en las posiciones +67 y +77 dentro de una región intrónica del mismo gen (Fig. 7, Sp) [51]. En *Saccharomyces cerevisiae* los genes de los snRNAs U1-U5 tienen una caja TATA que se ubica entre -85 y -100 [52], aunque su importancia no se ha determinado experimentalmente. Por otra parte, el promotor del snRNA U6 en *S. cerevisiae* consta de una caja TATA ubicada alrededor de la posición -34, y además contiene cajas A y B típicas de los promotores de tRNAs (ver promotor de la RNAP III tipo 2). La caja A se localiza en la posición -20 mientras que la caja B se encuentra fuera de la región codificante, cerca del extremo 3' del gen, a 200 pb río abajo del SIT (Fig. 7, Sc) [53-56].

1.8.3. Factores de transcripción generales de la RNAP III

Los factores de transcripción son un grupo de proteínas o complejos proteicos necesarios en el proceso de transcripción, ya que dirigen a las RNAPs al promotor y al SIT, ayudan a la formación de la burbuja transcripcional y permiten a las polimerasas iniciar la transcripción. La RNAP III necesita tres factores de transcripción generales: 1) el péptido TFIIIA; 2) el complejo TFIIIB compuesto por tres subunidades: TBP (proteína de unión a la caja TATA), Brf (factor relacionado a TFIIIB) y Bdp1; y 3) TFIIIC formado por los subcomplejos τ A y τ B. El subcomplejo τ A está compuesto a su vez por las subunidades τ 131 (Tfc4), τ 95 (Tfc1) y τ 55 (Tfc7), mientras que el subcomplejo τ B contiene a las subunidades τ 138 (Tfc3), τ 91 (Tfc6) y τ 60 (Tfc8) (Fig. 8) [57,58].

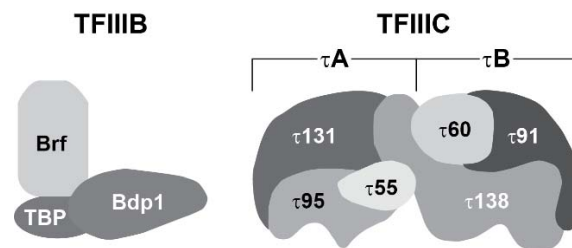


Figura 8. Factores de transcripción TFIIIB y TFIIIC. Se muestra una representación de TFIIIB (izquierda) que contiene a las subunidades TBP, Brf y Bdp1; y TFIIIC que está compuesto por los subcomplejos τ A (subunidades τ 131, τ 95 y τ 55) y τ B (subunidades τ 138, τ 91 y τ 60). Modificado de [57].

Los factores de transcripción reconocen las regiones promotoras en un orden determinado. En el promotor de tipo 1, TFIIIA se une primero a la caja A y al elemento intermedio, luego TFIIIC reconoce a TFIIIA y a la caja C, posteriormente TFIIIB se une al complejo previamente formado por medio de interacciones con TFIIIC. Finalmente, TFIIIB acarrea a la RNAP III para que se inicie la transcripción [59]. En los promotores de tipo 2 no se involucra TFIIIA, pues en este caso se lleva a cabo primero la unión de TFIIIC a las cajas A y B. Mientras que τ 95 interactúa con la caja A, la subunidad τ 138 reconoce a la caja B con una gran afinidad [60]. Posteriormente TFIIIC posiciona a TFIIIB en secuencias localizadas cerca del SIT para que sea reclutada la RNAPIII y se lleve a cabo la transcripción.

1.8.4. Factores de transcripción involucrados en la transcripción de los snRNAs

La transcripción de los snRNAs requiere de varios factores que se pueden dividir en aquellos que son requeridos por ambas polimerasas (RNAPs II y III) y los que son específicos de la transcripción llevada a cabo por cada una. Los factores compartidos son Oct1, Staf, SNAPc y TBP (Fig. 9). Oct1 y Staf funcionan como activadores de la transcripción al unirse al DSE [61,62]. SNAPc es un complejo formado por las subunidades: SNAP190, SNAP50, SNAP45, SNAP43 y SNAP19, y reconoce al PSE por medio de su subunidad SNAP190 [63-65] (Fig. 9).

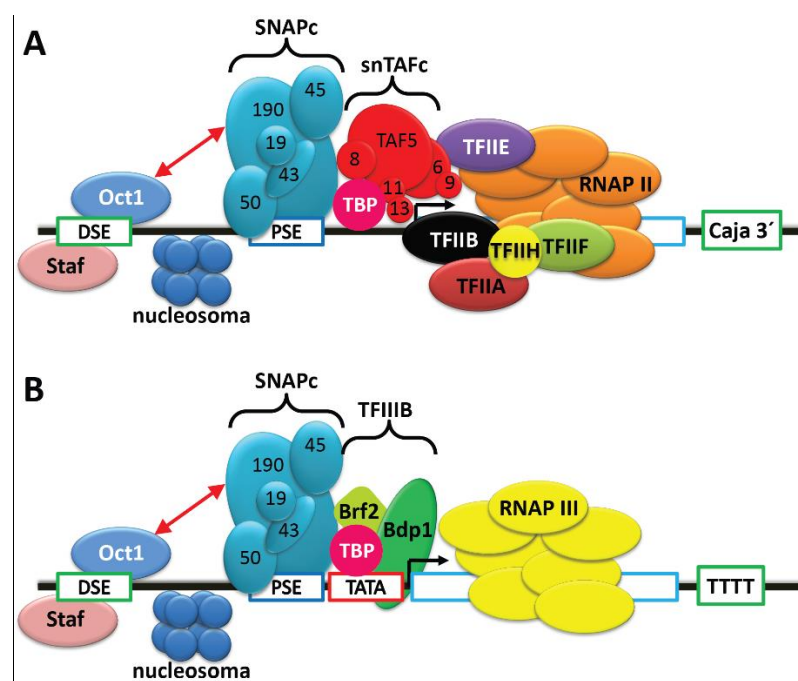


Figura 9. Factores involucrados en la transcripción de genes de snRNAs llevada a cabo por la RNAP II (panel A) o la RNAP III (panel B). La transcripción de los snRNAs involucra un conjunto de factores compartidos por los dos promotores: Oct1, Staf, SNAPc y TBP. Además se involucran factores especializados en la transcripción de cada polimerasa; para los snRNAs transcritos por la RNAP II estos factores son TFIIA, B, E, F, H y snTAFc. Los factores involucrados específicamente en la transcripción de la RNAP III incluyen a Bdp1 y Brf2, los cuales son parte del complejo TFIIB. La flecha roja representa las interacciones entre SNAP190 y Oct1. El término de la transcripción de la RNAP II es dirigido por la Caja 3', mientras que el término de la transcripción de la RNAP III es dirigido por un tracto de Ts. Modificado de [66].

Los factores específicos de la RNAP II involucrados en la transcripción de snRNAs son: TFIIA, TFIIB, TFIIE, TFIIF y TFIIH (Fig. 9, panel A) [67,68]. Generalmente la RNAP II necesita de TFIID para transcribir, el cual está formado por TBP y 13 proteínas TAFs [69]. De manera interesante, la RNAP II no necesita de TFIID para la transcripción de snRNAs, pues en su lugar es requerido un complejo

llamado snTAFc (complejo de TAFs específico de snRNAs) que está formado por TBP y únicamente las TAFs 5, 6, 8, 9, 11 y 13 (Fig. 9, panel A) [70]. La transcripción del snRNA U6 por la RNAP III necesita de manera específica a TFIIB y TFIIC [71-73]. El TFIIB que participa en la síntesis del snRNA U6 contiene la subunidad Brf2 (Fig. 9, panel B), a diferencia del TFIIB que recluta a la RNAP III en los genes de tRNAs, el cual contiene a Brf1 [72].

El proceso transcripcional de todos los snRNAs es parecido al principio, requiriéndose que el PSE y la caja TATA sean reconocidos por SNAPc y TBP, respectivamente [70]. Después de los eventos iniciales de reconocimiento del promotor, la ruta para el ensamblaje del complejo de pre-inicio de la transcripción de las RNAPs II y III es diferente. Para los promotores reconocidos por la RNAP II, el complejo formado por SNAPc y TBP recluta a snTAFc, quien a su vez recluta a los factores de transcripción generales TFIIA, B, E, F y H [67,68]. Finalmente la RNAP II reconoce al complejo formado en el promotor para iniciar la transcripción. En los promotores de snRNAs transcritos por la RNAP III, el complejo formado por SNAPc y TBP unido al promotor recluta a Brf2 [71,72,74,75]. Después, el complejo formado por SNAPc, TBP y Brf2 facilita el reclutamiento de Bdp1 [66,76]. Estos eventos conllevan al correcto posicionamiento de la RNAP III para que inicie la transcripción del snRNA U6. Los factores Oct1 y Staf aumentan en gran medida la transcripción a través de sus dominios de activación.

Varias proteínas reguladoras participan en el control de la expresión de los snRNAs. Entre éstas se encuentra el supresor de retinoblastoma (RB) y el supresor de tumores p53. RB ejerce su función reguladora uniéndose a las subunidades SNAP43 y SNAP50 de SNAPc [77], mientras que p53 interactúa con SNAP43 y SNAP190 [77,78]. Asimismo, se ha reportado que la proto-oncoproteína N-Myc activa globalmente la transcripción llevada a cabo por la RNAP III, pudiendo así regular la transcripción del snRNA U6 [79]. El factor Maf1 es otro regulador de la transcripción global llevada a cabo por las RNAP II y III [80,81]. Sin embargo, a pesar de que se ha probado su interacción directa con la RNAP III [82], no se ha demostrado si Maf1 tiene algún efecto en la transcripción de los snRNAs. Otros factores que pueden tener papeles reguladores en la formación del complejo de pre-inicio de la transcripción de

snRNAs son: HMGB1, el cual al unirse con TBP actúa como un represor transcripcional [83]; YY1, que se ha asociado con la activación y la represión de la transcripción [84]; β -actina, la cual está presente en el promotor de U6 transcripcionalmente activo [85]; y Sp1, proteína que al unirse a Oct1 actúa de manera sinérgica para estimular la transcripción de snRNAs [86].

Se sabe muy poco sobre el papel que juega la estructura de la cromatina en la regulación de la expresión de los snRNAs. En humanos, se ha demostrado la presencia de un nucleosoma posicionado entre el DSE y el PSE de los promotores de snRNAs transcritos por las RNAP II y III (Fig. 9) [87,88]. Este hecho facilita las interacciones entre SNAPc y Oct1 al acercar ambos elementos promotores. De manera similar, la unión de TFIIC y TFIIB al promotor de U6 en levadura conlleva al posicionamiento de un nucleosoma entre las cajas A y B ubicadas río abajo del SIT, así como río arriba de la caja TATA [89,90]. El posicionamiento de nucleosomas durante la transcripción activa de snRNAs sugiere que factores que modulan la estructura de la cromatina son importantes para la regulación de la transcripción de estos genes. De hecho, las acetil-transferasas de histonas p300 y CBP se han encontrado asociadas con promotores de snRNAs. Además, se ha observado una elevada proporción de la histona H3 acetilada en promotores de U6 en células que mantienen altos niveles de transcripción llevada a cabo por la RNAP III [66].

1.8.5. Procesamiento de los snRNAs

Durante la transcripción de los snRNAs se llevan a cabo dos procesos de maduración: la adición de una estructura *cap* a sus extremos 5' y el procesamiento de sus extremos 3'. Los snRNAs transcritos por la RNAP II, al igual que los mRNAs, adquieren co-transcripcionalmente una estructura *cap* monometilada m7GpppG (m7G) en su extremo 5' [91]. Posteriormente la estructura m7G se hipermetila a la forma 2,2,7-trimetil-guanosina (m3G) por la hipermetilasa del *cap* de snRNAs [92]. El snRNA U6, que es transcrito por la RNAP III, contiene una estructura *cap* diferente a la que se encuentra en los snRNAs transcritos por la RNAP II, llamada γ -mono-metil fosfato [93]. A diferencia del *cap* m7G, la

formación del *cap* del snRNA U6 es postranscripcional y depende de los primeros 25 nt del precursor del snRNA [94,95].

Debido a que la RNAP II continúa transcribiendo después de la región codificante, es necesario el procesamiento de los snRNAs primarios para generar su extremo 3' maduro (Fig. 10) [96]. Para ello, primero los residuos de serina 2, 5 y 7 del dominio carboxilo terminal (CTD) de la RNAP II, que está unida al promotor, son fosforilados por el factor elongador P-TEFb y la subunidad CDK7 del factor de transcripción TFIIH [97-99]. Posteriormente, el CTD fosforilado es reconocido por las enzimas que agregan el *cap* m⁷G y por un complejo multiproteico llamado Integrador [100]. Después, la RNAP II transcribe un elemento localizado de 9 a 19 pb río abajo de la región codificante, llamado Caja 3'. Este elemento es reconocido por el complejo integrador que corta al RNA en elongación, generando al extremo 3' maduro del snRNA [101]. El procesamiento conlleva al término de la transcripción, el cual puede ocurrir por mecanismos similares a los que causan el término de la transcripción de genes que codifican proteínas, como la desfosforilación del CTD y/o el ataque de exonucleasas como Xnr2 (Fig. 10) [90].

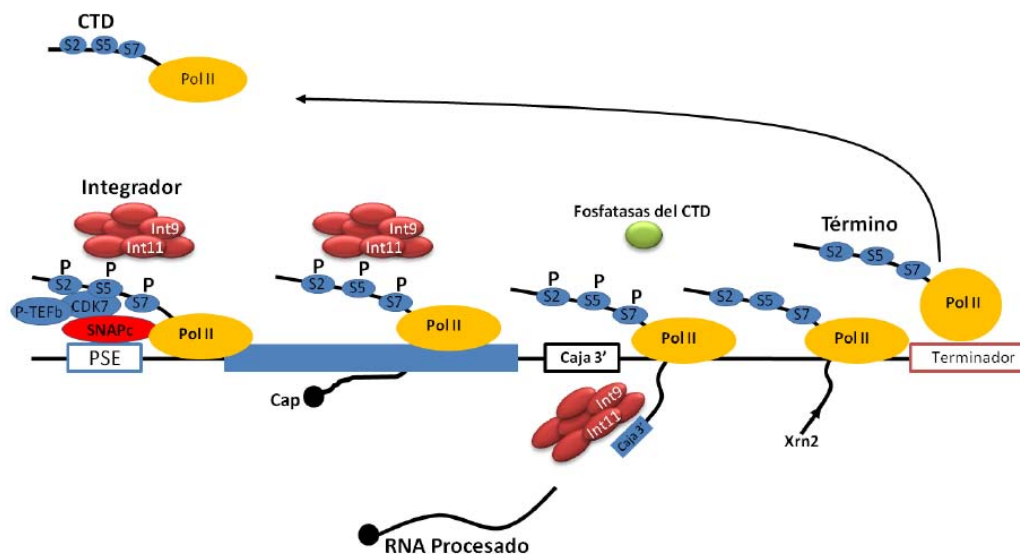


Figura 10. Formación del extremo 3' de snRNAs transcritos por la RNAP II. La RNAP II tiene fosforilados los residuos S2, S5 y S7 del CTD, gracias a P-TEFb y CDK7. El complejo integrador es reclutado al promotor al inicio de la transcripción y se asocia con la caja 3' después de que se transcribe para cortar al snRNA en el extremo 3'. La RNAP II se desfosforila para terminar la transcripción. Modificado de [96].

La transcripción por la RNAP III termina en un tracto de Ts localizado en el extremo 3' del gen del snRNA U6, lo que genera que el transcrito primario tenga un tracto de Us en el extremo 3'. En el humano, el snRNA U6 recién transcrito presenta cuatro Us en su extremo 3' que están unidos a la proteína "La" [102-105]. De manera interesante, este tracto no es estático pues está sujeto a la adición y remoción de Us por una transferasa terminal de uridinas específica de U6 (U6-TUT-ase) y una exoribonucleasa 3'→5' (USB1) [106-108]. Como resultado de los eventos de adición y remoción de Us, la mayoría de los snRNAs U6 terminan con cinco Us y un fosfato cíclico en su extremo 3' [109,110]. Esta modificación está involucrada en el reciclaje de los snRNAs U6 después de que éstos cumplen con su papel en el proceso denominado *splicing* [109].

1.9. Proceso de *splicing* y el empalmosoma

Después de la transcripción llevada a cabo por la RNAP II, los pre-mRNAs recién transcritos contienen secuencias no codificantes llamadas intrones, los cuales tienen que ser removidos antes de que los mRNAs puedan ser usados en la traducción. Al proceso de remoción de intrones y unión de exones o secuencias codificantes, se le llama *splicing*. Este proceso se lleva a cabo de manera coordinada por una elaborada maquinaria celular a la que se le ha nombrado empalmosoma. Los snRNAs U1, U2, U4, U5 y U6 y alrededor de 200 proteínas interactúan para formar dicho complejo [111,112]. Dentro de estas proteínas se encuentra un heptámero con forma de anillo compuesto por las proteínas Sm B, D3, D1, D2, F, E y G, las cuales reconocen una región de cadena sencilla conservada en los snRNAs U1, U2, U4 y U5 llamada sitio Sm [113]. Por otro lado, el snRNA U6 interactúa con un heptámero similar formado por las proteínas LSm. A diferencia de las proteínas Sm, las proteínas LSm reconocen un tracto de Us localizado en el extremo 3' del snRNA U6 [114]. El fosfato cíclico en el extremo 3' de U6 es necesario para la interacción con las proteínas LSm [115]. Existen también otras proteínas que reconocen a los diferentes snRNAs de manera específica. Por ejemplo, el snRNA U2 interactúa con las proteínas específicas U2A', U2B'', SF3a y SF3b [116]. A los complejos formados por las proteínas Sm o LSm y

los snRNAs se les denomina ribonucleoproteínas pequeñas nucleares (snRNPs). Dentro del empalmosoma, las snRNPs y el pre-mRNA cooperan dinámicamente para quitar a los intrones y unir a los exones de los pre-mRNAs [117].

En el *splicing*, regiones conservadas de los pre-mRNAs son reconocidas por las diferentes snRNPs. Estas regiones comprenden el sitio de empalme 5' (5'SS), el sitio de empalme 3' (3'SS), y la secuencia del punto de ramificación (BP). El ensamblaje del empalmosoma comienza con la formación del complejo A por medio de la unión de la snRNP U1 al 5'SS y de la snRNP U2 al sitio 3'SS (Fig. 11). Posteriormente el complejo formado por las snRNPs U4/U6•U5 se asocia con el complejo A, formando a su vez al complejo B. Luego el complejo NTC, que contiene a la proteína Prp19, se asocia establemente con el complejo B y causa la disociación de la proteína específica de U2 SF3b, y de las snRNPs U4 y U1, así como el apareamiento por pares de bases de U6 con U2, de U6 con el 5'SS (antes ocupado por U1) y de U5 con los nt del exón adyacentes al 5'SS. Estos últimos eventos activan al empalmosoma para que se lleven a cabo dos reacciones de transesterificación o reacciones de transferencia de fosfato, que median la escisión del intrón y la unión de los exones del pre-mRNA. Al complejo B activado se le ha llamado B^{act}. En la primera reacción, el grupo hidroxilo de un residuo de adenosina ubicado en el sitio de ramificación (al cual está unido U2), lleva a cabo un ataque nucleofílico a la unión fosfodiéster que se encuentra entre el último nt del exón 5' y el primer nt del intrón. Esto produce dos RNAs intermedios en la reacción: el exón 5' (unido al snRNA U5) que tiene un 3'-hidroxilo libre, y el intrón unido aún al exón río abajo. En este segundo RNA, el nt del extremo 5' del intrón está unido en una estructura de "lazo" al residuo de adenosina del sitio de ramificación. Al empalmosoma que contiene a ambos RNAs y ha realizado la primera reacción del *splicing*, se le denomina complejo C. En el segundo paso de la reacción, el 3'-hidroxilo del exón río arriba ataca al 3'SS en otra transesterificación nucleofílica, lo cual lleva a la ligación de los exones a través de un enlace fosfodiéster y a la formación del complejo P. En seguida, los exones ligados se liberan del complejo P dejando al complejo NTC unido a los snRNPs U6, U5 y U2 junto con el intrón en forma de lazo, a este último se le nombra complejo ILS. Finalmente,

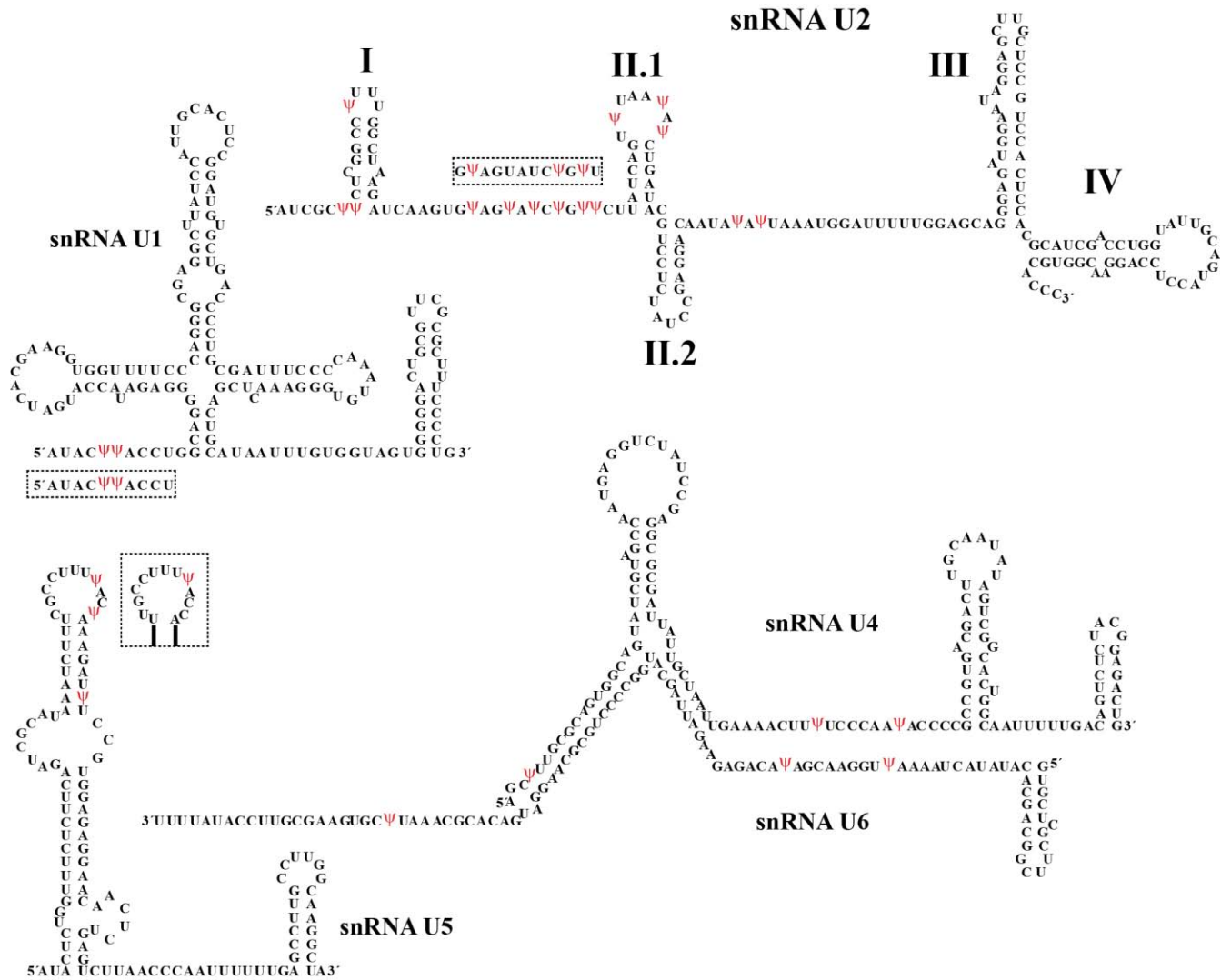


Figura 12. Secuencia, estructura secundaria y pseudouridinas de los snRNAs del empalmosoma. Se muestra en el esquema la secuencia y estructura secundaria de los snRNAs U1-U6 de humano. Los residuos pseudouridilados se muestran con un símbolo Ψ en rojo dentro de las secuencias. En rectángulos punteados se muestran los residuos de Ψ reportados en los snRNAs ortólogos de levadura. Modificado de [121]. Los tallos-asa I, II.1, II.2, III y IV se indican en el snRNA U2.

De manera interesante, todos los snRNAs del empalmosoma se modifican postranscripcionalmente. Existen dos tipos principales de modificaciones en los snRNAs: la pseudouridilación y la 2'-O-metilación, siendo la primera la más abundante (Fig. 12) [122]. La pseudouridilación es una modificación específica de los residuos de uracilo que resulta en el isómero de uridina 5'-ribosilado llamado pseudouridina (Ψ). Por otro lado, la 2'-O-metilación (2'-Ome) es una modificación en la que se agrega un grupo metilo al oxígeno de la posición 2' de la ribosa en los snRNAs. Esta modificación, a diferencia de la pseudouridilación, ocurre en las cuatro bases de los

snRNAs. La mayoría de los residuos modificados se encuentran en regiones funcionalmente importantes para el *splicing*, de hecho varios estudios han demostrado que las Ψ juegan un papel importante tanto en el *splicing* como en el ensamblaje de las snRNPs [121,123]. Las pseudouridilaciones pueden ocurrir por dos vías, una depende de RNA y la otra es independiente de RNA. En la primera se involucran RNAs pequeños no codificantes llamados RNAs tipo caja H/ACA que se unen a las proteínas Cbf5p, Nhp2p, Nop10p y Gar1p. El elemento con actividad catalítica es la proteína Cbf5p [124,125]. En la formación de Ψ independiente de RNA, las proteínas sintetasas de Ψ, como Pus1p y Pus7, son las encargadas de realizar la isomerización de la uridina [126,127].

1.11. Expresión genética en tripanosomátidos

La expresión genética en tripanosomátidos es, en diversos aspectos, diferente a los paradigmas establecidos por el estudio de la expresión genética de eucariontes superiores. Dentro de estas variantes encontramos la transcripción policistrónica de genes que codifican para proteínas que a su vez ocasiona que el número de promotores reconocidos por la RNAP II sea muy reducido, el procesamiento de pre-mRNAs largos por *trans-splicing*, la síntesis de tRNAs mitocondriales en el núcleo, la edición de mRNAs mitocondriales y la dependencia de la transcripción de genes de snRNAs en genes de tRNAs localizados río arriba [128,129]. A continuación revisaremos algunos aspectos de la expresión genética en *Leishmania* y otros tripanosomátidos.

1.11.1. Características del genoma nuclear de *Leishmania*

Los genomas de *L. major* y siete especies más de *Leishmania* se han secuenciado completamente hasta la fecha [130-136]. Estos estudios muestran que el tamaño del genoma nuclear de *Leishmania* es relativamente pequeño, pues mide alrededor de 32 megabases (Mb). Se considera que todas las especies de *Leishmania* del Viejo Mundo tienen 36 pares de cromosomas [137], mientras que las especies del subgénero *Viannia* tienen 35 cromosomas, ya que en estas especies los cromosomas 20 y 34 están

fusionados. Las especies del complejo de *L. mexicana* presentan 34 cromosomas, pues en este caso los cromosomas 8 y 29, así como 20 y 36, están fusionados [138]. El tamaño de los diferentes cromosomas de *Leishmania* oscila entre ~0.2 a ~3 Mb por lo que se consideran de un tamaño pequeño. La secuenciación de los genomas de *Leishmania* ha permitido determinar que en general, el orden genético y la secuencia genómica se encuentran altamente conservados en este género [139].

El genoma de *L. major* Friedlin tiene una longitud de 32.8 Mb, presenta una densidad de 260 genes por cada Mb y un contenido de G+C del 59.7%. El número total de genes predichos en los 36 cromosomas de *L. major* es de 8298, de los cuales 911 corresponden a genes de RNA no codificantes que engloban a los genes de snRNAs U1-U6, encontrados generalmente asociados con grupos de genes de tRNAs [133]. Los cromosomas en *Leishmania* y otros tripanosomátidos se caracterizan por el arreglo de sus genes codificadores de proteínas en unidades policistrónicas llamadas grupos de genes direccionados o DGCs, los cuales están compuestos por decenas a cientos de genes localizados en una misma hebra de DNA [140]. Los genes de proteínas dentro de los DGCs no están relacionados funcionalmente entre sí, a diferencia de los operones bacterianos. El genoma de *L. major* está organizado en 133 unidades policistrónicas, que pueden tener un tamaño de hasta 1259 kb y están separados por regiones de cambio convergentes o divergentes que van de 0.9 a 14 kb [133]. La RNAP II inicia la transcripción de manera bidireccional en las regiones de cambio divergentes [141] y termina dentro de las regiones de cambio convergentes, como en el caso del cromosoma 3 [142]. Esto sugiere que los sitios en donde la RNAP II inicia su transcripción son mucho menores que en eucariontes superiores, en donde la RNAP II inicia la transcripción en cada gen de mRNA.

1.11.2. Maduración de pre-mRNAs en tripanosomátidos

En tripanosomátidos, la gran mayoría de los mRNAs no son procesados por medio del *splicing* convencional. De hecho, hasta ahora sólo se han identificado dos genes de proteínas que contienen secuencias intrónicas homólogas a las de eucariontes superiores. Los genes dentro de las unidades

policistrónicas son transcritos juntos por la RNAP II, generando transcritos primarios largos que son procesados por medio de un mecanismo llamado *trans-splicing*, el cual está acoplado a la poliadenilación de los mRNAs [143]. En el *trans-splicing*, una secuencia de RNA de 39 nt de longitud, llamada mini exón o RNA del *spliced leader* (SL RNA), es agregada al extremo 5' de todos los mRNAs (Fig. 13) [144]. El precursor del SL RNA contiene una estructura *cap* en su extremo 5' y está formado por el mini exón (las 39 bases del extremo 5') y por un intrón de alrededor de 100 bases. Se ha estimado que existen entre 250 y 500 copias del gen del SL RNA por célula.

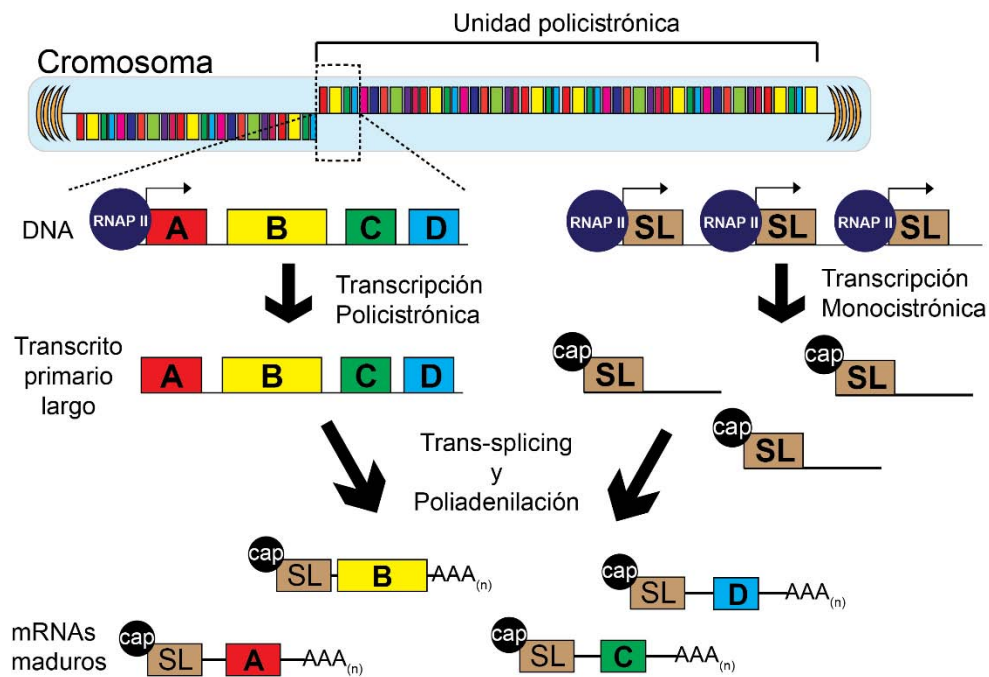


Figura 13. Producción de mRNAs maduros en tripanosomátidos. Se esquematiza en la parte superior de la figura un cromosoma conteniendo dos unidades policistrónicas. Se muestra con líneas punteadas una ampliación de un cistrón con los genes de proteínas A, B, C y D. La transcripción policistrónica de estos genes es llevada a cabo por la RNAP II. En la parte superior derecha de la figura se ejemplifica la transcripción monocistrónica del gen del SL RNA. El *cap* se ilustra con una esfera negra en el extremo 5' del exón. AAA(n) indica la cola de poli-A presente en el extremo 3' de los mRNAs maduros.

Al igual que en el *cis-splicing* o *splicing* convencional, en el *trans-splicing* ocurren dos reacciones de transesterificación. Sin embargo, en el *trans-splicing* se libera una estructura en forma de Y, en lugar de la estructura en forma de lazo del *cis-splicing* [143]. De forma similar al *cis-splicing*, el *trans-splicing* es llevado a cabo de manera coordinada por varios factores proteicos y los snRNAs U1, U2, U4, U5 y U6 (Fig. 14) [145].

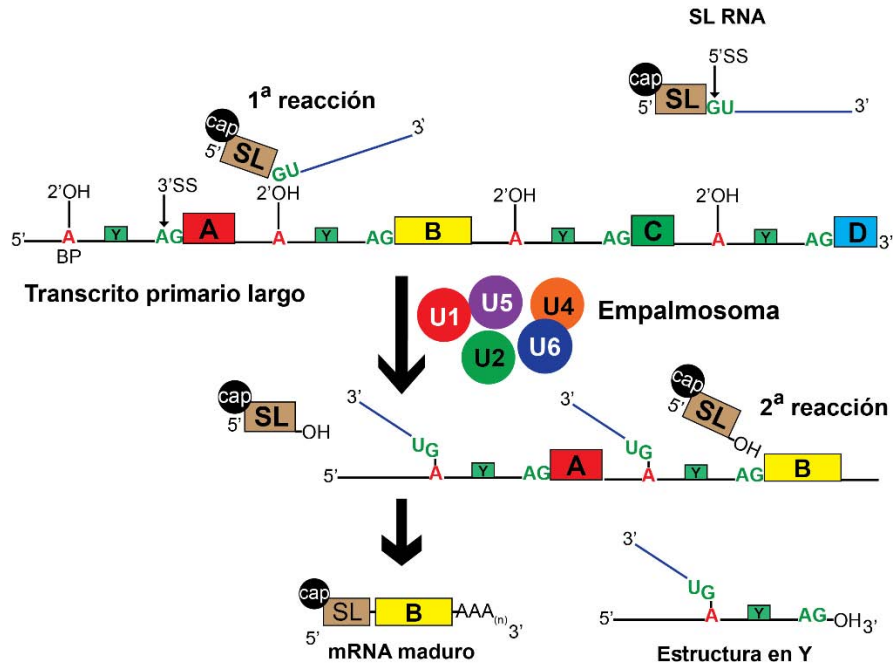


Figura 14. Trans-splicing. En la parte superior se esquematiza un transcrito primario largo con los genes A, B, C y D. En este transcrito se muestran los 3'SS y BPs para algunos de los genes. En la parte superior derecha se ilustra un transcrito del SL RNA, el cual contiene el 5'SS. Los círculos marcados como U1-U6 ejemplifican a las snRNPs del empalmosoma. Las cajas verdes marcadas con una Y representan las regiones ricas en pirimidinas.

A pesar de que todos los snRNAs tienen una participación esencial en el proceso de *splicing*, se ha demostrado que los snRNAs U2 y U6 juegan papeles clave en el desarrollo de este proceso, debido a que al unirse por apareamiento de bases forman el núcleo catalítico que lleva a cabo las dos reacciones de transesterificación [146].

2. Antecedentes

Otra peculiaridad interesante de la expresión genética en tripanosomátidos es el hecho de que todos los snRNAs que participan en el *trans-splicing* son transcritos por la RNAP III y no por dos diferentes RNAPs como en eucariontes superiores (Fig. 15).

No obstante, el promotor de clase III típico del gen del snRNA U6 de eucariontes superiores no está presente en los tripanosomátidos [147]. En contraste, todos los snRNAs en *T. brucei* presentan genes de tRNAs o secuencias parecidas a tRNAs (tRNAs-like) en su secuencia flanqueante 5', orientados en dirección opuesta a los snRNAs.

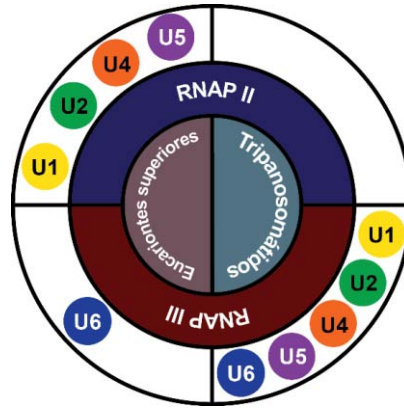


Figura 15. Los snRNAs son transcritos por diferentes RNAPs en eucariotes superiores y tripanosomátidos. Se ilustra el hecho de que todos snRNAs son transcritos por la RNAP III en tripanosomátidos mientras que en eucariotes superiores sólo el snRNA U6 es transcrito por esta RNAP y los demás snRNAs (U1-U5) son transcritos por la RNAP II.

De manera interesante, se ha demostrado que las cajas A y B presentes en estos tRNAs o tRNAs-like son esenciales para la transcripción de los snRNAs (Fig. 16) [147-149]. Notablemente, la distancia entre la caja A y el SIT de los snRNAs está muy conservada, pues siempre es de alrededor de 104 pb [150,151]. Adicionalmente, un elemento intragénico localizado en las primeras 21 pb de la región codificante de los snRNAs también es necesario para la expresión de estos genes [147,152]. En resumen, el promotor de snRNAs encontrado en *T. brucei* difiere en gran medida del presente en los genes de snRNAs de eucariotes superiores (Fig. 16).

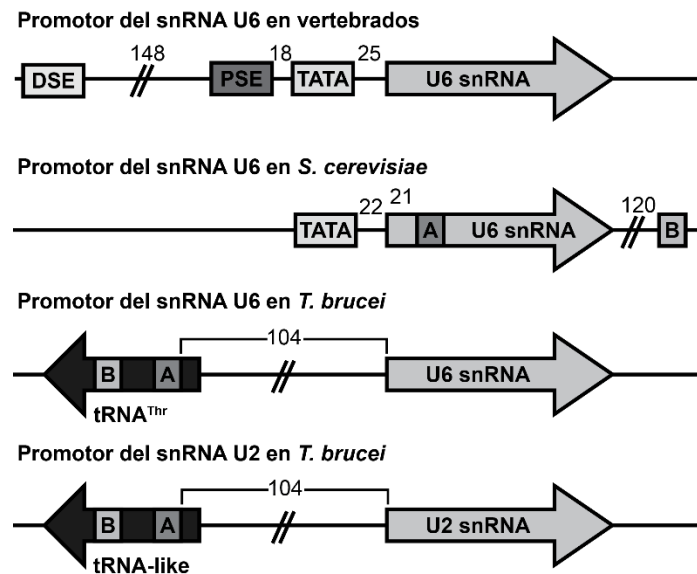


Figura 16. Los promotores de snRNAs de *T. brucei* difieren de los de eucariotes superiores. Se representan los promotores del snRNA U6 en vertebrados y *S. cerevisiae*, además de los promotores de los snRNAs U2 y U6 de *T. brucei*. Los elementos promotores del snRNA U6 se encuentran en un tRNA-Thr mientras que los del snRNA U2 están en un tRNA-like. El dibujo no está a escala.

Estos elementos son considerados como el promotor núcleo que dirige la transcripción de snRNAs en *T. brucei*. Sin embargo, no todos los genes de snRNAs necesitan de los tres elementos. Por ejemplo, la caja B no es requerida para la transcripción *in vitro* del snRNA U6 [152]. Además, la transcripción *in vivo* del snRNA U1 no requiere del elemento intragénico [149]. Por otra parte, la transcripción del snRNA U4 en el tripanosomátido *Leptomonas collosoma* no necesita de la caja A del gen del tRNA asociado [153].

Se ha comprobado que la distancia que separa a la caja A del snRNA es indispensable para la transcripción del gen, ya que el desplazamiento de dicho elemento 6 pb río arriba del SIT impide la expresión del snRNA U6 en *T. brucei* [152]. La locación de la caja B es menos esencial, pues sólo cuando este elemento se mueve 30 pb río abajo del SIT la expresión del snRNA U6 se ve drásticamente reducida [148,152].

3. Justificación de la investigación

A pesar de que algunos esfuerzos se han enfocado en el estudio de los snRNAs en diversos organismos, en *Leishmania* se sabe muy poco acerca de la expresión y modificaciones postranscripcionales de estas moléculas. De hecho, aún se desconocen los mecanismos moleculares que dirigen la transcripción de snRNAs en todos los tripanosomátidos. De manera más relevante, en *Leishmania* aún no se ha caracterizado ningún promotor que sea reconocido por la RNAP III. Debido a la importancia del snRNA U2 en el metabolismo celular de estos parásitos, nosotros estamos interesados en analizar funcionalmente la región promotora de dicho gen en *L. major*, así como en determinar la presencia y localización de Ψ en el snRNA U2 de este parásito. Esto aportará información importante que incrementará nuestro conocimiento sobre los procesos transcripcionales en *Leishmania*. Además, nos permitirá entender mejor el muy interesante proceso de *trans-splicing*.

4. Objetivos

Objetivo General:

Analizar funcionalmente la región promotora del gen del RNA pequeño nuclear U2 en el protozoo parásito *Leishmania major*.

Objetivos Particulares:

1. Identificar las secuencias importantes para la producción del snRNA U2, analizando *in vivo* la abundancia de transcritos producidos por vectores que contengan al gen del snRNA U2 y deleciones secuenciales de su región flanqueante 5'.
2. Localizar los nucleótidos esenciales para la síntesis del snRNA U2 mediante mutagénesis dirigida.
3. Determinar los sitios de inicio y de término de la transcripción del snRNA U2 mediante ensayos RACE y RT-PCR.
4. Determinar posiciones pseudouridiladas en el snRNA U2 por medio del tratamiento de RNA con el reactivo CMCT.
5. Analizar la expresión del tRNA-like asociado al snRNA U2 mediante ensayos tipo Northern blot y *primer extension*.

5. Materiales y métodos

5.1. Análisis *in silico*

Las secuencias genómicas de diferentes tripanosomátidos, incluyendo a *L. major*, se obtuvieron de bases de datos como NCBI (<http://www.ncbi.nlm.nih.gov/>) y TriTrypDB (Release 27) (<http://tritrypdb.org/tritrypdb/>). Para analizar las secuencias de las diferentes clonas generadas en este trabajo, así como para la identificación de los elementos localizados río arriba del gen del snRNA U2 y la construcción de mapas de sintenia, se utilizaron los programas bioinformáticos MUSCLE (<http://www.ebi.ac.uk/Tools/msa/muscle/>) y DNAMAN (versión 7). La búsqueda de genes de tRNAs se realizó usando el programa tRNAscan-SE (<http://lowelab.ucsc.edu/tRNAscan-SE>). Para obtener la estructura secundaria de las moléculas de RNA se usaron los programas RNAfold web Server (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>) y RNAstructure 5.0 de Mathews Lab. Las representaciones gráficas de patrones en alineamientos múltiples de secuencias se generaron usando el programa WebLogo application (<http://weblogo.threeplusone.com/>).

5.2. Cultivo de *L. major*

Promastigotes procíclicos de *Leishmania major* MHOM/IL/81/Friedlin fueron crecidos a 27 °C en medio BM 1× (infusión de cerebro y corazón 0.25×, M199 0.5×, suero fetal bovino 10%, hepes 40 mM pH 7.5, hemina 0.01 mg/ml, biotina 0.02%, biopterina 1.25 µg/ml, penicilina [100 u/ml]/estreptomicina [0.1 mg/ml]), hasta que éstos alcanzaron la fase media de crecimiento logarítmico. Los cultivos fueron analizados diariamente para medir la concentración de células. Para esto se tomaron 10 µl del cultivo diluido 90 µl de paraformaldehído al 2% (disuelto en PBS (NaCl 137 mM, KCl 2.7 mM, Na₂HPO₄ 10 mM, KH₂PO₄ 2 mM)) y los parásitos se contaron en una cámara de Neubauer.

5.3. Extracción de RNA total

Para extraer RNA total de las células de *L. major*, se centrifugó la cantidad deseada de células a 3000 g durante 10 min a 4°C, y si eran cultivos transfectados transitoriamente las células se lavaron 3

veces con 10 ml de PBS-G (PBS pH 7.2, con glucosa 6 mM). Posteriormente las células fueron lisadas adicionando 750 μ l de reactivo comercial TRIZOL-LS (Invitrogen) por cada 5×10^7 células. Las muestras homogenizadas se incubaron durante 5 min a temperatura ambiente (si se iba a extraer el RNA en ese momento) o a -70°C (el tiempo deseado) si se deseaba detener el experimento para continuarse posteriormente. Después se adicionaron 200 μ l de cloroformo y se agitó vigorosamente. La mezcla se incubó a temperatura ambiente durante 10 min para después centrifugarse a 13000 rpm durante 15 min a 4°C , luego se tomó la fase acuosa y el RNA se precipitó agregándose 500 μ l de alcohol isopropílico. La mezcla se incubó durante 10 min a temperatura ambiente y el precipitado se centrifugó a 13000 rpm durante 10 min. El RNA se lavó con 1 ml de etanol al 75% diluido con agua tratada con dietil pirocarbonato (DEPC) al 0.01% (Invitrogen). El precipitado final se resuspendió en el volumen deseado (dependiendo del número usado de células y de la concentración final deseada) de agua libre de RNAsas y se mantuvo a -70°C . Para determinar la concentración del RNA, se cuantificó en un espectrofotómetro (EPOCH, Biotek) a 260 nm.

5.4. Tratamiento de RNA con DNAsa I

Para eliminar DNA contaminante de las muestras de RNA, 5 μ g de RNA total se trataron con DNAsa I recombinante (Thermo Scientific). Primero se llevó el RNA a 40 μ l de volumen final con agua libre de RNAsas, posteriormente se adicionaron 5 μ l de buffer de reacción $10\times$ de la DNAsa I y 5 μ l de DNAsa I (1 U/ μ l, Thermo Scientific). Todo se mezcló y se incubó a 37°C por 1 h, y en seguida se agregaron otros 5 μ l de DNAsa I, se mezcló y la reacción se dejó nuevamente a 37°C por 1 h más. La reacción se detuvo haciendo una extracción con fenol: cloroformo: alcohol isoamílico (25:24:1). Para esto, la muestra se llevó a 100 μ l de volumen final con agua libre de RNAsas, luego se agregó 1 volumen de fenol: cloroformo: alcohol isoamílico (25:24:1) y se mezcló con vortex por 1 min. La mezcla se centrifugó a 13000 rpm por 2 min y la fase acuosa se recuperó en un tubo nuevo. Finalmente el RNA se precipitó con 0.1 volúmenes de acetato de sodio 3 M, pH 5.0 y 2.5 volúmenes de etanol absoluto, durante toda la noche a -20°C . Al otro día se centrifugó a 13000 rpm durante 15 min a 4°C , el sobrenadante se

retiró cuidadosamente y el RNA se lavó con 1 ml de etanol al 75% diluido con agua tratada con DEPC al 0.01% (Invitrogen). Se dejó secar la pastilla, el RNA se resuspendió en el volumen deseado (generalmente 50 μ l) de agua tratada con DEPC. Para verificar que no quedaran trazas de DNA contaminante, una alícuota del RNA tratado con DNAsa I se usó como molde en reacciones de PCR.

5.5. Ensayo 5'-RACE

El 5'-RACE es una técnica que facilita el aislamiento y caracterización de los extremos 5' de RNAs de bajo número de copias. Primero se sintetiza cDNA usando un oligonucleótido gen-específico (GSP1) y una transcriptasa reversa. Seguido a la síntesis del cDNA, el RNA es degradado y posteriormente el producto de la primera reacción se purifica de los dNTPs no incorporados y del GSP1. Luego, una deoxinucleotidil transferasa terminal (TdT) es usada para agregar colas homopoliméricas a los extremos 3' del cDNA. Como paso final se lleva a cabo un PCR con un oligonucleótido gen-específico anidado (GSP2), el cual se alinea del lado 3' del GSP1, y un oligonucleótido de anclaje que se une a la cola homopolimérica. Esto permite la amplificación de secuencias desconocidas entre el GSP2 y el extremo 5' del RNA, incluido el SIT. Para realizar esta técnica se utilizó el kit 5'-RACE versión 2.0 de Invitrogen™:

Síntesis de cDNA: A 5 μ g de RNA tratado con DNAsa I se adicionaron 2.5 pmol del oligonucleótido 3'U2GST1 (5'-AACGTGGAAGTCCAAGGAAA) y agua DEPC hasta alcanzar un volumen final de 15.5 μ l. Posteriormente se incubó a 80°C por 10 min para desnaturalizar al RNA; en seguida se pasó la reacción a hielo por 1 min y entonces se adicionaron 2.5 μ l del buffer PCR 10 \times (200 mM Tris-HCl pH 8.4, 500 mM KCl), 2.5 μ l de MgCl₂ 25 mM, 1 μ l de mezcla de desoxinucleósidos trifosfato (dNTPs) 10 mM y 2.5 μ l de DTT 0.1 M. Después de mezclar se incubó a 42°C por 1 min y se adicionó 1 μ l de SuperScript II RT 200 U/ μ l. Luego se incubó a 42°C por 30 min y a 70°C por 15 min para terminar la reacción. Para degradar las trazas de RNA se agregó 1 μ l de RNAsa mix y se incubó por 30 min a 37°C.

Purificación de cDNA: En este paso se usaron las columnas YM-30 de Microcon (Millipore). Para ello se hizo lo siguiente: primero se insertó el reservorio de la muestra en el vial, luego se depositó la muestra en el reservorio y se tapó, se centrifugó a 13000 g por 12 min; entonces el reservorio se pasó a un nuevo tubo poniéndolo al revés, se centrifugó nuevamente por 3 min a 1000 g para transferir el concentrado al nuevo vial.

Adición de colas homopoliméricas al extremo 3' del cDNA: En un tubo nuevo se agregaron 6.5 µl de agua DEPC, 5 µl de *tailing* buffer 5×, 2.5 µl de dCTP 2 mM y 10 µl del cDNA purificado, esta mezcla se incubó 3 min a 94°C y luego se enfrió en hielo por 1 min. Entonces se agregó 1 µl de la enzima TdT, se mezcló e incubó por 10 min a 37°C. La enzima se inactivó por calor, incubando a 65°C por 10 min.

Amplificación del cDNA: En un tubo de PCR se pusieron 26.5 µl de agua bidestilada estéril, 5 µl de buffer PCR 10×, 6 µl de MgCl₂ 25 mM, 1 µl de dNTPs mix 10 mM, 2 µl del oligonucleótido 3'U2GST2 (5'-CCTTGAGTTGTAATTTCTAT) 10 µM, 2 µl del oligonucleótido AAP (5'-GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG) 10 µM, 5 µl del cDNA con colas homopoliméricas y 2.5 µl de DNA polimerasa Red Taq (1U/ µl). Los productos del PCR se desalaron con el kit NucleoSpin Gel and PCR Clean-up (Macherey-Nagel), y se ligaron al vector pGEM-T (Promega) para su clonación y secuenciación.

5.6. Secuenciación automatizada de plásmidos

Todos los plásmidos obtenidos en este trabajo se enviaron a secuenciar al laboratorio de biología molecular de la Unidad de Biotecnología y Prototipos de la Facultad de Estudios Superiores Iztacala, UNAM. Para esta secuenciación el método empleado fue el de “BIG DYE Terminador Fluorescent Bases Sequencing” y se usó el equipo ABI 3100.

5.7. Mapeo de extremos 3' por RT-PCR

Los sitios de término de la transcripción del snRNA U2 fueron mapeados por RT-PCR en el que se usó RNA total al que se adicionaron colas de poli(A) en sus extremos 3'. Para esto se mezclaron 2 µg de RNA total, 1 µl de ATP 25 mM, 2 µl de buffer de reacción de la Poli (A) Polimerasa 5× (USB) y 1200 U de la Poli (A) Polimerasa de Levadura (USB), la reacción se llevó a un volumen final de 20 µl con agua libre de RNAsas. La mezcla se incubó a 37 °C por 20 min y se inactivó calentándola a 65 °C por 10 min. En seguida se realizó una síntesis de cDNA con el oligonucleótido Nested(dT) (5'-CCTCTGAAGGTTTCACGGATCCACATCTAGATTTTTTTTTTTTTTTTTTTTTVN), el cual se une a la cola de Poli (A) recién agregada. La reacción se hizo de la siguiente manera: La mitad del RNA poliadenilado (1 µg) se mezcló con 2 pmol del oligonucleótido Nested(dT) y 1 µl de una mezcla de dNTPs (dATP, dGTP, dCTP y dTTP 10 mM cada uno), luego se agregó agua tratada con DEPC a un volumen final de 13 µl. La mezcla se calentó a 65 °C por 5 min y se pasó a hielo por 1 min, después se agregó 4 µl de buffer First-Strand 5×, 1 µl de DTT 0.1 M, 1 µl de RiboLock RNase Inhibitor (40 U/µl, Thermo Scientific) y 1 µl de SuperScript™ III RT (200 U/µl). La reacción se llevó a cabo a 55 °C por 1 h y se inactivó calentando a 70 °C por 15 min. El cDNA se amplificó por medio de dos reacciones consecutivas de PCR realizadas de la siguiente manera: En el primer PCR se mezclaron 5 µl de buffer de reacción estándar adicionado con Mg²⁺ (10×), 1 µl de mezcla de dNTPs (10 mM), 1 µl del oligonucleótido U2Sa-5' (5'-ATATCTTCTCGGCTATTTAGC) (10 µM), 1 µl del oligonucleótido B1 (5'-CCTCTGAAGGTTTCACGGAT) (10 µM), 3 µl de cDNA y 1 U de la DNA polimerasa KAPATaq (KAPA BIOSYSTEMS), la reacción se llevó a 50 µl con agua y se mezcló. La temperatura de alineamiento usada en este PCR fue de 46°C. El segundo PCR se hizo esencialmente de la misma manera que el primero, lo único que se cambió fue que en este PCR se usaron los oligonucleótidos U2qFw2 (5'-CTATTTAGCTAAGATCATGTTTATA) y B2 (5'-CACGGATCCACATCTAGAT), y como molde se usaron 7.5 µl de la primer reacción de PCR; además la temperatura de alineamiento fue de 44 °C en esta reacción. El producto final de PCR se extrajo de un gel de agarosa y se purificó para su posterior

clonación en el vector pGEM-T Easy (Promega). Los sitios de término de la transcripción se determinaron mediante la secuenciación de varias clonas obtenidas en este ensayo.

5.8. Marcaje radioactivo de oligonucleótidos

Para marcar con ^{32}P los diferentes oligonucleótidos usados en este trabajo, se realizó la siguiente mezcla: 1 μl del oligonucleótido (200 ng/ μl), 5 μl de $\gamma\text{-}^{32}\text{P}\text{-ATP}$ (10 $\mu\text{Ci}/\mu\text{l}$, 7000 Ci/mmol), 3 μl de T4 cinasa (10 U/ μl), 1 μl de buffer 10 \times de la cinasa y agua bidestilada estéril hasta que se alcanzó una concentración final 1 \times del buffer. Se incubó la mezcla por 10 min a 37°C y en seguida se pasó a 90°C por 5 min para inactivar la cinasa. Para remover el $\gamma\text{-}^{32}\text{P}\text{-ATP}$ no incorporado al DNA se llevó a cabo una precipitación: la mezcla se llevó a un volumen de 50 μl con agua bidestilada estéril, luego se agregó 5 μl de glicógeno (10 mg/ml), 300 μl de acetato de amonio 5 M y 937.5 μl de etanol absoluto; la precipitación se dejó durante una noche a -20°C, luego se centrifugó a 13000 rpm por 15 min, se retiró el sobrenadante y se resuspendió el oligonucleótido en 20 μl de agua desionizada estéril. Se cuantificó 1 μl de la sonda en un contador de centelleo líquido (Wallac 1450 MicroBeta, Perkin Elmer) y el resto se almacenó a -20°C.

5.9. Ensayo de *primer extension*

Para hacer los diferentes ensayos de *primer extension*, de 20 a 50 μg de RNA total se mezclaron con 100,000 cpm de cualquiera de los siguientes oligonucleótidos específicos marcados radioactivamente con ^{32}P : 3'U2GST1 (5'-AACGTGGA ACTCCAAGGAAA), complementario a los nt 108 a 127 de U2; 3'endLmjU2 (5'-GAAAAAAGGAGTTGCTCCCCTGGAAACGTG), complementario a los nt 122 a 151 de U2; LmjU2tag-Rev (5'-ATTAGAGTCGAGGTCAGACC), complementario a los 15 nt de la secuencia *tag* insertada entre las posiciones 83 y 84 de U2; y 5'LmjU2DcajaB (5'-ATCCGGCCCTGGTCTCCAAA), complementario a los nt 58 a 77 del tRNA-like. Esta mezcla se precipitó con 0.1 volúmenes de acetato de sodio 3 M y 2.5 volúmenes de etanol, y la pastilla se resuspendió en 30 μl de solución de hibridación (formamida 80%, PIPES 40 mM pH 6.4, NaCl 0.4 M,

EDTA 1 mM pH 8.0) y se incubó a 30°C durante una noche. Posteriormente, la reacción se llevó a un volumen de 100 µl con agua tratada con DEPC y se volvió a precipitar con 0.1 volúmenes de acetato de sodio 3 M y 2.5 volúmenes de etanol. La pastilla se resuspendió en 12 µl de agua tratada con DEPC y se agregó 1 µl de mezcla de dNTPs (10 mM), la mezcla se incubó por 10 min a 80°C. Pasado este tiempo se enfrió en hielo por 2 min y se adicionó 4 µl del buffer de reacción de la transcriptasa reversa SuperScript III RT (5×), 1 µl de DTT 0.1 M, 1 µl de RNAsin (40 U/µl, Promega) y 1 µl de la SuperScript III RT (100 U/µl). Se mezcló cuidadosamente y se incubó durante 1 hr a 55 °C. La reacción se inactivó a 70°C por 15 min. Transcurrido este tiempo se adicionó 1 µl de RNAsa A (4 mg/ml) y se incubó a 37°C por 30 min. Hecho esto se llevó a un volumen de 100 µl con agua desionizada estéril, luego se agregaron 100 µl de fenol: cloroformo: alcohol isoamílico (25:24:1), se agitó en vortex y se centrifugó a 13000 rpm durante 2 min. La fase acuosa se transfirió a un tubo nuevo y se precipitó el DNA con 0.1 volúmenes de acetato de sodio 3 M pH 5.0 y 2.5 volúmenes de etanol absoluto. La pastilla se resuspendió en 20 µl de buffer TE (Tris-HCl 10 mM pH 7.6, EDTA 0.2 mM pH 8.0). Las muestras se analizaron en geles de poliacrilamida al 8% junto con reacciones de secuencia llevadas a cabo con el mismo oligonucleótido.

5.10. Mapeo de pseudouridinas

El ensayo de mapeo de Ψ se llevó a cabo como lo reportaron [154] y [155]. El RNA proveniente de 5×10^8 células se resuspendió en 50 µl de buffer CMCT (bicina 50 mM pH 8, urea 7 M, EDTA 4 mM) o en 50 µl de buffer con carbodiimida soluble (CMCT) (14 mg de CMCT en 200 µl de buffer CMCT). Entonces las muestras se incubaron a 37 °C por 20 min. La reacción se detuvo agregando 100 µl de acetato de sodio (300 mM, pH 5.2), 700 µl de etanol absoluto y 1 µl de glicógeno (2 µg/µl). El RNA se precipitó por 1 h a -70°C y se lavó con 500 µl de etanol al 70%. En seguida se retiró por completo todo el sobrenadante para evitar trazas de CMCT y las pastillas de las muestras que no sufrieron tratamiento alcalino se resuspendieron en 50 µl de agua tratada con DEPC. Para eliminar los grupos CMCT de los residuos de guanina y uridina se realizó un tratamiento alcalino resuspendiendo al RNA

en 80 μl de Na_2CO_3 (50 mM, pH 10.3) y dejando las muestras a 37 °C por 2h. Luego, el RNA se precipitó agregando 100 μl de acetato de sodio (300 mM, pH 5.2), 700 μl de etanol absoluto y 1 μl de glicógeno (2 $\mu\text{g}/\mu\text{l}$). El RNA se lavó con 500 μl de etanol al 70%. El RNA se resuspendió en 50 μl de agua DEPC. De 25 a 50 μg del RNA tratado se usaron como molde en ensayos de *primer extension* en los que se usó el oligonucleótido 3'endLmjU2. En este método, la transcriptasa reversa se para 1 nt antes de la posición pseudouridilada.

5.11. Secuenciación manual de DNA

La secuenciación nucleotídica involucra la síntesis *in vitro* de una cadena de DNA por una DNA polimerasa usando como molde una cadena sencilla de DNA. La síntesis de DNA se lleva a cabo en dos pasos: el primero es el de marcado, en el cual un oligonucleótido específico es extendido usando concentraciones limitadas de dNTPs que incluyen dATP marcado radioactivamente. Este paso continúa hasta la completa incorporación de los nucleósidos marcados en las cadenas de DNA. Estas extensiones iniciales del oligonucleótido tienen diferentes longitudes al azar y van de algunos pocos a cientos de nucleótidos. En el segundo paso, la concentración de todos los dNTPs es incrementada y un nucleósido de terminación análogo es agregado. Estos 2',3'-didesoxinucleósidos-5'-trifosfato (ddNTPs) carecen del grupo 3'-OH el cual es necesario para la elongación de DNA. La progresiva síntesis de DNA va dando extensiones de algunas docenas de nucleótidos hasta que todas las cadenas en crecimiento son terminadas con un ddNTP. Cuando se hacen mezclas apropiadas de dNTPs y uno de los cuatro ddNTPs, la polimerización terminará, en una fracción de la población de cadenas, en cada sitio donde el ddNTP pueda ser incorporado. Cuatro reacciones separadas, cada una con un diferente ddNTP, dan la información de la secuencia completa al correrse en un gel de poliacrilamida. Para la secuenciación se usó el kit “Sequenase 7-deaza-dGTP Sequencing kit” de USB, y se siguió el protocolo recomendado:

Reacciones de secuenciación: 10 μg de los plásmidos ptRNAU2 o pComp, que contienen al gen del snRNA U2 con o sin secuencia *tag*, respectivamente, y además su secuencia flanqueante 5' (con el

tRNA-like y el tRNA-ALA), se desnaturalizaron agregando 0.1 volúmenes de NaOH 2 M y EDTA 2 mM e incubando 30 min a 37°C. La mezcla se neutralizó con 0.1 volúmenes de acetato de sodio 3 M pH 5.0 y el DNA se precipitó con 2.5 volúmenes de etanol absoluto. El DNA se lavó con etanol al 70% y se resuspendió en 7 µl de agua, se agregaron 2 µl de buffer de reacción de la secuenasa 5× y 1 µl de alguno de los oligonucleótidos (3'U2GST1, 3'endLmjU2, LmjU2tag-Rev y 5'LmjU2DcajaB) (2 pmol/µl). Después, el oligo se alineó calentando la mezcla a 65°C por 2 min y se enfrió lentamente a temperatura ambiente en ~30 min, luego se pasó a hielo. Mientras se enfriaba la muestra se pusieron 2.5 µl de cada mezcla de terminación (80 µM 7-deaza-dGTP, 80 µM de los respectivos dNTPs, 8 µM del respectivo ddNTP, 50 mM NaCl) en tubos nuevos y se mantuvieron en hielo. Se diluyó la mezcla de marcado (1:15) con agua bidestilada estéril. Se precalentaron a 37°C los 4 tubos con las diferentes mezclas de reacción. Posteriormente, a la mezcla que contenía el DNA y el oligonucleótido se le agregó 1 µl de DTT 0.1 M, 2 µl de la mezcla de marcado 1×, 0.5 µl de [α -³²P]-dATP (3000 Ci/mmol) y 3.25 U de la DNA polimerasa Secuenasa version 2.0. Esta mezcla se incubó 5 min a temperatura ambiente y luego se transfirieron 3.5 µl a cada tubo con las mezclas de terminación; entonces se mezcló e incubó por 5 min a 37°C. Las reacciones se detuvieron agregando 4 µl de solución de paro, y se calentaron a 75°C por 2 min justo antes de cargarse en el gel de poliacrilamida.

5.12. Electroforesis en geles desnaturalizantes

Las reacciones de *primer extension* se corrieron en geles de poliacrilamida bajo condiciones desnaturalizantes y junto a escaleras de secuencia que se hicieron con los mismos oligonucleótidos usados en el *primer extension*. Para preparar el gel de poliacrilamida al 8% se preparó primero un stock de acrilamida/bis-acrilamida al 30% (19:1) como sigue: en agua desionizada se disolvieron 28.5 g de acrilamida y 1.5 g de bis-acrilamida, la mezcla se aforó a 100 ml con agua. Después se mezclaron 34.89 g de urea, 7 ml de TBE 10× y 18.67 ml del stock de poliacrilamida (30%), se agregó agua a un volumen final de 70 ml y se filtró la solución. En seguida, se montaron los vidrios de una cámara de electroforesis hecha para correr reacciones de secuencia (Sequi-Gen GT Nucleic Acid Electrophoresis Cell de Bio-

Rad). Entonces se procedió a polimerizar la mezcla de poliacrilamida filtrada agregando 520 µl de persulfato de amonio al 10% y 18.6 ml de TEMED. Inmediatamente después de que se agregaron estos reactivos se mezcló y vació en la cámara de secuenciación. Una vez que el gel se polimerizó, se realizó una precorrida a 65 W por 50 min usando TBE 1× (Tris base 89 mM; ácido bórico 89 mM; EDTA 2 mM) hasta que el gel alcanzó una temperatura de 50°C. Antes de cargar las muestras, éstas se calentaron en un baño a 75°C durante 5 min y se colocaron en hielo durante 1 min. La corrida de las muestras se realizó a 50 W. Finalmente el gel secó sobre papel Whatman y se envolvió en plástico adherente transparente. El gel se expuso en placas de un equipo tipo Phosphorimager (FLA-5000, Fujifilm), que se escanearon después de ~24 h de exposición.

5.13. Ensayos tipo Northern blot

De 10 a 30 µg de RNA total aislado con el reactivo TRI (Sigma) se usaron para hacer ensayos tipo Northern blot. El RNA se corrió en un gel de poliacrilamida al 10% con Urea 8 M en TBE 1× a 120 V. Después de la electroforesis, el RNA se transfirió a membranas Hybond N⁺ (Amersham) usando el equipo Trans-Blot Semy Dry System (Bio-Rad) a 25 V por 1 h. El RNA se fijó en la membrana con luz UV en una cámara UV Stratalinker 1800 (Stratagene). Las membranas se hibridaron con los siguientes oligonucleótidos marcados radioactivamente con ³²P: para identificar al tRNA-ALA se usó Ala-Rvs (5'-TCGATCCCAATACCTACCGC), para el tRNA-like fue usado 5'LmjU2DcajA, para el snRNA U2 endógeno se empleó 3'endLmjU2; la sonda que reconoció a la secuencia *tag* insertada en el snRNA U2 fue el oligonucleótido LmjU2tag-Rev. Las hibridaciones se hicieron en solución SSPE 6× (Na₂HPO₄ 60 mM, NaCl 0.9 M y EDTA 6 mM), solución de Denhardt 5× y SDS al 1% a 42 °C por 18 h con 5 × 10⁶ cpm de la sonda. Los lavados se realizaron con SSC 0.2×/SDS 0.1 % a 55 °C. Las membranas se expusieron en placas de un equipo tipo Phosphorimager (FLA-5000, Fujifilm), escaneándose en el aparato después de ~24 h de exposición. Composición de la solución de Denhardt 50×: ficoll 1%, polivinilpirrolidona (PVP) 1%, albúmina sérica bovina (BSA) 1%. Composición del SSC 1×: NaCl 3 M, acetato de sodio 0.3 M pH 7.0.

5.14. Extracción de DNA genómico

Para realizar la extracción de DNA genómico (gDNA) se centrifugaron $\sim 3 \times 10^8$ células a 3000 g, durante 10 min a 4 °C. Posteriormente la pastilla celular se resuspendió en 5 ml de la siguiente solución: 5 ml de SSC 1×, 4.5 ml de TNE (Tris 40 mM pH 7.6, EDTA 1 mM pH 8.0, NaCl 15 mM), 0.5 ml de sarcosyl al 10%. Después se adicionaron 100 µg de proteinasa K (Invitrogen), se agitó en vortex y se incubó a 55°C por 2 h. Posteriormente se adicionó un volumen de fenol:cloroformo:alcohol isoamílico (25:24:1), se agitó en vortex y se centrifugó a 1500 g durante 8 min a 4°C. La fase acuosa se transfirió a un tubo para microcentrífuga de 1.5 ml y se precipitó el DNA con 0.1 volúmenes de acetato de sodio 3 M pH 5.0 y 2.5 volúmenes de etanol absoluto. Después se centrifugó a 7000 g por 15 min a 4°C, se lavó la pastilla con etanol al 70 % y finalmente se resuspendió en buffer TE o agua desionizada estéril. La concentración del DNA fue estimada en un espectrofotómetro a 260 nm.

5.15. Amplificación por PCR

Diferentes fragmentos del locus del snRNA U2 (Comp, Δ tRNA-Ala, Δ Box-B, Δ Box-A y Δ -293/-1) fueron amplificados por PCR. La construcción ptRNAU2 sirvió como molde para la obtención del plásmido ptRNAU2M (ver sección de mutagénesis dirigida), que a su vez se usó como molde para amplificar los insertos de las construcciones que presentan regiones del promotor deletadas. Se diseñaron oligonucleótidos específicos complementarios a los extremos de la región que se deseaba amplificar para cada plásmido. El protocolo estándar utilizado para las reacciones de PCR se indica a continuación.

Mezcla de reacción: como molde se usó gDNA (para la construcción ptRNAU2), o los plásmidos ptRNAU2M (para la construcción pComp) o pComp (para las construcciones con regiones deletadas) a una concentración final de 2 ng/µl, buffer de amplificación 1× (Invitrogen), 0.2 mM de una mezcla de dNTPs, 1.5 mM de sulfato de magnesio (Invitrogen), los oligonucleótidos a 1 µM (Sigma o Invitrogen) y la polimerasa Easy A o *Pfu* o Herculasa II. El volumen final se ajustó a 50 µl con agua bidestilada estéril.

El protocolo que siguió el termociclador para la reacción de PCR consistió de siete pasos: una desnaturalización inicial de 2 min a 95 °C (paso 1), seguida por otra desnaturalización corta de 1 min a 95 °C (paso 2), renaturalización por 30 seg a 55-60 °C (paso 3), una extensión a 72 °C por 1 min (paso 4), repetir 29 veces los pasos del 2 al 3 (paso 5), una extensión final a 72 °C por 5 min (paso 6) y un enfriamiento a 4 °C (paso 7). La secuencia de los oligonucleótidos utilizados para cada construcción se muestra en la tabla 1.

Nombre del plásmido	Oligonucleótidos sentido	Oligonucleótidos antisentido
ptRNAU2	5' LmjtRNAala 5'-TGAAAAAGTTGGAGAAGTTG	3' U2Lmjend 5'-TCCCCTGGAAACGTGGAAC
pComp	LmjtRNAala-XmaI-F 5'-ACCCGGG TGAAAAAGTTGGAGAAGTTG	endLmjU2-XbaI-R 5'-ATCTAGAGAAAAAAGGA GTTGCTCCCCTGGAAACGTG
pΔtRNA-Ala	U2tRNA-like-XmaI-F 5'-ACCCGGG AGGAAAAGATGCTTTCGACGAG	endLmjU2-XbaI-R 5'-ATCTAGAGAAAAAAGGA GTTGCTCCCCTGGAAACGTG
pΔBox-B	LmjU2DcajaB-XmaI-F 5'-ACCCGGG ATCCCGCCCTGGTCTCCAAA	endLmjU2-XbaI-R 5'-ATCTAGAGAAAAAAGGA GTTGCTCCCCTGGAAACGTG
pΔBox-A	LmjU2DcajaA-XmaI-F 5'-ACCCGGG GTAAGCGGCACGGCAGTGAG	endLmjU2-XbaI-R 5'-ATCTAGAGAAAAAAGGA GTTGCTCCCCTGGAAACGTG
pΔ-293/-1	U2Sa-XmaI-F 5'-ACCCGGG ATATCTTCTCGGCTATTTAGC	endLmjU2-XbaI-R 5'-ATCTAGAGAAAAAAGGA GTTGCTCCCCTGGAAACGTG

Tabla 1. Oligonucleótidos usados para generar los plásmidos con deleciones de la región flanqueante 5' del snRNA U2 de *L. major*. En la tabla se muestran los nombres de los diferentes plásmidos con deleciones de elementos río arriba del snRNA U2, en la 2ª y 3ª columnas se muestran los oligonucleótidos utilizados para amplificar los insertos en cada una de las construcciones.

5.16. Electroforesis en geles de agarosa

Las electroforesis de DNA y RNA se llevaron a cabo en geles de agarosa (agarose D-1 low EEO, Pronadisa, Laboratorios Conda) al 0.8-2% con buffer TBE 1× preparado con agua tratada con DEPC (para el caso de geles de RNA) y 0.1 µg/ml de bromuro de etidio. El buffer de corrida usado fue TBE 1×. El voltaje al que fueron sometidos los geles para la separación de las bandas varió dependiendo de los resultados esperados. Si se deseaba la separación de algún inserto para su posterior purificación, el voltaje fue generalmente de 50 V. En cambio, si lo que se quería era únicamente visualizar un patrón de bandeo en algún análisis de restricción específico o si se quería determinar la cantidad y pureza de una muestra dada, el voltaje empleado fue generalmente mayor (75-100 V). De igual manera, el tiempo de corrida varió de 30 min a ~2 h. Todas las muestras fueron resuspendidas en buffer de carga 1× (azul de bromofenol 0.25%, xilen cianol 0.25%, glicerol 30%) y se corrieron paralelamente con un marcador de peso molecular (escaleras de 1 Kb o de 100 pb, Invitrogen).

5.17. Purificación de fragmentos de DNA

Para realizar la purificación de fragmentos de DNA de geles de agarosa se empleó el kit QIAEX II Gel Extraction (QIAGEN) o el NucleoSpin Extract II (Macherey-Nagel), siguiéndose las instrucciones y recomendaciones especificadas por los fabricantes. Para ello, inicialmente se cortó con un escalpelo limpio la banda que contenía el DNA de interés, procurando en todo momento limitar al mínimo la cantidad de agarosa que se tomaba. El DNA contenido en el fragmento de agarosa se purificó con los kits antes mencionados. Algunos fragmentos se desalaron y purificaron directamente de la reacción de PCR, sin ser corridos en un gel. Para ello se usó el kit QIAEX II (QIAGEN) y se siguieron las instrucciones del protocolo para desalar y concentrar soluciones de DNA especificadas por la casa comercial.

5.18. Ligaciones y transformaciones

Los fragmentos purificados fueron ligados en los vectores pGEM-T (para el caso del fragmento tRNAU2) o pGEM-T Easy (para los fragmentos Comp, Δ tRNA-Ala, Δ Box-B, Δ Box-A y Δ -293/-1) (Promega). La reacción incluyó 2 μ l del inserto (producto del purificado obtenido por PCR), 50 ng del vector, 1 μ l de ligasa T4 (3 U/ μ l, Promega) y 5 μ l de buffer de ligación rápida (2 \times), en un volumen final de 10 μ l. Se mezcló muy bien por pipeteo y se dejó que la reacción de ligación ocurriera a 4°C durante toda la noche. Previo a la transformación, el DNA de la reacción de ligación fue precipitado con acetato de sodio 3 M pH 5.0, extraído con etanol absoluto y resuspendido en 5-10 μ l de agua desionizada estéril. Para realizar las transformaciones necesarias se trabajó con diferentes cepas de células competentes y electrocompetentes de *E. coli*, tanto comerciales como no comerciales. En el caso de las células comerciales, se siguieron las indicaciones y recomendaciones marcadas por el fabricante. En lo que respecta a las células no comerciales, se siguieron las siguientes indicaciones: después de adicionar el DNA a las células, éstas se incubaron en hielo por 30 min; posteriormente se dio un choque térmico a 42°C por 40 seg y se añadió medio SOC (triptona 2%, extracto de levadura 0.5%, NaCl 1%, KCl 2.4 mM, MgCl₂ 1%, glucosa 2%) a un volumen final de 1 ml y se incubó durante 1 h a 37 °C en agitación rápida (250 rpm). Las células fueron centrifugadas a 13000 rpm por 30 seg y la pastilla fue resuspendida en 100 μ l de medio LB (triptona 1%, extracto de levadura 0.5%, NaCl 0.5%, ajustando a pH 7.0) y sembrada en cajas Petri con medio LB agar (15 g de agar-agar en 1 L de LB) con ampicilina (0.1 mg/ μ l), IPTG (100 mM) y X-Gal (50 mg/ml). Las células se dejaron crecer durante toda la noche a 37 °C. Las células competentes comerciales empleadas fueron: DH5 α (Invitrogen), JM109 (Promega) y XL10-Gold (Stratagene).

5.19. Extracción de DNA plasmídico

Se realizaron minipreparaciones de plásmidos mediante el método de lisis alcalina. Para ello, un cultivo de 2-5 ml de LB inoculado con la clona deseada e incubado a 37 °C por 12 h se centrifugó a

velocidad máxima durante 30-60 seg a 4 °C y posteriormente se retiró el sobrenadante. La pastilla bacteriana se resuspendió en 100 µl de Sol. I (glucosa 50 mM, Tris-HCl 25 mM pH 8.0, EDTA 10 mM pH 8.0) con un vortex. Se adicionaron 200 µl de Sol. II preparada al momento (NaOH 0.2 N, SDS 1 %), se mezcló rápidamente invirtiendo 5 veces y se mantuvo en hielo. Se adicionaron 150 µl de Sol. III fría (60 ml de acetato de sodio 3 M pH 4.8, 11.5 ml de ácido acético glacial, 28.5 ml de agua), se mezcló invirtiendo varias veces, y se mantuvo en hielo 3-5 min. Se centrifugó a velocidad máxima por 5 min a 4 °C y se pasó el sobrenadante a un tubo para microcentrífuga de 1.5 ml. Se añadió un volumen de fenol: cloroformo (1:1), se vortexeó y centrifugó a velocidad máxima 5 min a 4 °C; posteriormente se pasó la fase líquida a un tubo nuevo. Se precipitó el DNA con dos volúmenes de etanol absoluto, se mezcló por vortex y se incubó en hielo por 2 min. Se centrifugó a velocidad máxima 5 min a 4 °C y se retiró el sobrenadante. Se realizó un lavado con 1 ml de etanol al 70 %, se dejó secar la pastilla y se resuspendió en 50 µl de buffer TE con RNAsa A (20 µg/µl) y se incubó a 37°C por 1 h.

Para la obtención de los plásmidos que se emplearon en las transfecciones transitorias de los parásitos, se llevaron a cabo midi y maxi-preparaciones, utilizando los kits QIAfilter Plasmid Midi Kit y QIAfilter Plasmid Maxi Kit (QIAGEN) o NucleoSpin Plasmid (Macherey-Nagel), siguiendo las instrucciones y recomendaciones marcadas por el fabricante.

5.20. Ensayo de mutagénesis dirigida

Con la finalidad de insertar una secuencia *tag* dentro del gen del snRNA U2 de *L. major*, así como para obtener los plásmidos con mutaciones de la región codificante y región flanqueante 5' del mismo gen se usó el kit comercial QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies). El procedimiento básico utiliza un vector de DNA de doble cadena (dsDNA) que tenga el inserto de interés, y dos oligonucleótidos con la mutación deseada. Los oligonucleótidos, cada uno complementario a hebras opuestas del vector, son extendidos durante ciclos de diferentes temperaturas por la DNA polimerasa de alta fidelidad *Pfu Ultra* sin que los oligonucleótidos se desalineen. La

extensión de los oligonucleótidos produce un plásmido mutado que tiene extremos romos. Después de esta fase, el producto se digiere con *Dpn* I, endonucleasa específica de DNA metilado y hemimetilado. Así, la enzima se usa para digerir el DNA molde (parental), pues el DNA aislado de casi todas las cepas de *E. coli* está metilado y por eso es susceptible a la digestión con *Dpn* I. Después de esto, sólo el vector con la mutación deseada queda disponible para transformarse en células ultracompetentes XL10-Gold. En este trabajo se usaron como molde los plásmidos ptRNAU2, para insertar la secuencia *tag*; pComp, para mutar la región codificante y 5' flanqueante de U2; y pBS-105/-116, para generar la construcción pDBS, que tiene mutación en ambas cajas del tRNA-like. Los oligonucleótidos y los plásmidos que se usaron como molde en las reacciones de mutagénesis se muestran en la tabla 2. Las reacciones se llevaron a cabo siguiendo las recomendaciones del fabricante:

Nombre del plásmido	Oligonucleótidos sentido	Oligonucleótidos antisentido	Plásmido usado como molde en la mutagénesis
ptRNAU2M	5'U2oligoTag 5'-GCCTTCGGGCAAAGG TCTGACCTCGACTCTAA TAGAAATTACAAC	3'U2oligoTag 5'-GTTGTAATTTCTATT AGAGTCGAGGTCAGACC TTGCCCCGAAGGC	ptRNAU2
pBS+6/+12	5U2LmjmutFw 5'-TGGTACTAACATATC CATGGA ACTATTTAGCT AAGAT	5U2LmjmutRv 5'-ATCTTAGCTAAATAG TTCCATGGATATGTTAG TACCA	pComp
pBS-128/-138	TLBTWU2LmjmutFw 5'-CAGGTACCGTTTGGA TAGGATCGCCCGATTCC TTCGGGGTT	TLBTWU2LmjmutRv 5'-AACCCCGAAGGAATC GGGCGATCCTATCCAAA CGGTACCTG	pComp
pBS-263/-269	BalaU2LmjmutFw 5'-ATGCGGTAGGTATTG TTGGACGTACCCA ACTT CTCCA	BalaU2LmjmutRv 5'-TGGAGAAGTTGGGTA CGTCCAACAATACCTAC CGCAT	pComp

pBS-219/-230	AalaU2LmjmutFw 5'-ACCGCGTCGGGATG GATATCATAGCCTAGAG CGACCGCTTA	AalaU2LmjmutRv 5'-TAAGCGGTCGCTCTA GGCTATGATATCCATCC CCGACGCGGT	pComp
pBS-150/-156	BlikeU2LmjmutFw 5'-GCGGGATTCTTCGG TTGGACGACCTCGTCGA AAGCA	BlikeU2LmjmutRv 5'-TGCTTTCGACGAGGT CGTCCAACCGAAGGAAT CCCGC	pComp
pBS-105/-116	AlikeU2LmjmutFw 5'-CGCGGCGTTTTTTTT GATACCATCGCTACCG TTTGGAGACC	AlikeU2LmjmutRv 5'-GGTCTCCAAACGGTA GGCGATGGTATCAAAAA AAACGCCGCG	pComp
pDBS	BlikeU2LmjmutFw 5'-GCGGGATTCTTCGG TTGGACGACCTCGTCGA AAGCA	BlikeU2LmjmutRv 5'-TGCTTTCGACGAGGT CGTCCAACCGAAGGAAT CCCGC	pBS-105/-116

Tabla 2. Oligonucleótidos usados para generar los plásmidos con mutaciones en la región codificante y flanqueante 5' del snRNA U2 de *L. major*. En la 1^{er} columna se muestran los nombres de los plásmidos generados por mutagénesis dirigida en este trabajo. En la 2^a y 3^{er} columnas se muestran los oligonucleótidos utilizados en las reacciones. En la última columna se muestra el plásmido que se empleó como molde para cada caso.

5.21. Digestión de plásmidos con enzimas de restricción

Las digestiones con enzimas de restricción se llevaron a cabo con las condiciones recomendadas por el fabricante (New England BioLabs). El buffer de reacción proporcionado se usó a una concentración final de 1×, la temperatura de digestión fue de 37°C. Los diversos análisis de restricción para caracterizar o confirmar la identidad de las clonas obtenidas se llevaron a cabo en un volumen final de 20 µl, digiriendo ~1 µg de DNA. Finalmente una alícuota de la digestión se corrió un gel de agarosa al 1% para visualizar el patrón de bandas de DNA esperado.

5.22. Transfecciones transitorias

Promastigotes de *L. major* fueron crecidos en medio BM y cultivados en fase media logarítmica (30-40 × 10⁶ cels/ml) para llevar a cabo transfecciones por medio de electroporación. Para ello, 1 × 10⁸ células fueron centrifugadas a 2500 g, por 12 min a 4 °C; posteriormente se realizaron dos lavados empleando 10 ml de PBS-G (PBS pH 7.2, glucosa 6 mM) y buffer cytomix pH 7.6 (Hepes 25 mM pH

7.6, KCl 120 mM, CaCl₂ 0.15 mM, K₂HPO₄/KH₂PO₄ 10 mM pH 7.6, EDTA 2 mM y MgCl₂ 5 mM). La pastilla celular fue resuspendida en 400 µl de buffer cytomix y las células fueron transferidas a una cubeta de electroporación de 4 mm (Electroporation Cuvettes Plus, BTX); dentro de la misma cubeta se adicionaron de 25 a 50 µg del plásmido a transfectar (pComp, pΔtRNA-Ala, pΔBox-B, pΔBox-A, Δ-293/-1, pBS+6/+12, pBS-128/-138, pBS-263/-269, pBS-219/-230, pBS-150/-156, pBS-105/-116 y pDBS) y 50 µg de los vectores pB6-TAP o pLMRIB, usados como plásmidos de cotransfección. Se mezcló suavemente y se incubó 10 min en hielo. Para llevar a cabo la electroporación se empleó un equipo Electro Cell Manipulator (BTX). El pulso conferido fue de 1600 volts, 25 Ω y 50 µF; las células fueron transferidas inmediatamente a hielo y se incubaron por 10 min adicionales. Las células electroporadas fueron resuspendidas en 10 ml de medio BM y se dejaron a 28 °C hasta que el RNA fue extraído (24 h postransfección).

5.23. RT-qPCR

Para la síntesis de cDNA, se tomó 1 µg de RNA total tratado con DNasa I y se agregaron 2 pmol del oligonucleótido 3'U2Lmjend o Alfa-tub-3' (5'-GTAGTTGATGCCGCACTTGAAG), 1 µl de mezcla de dNTPs (10 mM) y agua DEPC hasta alcanzar un volumen final de 13 µl; entonces se calentó a 65 °C por 5 min y en seguida se pasó a hielo por 2 min más. Después se adicionaron 4 µl del buffer first-strand (FS) 5× de la transcriptasa reversa (Invitrogen), 1 µl de DTT 0.1 M, 1 µl del inhibidor de RNAsas (RNAsin, 40 U/µl) y 1 µl de la transcriptasa reversa SuperScript III (200 U/µl, Invitrogen). Después de mezclar por pipeteo se incubó por 1 h a 55°C y se inactivó la enzima a 70°C por 15 min.

Los ensayos de PCR cuantitativos para amplificar al gen del snRNA U2 exógeno o a α-tubulina (*LmjF.13.0280*), se realizaron de la siguiente manera: en un tubo de PCR se pusieron 7.5 µl del reactivo Platinum® SYBR® Green qPCR SuperMix-UDG (Invitrogen), 1.5 µl de cDNA, 0.3 µl de los oligonucleótidos sentido (U2Sa-5' o Alfa-tub-5': 5'-AGAAGTCCAAGCTCGGCTACAC) y antisentido (LmjU2tag-Rev o a-tubLmjrvs: 5'-GGTCGTAGATGGCCTCATTG) a una concentración de 10 µM, y

agua hasta alcanzar un volumen final de 15 μ l. Para analizar el RNA de células transfectadas con la construcción pBS+6/+12, se usaron exactamente las mismas condiciones descritas, excepto que en este caso se usó al oligonucleótido sentido U2qFw2 (en lugar de U2Sa-5').

El protocolo que siguió el termociclador (Rotor-Gene 3000 cycler, Corbett Research) consistió de siete pasos: incubación a 50 °C por 2 min (paso 1), una desnaturalización inicial de 2 min a 95 °C (paso 2), seguida por otra desnaturalización corta de 15 seg a 95 °C (paso 3), renaturalización por 30 seg a 55 °C para amplificar a U2; o a 63 °C para analizar a α -tubulina (paso 4), una extensión a 72°C por 30 seg y lectura de fluorescencia (paso 5); se repitieron los pasos del 2 al 4, hasta 40 veces (paso 6), en el último paso se obtuvo una curva de desnaturalización del producto amplificado. Para hacer dicha curva, la muestra se calienta gradualmente desde 72 hasta 95 °C midiéndose la fluorescencia cada que se aumentaba 1 °C. Como resultado se obtiene una gráfica que muestra la pérdida de fluorescencia en función del cambio de temperatura. En las reacciones específicas solo un pico de desnaturalización es obtenido. La especificidad de las reacciones se corroboró corriendo una alícuota de los productos amplificados en geles de agarosa. Cuando se transfectó al plásmido pBS+6/+12, la temperatura de renaturalización (paso 4) usada fue de 46 °C. Debido a que las eficiencias de amplificación para cada par de oligonucleótidos usados fueron cercanas a al 100%, los datos se analizaron por el método de cuantificación relativa $2^{-\Delta\Delta C_q}$. Para la normalización de los datos se usó como gen de referencia a α -tubulina. Los valores se presentaron en relación a los datos obtenidos de células transfectadas con la construcción pComp.

6. Resultados

6.1. Contexto genómico del snRNA U2 en *L. major* y otros tripanosomátidos

La búsqueda del gen que codifica para el snRNA U2 en las bases de datos del genoma de *L. major* reveló que es un gen de copia única localizado en la hebra antisentido del cromosoma 31. Está anotado en las bases de datos de TriTrypDB con el nombre sistemático de *LmjF.31.snRNA.01*. Además, se encuentra embebido en una unidad policistrónica de genes que codifican para proteínas. Dentro de esta unidad se encuentra flanqueado por dos genes de transportadores de aminoácidos diferentes: mientras que el transportador de aminoácidos anotado con el nombre *LmjF.31.0580* se localiza 3694 bp río arriba del gen del snRNA U2, el gen *LmjF.31.0570* se posiciona 571 pb río abajo (Fig. 17). Se presenta además un gen de tRNA de alanina (tRNA-Ala) ubicado a 211 pb río arriba del snRNA U2 que a diferencia de los genes de proteínas que lo flanquean, presenta una orientación contraria al gen del snRNA U2 (Fig. 17).

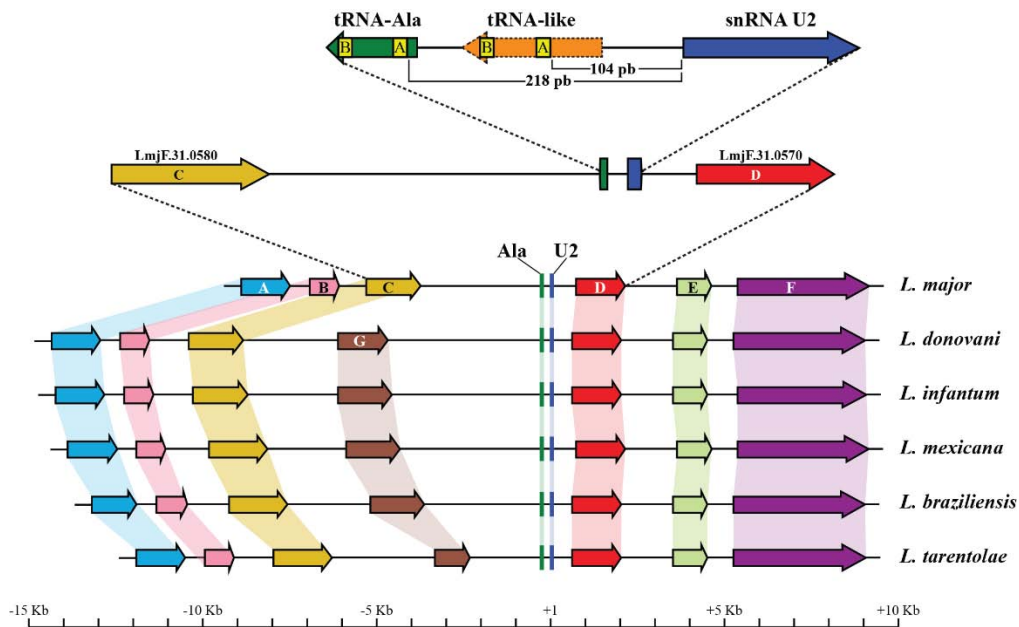


Figura 17. Contexto genómico del snRNA U2 en *Leishmania* spp. En la parte superior se muestran dos ampliaciones del gen del snRNA U2 de *L. major*. Se representan con flechas los genes del snRNA U2 (azul), tRNA-like (naranja) y tRNA-Ala (verde). Las cajas A y B dentro de los genes de tRNA-like y tRNA-Ala se muestran en amarillo. Se indican las distancias entre el snRNA U2 y las cajas A del tRNA-like (104 pb) y tRNA-Ala (218 pb). En la parte inferior de la figura se ilustra la sintenia de los genes del snRNA U2 en diferentes especies de *Leishmania*. Los genes ortólogos están unidos por líneas de colores. En *L. major* el gen A (azul claro) corresponde a *LmjF.31.0600*, el gen B (rosa) a *LmjF.31.0590*, el gen C (mostaza) a *LmjF.31.0580*, el gen D (rojo) a *LmjF.31.0570*, el gen E (verde claro) a *LmjF.31.0560* y el gen F (púrpura) a *LmjF.31.0550*. El gen G (café) no está presente en el genoma de *L. major* pero corresponde a *LdBPK_310600.1* en *L. donovani*. La figura está a escala.

Para determinar el contexto genómico del snRNA U2 en otras especies de *Leishmania*, se analizaron las bases de datos de los genomas secuenciados de *L. donovani* (cepa BPK282A1) [135], *L. infantum* (JPCM5), *L. braziliensis* (M2904) [132,134], *L. mexicana* (U1103) [132] y *L. tarentolae* (Parrot-TarII) [136]. Se encontró que al igual que en *L. major*, el gen del snRNA U2 es de copia única y está insertado en la misma unidad policistrónica en todas las especies de *Leishmania* analizadas (Fig. 17). El gen se ubicó en el cromosoma 31 de *L. donovani*, *L. infantum*, *L. braziliensis* y *L. tarentolae*. Sin embargo, en *L. mexicana*, debido a que los cromosomas 8 y 29, así como 20 y 36, están fusionados [138], el gen del snRNA U2 se localizó en el cromosoma 30. Adicionalmente, el gen del tRNA-Ala se encontró orientado en dirección opuesta y a una distancia muy similar del snRNA U2 en todas las especies de *Leishmania*. Por lo tanto, el análisis del contexto genómico del snRNA U2 en *L. major* y otras especies de *Leishmania* indicó que el locus de dicho gen es altamente sinténico en este género de tripanosomátidos. De manera relevante, un gen que codifica para un transportador de aminoácidos adicional (gen G), está presente en cinco de los genomas de *Leishmania* analizados, pero está ausente en el genoma de *L. major* (Fig. 17).

La ubicación y número de copias del gen del snRNA U2 también se determinó en otros tripanosomátidos. Para esto, las secuencias genómicas de algunos miembros de los géneros *Trypanosoma*, *Crithidia*, *Endotrypanum* y *Leptomonas* se examinaron usando las bases de datos de TriTrypDB. Al igual que en *Leishmania* spp., sólo una copia del gen del snRNA U2 fue identificada en *T. brucei* (cepas Lister 427, DAL972 y TREU927), *T. evansi* (STIB 805), *T. vivax* (Y486), *C. fasciculata* (Cf-C1), *E. monterogeii* (LV88), *L. pyrrocoris* (H10) y *L. seymouri* (ATCC 30220). De manera interesante, se encontraron tres copias del snRNA U2 en dos cepas diferentes de *T. cruzi*: CL Brener y Dm28_c. Entonces, al analizar las secuencias genómicas de diversas especies de tripanosomátidos, se determinó que el gen del snRNA U2 está presente en la mayoría de estos organismos como un gen de copia única, con la excepción de la especie *T. cruzi* que contiene 3 copias de U2.

6.2. Mapeo de sitios de inicio y término de la transcripción del snRNA U2 de *L. major*

El tamaño del snRNA U2 ha sido determinado en algunas especies de tripanosomátidos mediante ensayos de secuenciación directa de RNA, estimaciones electroforéticas y el mapeo de sus extremos 5' y 3' usando protocolos basados en la técnica de RT-PCR [156-163]. Estos resultados muestran que el tamaño del snRNA U2 difiere entre especies, ya que mientras el tamaño máximo de U2 en tripanosomátidos se ha reportado como de 152 nt en *L. collosoma*, el tamaño mínimo de este transcrito se ha encontrado de 141 nt en *L. amazonensis*. De acuerdo con las bases de datos de TriTrypDB, la longitud del gen del snRNA U2 en *L. major* es de 134 pb (Fig. 18). Sin embargo, la comparación de la secuencia 3' flanqueante del snRNA U2 de *L. major* con las de *T. brucei* y *T. cruzi*, cuyos extremos 3' maduros han sido mapeados en el snRNA U2, sugirió que el extremo 3' del gen del snRNA U2 de *L. major* estaba mal anotado en las bases de datos (Fig. 18, panel A). Por lo tanto, para determinar el tamaño correcto del snRNA U2 en *L. major*, tanto el sitio de inicio (extremo 5') como los sitios de término de la transcripción (extremo 3') se mapearon experimentalmente.

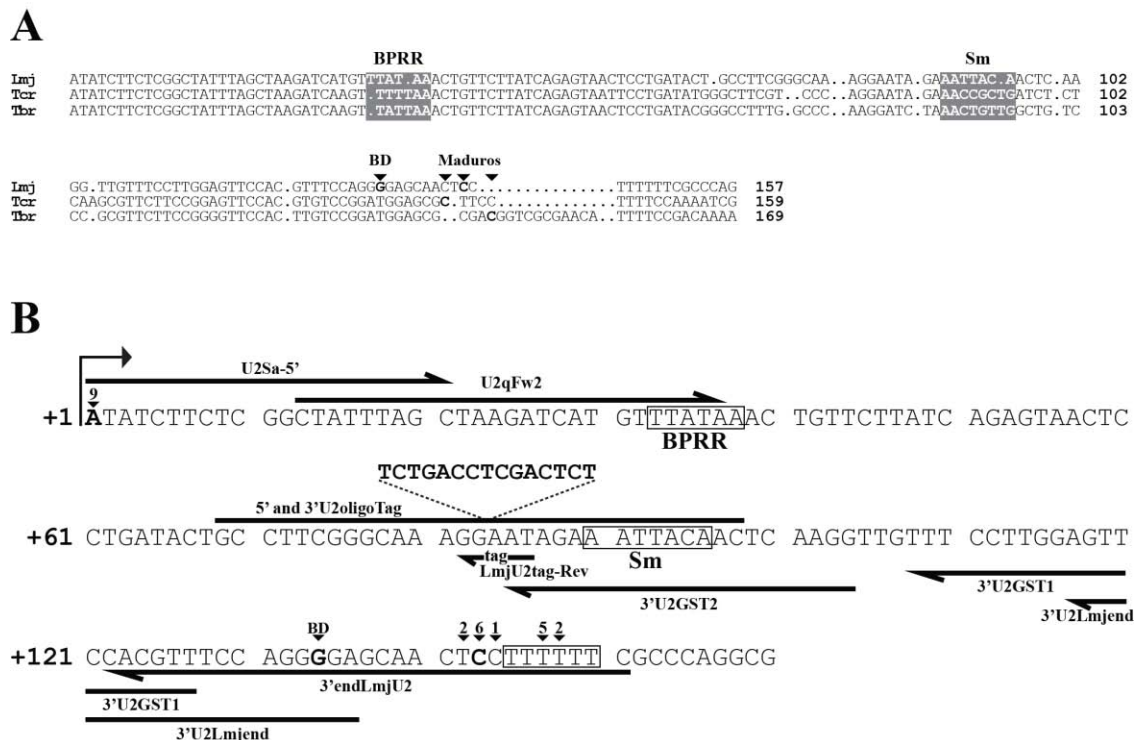


Figura 18. Secuencia codificante y 3' flanqueante del gen del snRNA U2 en *L. major*. A) Alineamiento de las secuencias de U2 de *L. major* Friedlin (Lmj), *T. cruzi* CL Brener Non-Esmeraldo-like, copia del cromosoma 23 (Tcr) y *T. brucei* DAL972 (Tbr). El alineamiento se obtuvo con el software DNAMAN. Se resaltan con negritas los extremos 3' maduros de U2

mapeados experimentalmente en *L. major* (este trabajo), *T. cruzi* [163] y *T. brucei* [162]. En los dos paneles, el extremo 3' anotado erróneamente en la posición 134 (G) se muestra en negritas e indicado con BD: base de datos de *L. major*. B) El residuo A en negritas (+1) corresponde al SIT mapeado por ensayos de *primer extension* y 5'-RACE (el número de clonas encontradas en esa posición está indicado arriba de la A). En el extremo 3', el tracto de residuos de Ts se muestra encerrado en una caja. Los términos de la transcripción mapeados por RT-PCR se señalan con triángulos que especifican el número de clonas identificadas en cada posición. La región de reconocimiento del sitio de ramificación (BPRR) y el sitio Sm se muestran en cajas. Se indica también una secuencia *tag* insertada entre las posiciones 83 (G) y 84 (A) que fue utilizada para el análisis transcripcional del gen. Se señala la secuencia y localización de algunos oligonucleótidos usados en este trabajo. Para facilitar el análisis de estos resultados, la secuencia se muestra como la reversa complementaria de la secuencia del cromosoma 31 anotada en las bases de TriTrypDB.

6.2.1. Mapeo del sitio de inicio de la transcripción del snRNA U2 en *L. major*

Para localizar el primer nucleótido transcrito del gen, se llevó a cabo un ensayo de *primer extension* usando un oligonucleótido (3'U2GST1) marcado radioactivamente con P³² y complementario al extremo 3' del snRNA U2 (Fig. 18, panel B). Como primer paso, se obtuvo RNA total de *L. major* utilizando el reactivo TRIZOL-LS de Invitrogen, el cual fue analizado mediante electroforesis en un gel de agarosa para determinar su calidad (Fig. 19). La abundancia y nitidez de las bandas de rRNA 18S, 24S α y 24S β indicó que el RNA presentó una calidad muy buena.

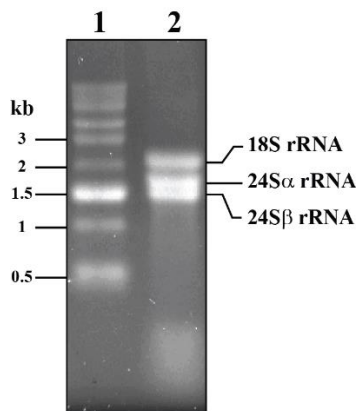


Figura 19. Electroforesis de RNA total extraído de *L. major*. Se presenta al RNA total (carril 1) y al patrón de peso molecular de RNA (RNA ladder, Invitrogen) (carril 2). Se observan tres bandas abundantes en la muestra de RNA total que corresponden a las moléculas del rRNA 18S, 24S α y 24S β . Gel de agarosa al 1% teñido con bromuro de etidio.

Posteriormente el RNA obtenido se usó como molde en un ensayo de *primer extension*. El cDNA generado en este ensayo se analizó mediante electroforesis en un gel de poliacrilamida junto con reacciones de secuenciación manual del gen snRNA U2, las cuales se hicieron con el mismo oligonucleótido 3'U2GST1. Como se muestra en la figura 20, la reacción de *primer extension* migró a la

altura de una A (T en la secuencia reversa complementaria) y por lo tanto, el SIT del snRNA U2 corresponde a un residuo de A, el cual es equivalente a los SIT reportados para el mismo gen en *T. cruzi*, *T. brucei* y *L. collosoma* [156,159]. Debido a que las reacciones de secuencia se hicieron con el reactivo 7-deaza dGTP, y el *primer extension* con dGTP, existe una diferencia de alrededor de 1 nt entre la escalera de secuencia y el producto de *primer extension* [164].

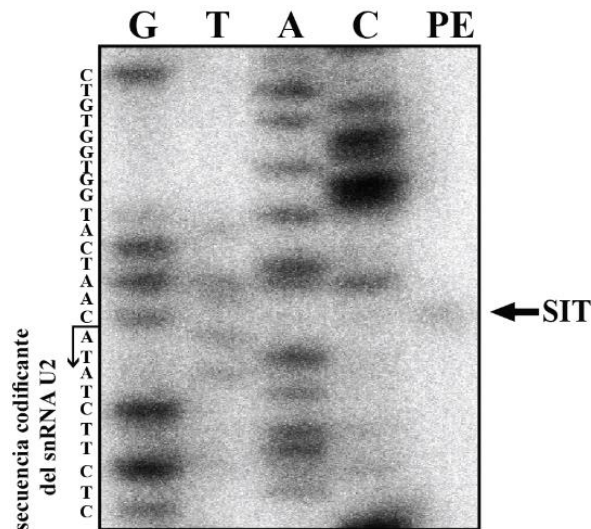


Figura 20. Sitio de inicio de la transcripción del snRNA U2 de *L. major* determinado mediante *primer extension*. El producto de la reacción de *primer extension* (carril PE) se señala con una flecha (SIT). Las reacciones de secuencia (G, T, A y C) se corrieron junto a la reacción del PE. Del lado izquierdo se muestra la secuencia obtenida; con una flecha se señala el inicio y la dirección de la transcripción del gen del snRNA U2. Gel de poliacrilamida al 6%.

Para corroborar el resultado obtenido mediante *primer extension*, se localizó el sitio de inicio de la transcripción del gen esta vez mediante ensayos de 5'-RACE. Para ello, RNA total de *L. major* fue usado como molde en una transcripción reversa en la que se usó el mismo oligonucleótido que en el *primer extension* (3'U2GST1). En seguida, al cDNA generado se le agregó una cola de citocinas en su extremo 3', con la finalidad de realizar una reacción de PCR posterior en la que se usarían los oligonucleótidos AAP (complementario a la cola de Cs) y 3'U2GST2, que es complementario a una secuencia específica del lado 3' del snRNA U2 (Fig. 18, panel B). La reacción de PCR generó un producto de ~140 pb (Fig. 21, panel A), el cual fue purificado usando un kit comercial (Fig. 21, panel B).

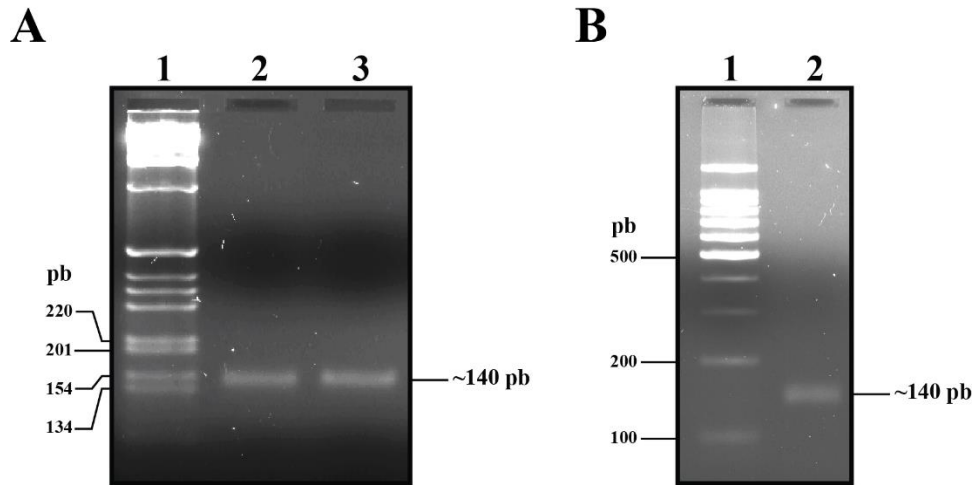


Figura 21. Productos de PCR del ensayo 5'-RACE. A) Los carriles 2 y 3 muestran los productos de dos PCRs que amplificaron al gen del snRNA U2 en el ensayo de 5'-RACE. En el carril 1 se cargó un patrón de peso molecular de 1 kb (Invitrogen). B) Se muestra el inserto purificado de ~140 pb, producto del ensayo de 5'-RACE con el snRNA U2 (carril 2). En el carril 1 se muestra el patrón de peso molecular de 100 pb (Invitrogen). Geles de agarosa al 2% teñidos con bromuro de etidio.

Posteriormente, el producto purificado fue clonado en el vector pGEM-T. Un total de 10 clonas fueron analizadas mediante restricción con las enzimas *Nco* I y *Sac* I, las cuales reconocen secuencias que flanquean el sitio de clonación del vector. Mediante esta digestión se liberó un fragmento de ~200 pb. Todos los plásmidos presentaron la banda del peso esperado (Fig. 22).

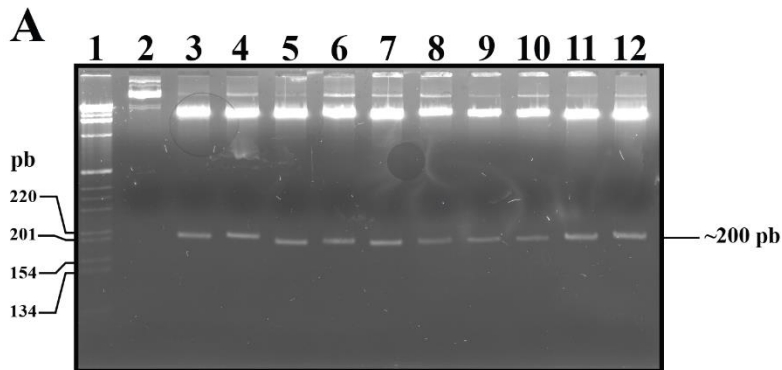


Figura 22. Digestión con las enzimas *Nco* I y *Sac* I de las clonas obtenidas del ensayo 5'-RACE. A) Los carriles 3-12 muestran las digestiones de las clonas 1-10, respectivamente. Se señala del lado derecho el fragmento liberado de ~200 pb. En el carril 2 se analizó la clona 1 sin digerir. En el carril 1 se cargó un patrón de peso molecular de 1 kb (Invitrogen). Gel de agarosa al 2% teñido con bromuro de etidio.

Del total de clonas analizadas con enzimas de restricción, 9 se mandaron a secuenciar. El análisis de secuencia de las clonas confirmó que la transcripción del snRNA U2 inicia en el residuo A mapeado

previamente por *primer extension* (Fig. 18, panel B). Así, estos resultados sugieren que el SIT del snRNA U2 en todos los tripanosomátidos corresponde a un residuo de A el cual está muy conservado.

6.2.2. Mapeo del extremo 3' maduro del snRNA U2 en *L. major*

Para determinar los sitios de término de la transcripción del gen del snRNA U2 en *L. major*, se realizó un ensayo de RT-PCR con RNA al cual se le agregaron colas de poli(A) *in vitro*. Después de adicionar las colas de poli(A) a los extremos 3' de todos los RNAs de la célula, se llevó a cabo una transcripción reversa con el oligonucleótido Nested(dT), el cual reconoce a la cola de poli(A) recién adicionada y además agrega el sitio de anclaje de los oligonucleótidos antisentido usados en reacciones de PCR subsecuentes. El cDNA obtenido fue usado como molde en una primera reacción de PCR en la que se usaron al oligonucleótido específico U2Sa-5' (Fig. 18, panel B), y al oligonucleótido B1 que se une a la secuencia agregada por el oligo Nested(dT). La reacción de PCR generó un producto de ~200 pb (Fig. 23, panel A). Posteriormente, el producto amplificado se usó como molde en una segunda reacción de PCR en la que se usaron los oligonucleótidos U2qFw2, que es complementario a la secuencia del snRNA U2 (Fig. 18, panel B), y B2, que se une a una secuencia anidada dentro del oligo Nested(dT). La reacción de PCR anidado generó un amplificado de ~170 pb, el cual fue purificado usando un kit comercial (Fig. 23, panel B).

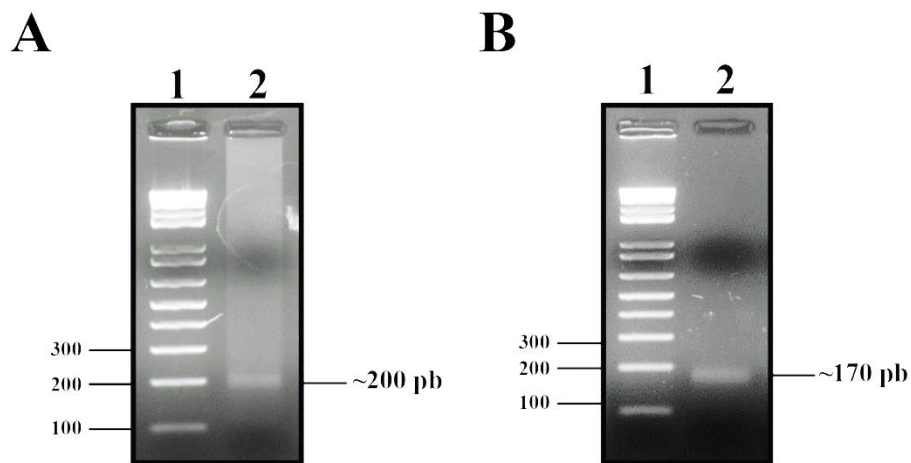


Figura 23. Productos de RT-PCR del ensayo para mapear el extremo 3' del snRNA U2. A) El carril 2 muestra al producto de la primera reacción de PCR que se usó como molde en un PCR anidado. B) Se muestra el inserto purificado de ~170 pb, producto del ensayo para mapear al extremo 3' del snRNA U2 (carril 2). En los carriles 1 se cargó un patrón de peso molecular de 100 pb (Invitrogen). Geles de agarosa al 2% teñidos con bromuro de etidio.

Después, el DNA purificado fue clonado en el vector pGEM-T easy. Un total de 24 clonas fueron analizadas mediante restricción con la enzima *Eco* RI, la cual reconoce secuencias que flanquean el sitio de clonación del vector. Mediante esta digestión se liberó un fragmento de ~200 pb. Todos los plásmidos presentaron la banda del peso esperado (Fig. 24).

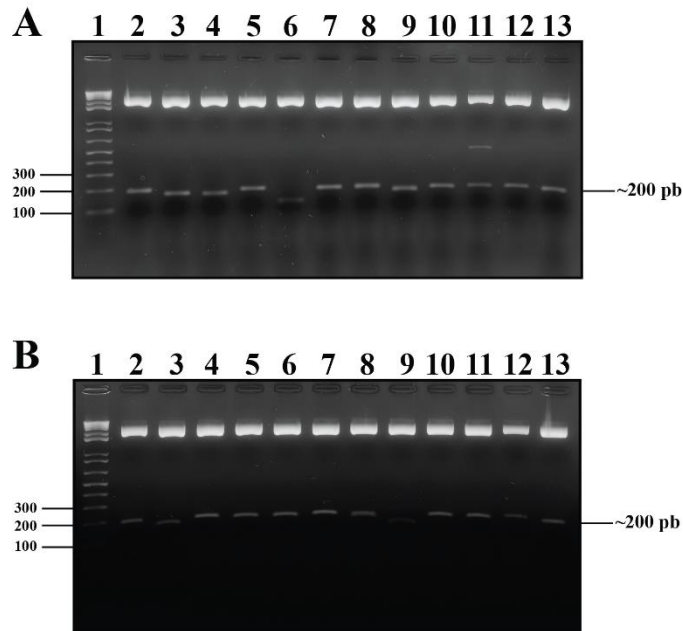


Figura 24. Digestión con *Eco* RI de las clonas obtenidas en el ensayo para mapear el extremo 3' del snRNA U2. A) Los carriles 2-13 muestran las digestiones de las clonas 1-12, respectivamente. B) Los carriles 2-13 muestran las digestiones de las clonas 13-24, respectivamente. En ambos paneles se señala del lado derecho el fragmento liberado de ~200 pb. En los carriles 1 se cargó un patrón de peso molecular de 100 pb (Invitrogen). Geles de agarosa al 2% teñidos con bromuro de etidio.

Del total de clonas analizadas con enzimas de restricción, 16 se mandaron a secuenciar. El análisis de secuencia de las clonas indicó que la transcripción del gen del snRNA U2 en *L. major* termina en un tracto de residuos de Ts ubicado de 145 a 150 pb río abajo del SIT. En dos de las clonas analizadas, la cola de poli(A) se agregó después de la cuarta T y en cinco clonas se localizó después de la tercer T (Fig. 18, panel B). Además, en una clona la cola de poli(A) se encontró después de una C ubicada en la posición +144 con respecto al SIT, seis clonas terminaron en la posición +143 (C), y dos clonas en la base +142 (T) (Fig. 18, panel B). Por lo tanto, estos resultados sugieren que durante el procesamiento del extremo 3' del snRNA U2 existen varias poblaciones con diferentes extensiones 3' que tienen que ser removidas para generar el extremo 3' maduro. Debido a que la mayoría de las clonas terminaron en el residuo C

situado en la posición +143, este nucleótido se consideró como el extremo 3' maduro del snRNA U2 en *L. major* (Fig. 18). De esta manera, como consecuencia del mapeo de los extremos 5' y 3' del snRNA U2 en *L. major*, se determinó que el transcrito maduro mide 143 nt.

6.3. Localización de residuos de pseudouridina en el snRNA U2 de *L. major*

Se ha reportado que la pseudouridilación del snRNA U2 es esencial para su unión con diversas proteínas durante el ensamblaje de su snRNP, así como para el proceso de *splicing* en vertebrados y levaduras [123]. Para determinar la presencia y localización de residuos de pseudouridina (Ψ) en el snRNA U2 de *L. major*, el RNA total de los parásitos se trató con el reactivo CMCT, el cual se une covalentemente a todos los residuos de G, U y Ψ presentes en el RNA [154]. Después, para remover los grupos CMCT de los residuos G y U se llevó a cabo un tratamiento alcalino. El RNA sometido a este tratamiento se usó como molde en un ensayo de *primer extension* en donde la síntesis de cDNA se detiene 1 nt antes de las posiciones de Ψ debido a la presencia del reactivo CMCT. Por medio de este ensayo se identificó la presencia de 6 residuos de Ψ en la secuencia del snRNA U2 de *L. major* localizados en las posiciones 14, 34, 36, 55, 86 y 93 (Fig. 25, carril 6). Un control del experimento se hizo sin el reactivo CMCT para detectar pausas naturales de la transcriptasa reversa, probablemente debido a la degradación de RNA o la presencia de estructuras secundarias estables en el snRNA U2 (Fig. 25, carril 5). Otro control se hizo sin el tratamiento alcalino para identificar todas las posiciones en donde se unió el reactivo CMCT (Fig. 25, carril 7).

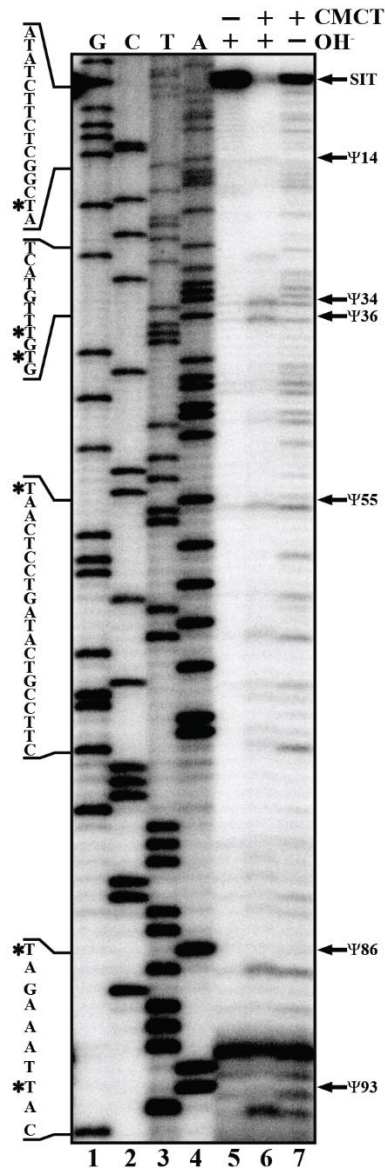


Figura 25. Mapeo de pseudouridinas en el snRNA de *L. major*. Las reacciones de *primer extension* se llevaron a cabo usando 25 μ g de RNA total de *L. major* tratado (+) o no (-) con CMCT. El RNA tratado fue sometido a una hidrólisis alcalina (OH, +). Un control del experimento se hizo sin el tratamiento alcalino (OH, -). Los productos del *primer extension* (carriles 5-7) se corrieron junto a una escalera de secuencia del snRNA U2 (carriles 1-4). Tanto las reacciones de *primer extension* como las reacciones de secuencia se hicieron con el oligonucleótido 3'endLmjU2. Las flechas señalan los residuos de Ψ identificados, por lo que se ubican una posición arriba de las pausas de la transcriptasa reversa. Una secuencia parcial del gen del snRNA U2 se muestra a la izquierda, indicando con asteriscos las posiciones de los nucleótidos pseudouridilados.

En el experimento mostrado en la figura 25, la banda que corresponde a la Ψ ubicada en la posición 14 es muy tenue. En dicho experimento se usaron 25 μ g de RNA tratado con CMCT; sin embargo, al aumentar la cantidad de RNA a 50 μ g en los ensayos de *primer extension*, la banda que corresponde a la Ψ 14 se hizo claramente visible (Fig. 26).

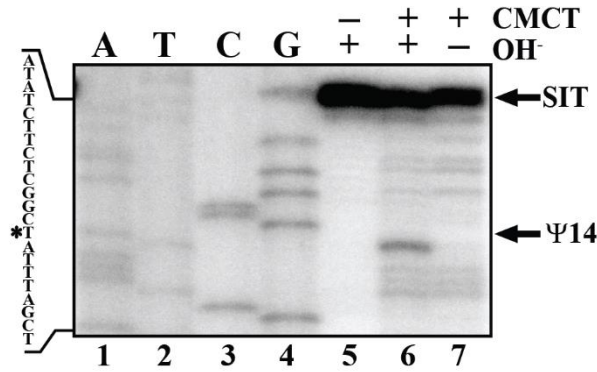


Figura 26. Mapeo de la Ψ 14 en el snRNA U2 de *L. major*. El ensayo de mapeo de pseudouridinas se llevó a cabo como en el experimento de la figura 25, excepto que en el experimento mostrado se usaron 50 μ g de RNA tratado con el reactivo CMCT con el objeto de hacer visible a la Ψ 14. Las etiquetas corresponden a las mismas mostradas en la figura 25.

6.4. Localización de residuos de pseudouridina en dominios importantes del snRNA U2 de *L. major*

Para determinar si las Ψ identificadas están localizadas en dominios importantes para la función del snRNA U2 en *L. major*, se generó la estructura secundaria probable de dicho RNA usando los programas RNAfold web Server y RNAstructure 5.0. La estructura obtenida presenta 3 regiones tallo-asa (I, II y IV) y dos regiones de cadena sencilla. La región tallo-asa II está formada por dos regiones (II.1 y II.2) (Fig. 27).

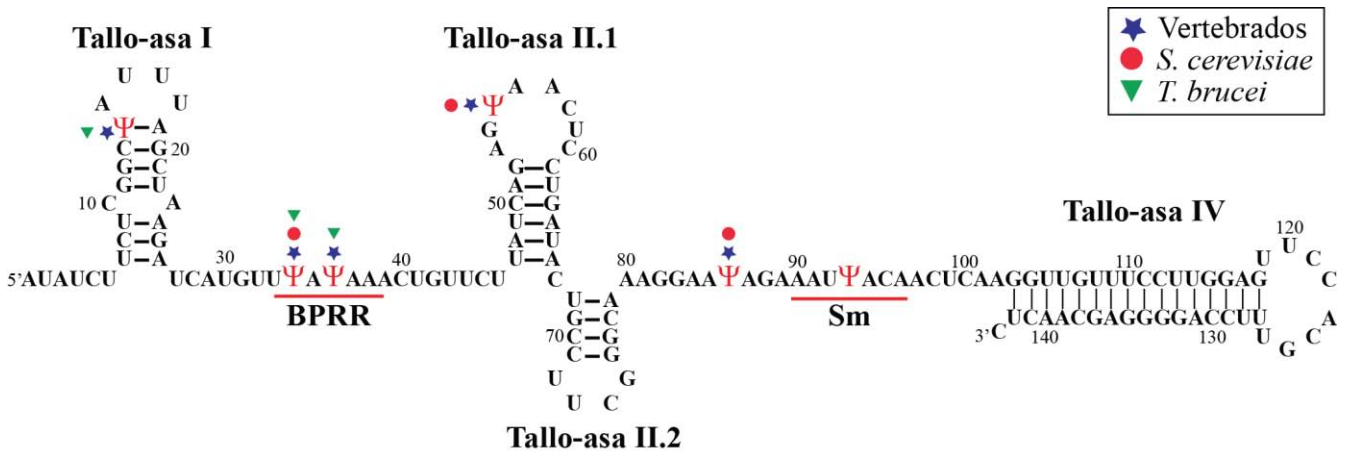


Figura 27. Posición de las Ψ en la estructura secundaria del snRNA U2 de *L. major*. Se muestra la secuencia y la probable estructura secundaria del snRNA U2 de *L. major*. Los tallos-asa I, II.1, II.2 y IV, así como la región de reconocimiento del sitio de ramificación (BPRR) y el sitio Sm, están indicados. Los residuos de Ψ identificados en este estudio se muestran en rojo. Las Ψ conservadas entre *L. major* y vertebrados (estrella azul), *S. cerevisiae* (círculo rojo) y *T. brucei* (triángulo verde) también están indicadas.

Esta estructura secundaria es muy similar a la reportada para *T. brucei*, y parecida a la de eucariontes superiores, donde el snRNA U2 se dobla formando 4 estructuras tallo-asa (I-IV) separadas por dos regiones de cadena sencilla, localizadas entre los tallos-asa I y II y los tallos-asa II y III (ver al snRNA U2 en la Fig. 12). Así, a diferencia de los snRNAs U2 de otros eucariontes, el tallo-asa III no está presente en el U2 de tripanosomátidos [159]. Después del tallo-asa I se encuentra la región de cadena sencilla I, la cual contiene la secuencia GUAGUA que corresponde a la secuencia de unión al sitio de ramificación del pre-mRNA (BPRR, por sus siglas en inglés). Se ha reportado que la secuencia del BPRR está muy conservada de levadura a eucariontes superiores. Sin embargo, la secuencia típica BPRR no está presente en *L. major*. En esta especie, en su lugar se encontró la secuencia UUAUAA (Fig. 27). La secuencia de unión de las proteínas Sm (sitio Sm), cuya secuencia consenso es RAU₄₋₆GR en otros eucariontes, se localiza en la región de cadena sencilla II. En *L. major* la secuencia del sitio Sm reportada difiere de la encontrada en eucariontes superiores, pues en este organismo dicha secuencia corresponde a AAUUACA (Fig. 27) [165].

De manera interesante, las Ψ en las posiciones 34 y 36 fueron localizadas dentro de la región BPRR del snRNA U2 (Fig. 27). La Ψ 34 está muy conservada, pues también está presente en los U2 de vertebrados, levadura y *T. brucei*. La Ψ 36 ha sido identificada en vertebrados y en *T. brucei*, pero está ausente en levadura. Otras tres Ψ están conservadas entre *L. major* y vertebrados (Ψ 14, Ψ 55 y Ψ 86 en *L. major*), dos Ψ adicionales están conservadas entre *S. cerevisiae* y *L. major* (Ψ 55 y Ψ 86 en *L. major*), y una más está presente en *T. brucei* y *L. major* (Ψ 14) (Fig. 27). Es necesario hacer notar que el snRNA U2 de *L. major* contiene una Ψ dentro del sitio Sm en la posición 93 (Fig. 27), la cual no ha sido reportada en ningún otro organismo hasta la fecha.

6.5. Análisis comparativo de las secuencias de snRNAs U2 de tripanosomátidos

Para determinar la homología de secuencia del snRNA U2 entre tripanosomátidos, se llevaron a cabo alineamientos de algunos miembros de esta familia de organismos. Como era de esperarse, la

secuencia del snRNA U2 se encontró muy conservada entre tripanosomátidos (Figs. 28 y 29). La región 5' del gen resultó ser la más conservada, siendo casi idéntica entre las diferentes especies. En cambio, varios cambios se observaron en la región 3' del U2 (Figs. 28 y 29). No obstante, la homología general entre el snRNA U2 de *L. major* y sus ortólogos en otras especies de *Leishmania* es muy alta, pues los porcentajes de identidad obtenidos varían entre un 95.1% con respecto al U2 de *L. braziliensis* y un 98.6% con respecto al U2 de *L. tarentolae* (Tabla 3). De manera similar, el snRNA U2 de *L. major* muestra 69.4% y 76.2% de identidad cuando se compara con sus ortólogos de *T. cruzi* y *T. brucei*, respectivamente (Tabla 3). La identidad entre *L. major* y tres especies de *Leptomonas* fue de 82.6-93.1%. Asimismo, una identidad de 92.36% se encontró entre *L. major* y *C. fasciculata* (Tabla 3).

De forma notable, se encontró que todos los genes del snRNA U2 en las 21 especies de tripanosomátidos analizadas contienen un residuo T (U en la secuencia de RNA) en cinco de las seis posiciones pseudouridiladas en el snRNA U2 de *L. major* (Figs. 28 y 29). La excepción corresponde a la posición 93, la cual es una C o una G en cinco de las especies analizadas. Resulta interesante el hecho de que esta posición se ubica en el sitio de anclaje de las proteínas Sm en el snRNA U2 de *L. major*. Así, estos resultados indican que la gran mayoría de las Us que son pseudouridiladas en el U2 de *L. major* están conservadas en otros tripanosomátidos, lo que abre la posibilidad de que dichas posiciones sean pseudouridiladas también en otros tripanosomátidos.

	<i>L. collosoma</i>	<i>E. monterogeii</i>	<i>L. enrietti</i>	<i>L. braziliensis</i>	<i>L. panamensis</i>	<i>L. major</i>	<i>L. tarentolae</i>	<i>L. mexicana</i>	<i>L. amazonensis</i>	<i>L. donovani</i>	<i>L. aethiopica</i>	<i>L. arabica</i>	<i>L. infantum</i>	<i>C. fasciculata</i>	<i>L. pyrrhocoris</i>	<i>L. seymouri</i>	<i>T. brucei</i>	<i>T. evansi</i>	<i>F. cruzi</i>	<i>T. congolense</i>	<i>T. vivax</i>
<i>L. collosoma</i>	100.00	81.94	83.33	82.64	83.33	82.64	83.92	83.22	83.92	84.03	83.33	82.64	83.22	84.14	84.83	84.14	74.48	75.86	79.17	76.39	74.66
<i>E. monterogeii</i>	81.94	100.00	94.52	94.44	95.14	93.75	95.80	94.41	95.10	95.14	95.14	95.14	96.50	90.97	92.36	90.28	70.07	71.23	73.61	69.18	71.43
<i>L. enrietti</i>	83.33	94.52	100.00	95.83	96.53	93.75	95.80	94.41	95.10	96.53	95.83	95.14	96.50	93.06	94.44	92.36	68.49	69.66	74.13	73.10	73.29
<i>L. braziliensis</i>	82.64	94.44	95.83	100.00	99.31	95.14	97.20	95.80	96.50	95.14	95.83	95.14	96.50	92.36	93.75	92.36	70.14	71.33	74.83	73.43	72.22
<i>L. panamensis</i>	83.33	95.14	96.53	99.31	100.00	95.83	97.90	96.50	97.20	95.83	96.53	95.83	97.20	92.36	93.75	92.36	70.14	71.33	74.83	72.73	72.92
<i>L. major</i>	82.64	93.75	93.75	95.14	95.83	100.00	98.60	97.20	97.90	97.22	97.22	98.61	97.90	92.36	93.06	91.67	69.44	70.63	76.22	72.73	72.92
<i>L. tarentolae</i>	83.92	95.80	95.80	97.20	97.90	98.60	100.00	98.60	99.30	98.60	98.60	98.60	99.30	93.71	94.41	93.01	71.33	72.54	77.46	73.94	74.13
<i>L. mexicana</i>	83.22	94.41	94.41	95.80	96.50	97.20	98.60	100.00	99.30	97.20	97.20	97.20	97.90	93.01	93.71	92.31	72.03	73.24	77.46	74.65	74.13
<i>L. amazonensis</i>	83.92	95.10	95.10	96.50	97.20	97.90	99.30	99.30	100.00	97.90	97.90	97.90	98.60	93.71	94.41	93.01	71.33	72.54	76.76	73.94	73.43
<i>L. donovani</i>	84.03	95.14	96.53	95.14	95.83	97.22	98.60	97.20	97.90	100.00	98.61	98.61	99.30	94.44	94.44	92.36	69.44	70.63	76.22	73.43	74.31
<i>L. aethiopica</i>	83.33	95.14	95.83	95.83	96.53	97.22	98.60	97.20	97.90	98.61	100.00	98.61	99.30	93.06	95.14	93.75	69.44	70.63	76.22	74.13	75.00
<i>L. arabica</i>	82.64	95.14	95.14	95.14	95.83	98.61	98.60	97.20	97.90	98.61	98.61	100.00	99.30	93.06	94.44	92.36	69.44	70.63	76.22	73.43	74.31
<i>L. infantum</i>	83.22	96.50	96.50	96.50	97.20	97.90	99.30	97.90	98.60	99.30	99.30	99.30	100.00	93.71	95.10	93.01	70.63	71.83	76.76	73.94	74.83
<i>C. fasciculata</i>	84.14	90.97	93.06	92.36	92.36	92.36	93.71	93.01	93.71	94.44	93.06	93.06	93.71	100.00	94.48	95.17	68.03	69.18	75.00	71.72	69.39
<i>L. pyrrhocoris</i>	84.83	92.36	94.44	93.75	93.75	93.06	94.41	93.71	94.41	94.44	95.14	94.44	95.10	94.48	100.00	95.86	68.97	70.14	77.08	72.92	71.72
<i>L. seymouri</i>	84.14	90.28	92.36	92.36	92.36	91.67	93.01	92.31	93.01	92.36	93.75	92.36	93.01	95.17	95.86	100.00	70.34	71.53	78.47	73.61	72.41
<i>T. brucei</i>	74.48	70.07	68.49	70.14	70.14	69.44	71.33	72.03	71.33	69.44	69.44	69.44	70.63	68.03	68.97	70.34	100.00	98.06	83.45	82.35	80.77
<i>T. evansi</i>	75.86	71.23	69.66	71.33	71.33	70.63	72.54	73.24	72.54	70.63	70.63	70.63	71.83	69.18	70.14	71.53	98.06	100.00	85.52	83.01	80.77
<i>T. cruzi</i>	79.17	73.61	74.13	74.83	74.83	76.22	77.46	77.46	76.76	76.22	76.22	76.22	76.76	75.00	77.08	78.47	83.45	85.52	100.00	84.14	86.21
<i>T. congolense</i>	76.39	69.18	73.10	73.43	72.73	72.73	73.94	74.65	73.94	73.43	74.13	73.43	73.94	71.72	72.92	73.61	82.35	83.01	84.14	100.00	82.35
<i>T. vivax</i>	74.66	71.43	73.29	72.22	72.92	72.92	74.13	74.13	73.43	74.31	75.00	74.31	74.83	69.39	71.72	72.41	80.77	80.77	86.21	82.35	100.00

Tabla 3. Porcentajes de identidad de la secuencia del snRNA U2 entre diferentes tripanosomátidos.

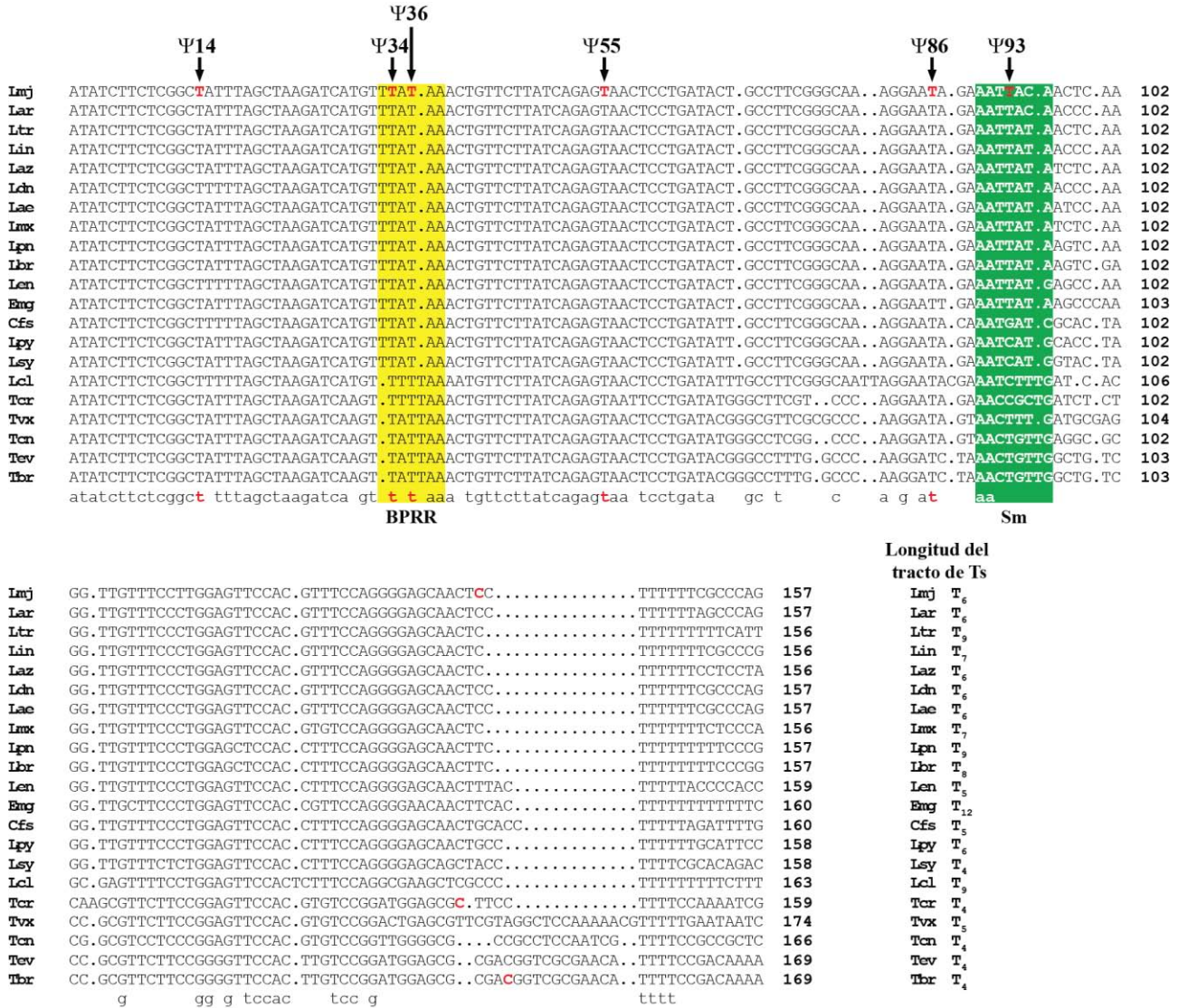


Figura 28. Alineamiento múltiple de los genes del snRNA U2 de tripanosomátidos. La secuencias de las siguientes especies de tripanosomátidos se obtuvieron de la base de datos de TriTrypDB: *L. major* Friedlin (Lmj), *L. arabica* LEM1108 (Lar), *L. tarentolae* Parrot-TarII (Ltr), *L. infantum* JPCM5 (Lin), *L. amazonensis* MHOMBR/71973/M2269 (Laz), *L. donovani* BPK282A1 (Ldn), *L. aethiopica* L147 (Lae), *L. mexicana* U1103 (Lmx), *L. panamensis* MHOM/COL/81/L13 (Lpn), *L. braziliensis* M2904 (Lbr), *L. enriettii* LEM3045 (Len), *E. monterogei* LV88 (Emg), *C. fasciculata* Cf-CI (Cfs), *L. pyrrocoris* H10 (Lpy), *L. seymouri* ATCC 30220 (Lsy), *L. collosoma* (cepa no especificada) [159] (Lcl), *T. cruzi* CL Brener Non-Esmeraldo-like, copia del cromosoma 23 (Tcr), *T. vivax* Y486 (Tvx), *T. congolense* IL3000 (Tcn), *T. evansi* STIB 805 (Tev) y *T. brucei* DAL972 (Tbr). El alineamiento se obtuvo usando el programa DNAMAN. Las flechas indican las posiciones de las Ψ identificadas en *L. major* (este trabajo). Los extremos 3' (residuos C) que han sido mapeados experimentalmente se muestran en rojo para *L. major* (este trabajo), *T. cruzi* [163] y *T. brucei* [162]. La longitud del tracto de Ts es variable en las diferentes especies analizadas, pues fluctúa entre 4 y 12 residuos de Ts. Los números en la secuencia son relativos al SIT (+1) del snRNA U2. Este alineamiento se usó para generar el logo de la figura 29.

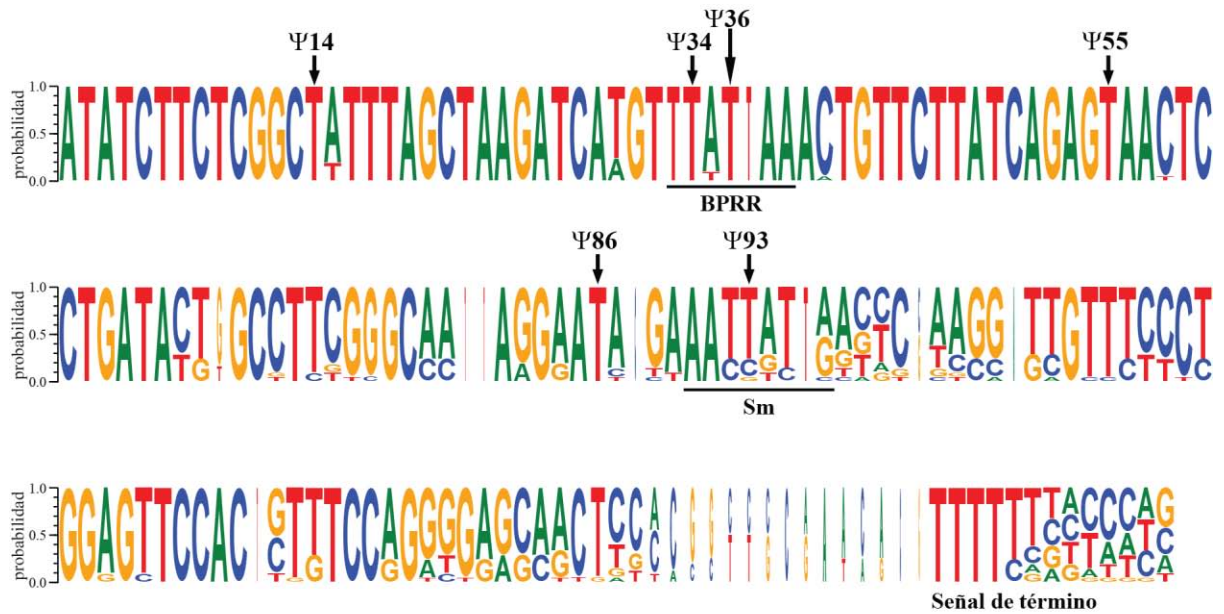


Figura 29. Secuencia consenso del gen del snRNA U2 en tripanosomátidos. Las secuencias de los genes del snRNA U2 y sus regiones flanqueantes 3' de 21 especies de tripanosomátidos (ver figura 28) fueron alineadas con el programa DNAMAN. Posteriormente, usando la aplicación WebLogo se generó un patrón representado como un logo por la distribución cuantitativa de los cuatro nucleótidos en cada posición. Las regiones BPRR, Sm y la señal de término de la transcripción de la RNAP III están indicadas. La posición de las Ψ mapeadas en *L. major* se muestran con flechas.

6.6. Identificación de un tRNA-like asociado al snRNA U2 de *L. major*

Estudios realizados en *T. brucei* y *Leptomonas* demostraron que la transcripción de los snRNAs es mediada por secuencias que se asemejan a las cajas A y B de genes de tRNAs orientados en dirección opuesta a los genes de snRNAs [166]. De manera interesante, la caja A está localizada consistentemente a una distancia de 104 a 106 pb río arriba del snRNA. Como se mencionó anteriormente, río arriba del gen del snRNA U2 de *L. major* se localizó un gen de tRNA-Ala; sin embargo la caja A de dicho gen se localiza a 218 pb del snRNA U2 (Fig. 30), lo cual hace poco probable que las cajas A y B del tRNA-Ala funcionen como la región promotora núcleo del gen del snRNA U2. Debido a este hecho, se decidió buscar un tRNA-like entre el snRNA U2 y el tRNA-Ala, el cual pudiera funcionar como región promotora del snRNA U2.

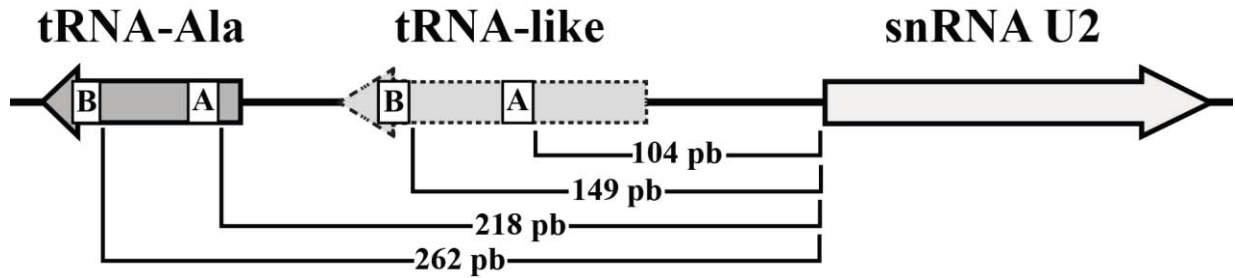


Figura 30. Esquema del snRNA U2 de *L. major* y sus tRNAs asociados. Los genes del snRNA U2, tRNA-like y tRNA-Ala se representan con flechas. Se muestran las cajas A y B dentro de sus respectivos genes. Las distancias de las cajas A y B del tRNA-like (104 pb y 149 pb, respectivamente) y tRNA-Ala (218 pb y 262 pb, respectivamente) están indicadas. La figura está a escala.

Al analizar la secuencia ubicada entre el snRNA U2 y el tRNA-Ala se encontró una molécula semejante a un tRNA (tRNA-like) en la cadena opuesta al snRNA (Fig. 30). Las cajas A y B del gen del tRNA-like están separadas por 104 y 149 pb del snRNA U2, respectivamente. Estos resultados sugirieron que las cajas del tRNA-like eran candidatas a secuencias importantes para la transcripción del gen del snRNA U2 en *L. major*.

6.7. Caracterización de elementos promotores del gen del snRNA U2 de *L. major*

Para localizar las regiones de DNA que regulan la síntesis del snRNA U2 de *L. major*, se llevó a cabo la siguiente estrategia experimental: 1) mediante PCR fue amplificado un fragmento de 429 pb que contiene el gen del snRNA U2 y 293 pb de la región flanqueante 5', que incluye a los genes de tRNA-Ala y tRNA-like; el fragmento de DNA fue clonado en el vector pGEM-T. 2) Se introdujo una secuencia de *tag* de 15 pb entre las posiciones 83 y 84 del gen del snRNA U2 (Fig. 31), mediante mutagénesis dirigida; dicho *tag* fue usado para diferenciar entre el snRNA U2 endógeno (que no contiene el *tag*), de aquel que se sintetizó a partir de los plásmidos (snRNA U2 exógeno, que sí contiene el *tag*). 3) Se generaron dos conjuntos de construcciones: en el primer conjunto se removieron secuencialmente las regiones 5' flanqueantes del snRNA U2; en el segundo conjunto se introdujeron mutaciones en las cajas A y B de ambos genes de tRNA asociados al snRNA U2. 4) Los plásmidos obtenidos fueron transfectados en *L. major*, y se purificó RNA total 24 horas después. 5) La abundancia del snRNA U2 se analizó por

6.7.1. Amplificación y clonación del gen del snRNA U2

El paso inicial hacia la identificación de la región promotora del snRNA U2 consistió en la amplificación por PCR del gen del snRNA U2 de *L. major* y las secuencias flanqueantes 5' que incluyen el tRNA-Ala y el tRNA-like. La amplificación se llevó a cabo con los oligonucleótidos 5' LmjtRNAala y 3' U2Lmjend (Fig. 31), observándose el fragmento esperado de 429 pb (Fig. 32, panel A). Dicho fragmento se desaló y purificó con la ayuda de un kit comercial, y se clonó en el vector pGEM-T. Tres de las clonas que se obtuvieron fueron digeridas con las enzimas *Nco* I y *Sac* I, las cuales flanquean el sitio de clonación del vector. Las tres clonas presentaron el fragmento de 487 pb, que incluye el inserto de 429 pb del locus del snRNA U2 y secuencias del *polilinker* del vector (Fig. 32, panel B). La clona 3 se envió a secuenciar, y se confirmó que el inserto clonado corresponde al locus del snRNA U2 (Fig. 33). El vector resultante fue llamado ptRNAU2.

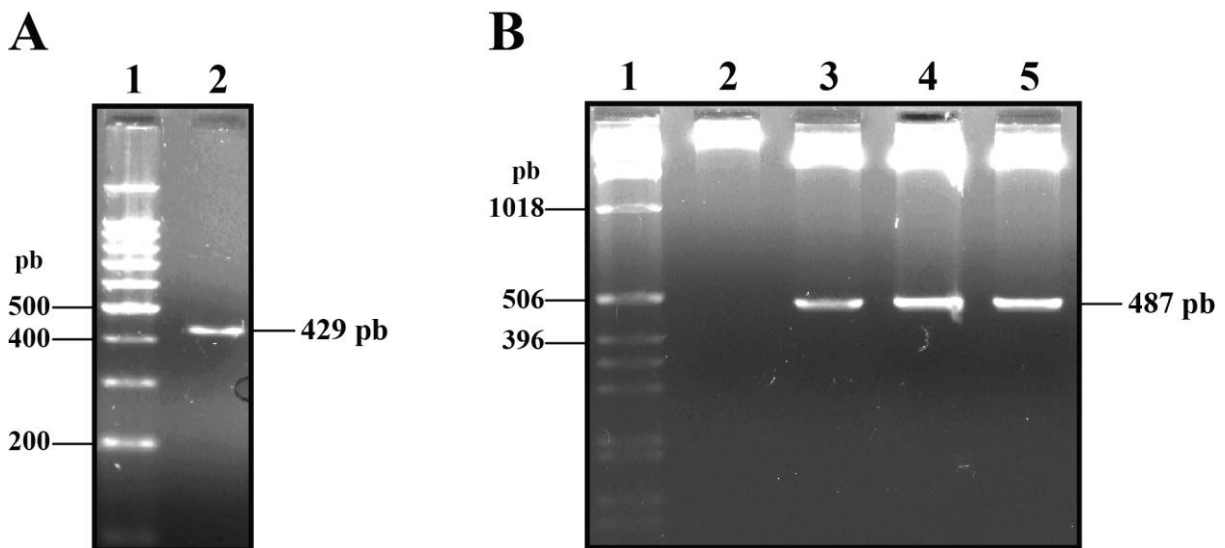


Figura 32. Obtención de ptRNAU2. A) Amplificación por PCR del gen del snRNA U2 y secuencias flanqueantes 5'. Un producto del tamaño esperado (429 pb) pudo observarse (carril 2). En el carril 1 se corrió el patrón de peso molecular de 100 pb (invitrogen). B) Digestión del plásmido ptRNAU2 con las enzimas *Nco* I y *Sac* I. Carril 1: patrón de peso molecular de 1 kb; carril 2: Clona 3 de ptRNAU2 sin digerir. Carriles 3, 4 y 5: clonas 1, 2 y 3 de ptRNAU2 digeridas. Geles de agarosa al 2%.

Caja B

-293 TGAAAAAGTT **GGAGAAGTTG** GGTATCGATC CCAATACCTA CCGCATGCTA AGCGGTCGCT CTACCCATCTG Caja A

tRNA-Ala ←

-223 AGCTACATCC **CCGACGCGGT** AACGTCCCAA TACTACTAGAA TCTAGGAAAA **GATGCTTTTCG** ACGAGGTTTCG Caja B

Caja A

-153 AACCCGAAG GAATCCCGCC CTGGTCTCCA AACGGTACCT GTTGGGCTAA AAAAAAACGC CGCGGTTTCC

tRNA-like ←

-83 **GCGGTGCGTA** **AGCGGCACGG** CAGTGAGGAA ACCGCTGATT ATAAGTAGTG CATGTGGTTT CTTTTGTCTG

snRNA U2

-13 TGGTGGTACT AACATATCTT **CTCGGCTATT** TAGCTAAGAT CATGTTTATA AACTGTTCTT ATCAGAGTAA

+57 CTCCTGATAC TGCCTTCGGG CAAAGGAATA GAAATTACAA CTCAAGGTTG TTTCTTGGG GTTCCACGTT

+127 TCCAGGGGA

Figura 33. Análisis de secuencia de ptRNAU2. Se muestra la secuencia del inserto del locus del snRNA U2 de *L. major* que fue clonado en pGEM-T para generar el plásmido ptRNAU2. Las secuencias codificantes del snRNA U2, tRNA-like y tRNA-Ala están en negritas. Las cajas A y B dentro del tRNA-like y tRNA-Ala están encerradas en un recuadro. Las flechas indican el inicio y dirección de la transcripción. En esta construcción, la región codificante del snRNA U2 carece de 7 nt en el extremo 3' del gen, además está ausente la señal de término de la transcripción.

6.7.2. Inserción de una secuencia *tag* en la región codificante de U2

Como siguiente paso, se introdujo una secuencia *tag* de 15 pb entre las posiciones 83 y 84 del gen del snRNA U2 clonado en el plásmido ptRNAU2. Esto se hizo para de diferenciar por tamaño y secuencia los transcritos del snRNA U2 endógenos, generados a partir del gen localizado en el cromosoma 31, de los exógenos, producidos a partir de los plásmidos. Esto se logró con un kit de mutagénesis dirigida (QuikChange Lightning Site-Directed Mutagenesis Kit) y utilizando a los oligonucleótidos 5'U2oligoTag y 3'U2oligoTag (Fig. 31) junto con el vector ptRNAU2. El plásmido parental se digirió con la enzima *Dpn* I y el plásmido con la secuencia *tag* se transformó en células ultracompetentes XL 10-Gold®. Las clonas obtenidas se digirieron con las enzimas *Nco* I y *Sac* I, y las restricciones se analizaron mediante electroforesis. Todas las clonas presentaron el patrón de digestión esperado: una banda de ~500 pb y otra de ~3 kb (que corresponde al vector) (Fig. 34). El análisis de la secuencia de una de las clonas confirmó la presencia de la secuencia *tag* dentro del gen del snRNA U2 (Fig. 35). Además un análisis de estructura secundaria con el programa RNAfold mostró que la introducción de la

secuencia “tag” no generó cambios dramáticos en la estructura secundaria de U2 (datos no mostrados).

De esta manera se generó el vector ptRNAU2M.

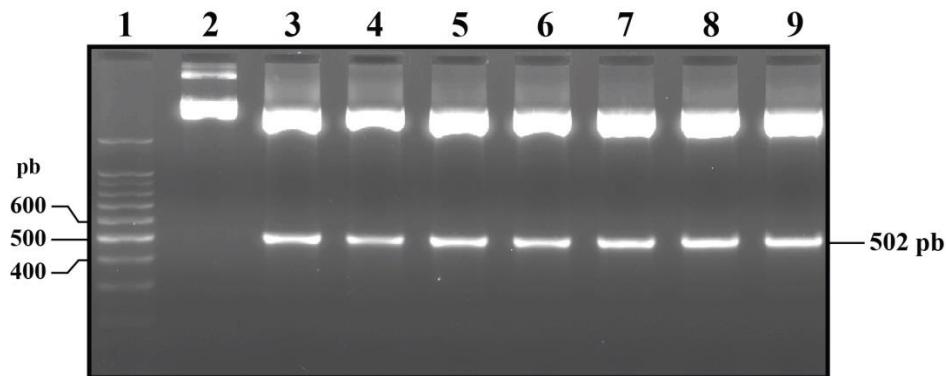


Figura 34. Digestión de siete clones de ptRNAU2M con *Nco I* y *Sac I*. Los carriles 3-9 muestran las digestiones de las clonas 1-7, respectivamente, señalándose el inserto de 502 pb. En el carril 2 se analizó la clona 1 sin digerir. Como marcador de peso molecular se usó una escalera de 100 pb (Invitrogen). Gel de agarosa al 2%.

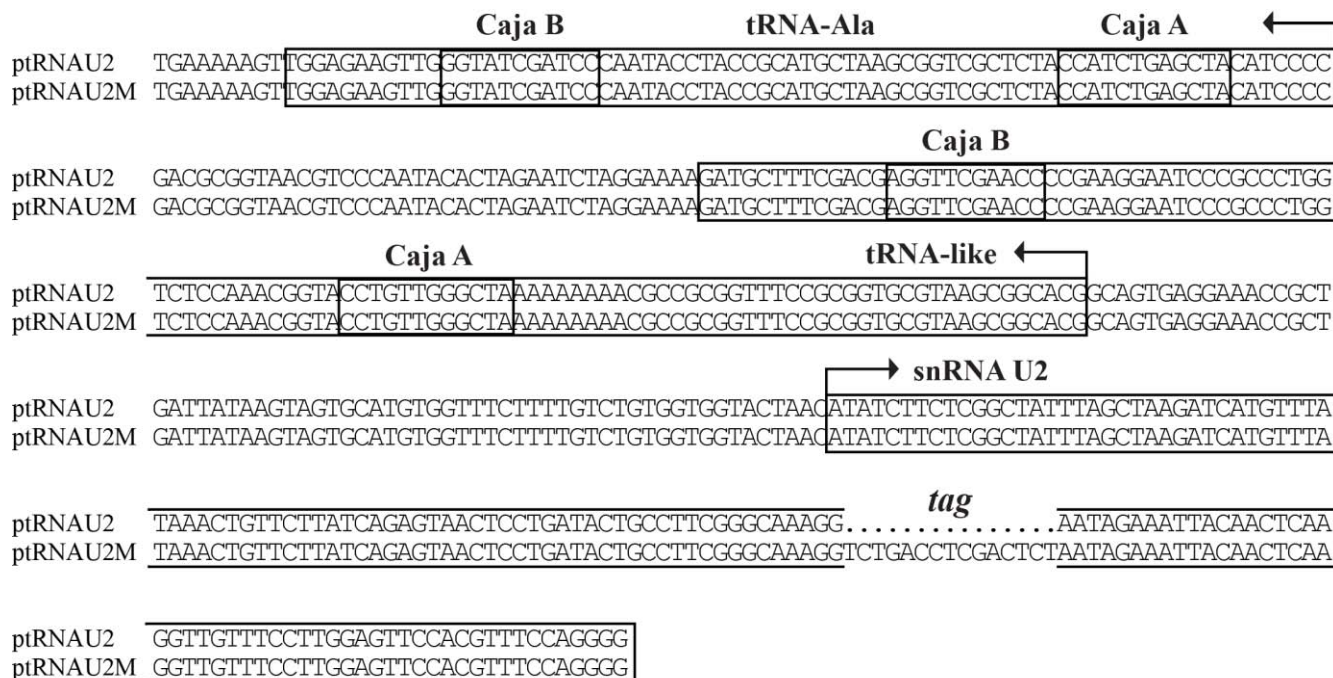


Figura 35. Comparación de las secuencias de los vectores ptRNAU2M y ptRNAU2. Los puntos señalan que en la secuencia de ptRNAU2 no se encuentra la secuencia tag. Los genes de snRNA U2, tRNA-like, tRNA-Ala, así como las cajas A y B se muestran dentro de recuadros. En ambas construcciones, la región codificante del snRNA U2 carece de 7 nt en el extremo 3' del gen, además está ausente la señal de término de la transcripción. Las flechas indican el inicio y dirección de la transcripción.

6.7.3. Adición de la señal de término de la transcripción del gen del snRNA U2

Debido a que el plásmido ptRNAU2M no contiene la señal de término de la transcripción ubicada río abajo del gen del snRNA U2, se decidió agregar esta secuencia por medio de una reacción de PCR en la que se usaron los oligonucleótidos LmjtRNAala-XmaI-F y endLmjU2-XbaI-R (Fig. 31). Estos oligonucleótidos, además de agregar las secuencias faltantes de U2, agregaron sitios de corte de las enzimas de restricción *Xma* I y *Xba* I para facilitar la caracterización de los vectores generados (Fig. 31). El producto de PCR que presentó el tamaño esperado de 473 pb (Fig. 36, panel A), se purificó, ligó al vector pGEM-T easy, y posteriormente se transformó en células de *E. coli*. Una de las clonas obtenidas se analizó por medio de la digestión con las enzimas de restricción *Xma* I y *Xba* I. El fragmento liberado presentó el tamaño esperado de 473 pb (Fig. 36, panel B). El análisis de la secuencia de una de las clonas confirmó la presencia de la señal de término de la transcripción río abajo del gen del snRNA U2 (Fig. 37). Fue así como se obtuvo al plásmido pComp.

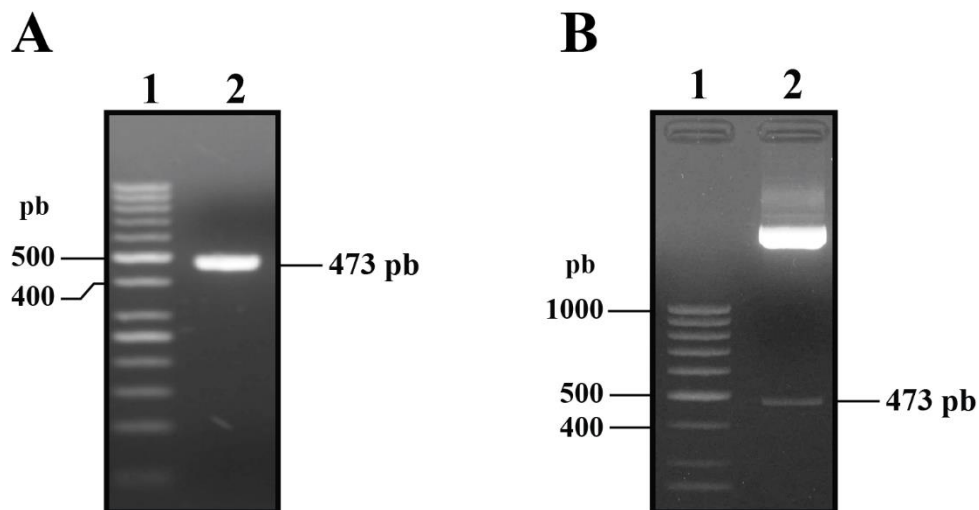


Figura 36. Obtención de pComp. A) Amplificación por PCR en la que se agregaron secuencias del extremo 3' codificante de U2 y una señal de término de la transcripción. Un producto del tamaño esperado (429 pb) pudo observarse (carril 2). B) Digestión del plásmido pComp con las enzimas *Xma* I y *Xba* I. El carril 2 muestra la digestión de una de las clonas de pComp, señalándose el inserto de 473 pb. En los carriles 1 se corrió el patrón de peso molecular de 100 pb (invitrogen). Geles de agarosa al 2%.

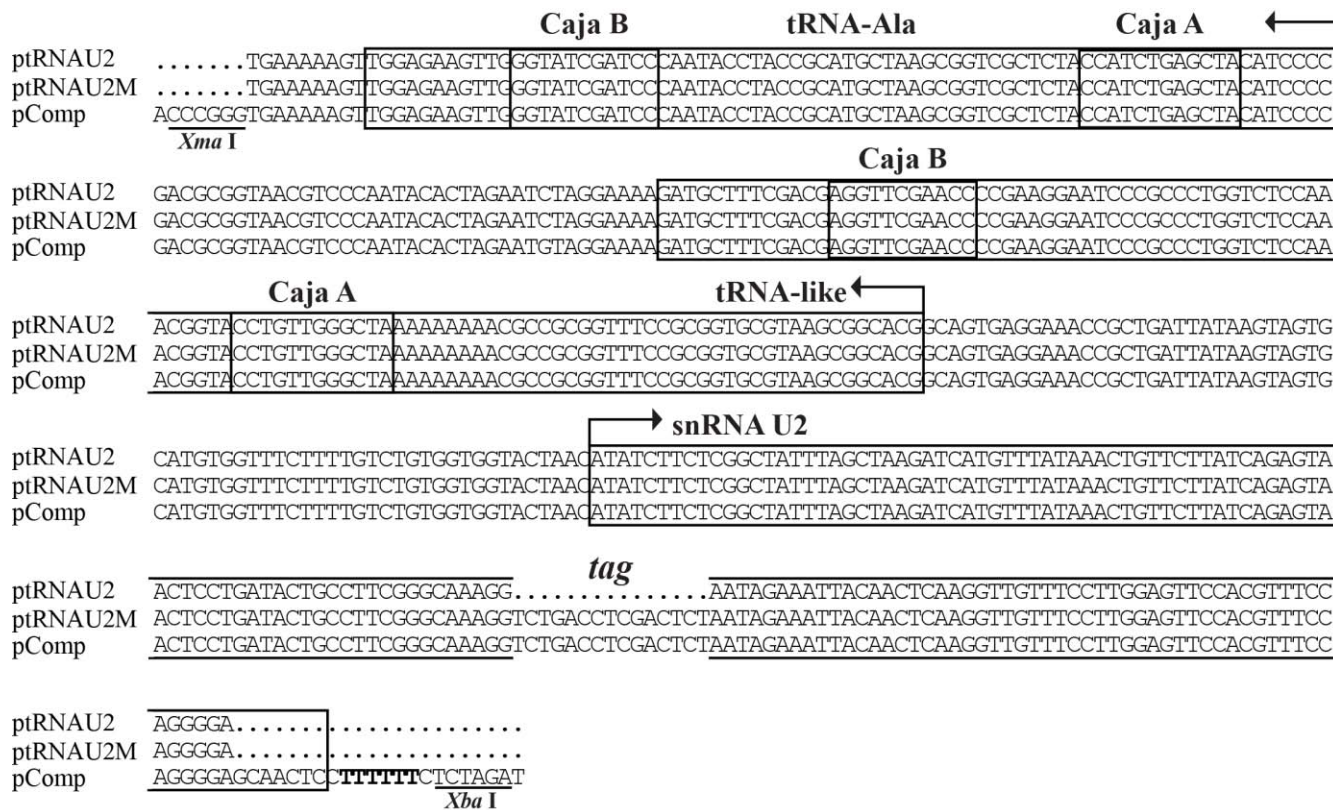


Figura 37. Comparación de las secuencias de los vectores pComp, ptRNAU2M y ptRNAU2. Los puntos señalan que en la secuencia de ptRNAU2 no se encuentra la secuencia *tag*, también indican la ausencia de 7 nt del extremo 3' de U2 y su señal de término de la transcripción en los vectores ptRNAU2 y ptRNAU2M. La señal de término de la transcripción y los sitios de corte de las enzimas *Xma I* y *Xba I* agregados en la secuencia de pComp se muestran en negritas y en cursivas, respectivamente. Los genes de snRNA U2, tRNA-like, tRNA-Ala, así como las cajas A y B se muestran dentro de recuadros. Las flechas indican el inicio y dirección de la transcripción.

6.7.4. Obtención de vectores con deleciones de la región flanqueante 5' de U2

Después de haber insertado la secuencia *tag* en el gen del snRNA U2, así como la señal de término de la transcripción, se generaron cuatro construcciones más en el vector pGEM-T Easy. En los nuevos vectores se eliminaron progresivamente las secuencias localizadas río arriba del gen del snRNA U2, las cuales contienen los elementos promotores probables (Fig. 38).

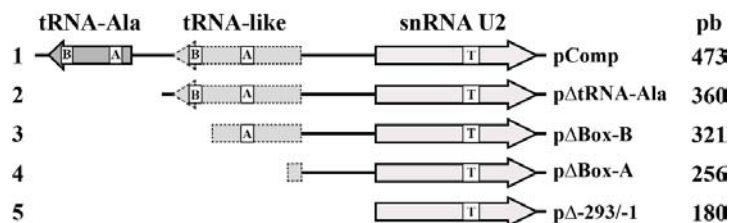


Figura 38. Esquema de los insertos de los plásmidos con deleciones de la región flanqueante 5' del snRNA U2. Los genes se representan con flechas. Las cajas A y B se muestran dentro de los genes de tRNA-Ala y tRNA-like. La secuencia *tag* introducida en el snRNA U2 se indica con una T. Con fines prácticos, los plásmidos se enumeraron del 1 al 5. Los tamaños de los insertos (pb) se muestran del lado derecho de la figura. El esquema se encuentra a escala.

Los plásmidos con deleciones de la región flanqueante 5' del snRNA U2 se nombraron p Δ tRNA-Ala, p Δ Box-B, p Δ Box-A y p Δ -293/-1 (Fig. 38). En el vector p Δ tRNA-Ala se deletó el gen del tRNA-Ala, quedando el tRNA-like completo y el gen del snRNA U2. En el plásmido p Δ Box-B se eliminó la caja B del tRNA-like, además del tRNA-Ala. Asimismo, en el plásmido p Δ Box-A se deletó la secuencia completa del tRNA-like (además del tRNA-Ala), quedando sólo el gen del snRNA U2 y un fragmento de 76 pb río arriba de éste. Finalmente el vector p Δ -293/-1 contiene únicamente el gen del snRNA U2 sin ninguna secuencia flanqueante 5'. Para la obtención de estos plásmidos se amplificaron por PCR diferentes fragmentos de DNA usando el vector pComp como molde y el oligonucleótido antisentido endLmjU2-XbaI-R (Fig. 31). Los oligonucleótidos sentido fueron: U2tRNA-like-XmaI-F para p Δ tRNA-Ala, LmjU2DcajaB-XmaI-F para p Δ Box-B, LmjU2DcajaA-XmaI-F para p Δ Box-A y U2Sa-XmaI-F para p Δ -293/-1 (Fig. 31). Los productos de PCR mostraron los tamaños esperados de 180 pb para el inserto Δ -293/-1, 256 pb para Δ Box-A, 321 pb para Δ Box-B y 360 pb para Δ tRNA-Ala (Fig. 39).

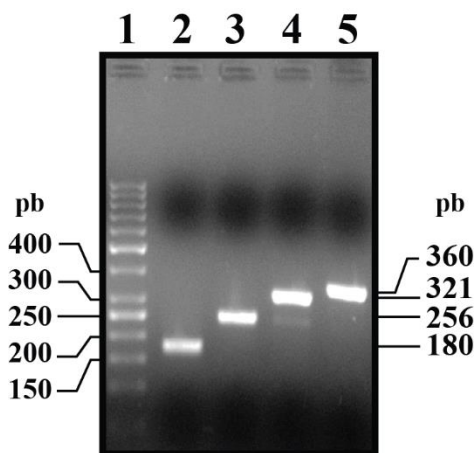


Figura 39. Delección por PCR de secuencias flanqueantes 5' del snRNA U2 de *L. major*. Insertos Δ -293/-1 (carril 2), Δ Box-A (carril 3), Δ Box-B (carril 4) y Δ tRNA-Ala (carril 5). En el carril 1 se corrió un patrón de peso molecular (GeneRuler 50 bp DNA Ladder, Fermentas). Gel de agarosa al 1.5%.

Los insertos amplificados por PCR se purificaron de geles de agarosa y se ligaron al vector pGEM-T Easy. Con las ligaciones se transformaron células competentes de *E. coli*, y se aislaron plásmidos de varias colonias. Los plásmidos se digirieron con las enzimas *Xma* I y *Xba* I y se corrieron en un gel de agarosa. En todos los casos se observaron bandas con los tamaños esperados de: 360 pb

(pΔtRNA-Ala), 321 pb (pΔBox-B), 256 pb (pΔBox-A) y 180 pb (pΔ-293/-1) (Fig. 40). Posteriormente, los plásmidos se mandaron a secuenciar. El análisis de la secuencia de los plásmidos confirmó que el fragmento clonado era el esperado (Fig. 41).

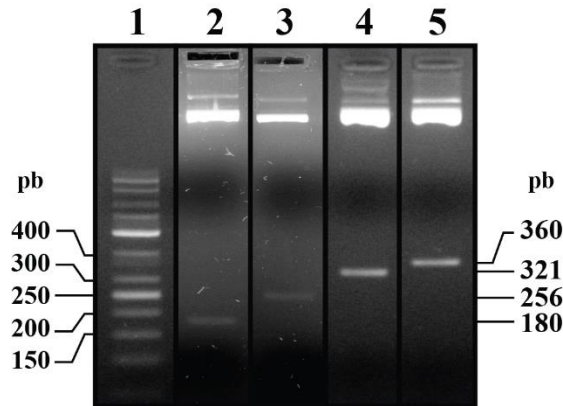


Figura 40. Análisis por restricción de los vectores con deleciones de la región flanqueante 5' del snRNA U2. Se muestran digestiones con *Xma* I y *Xba* de los vectores pΔ-293/-1 (carril 2), pΔBox-A (carril 3), pΔBox-B (carril 4) y pΔtRNA-Ala (carril 5). En el carril 1 se cargó un marcador de peso molecular (GeneRuler 50 bp DNA Ladder, Fermentas) Gel de agarosa al 1.5%.

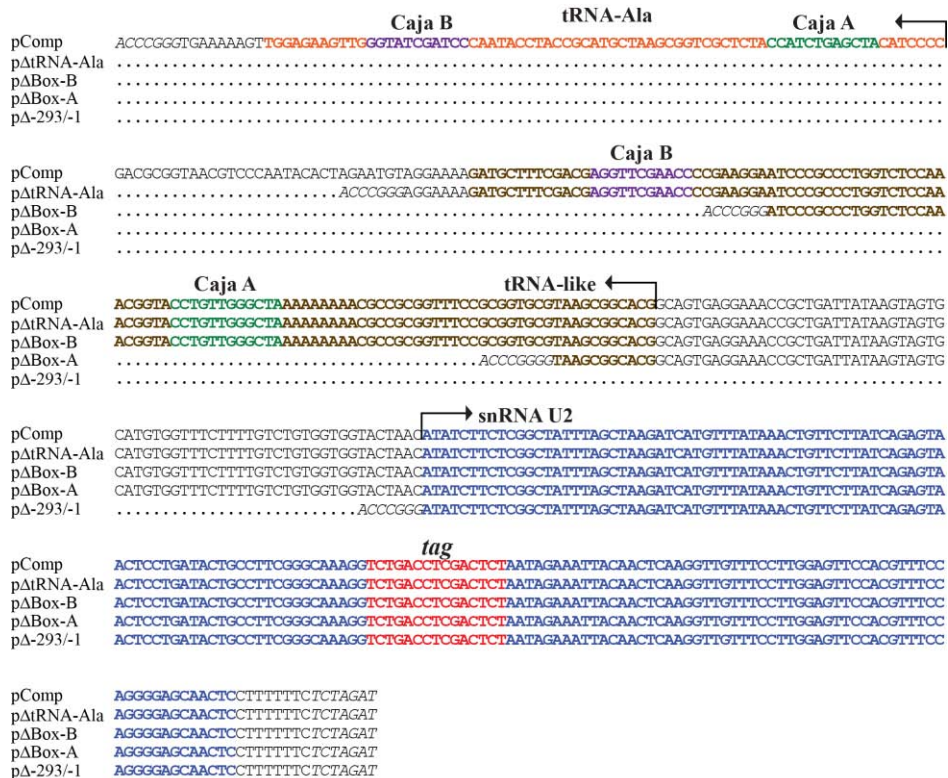


Figura 41. Comparación de la secuencia de los vectores pComp, pΔtRNA-Ala, pΔBox-B, pΔBox-A y pΔ-293/-1. Se comparan los insertos de los vectores que contienen deleciones de la región flanqueante del snRNA U2 (secuencia en azul). En café se muestra la secuencia del tRNA-like, en naranja al tRNA-Ala, en verde y morado a las cajas A y B de ambos genes, respectivamente. La secuencia tag introducida en el snRNA U2 se muestra en rojo. En cursivas se muestran los sitios de restricción reconocidos por las enzimas *Xma* I y *Xba* I, los cuales fueron añadidos en las reacciones de PCR. Los puntos indican que en esa posición no se encuentra el nucleótido comparado.

6.7.5. Obtención de vectores con mutaciones

Para analizar independientemente la participación de las cajas A y B (ubicadas dentro de los genes de tRNA-Ala y tRNA-like) en la transcripción del snRNA U2, se crearon siete construcciones más que contenían mutaciones en las cajas A y B de los genes de tRNA-Ala o tRNA-like, o bien en los primeros nucleótidos de la región codificante del snRNA U2 (Fig. 42).

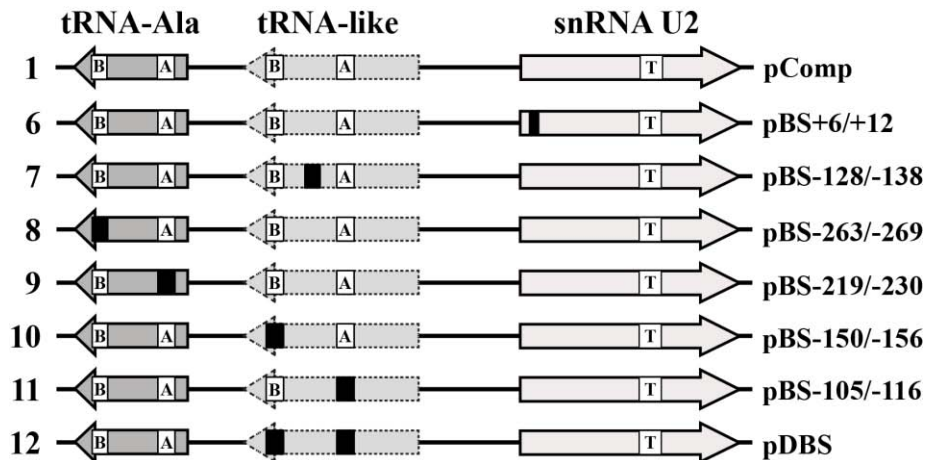


Figura 42. Esquema de los insertos de los plásmidos con mutaciones por sustitución de bases. Los genes se representan con flechas. Las cajas A y B se muestran dentro de los genes de tRNA-Ala y tRNA-like. Las regiones mutadas se representan con cuadros negros. La secuencia *tag* introducida en el snRNA U2 se indica con una T. Con fines prácticos, los plásmidos se enumeraron: 1 a pComp; y del 6 al 12 a los plásmidos con mutaciones. El tamaño de los insertos en todos los plásmidos es de 473 pb. El esquema se encuentra a escala.

Dichos plásmidos se nombraron pBS+6/+12, pBS-128/-138, pBS-263/-269, pBS-219/-230, pBS-150/-156, pBS-105/-116 y pDBS. En el vector pBS+6/+12 (6 en la Fig. 42) se mutaron las bases +6 a +12 (en relación al SIT, +1) de la región codificante del snRNA U2, mientras que la región flanqueante 5' se mantuvo con la secuencia silvestre; en el plásmido pBS-128/-138 (7 en la Fig. 42) se mutó una región ubicada entre las cajas A y B del tRNA-like (posiciones de -128 a -138); en los plásmidos pBS-263/-269 y pBS-219/-230 (8 y 9 en la Fig. 42) se mutaron las secuencias de la caja B y A del tRNA-Ala, respectivamente; en los plásmidos pBS-150/-156 y pBS-105/-116 (10 y 11 en la Fig. 42) las cajas B y A del tRNA-like se mutaron por sustitución de bases. Finalmente el plásmido pDBS (12 en la Fig. 42) contiene una doble mutación, tanto en la caja A como en la caja B del tRNA-like.

Para generar estos plásmidos se usó un kit de mutagénesis dirigida (QuikChange Lightning Site-Directed Mutagenesis Kit) además del vector pComp y los siguientes oligonucleótidos: 5U2LmjmutFw y 5U2LmjmutRv se usaron para generar a pBS+6/+12, TLBTWU2LmjmutFw y TLBTWU2LmjmutRv para pBS-128/-138, BalaU2LmjmutFw y BalaU2LmjmutRv para pBS-263/-269, AalaU2LmjmutFw y AalaU2LmjmutRv para pBS-219/-230, BlikeU2LmjmutFw y BlikeU2LmjmutRv para pBS-150/-156, y finalmente a AlikeU2LmjmutFw junto con AlikeU2LmjmutRv para pBS-105/-116 (Fig. 31). La construcción pDBS, que contiene una doble mutación, se generó usando al plásmido pBS-105/-116 como molde y a los oligonucleótidos BlikeU2LmjmutFw y BlikeU2LmjmutRv (Fig. 31). El plásmido parental se digirió con la enzima *Dpn* I y las siete construcciones recién generadas se transformaron en células ultracompetentes XL 10-Gold®. Las clonas obtenidas se digirieron con las enzimas *Xma* I y *Xba* I, y las restricciones se analizaron mediante electroforesis. Todas las clonas presentaron el patrón de digestión esperado: una banda de ~3 kb que corresponde al vector, y una banda de 473 pb que corresponde al fragmento liberado (Fig. 43). El análisis de secuencia de las clonas confirmó que se habían insertado las mutaciones deseadas en los sitios esperados (Fig. 44).

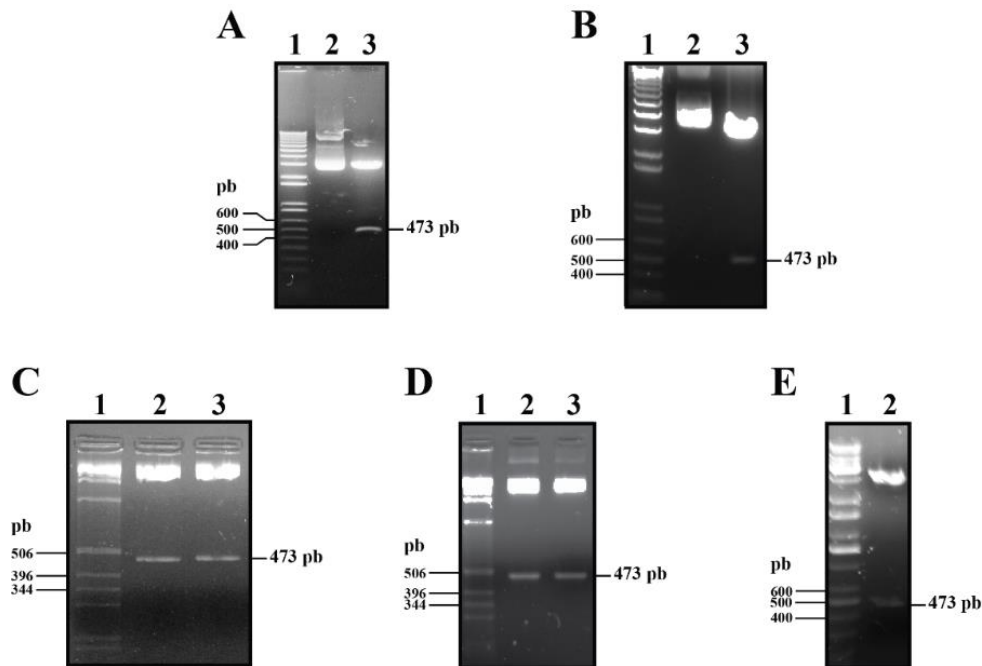


Figura 43. Análisis por restricción de los vectores que contienen mutaciones por sustitución de bases. Se muestran digestiones con *Xma* I y *Xba* I de los siguientes vectores: A) pBS+6/+12, carril 3; B) pBS-128/-138, carril 3, C) pBS-263/-269 y pBS-219/-230, carriles 2 y 3, respectivamente; D) pBS-150/-156 y pBS-105/-116, carriles 2 y 3, respectivamente; y E) pDBS, carril 2. En el carril 2 de los paneles A y B se analizaron los plásmidos sin digerir pBS+6/+12 y pBS-128/-138,

respectivamente. En el carril 1 de todos los paneles se cargaron marcadores de peso molecular: A) y B) 1 Kb Plus DNA Ladder, Invitrogen; C) y D) 1kb DNA Ladder, Invitrogen; E) KAPA Universal Ladder, Kapa Biosystems). Geles de agarosa al 1.5%.

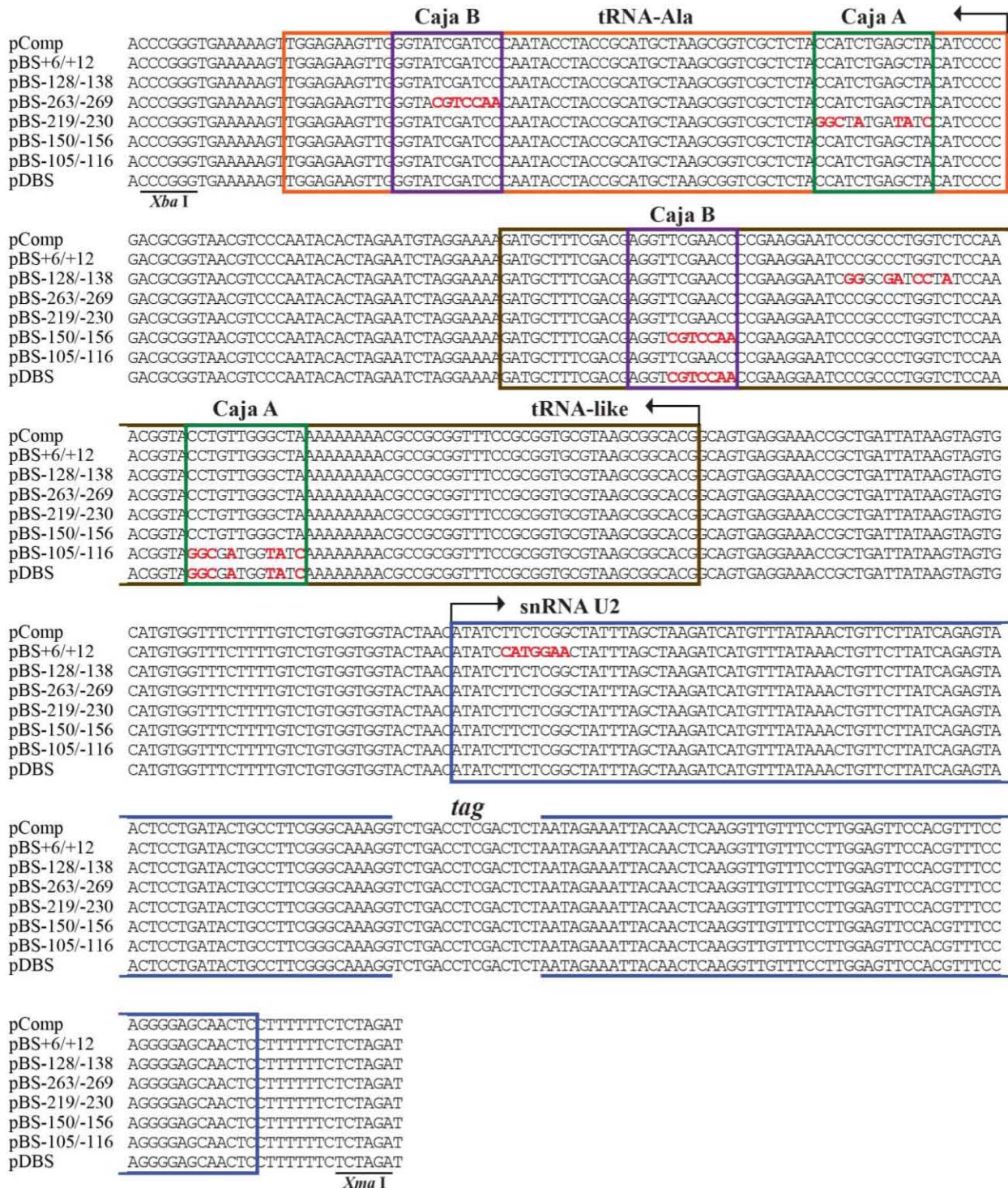


Figura 44. Comparación de las secuencias de los vectores que contienen mutaciones. Se comparan los insertos de los vectores pBS+6/+12, pBS-128/-138, pBS-263/-269, pBS-219/-230, pBS-150/-156, pBS-105/-116 y pDBS. Los nucleótidos substituidos mediante mutagénesis dirigida se muestran en rojo dentro de las secuencias de los vectores. Con cajas de colores se muestran los genes del snRNA U2 (azul), tRNA-like (café) y tRNA-Ala. Las cajas A y B (del tRNA-like y tRNA-Ala) se muestran con recuadros de color verde y morado, respectivamente. Las flechas indican el inicio y dirección de la transcripción. Los sitios reconocidos por *Xma* I y *Xba* I que flanquean a todos los insertos están indicados.

6.7.6. Actividad transcripcional *in vivo* del promotor del snRNA U2 en *L. major*

6.7.6.1. Análisis de la abundancia del snRNA U2 exógeno mediante *primer extension*

Una vez que se contó con las diferentes construcciones que contienen elementos de la probable región promotora del gen del snRNA U2, se decidió primero caracterizar la actividad transcripcional de los plásmidos que contienen deleciones de la región flanqueante 5' por medio de ensayos de *primer extension*. Para evaluar la presencia de un posible elemento promotor intragénico, también se analizó el efecto causado por las mutaciones de los primeros nucleótidos del gen (posiciones +6 a +12), en la abundancia de transcritos exógenos. Para ello, se realizaron transfecciones transitorias en las que se introdujeron los plásmidos pComp, p Δ tRNA-Ala, p Δ Box-B, p Δ Box-A, p Δ -293/-1, pBS+6/+12 y pBS-128/-138 a promastigotes de *L. major* que se encontraban en fase media de crecimiento logarítmico. Después de 24 horas se extrajo el RNA total de las células y éste se usó como molde en ensayos de *primer extension* realizados con el oligonucleótido LmjU2tag-Rev. Debido a que este oligonucleótido es complementario a la secuencia *tag* (Fig. 31), se analizaron exclusivamente a los transcritos producidos a partir de los vectores transfectados (transcritos exógenos). El cDNA generado se analizó mediante electroforesis en un gel de poliacrilamida junto con reacciones de secuenciación manual del gen del snRNA U2 exógeno, las cuales se hicieron con el mismo oligonucleótido (LmjU2tag-Rev) (Fig. 45).

Cuando se transfectó el plásmido pComp, el cual contiene la región 5' flanqueante completa, dos productos del *primer extension* fueron detectados (Fig. 45, carril 1). El producto más grande se mapeó en una A (T en la secuencia complementaria) que corresponde al primer nucleótido transcrito del gen (SIT, +1), mientras que el otro producto corresponde a una A localizada en la posición +3 (marcado con un asterisco en la Fig. 45) que probablemente represente una pausa de la transcriptasa reversa. De manera esperada, estas bandas no se presentaron en células transfectadas sin ningún plásmido (control mock), ni en células transfectadas con un plásmido no relacionado (pLMRIB). También se observó una banda inespecífica presente en todas las reacciones, la cual sirvió como control de carga (marcada con # en Fig. 45). La deleción del gen tRNA-Ala (p Δ tRNA-Ala) no afectó la abundancia de las bandas del snRNA U2

(Fig. 45, carril 2). En contraste, la eliminación de la caja B del tRNA-like (p Δ Box-B, carril 3), la delección de la caja A del tRNA-like (p Δ Box-A, carril 4) y la remoción de la secuencia flanqueante 5' completa (p Δ -293/-1, carril 5) redujeron drásticamente los transcritos del snRNA U2 (Fig. 45). Por lo tanto, estos resultados demostraron que secuencias ubicadas dentro del tRNA-like son necesarias para la transcripción del snRNA U2 en *L. major*.

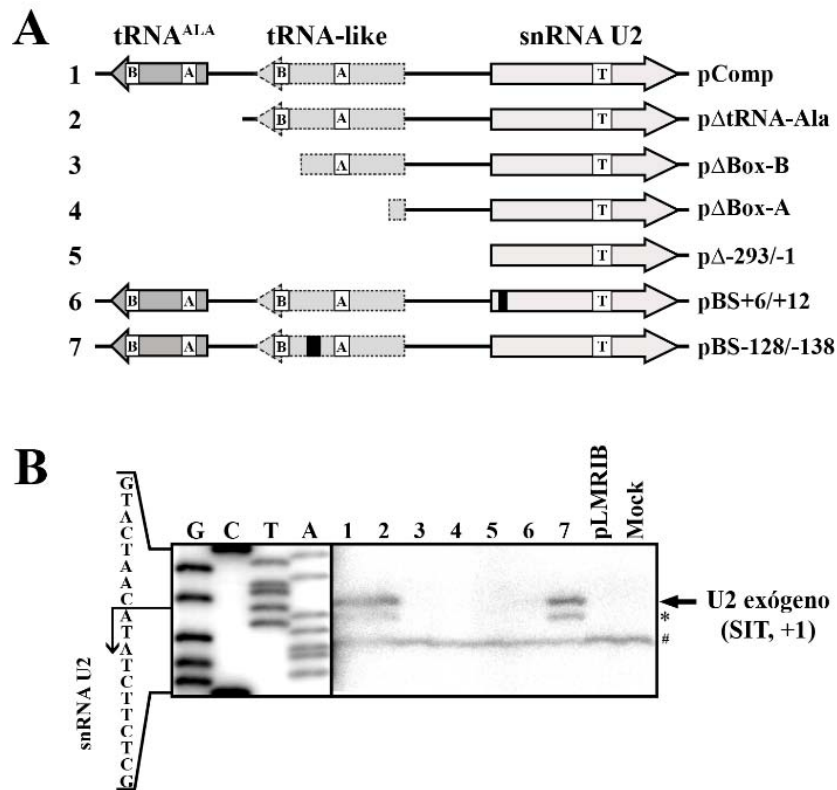


Figura 45. Identificación de elementos de secuencia requeridos para la expresión de U2 por medio de *primer extension*. A) Representación esquemática de las construcciones usadas para el análisis por *primer extension*. Las etiquetas corresponden a las de las figuras 38 y 42. B) Análisis de elementos promotores del snRNA U2 por *primer extension*. Las construcciones 1-7 se transfectaron en células de *L. major* y su RNA total fue extraído 24 horas después. El análisis por *primer extension* se realizó con el oligonucleótido LmjU2tag-Rev, el cual reconoce a la secuencia *tag* dentro de los transcritos de U2 exógenos (T en panel A). Los productos fueron separados en un gel de poliacrilamida al 8% junto con una escalera de secuencia obtenida con el mismo oligonucleótido. Los carriles 1 a 7 corresponden al RNA de células de células transfectadas con los vectores 1 a 7, respectivamente. Como control negativo se usó RNA de células transfectadas sin plásmido (mock), o con un vector no relacionado (pLMRIB [142]). Con un asterisco se señalan productos de *primer extension* mapeados en la posición +3. Con # se señala una banda inespecífica presente en todas las muestras, la cual se usó como control de carga.

De manera similar, cuando se mutaron los nucleótidos de las posiciones +6 a +12 dentro del gen del snRNA U2 la abundancia de transcritos exógenos se vio disminuida en gran medida (Fig. 45, carril 6). Este resultado indica que los primeros nucleótidos del snRNA U2 son requeridos para la expresión

del gen. De forma opuesta, las mutaciones generadas en una región ubicada entre las cajas A y B del tRNA-like no afectaron la abundancia del snRNA U2 (Fig. 45, carril 7).

Para determinar de manera más precisa la abundancia de los transcritos del snRNA U2 exógenos, el RNA de los parásitos transfectados transitoriamente se analizó por medio de RT-qPCR. Al igual que en los ensayos de *primer extension*, para amplificar únicamente a los transcritos marcados con la secuencia *tag*, se usó al oligonucleótido LmjU2tag-rev (Fig. 31).

Como lo muestra la figura 46, los niveles del snRNA U2 se redujeron drásticamente (90-93%) cuando la caja B del tRNA-like y las secuencias río abajo de ésta fueron eliminadas (comparar barra 1 con 3 a 5). Estos resultados concuerdan con los resultados obtenidos por *primer extension* (Fig. 46, carriles 3 a 5) y por lo tanto, confirman que el gen del tRNA-like es necesario para la expresión *in vivo* del snRNA U2 en *L. major*. La importancia de los primeros nucleótidos del gen también fue evaluada por RT-qPCR. Los niveles de U2 se vieron fuertemente afectados cuando dichos nucleótidos (de +6 a +12) se mutaron ya que la transcripción del gen cayó a un 17% (comparar barra 1 con 6). Este resultado también concuerda con los resultados obtenidos por *primer extension* e indica que los primeros nucleótidos del gen son necesarios para su transcripción *in vivo*.

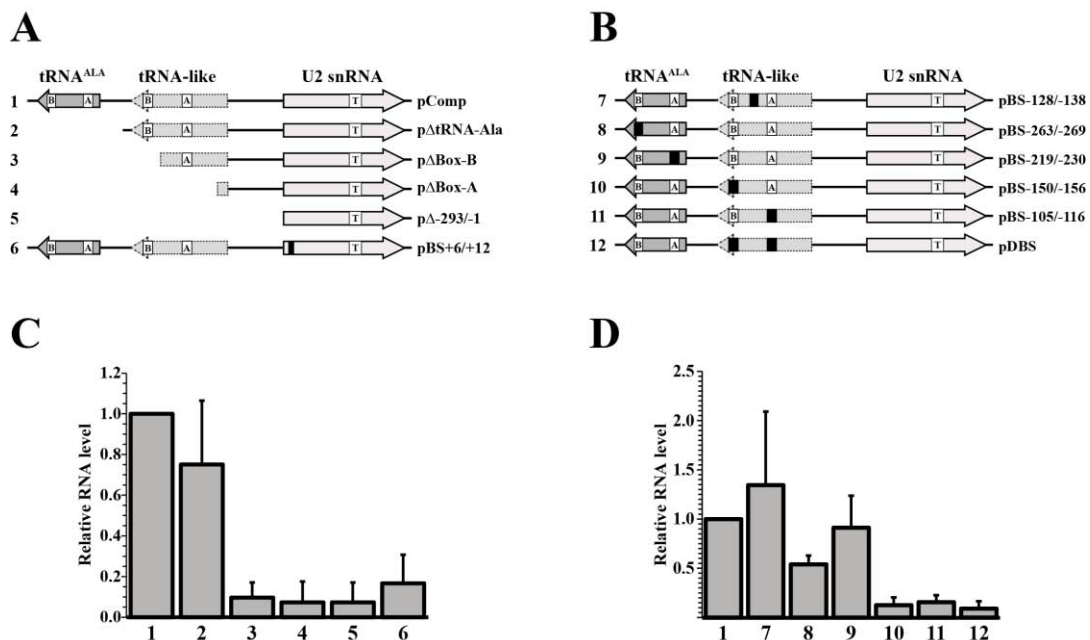


Figura 46. Identificación de elementos de secuencia requeridos para la expresión de U2 por medio de RT-qPCR. A) Esquemas de los vectores usados para el análisis por RT-qPCR en C). B) Esquemas de los vectores usados para el análisis

por RT-qPCR en D). Las etiquetas en A) y B) corresponden a las de las figuras 38 y 42. C) Efecto transcripcional de deleciones progresivas de la región flanqueante 5' del gen del snRNA U2 por RT-qPCR. Los niveles relativos de U2 fueron analizados en células transfectadas con las construcciones 1 a 6. D) Ensayos de RT-qPCR para determinar el efecto de mutaciones por sustitución de bases en la expresión del snRNA U2. Los niveles relativos de U2 fueron analizados en células transfectadas con las construcciones 1 y 7 a 12, las cuales contienen sustitución de bases en las cajas A y B del tRNA-like y tRNA-Ala. Las barras de error indican desviaciones estándar de tres o más réplicas biológicas.

Para analizar de manera independiente el papel que juegan las cajas A y B del tRNA-like en la transcripción del snRNA U2, los plásmidos que contienen mutaciones en dichas cajas se transfectaron transitoriamente en promastigotes procíclicos de *L. major*. La abundancia de los transcritos de U2 exógenos, presentes en estos parásitos, se analizó por medio de RT-qPCR. Se observó que las mutaciones en la caja B del tRNA-like (Fig. 46, barra 10), en la caja A del tRNA-like (Fig. 46, barra 11), y en ambas cajas del mismo gen (Fig. 46, barra 12), disminuyen drásticamente los niveles del snRNA U2 a un 13, 16 y 9%, respectivamente. En contraste, un ligero aumento en la transcripción fue detectado cuando se usó al plásmido control que contiene mutaciones en la secuencia ubicada entre las cajas A y B del tRNA-like, por lo tanto, esta secuencia no contiene elementos promotores (Fig. 46, barra 7).

Un resultado inesperado fue que la deleción del gen del tRNA-Ala, localizado a 211 pb río arriba de U2, produjo una reducción del 25% en los niveles del snRNA U2 (Fig. 46, barra 2). Para explorar este efecto, los plásmidos que contienen mutaciones en las cajas A y B del tRNA-Ala fueron transfectados y su efecto en la transcripción de U2 se analizó por medio de RT-qPCR. De manera interesante, cuando la caja B del tRNA-Ala fue mutada, la expresión del snRNA U2 se redujo considerablemente a un 54% (Fig. 46, barra 8). Sin embargo, las mutaciones introducidas en la caja A del mismo gen solamente tuvieron un pequeño impacto en los niveles del snRNA U2 (Fig. 46, barra 9). Por lo tanto, estos resultados demuestran que la caja B del gen de tRNA-Ala, localizado río arriba del tRNA-like, es importante para la transcripción *in vivo* del snRNA U2 de *L. major*.

6.7.6.2. Integridad de los transcritos de U2 exógenos.

Debido a que los transcritos del snRNA U2 exógenos contienen una secuencia *tag*, se exploró la posibilidad de que esta inserción propiciara la degradación de los transcritos de U2 exógenos. Para ello se realizaron ensayos tipo Northern blot usando RNA proveniente de células transfectadas con los

plásmidos pComp, que contiene la región completa 5' flanqueante de U2, y pBS+6/+12, el cual contiene mutaciones en los primeros nucleótidos del gen. Como sonda se utilizó el oligonucleótido LmjU2tag-Rev, el mismo que se usó en ensayos de *primer extension* y RT-qPCR para analizar la abundancia del U2. Cuando el plásmido pComp fue transfectado, se obtuvo una señal claramente visible del tamaño esperado de alrededor de ~157 nt (Fig. 47, carril 3). Esta banda es específica del transcrito del snRNA U2 exógeno, pues en células transfectadas con sin plásmido (control mock) o con el plásmido no relacionado pLMRIB no se observó dicha señal (Fig. 47, carriles 1 y 2). Por lo tanto, el transcrito de U2 exógeno mantiene su integridad *in vivo*, y a juzgar por su tamaño estimado, parece que su transcripción termina en el tracto de Ts ubicado río abajo del gen, al igual que el snRNA U2 endógeno.

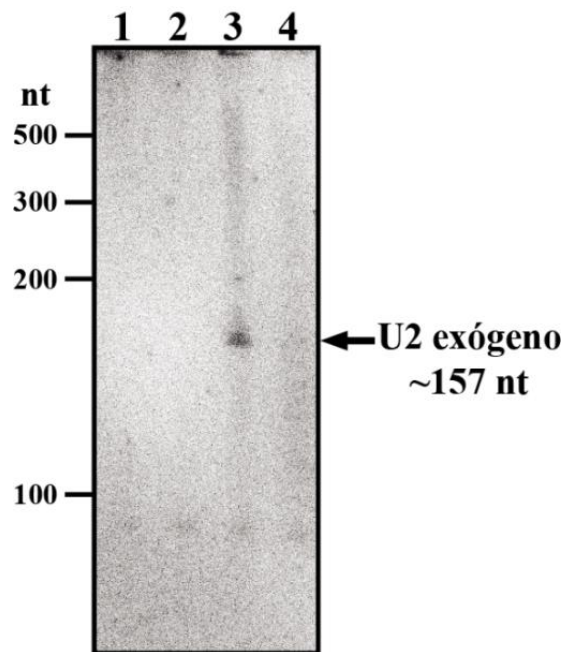


Figura 47. Análisis tipo Northern blot del snRNA U2 exógeno. El experimento se realizó con RNA de células transfectadas sin ningún vector (control mock) (carril 1), o con pLMRIB (carril 2), pComp (carril 3) o pBS+6/+12 (carril 4). La sonda utilizada fue el oligonucleótido LmjU2tag-Rev, que reconoce específicamente la secuencia *tag*.

Cuando se transfectó al plásmido que contiene mutaciones en los primeros nucleótidos de U2, se observó un efecto similar al obtenido con los ensayos de *primer extension* y RT-qPCR, pues la abundancia del snRNA U2 exógeno se vio severamente reducida (Fig. 47). Por lo tanto, este resultado confirma que los primeros nucleótidos son importantes para la transcripción del snRNA U2.

6.8. Análisis transcripcional del tRNA-like asociado a U2 en *L. major*

La secuencia del tRNA-like presenta algunas similitudes con genes de tRNAs clásicos; entre éstas se encuentra: la presencia de cajas A y B, una distancia conservada entre ambas cajas de ~30 pb, la presencia de la señal de término de la transcripción de la RNAP III río abajo del gen y de manera más relevante, sus cajas juegan un papel importante en la síntesis de moléculas transcritas por la RNAP III. Estas características comunes sugieren que el tRNA-like podría ser transcrito *in vivo*.

Para determinar si el tRNA-like asociado al snRNA U2 de *L. major* es transcrito *in vivo*, se llevaron a cabo experimentos tipo Northern blot con RNA total extraído de promastigotes silvestres (WT) de *L. major*. La sonda utilizada en este experimento consistió de un oligonucleótido específico (5' LmjU2DcajaB) que es complementario a una región ubicada entre las cajas A y B del tRNA-like (Fig. 48, panel C). Una señal de alrededor de 109 nt fue detectada, indicando que el tRNA-like es transcrito en *L. major* (Fig. 48, panel A). Se observaron también otras bandas más grandes de menor intensidad que pueden representar transcritos largos del tRNA-like o bandas inespecíficas. Como controles del experimento se usaron sondas complementarias al snRNA U2 y al tRNA-Ala, detectándose bandas del tamaño esperado de alrededor de 143 y 147 nt, respectivamente (Fig. 48, panel A). El análisis por densitometría de las bandas observadas mostró que el tRNA-like es al menos 10 veces menos abundante que snRNA U2.

El sitio de inicio de la transcripción del tRNA-like se mapeó mediante *primer extension* (Fig. 48 panel B). La banda más abundante correspondió a un residuo de C (indicado por un asterisco rojo en la Fig. 48, panel B) localizado a 65 pb río arriba del snRNA U2 (Fig. 48, panel C). Si se asume que la transcripción del tRNA-like termina en el tracto de Ts encontrado río abajo de este SIT (Fig. 48, panel C), el transcrito del tRNA-like debería de ser de 109 nt, lo cual concuerda con el tamaño del transcrito observado en el ensayo tipo Northern blot (Fig. 48, panel A). Las bandas más pequeñas detectadas en el *primer extension* (73, 80 y 94 nt) pueden representar paros naturales de la transcriptasa reversa, mientras

que las bandas más grandes (120, 122 y 127 nt) pueden corresponder a SIT alternativos (asteriscos negros en la Fig. 48, panel B; flechas en la Fig. 48, panel C).

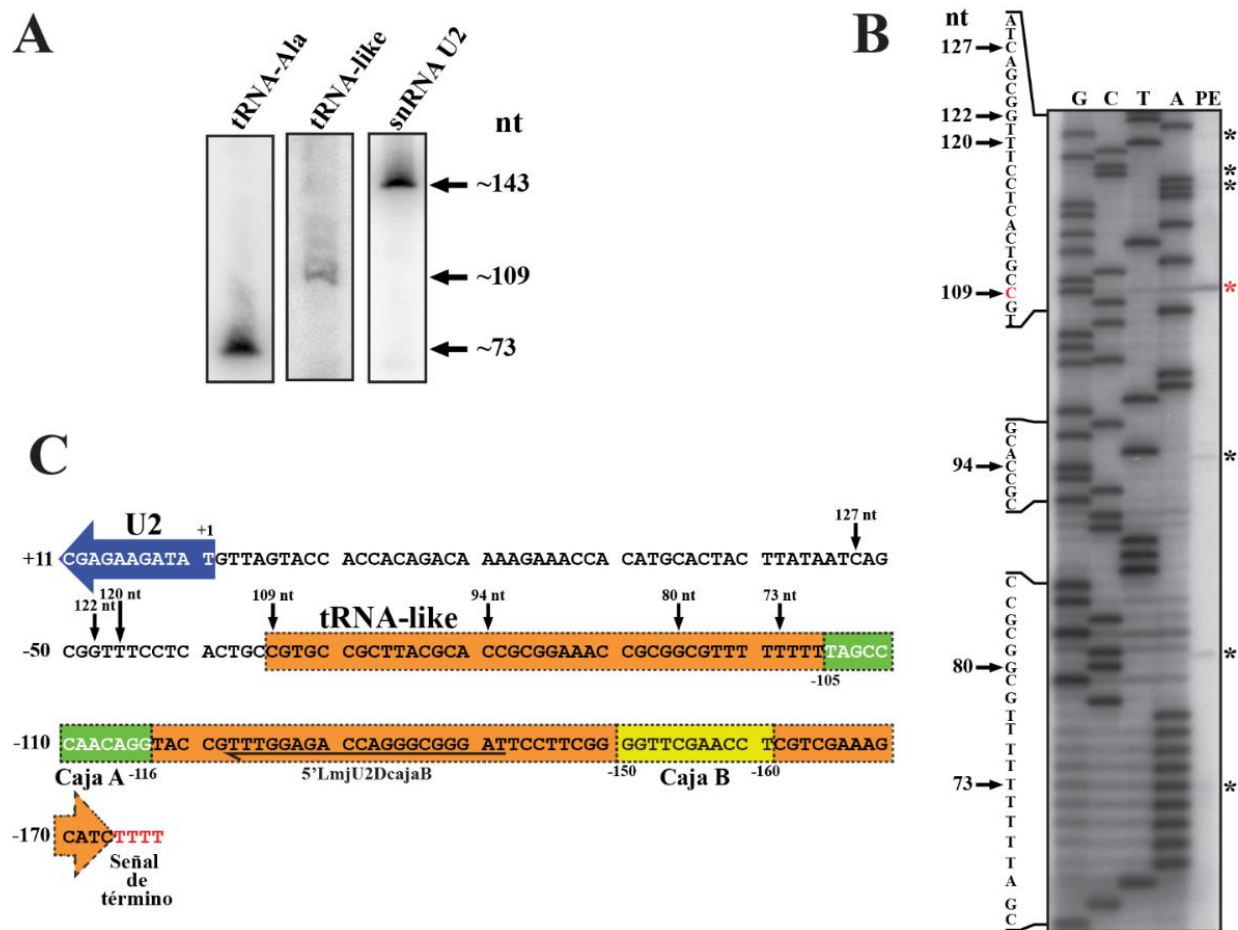


Figura 48. Transcripción del gen del tRNA-like asociado al snRNA U2. A) Análisis tipo Northern blot con sondas específicas para el tRNA-Ala (imagen izquierda), tRNA-like (imagen central) y snRNA U2 (imagen derecha). B) Mapeo del extremo 5' de moléculas de tRNA-like en células silvestres de *L. major*. RNA total fue sometido a un análisis de *primer extension* con el oligonucleótido 5'LmjU2DcajaB (mostrado en el panel C), los productos fueron separados en un gel de poliacrilamida al 8% (carril PE) junto con una escalera de secuencia obtenida con el mismo oligonucleótido (carriles G, C, T y A). Los productos de *primer extension* se marcan con asteriscos a la derecha; el asterisco rojo señala al producto más abundante. Secuencias parciales del tRNA-like se muestran a la izquierda, indicándose con flechas las posiciones de los extremos 5' y los tamaños esperados de las moléculas de tRNA-like (del extremo 5' mapeado a la señal de término). C) Secuencia del tRNA-like. El gen se muestra con una flecha naranja con el contorno punteado que indica la dirección de la transcripción. Los extremos 5' mapeados por *primer extension* se señalan con flechas, indicándose la longitud esperada del transcrito. El producto de *primer extension* más abundante fue considerado el SIT del gen (C, 109 nt). Las cajas A y B se muestran en verde y amarillo, respectivamente. Los primeros once nt del gen del snRNA U2 se muestran como referencia. Los números en la secuencia se muestran en relación al SIT (+1) del snRNA U2. Para facilitar el análisis de los resultados, la secuencia se presenta como aparece en la base de datos de TriTrypDB, la cual es la reversa complementaria de la mayoría de las secuencias mostradas en esta Tesis.

Para determinar si el tRNA-like de 109 nt se puede plegar en la estructura de hoja de trébol característica de los tRNAs clásicos, su estructura secundaria probable se obtuvo usando el programa

RNAfold. Como se puede observar en la figura 49, la estructura secundaria predicha no se asemeja a la estructura secundaria de los tRNAs clásicos, y parece ser menos estable.

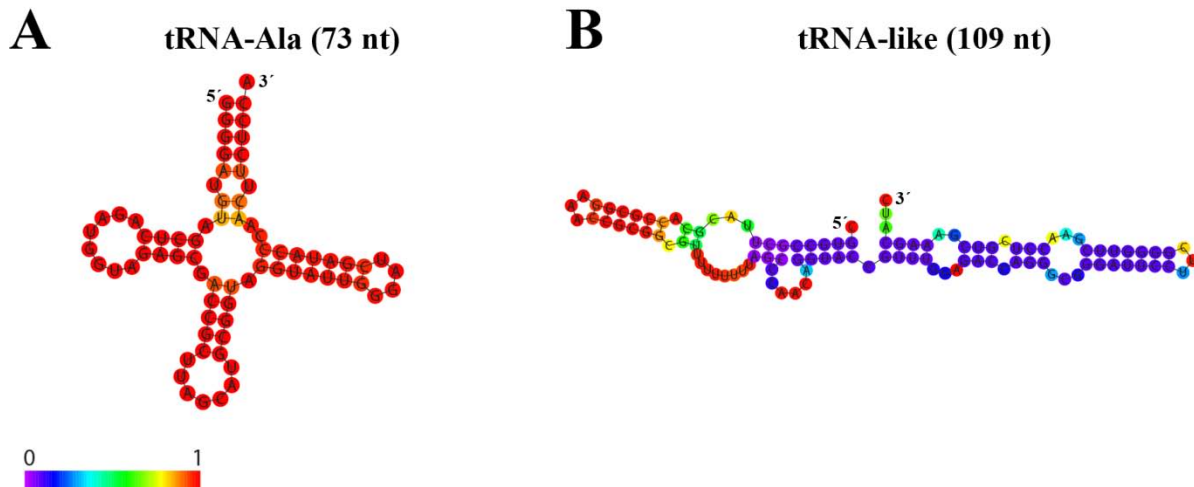


Figura 49. Posible estructura secundaria del tRNA-like. Las posibles estructuras secundarias del tRNA-Ala (A) y tRNA-like (B) se obtuvieron con el programa RNAfold. La escala de color indica las probabilidades de apareamiento de bases en donde rojo es muy probable y azul es poco probable.

6.9. Análisis de las secuencias de las cajas A y B de tRNAs-like asociados a snRNAs U2 de tripanosomátidos

El análisis de las secuencias de las cajas A (TAGCCCAACAGG) y B (GGTTCGAACCT) del tRNA-like asociado al snRNA U2 en *L. major* reveló que son similares a las cajas A (TAGCTCAGATGG) y B (GGATCGATACC) del tRNA-Ala. Para comparar las secuencias de estos elementos de otros tripanosomátidos, se realizó una búsqueda de cajas A y B localizadas río arriba de genes del snRNA U2 en las secuencias anotadas de *L. arabica*, *L. aethiopica*, *L. donovani*, *L. amazonensis*, *L. mexicana*, *L. tarentolae*, *L. infantum*, *L. panamensis*, *L. braziliensis*, *L. enriettii*, *T. brucei*, *T. evansi*, *T. cruzi*, *T. vivax*, *T. congolense*, *E. monterogeii*, *L. pyrrocoris* y *L. seymouri*. Todas las especies de tripanosomátidos analizadas mostraron cajas A y B en las secuencias flanqueantes 5' del snRNA U2, las cuales fueron alineadas independientemente (Fig. 50, panel A) y estos alineamientos se usaron para generar logos de secuencia de ambos elementos (logos superiores en la Fig. 50, panel B). Con fines comparativos, las secuencias de las cajas A y B de los 258 genes de tRNAs clásicos

encontrados en *T. brucei*, *T. cruzi* y *L. major* (TriTryps) [167] fueron alineadas y sus logos de secuencia fueron generados (logos inferiores en la Fig. 50, panel B).

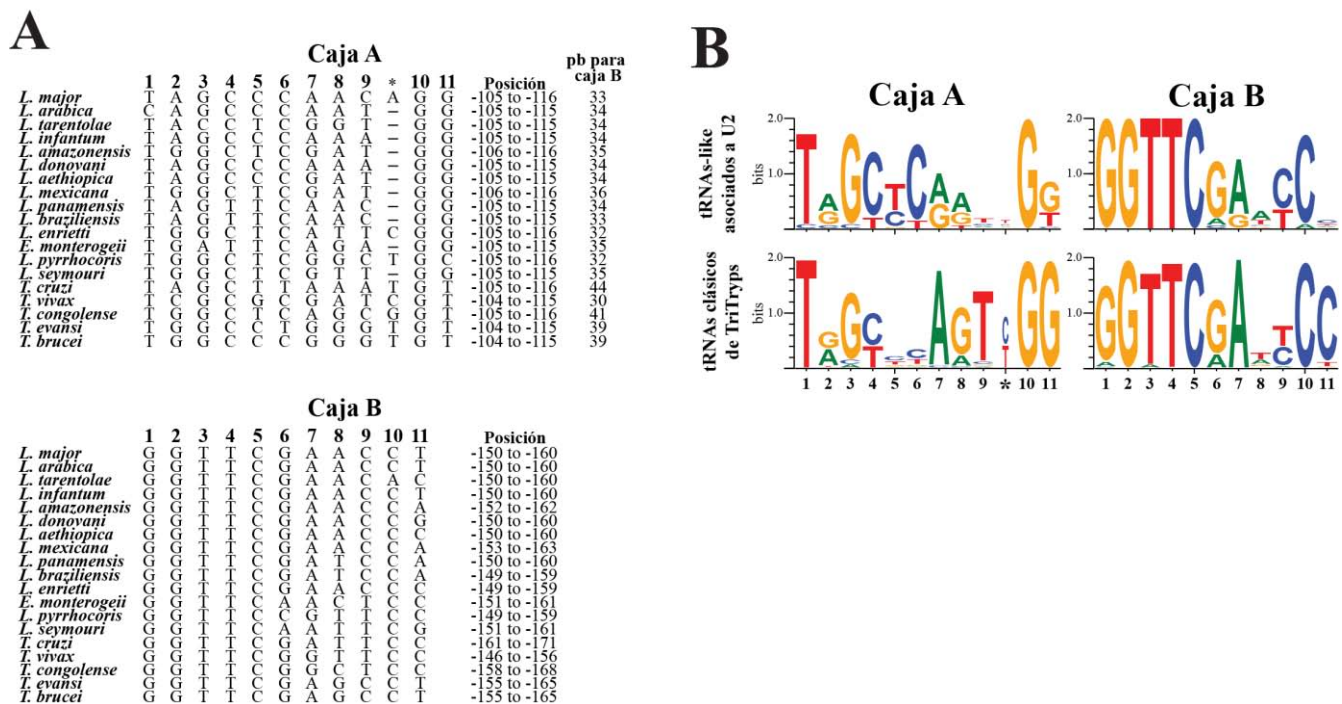


Figura 50. Análisis de secuencias de las cajas A y B de genes de tRNA-like asociados a genes del snRNA U2 en tripanosomátidos. A) Alineamientos de secuencias de cajas A y B de genes de tRNA-like de las siguientes especies de tripanosomátidos: *L. major* (Friedlin), *L. arabica* (LEM1108), *L. tarentolae* (Parrot-TarII), *L. infantum* (JPCM5), *L. amazonensis* (MHOMBR/71973/M2269), *L. donovani* (BPK282A1), *L. aethiopica* (L147), *L. mexicana* (U1103), *L. panamensis* (MHOM/COL/81/L13), *L. braziliensis* (M2904), *L. enrietti* (LEM3045), *E. monterogeei* (LV88), *L. pyrrhocoris* (H10), *L. seymouri* (ATCC 30220), *T. cruzi* (CL Brener Non-Esmeraldo-like, copia del cromosoma 23), *T. vivax* (Y486), *T. congolense* (IL3000), *T. evansi* (STIB 805) y *T. brucei* (DAL972). Todas las secuencias se obtuvieron de la base de datos de TriTrypBD. La posición señalada con un asterisco indica que algunas cajas A contienen un nucleótido extra entre las posiciones 9 y 10. Para cada caja, se indica su posición en relación al SIT del snRNA U2 asociado. Las distancias entre las cajas A y B también se muestran. B) Secuencias consenso de las cajas A y B de genes de tRNA-like asociados a genes del snRNA U2 en tripanosomátidos. Los alineamientos mostrados en el panel A se usaron para generar logos de secuencia con la aplicación WebLogo (logos de arriba). Con fines comparativos, se obtuvieron logos de secuencia adicionales para las cajas A y B de los 258 genes de tRNAs clásicos encontrados en TriTryps [167]. Los cuatro logos se presentan por su contenido de información.

En lo que respecta a los elementos de tRNAs-like, el análisis de secuencia del logo generado para las cajas A muestra que existe un alto grado de conservación de la mayoría de sus bases, con la excepción de la posición marcada con un asterisco (la cual está ausente en la mayoría de especies analizadas) y las bases 8 y 9 (logo superior en la Fig. 50, panel B). La secuencia de la caja B del tRNA-like está aún más conservada, pues las posiciones 1 a 5 (GGTTC) se encuentran presentes en todos los

tripanosomátidos analizados. El resto de las bases también están conservadas aunque en menor grado; las únicas posiciones variables fueron la 8 y 11 (logo superior en la Fig. 50, panel B).

En general se observó una alta conservación de secuencia entre las cajas B de tRNAs-like y tRNAs clásicos, pues la posición 11 es la única que no está conservada (logos de la derecha en la Fig. 50, panel B). En contraste, las cajas A de tRNAs like no están tan conservadas como las cajas B pues las posiciones 5, 6, 7, 9, 11 y (*) son diferentes entre tRNAs-like y tRNAs clásicos (logos de la izquierda en la Fig. 50, panel B).

Es interesante el hecho de que 104 pb separan al snRNA U2 de la caja A del tRNA-like en la mayoría de los tripanosomátidos analizados (Fig. 50, panel A). Las excepciones son *T. brucei*, *T. evansi* y *T. vivax* (donde la distancia es de 103 pb), y *L. amazonensis* junto con *L. mexicana* (105 pb) (Fig. 50, panel A). Adicionalmente, la distancia que separa a las cajas A de las cajas B de tRNAs-like está conservada en *Leishmania*, ya que la distancia fluctúa de 33 pb (en *L. major* y *L. braziliensis*) a 36 pb (en *L. mexicana*). En contraste, las distancias entre las cajas A y B no están conservadas en los tRNAs-like del género *Trypanosoma*, en donde las distancias varían de 30 pb (*T. vivax*) a 44 pb (*T. cruzi*) (Fig. 50, panel A). De esta manera, el análisis de las cajas A y B de los tRNAs-like de tripanosomátidos muestra que mientras las cajas A son más divergentes en secuencia, su posición en relación al SIT de U2 es casi invariable. En cambio, las secuencias de las cajas B de tRNAs-like se encuentran muy conservadas pero su localización con respecto a la caja A (o al snRNA U2) es variable en tripanosomátidos.

Otro hecho interesante, es que al buscar las cajas A y B de tRNAs-like de tripanosomátidos, el único miembro de esta familia que no contiene a dichas cajas en un tRNA-like fue *C. fasciculata*. En cambio, las cajas A y B asociadas al gen del snRNA U2 en este organismo se encontraron dentro de un gen de tRNA-Ala (Fig. 51). Este resultado, sugiere que el tRNA-like pudo haber surgido de un gen de tRNA-Ala verdadero y mutaciones no deletéreas se acumularon a través de la evolución en las secuencias actuales de los tRNAs-like presentes en otros tripanosomátidos.

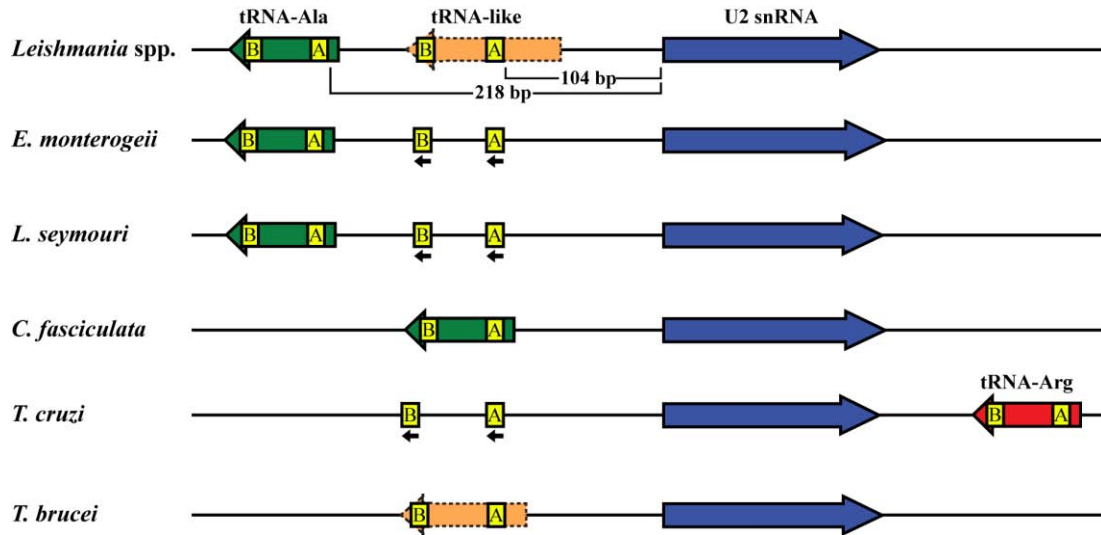


Figura 51. Contexto genómico de los genes del snRNA U2 en tripanosomátidos. Representación esquemática del locus del snRNA U2 en *Leishmania* spp., *E. monterogei* (LV88), *L. seymouri* (ATCC 30220), *C. fasciculata* (Cf-C1), *T. cruzi* (CL Brener Non-Esmeraldo-like, copia del cromosoma 23) y *T. brucei* (DAL972). Todos los genes del snRNA U2 en tripanosomátidos contienen cajas A y B orientadas divergentemente y localizadas río arriba del snRNA U2 en genes de tRNA-like. La caja A presenta una distancia muy conservada de alrededor de 104 pb con respecto al SIT de U2. La excepción es *C. fasciculata*, en donde las cajas A y B están localizadas dentro de un gen de tRNA-Ala. En *T. cruzi*, un gen de tRNA-Arg se localiza río abajo del snRNA U2. La figura está a escala.

7. Discusión

7.1. Número de copias del snRNA U2

En el presente trabajo se estudió la secuencia, estructura y contexto genómico del gen del snRNA U2 en *Leishmania* y otros tripanosomátidos, y además se identificaron las secuencias que dirigen la transcripción del snRNA U2 en *L. major*. Los datos obtenidos muestran que todas las especies de *Leishmania* contienen un gen único del snRNA U2 el cual es altamente sinténico (Fig. 17).

De manera similar, en otros tripanosomátidos el snRNA U2 también es codificado por un sólo gen. Las únicas excepciones son *L. collosma* y *T. cruzi*, en los cuales existen tres copias de este snRNA (este trabajo y [159]). De manera interesante, de las tres copias presentes en el genoma de *T. cruzi*, sólo una es expresada [159]. La búsqueda de cajas A y B en genes de tRNAs-like asociados a genes del snRNA U2 en tripanosomátidos, permitió identificar que 104 pb río arriba de las tres copias de U2 de *T. cruzi* se localiza una caja A. Sin embargo, nuestro análisis reveló que en las copias no expresadas de U2, el gen del tRNA-like asociado carece de la caja B (Fig. 52). Por lo tanto, la ausencia de la caja B puede contribuir al silenciamiento de dos copias del gen del snRNA U2 en *T. cruzi*.

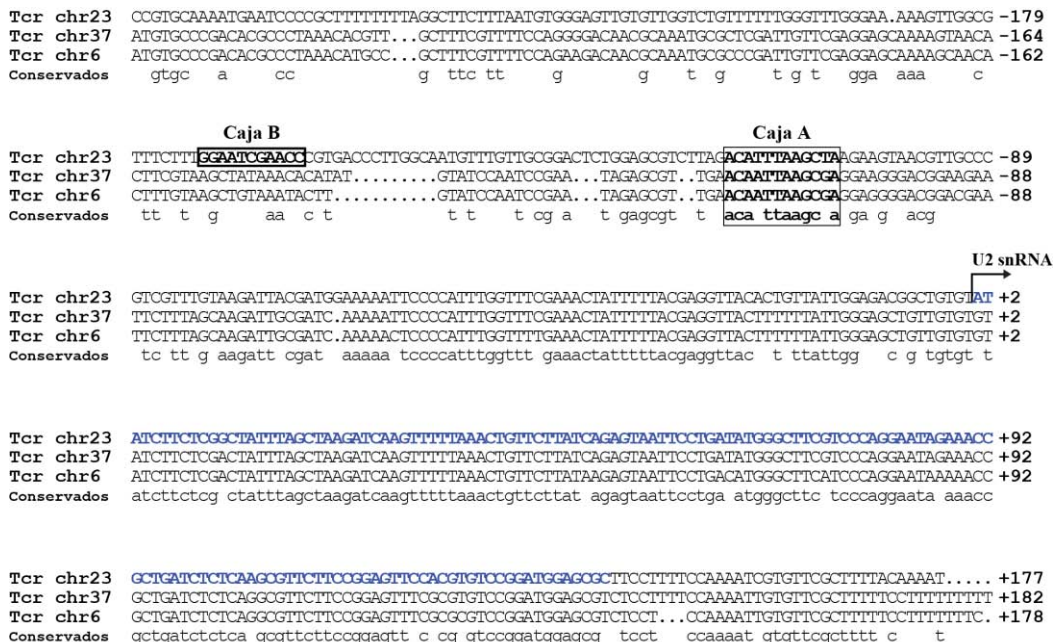


Figura 52. Comparaciones de las secuencias de los genes del snRNA U2 y sus secuencias flanqueantes presentes en *T. cruzi* (CL Brener Non-Esmeraldo-like). Se muestra la secuencia de los genes de U2 localizados en los cromosomas 23, 37 y 6. El snRNA U2 del cromosoma 23 se muestra en azul. Las posiciones de las cajas A y B están indicadas. Los números de secuencia se muestran en relación al SIT (+1) del snRNA U2.

Varios reportes indican que al igual que el snRNA U2, los demás snRNAs en tripanosomátidos se presentan generalmente como genes de copia única [153,158,168,169]. Dado que los tripanosomátidos requieren de una elevada cantidad de snRNAs para llevar a cabo el *trans-splicing*, la tasa de transcripción de estos genes de copia única debe ser muy elevada en estos organismos. En contraste, en eucariontes superiores los genes de snRNAs se presentan como genes multicopia. Por ejemplo, en humanos se reportan 3 genes de U2, mientras que el género *Drosophila* presenta de 6 a 9 genes; asimismo, en el género *Caenorhabditis* se presentan de 10 a 19 copias [170]. En general, los genes de snRNAs no aparecen en clusters en los genomas de metazoarios [170]. Sin embargo, algunas veces se han encontrado asociados a genes del rRNA 5S [171-175]. De manera interesante, algunos genomas, más notablemente los de mamíferos, contienen un número considerable de pseudogenes de snRNAs [170-173,175,176].

7.2. Término de la transcripción del snRNA U2

En general, la secuencia del snRNA U2 se encontró muy conservada en tripanosomátidos (Tabla 3). Sin embargo, la secuencia del extremo 3' es variable, al igual que lo es la longitud del snRNA U2 (Figs. 28 y 29). El mapeo de los extremos 5' y 3' del snRNA U2 en *L. major* mostró que la longitud del transcrito maduro es de 143 nt (Fig. 18), la cual es muy similar a las longitudes de 143 y 145 nt reportadas para el snRNA U2 de *T. cruzi* y *T. brucei*, respectivamente [162,163]. En otros organismos, este snRNA tiene un tamaño que varía entre 141 (en *L. amazonensis*) y 152 (en *L. collosoma*) nt [159,160]. Así, las diferentes especies de tripanosomátidos presentan el snRNA U2 más pequeño hasta ahora descrito.

Se encontró que la transcripción del snRNA U2 en *L. major* termina en un tracto de Ts localizado río abajo del gen, obteniéndose extensiones que tienen que ser procesadas para generar el transcrito maduro de 143 nt mapeado (Fig. 18). El tracto de Ts, que constituye una señal de término de la transcripción llevada a cabo por la RNAP III, está presente río abajo del gen del snRNA U2 en todos los tripanosomátidos (Figs. 28 y 29). Sin embargo, mientras que en *L. major* y *T. cruzi* el tracto comienza 2 y 4 nt río abajo del extremo 3' maduro, respectivamente, en *T. brucei* la señal de término se ubica a 12

nt del extremo 3' (Fig. 28). De hecho, se ha reportado que en *T. brucei* la transcripción del snRNA U2 genera transcritos con extensiones de justamente 12 nt del lado 3', las cuales son procesadas para generar al snRNA U2 maduro. Por lo tanto, estos resultados son similares a los obtenidos por el mapeo del extremo 3' del snRNA U2 de *L. major*.

La señal de término de la RNAP III también está presente río abajo del gen del snRNA U6 de eucariontes superiores. Sin embargo, el extremo 3' del transcrito primario sufre de varios pasos de adición y remoción de Us, de tal forma que el transcrito maduro mantiene un tracto de 5 Us y un fosfato cíclico en su extremo 3' [106,114]. Los snRNAs en tripanosomátidos (exceptuando el snRNA U6) no mantienen un tracto de Us en su extremo 3' y por lo tanto, el procesamiento del transcrito primario difiere del encontrado en eucariontes superiores.

7.3. Pseudouridinas en el snRNA U2

Debido a que los residuos de Ψ juegan un papel importante en la maduración de pre-mRNAs y el ensamblaje de snRNPs, se decidió identificar y localizar a estos nucleótidos modificados en la secuencia del snRNA U2 de *L. major*. Se localizaron seis residuos de Ψ , de los cuales cinco también están presentes en vertebrados (Ψ 14, 34, 36, 55 y 86 en *L. major*), tres se han encontrado en *S. cerevisiae* (Ψ 34, 55 y 86) [121,123] y tres están presentes en *T. brucei* (Ψ 14, 34 y 36) [155] (Fig. 27).

Notablemente, la única posición conservada en todos los organismos analizados corresponde a la Ψ 34, la cual se encuentra en la región involucrada en el reconocimiento del sitio de ramificación del pre-mRNA (BPRR) (Fig. 27). La Ψ 36 también se localiza en el BPRR. Las Ψ encontradas en estas posiciones favorecen la unión por apareamiento de bases entre el snRNA U2 y el pre-mRNA en levadura, lo que a su vez promueve que se realice el primer ataque nucleofílico del *splicing* [177]. Esto sugiere un posible papel de las Ψ 34 y 36 del snRNA U2 de *L. major* en el trans-*splicing*.

Un total de 14 Ψ s se han reportado en el snRNA U2 de vertebrados (tres de ellas en el BPRR), 5 se han encontrado en *S. cerevisiae* (una en el BPRR) [121] y seis fueron mapeadas en *T. brucei* (cuatro en el BPRR) [155]. De manera interesante, el snRNA U2 de *L. major* contiene tres pseudouridinas que no han sido reportadas en ningún otro organismo, incluyendo una Ψ localizada en dentro del sitio Sm en la posición 93 (Fig. 27). En eucariontes superiores, esta región es reconocida por un heptámero en forma de dona compuesto por las proteínas Sm B, D3, D1, D2, F, E y G. Se ha reportado que los sitios Sm dentro de los snRNAs U2 y U4 de *T. brucei* son reconocidos por heptámeros específicos, a diferencia de los otros snRNAs que son reconocidos por las proteínas Sm canónicas [165,178]. El hecho de que el snRNA U2 presente un complejo Sm específico sugiere que la Ψ 93 del snRNA U2 de *L. major*, ubicada en el sitio Sm, podría estar involucrada en la formación de una snRNP específica de U2 en este organismo. Así, es probable que todas las Ψ s encontradas en el snRNA U2 de *L. major* jueguen algún papel importante tanto en el ensamblaje de la snRNP como en el *trans-splicing*.

De manera notable, se encontraron nucleótidos conservados en tripanosomátidos que podrían ser pseudouridilados *in vivo*, pues la secuencia del gen del snRNA U2 en las 21 especies de tripanosomátidos analizadas, presentó un residuo T (U en el RNA) en cinco de las seis posiciones que se identificaron como pseudouridiladas en el snRNA U2 de *L. major* (Figs. 28 y 29).

7.4. Promotor del snRNA U2 en *L. major* y otros organismos

El análisis de secuencia del locus del snRNA U2 permitió la identificación de cajas A y B localizadas dentro de un gen de tRNA-like ubicado entre el snRNA U2 y un gen de tRNA-Ala en *L. major*. El análisis por medio de *primer extension* y RT-qPCR mostró que la transcripción del snRNA U2 en *L. major* depende de las cajas A y B del tRNA-like, ya que tanto la eliminación como la mutación de estos elementos provocó la abolición de la transcripción del gen (Figs. 45 y 46). Adicionalmente, un elemento ubicado en los primeros nucleótidos del gen del snRNA U2 (+6 a +12) también es requerido para la expresión de U2 en *L. major* (Figs. 45 y 46). Elementos de secuencia similares dirigen la expresión del snRNA U2 en *T. brucei* [147]. De manera similar, se ha reportado que las cajas A y/o B localizadas

en genes de tRNAs o tRNAs-like controlan la transcripción de los genes de snRNAs U1, U3 y U6 en *T. brucei*, U4 en *L. collosoma* y los genes del RNA 7SL en ambas especies (*T. brucei* y *L. collosoma*) [148,149,152,153,179]. Sin embargo, a diferencia de cualquier otro gen de snRNAs o RNA 7SL en tripanosomátidos, nuestros resultados muestran que la caja B de un tRNA-Ala ubicada río arriba (a 263 pb del SIT), también participa en la regulación de la transcripción del gen del snRNA U2 en *L. major*. Mientras que la mutación de la caja B del tRNA-like redujo 87% la transcripción de U2, la sustitución de bases de la caja B del tRNA-Ala causó un decremento del 46% en los niveles del snRNA U2 (Fig. 46, comparar barras 8 y 10). Así, es probable que la caja B del tRNA-Ala no forme parte del promotor basal del U2, sino que funcione como un elemento activador involucrado en la regulación fina de la expresión del snRNA U2 de *L. major*.

En eucariontes superiores también se ha reportado la participación de cajas A y B de genes de tRNAs en la transcripción de otros genes no involucrados en el *splicing*. Por ejemplo, en plantas se ha observado que la expresión de algunos RNAs pequeños nucleolares (snoRNAs) depende de un gen de tRNA localizado río arriba del snoRNA [180]; en este caso, el tRNA y el snoRNA se localizan en la misma cadena de DNA. Además, la transcripción de la región que codifica al snoRNA SNR52 de levadura es conducida por un promotor compuesto por las cajas A y B de un tRNA-like localizado río arriba [181]. Una situación similar se reportó para la expresión del RNA de la RNasa P de levadura [182]. Asimismo, en *C. elegans* alrededor de 40 snoRNAs son transcritos desde un promotor tipo tRNA-like ubicado río arriba [183]. Esta estrategia también se ha encontrado para la expresión de micro RNAs (miRNAs) de virus y humanos [184,185]. Varios miRNAs del virus del herpes murino son expresados a partir de promotores tRNA-like (con las cajas A y B) localizados río arriba y en la misma cadena de DNA; los tRNAs-like han perdido el elemento terminador, permitiendo que la transcripción proceda a través de la región codificadora del miRNA [184]. El procesamiento del transcrito primario produce moléculas de tRNA-like maduras (además de miRNAs) que mostraron no estar aminoaciladas [186].

Resulta interesante que el promotor del snRNA U6 de *S. cerevisiae* está formado por una caja A dentro de la región codificante y una caja B localizada río abajo del gen [53-56]. De manera similar, se ha reportado la presencia de una caja B ubicada en un intrón del snRNA U6 de *S. pombe* [51]. Estos resultados indican que el uso de cajas A y B (o de secuencias similares) como regiones promotoras de genes transcritos por la RNAP III está ampliamente distribuido entre organismos muy poco relacionados; lo que sugiere que este modo de transcripción ha surgido varias veces a través de la evolución, o que surgió tempranamente en un organismo ancestral y se ha mantenido entre los distintos grupos de organismos.

7.5. Transcripción de tRNAs-like asociados a snRNAs en tripanosomátidos

Ensayos de tipo Northern blot y *primer extension* permitieron determinar que el gen del tRNA-like asociado al snRNA U2 de *L. major* es transcrito *in vivo* generando un RNA de ~109 nt (Fig. 48). Se ha reportado que los tRNA-like asociados al snRNA U2 de *T. brucei* y al snRNA U4 de *L. collosoma*, también se transcriben [151,153]. Sin embargo, el tamaño estimado de las moléculas de tRNAs-like fue de ~89 nt para el que está asociado al snRNA U4 y de ~90 nt para el que está asociado con el snRNA U2. Por lo tanto, el tRNA-like asociado al snRNA U2 de *L. major* es alrededor de 19 nt más largo que los reportados en *T. brucei* y *L. collosoma*. Un análisis densitométrico de las señales detectadas por Northern blot mostró que el tRNA-like es 10 veces menos abundante que el snRNA U2. Por lo tanto, si el snRNA U2 y el tRNA-like son transcritos al mismo tiempo y con la misma eficiencia, el tRNA-like debería ser mucho menos estable. La probable estructura secundaria obtenida para el tRNA-like es muy diferente a la estructura de hoja de trébol presente en los tRNAs clásicos (Fig. 49). En consecuencia, es posible que esta estructura secundaria atípica pudiera contribuir a la aparente inestabilidad del transcrito del tRNA-like. No se descarta la posibilidad alterna de que las tasas de transcripción en direcciones opuestas sean diferentes y por lo tanto, se produzcan menos transcritos del tRNA-like.

7.6. Modelo de la transcripción del snRNA U2 en *L. major*

Hasta la fecha se desconoce el mecanismo por el cual las cajas A y B de genes de tRNAs o tRNAs-like controlan la expresión de snRNAs en tripanosomátidos. Las cajas A y B de promotores de tRNAs en eucariontes superiores son en primer lugar reconocidas por TFIIC; mientras que la subunidad $\tau 95$ interacciona con la caja A, la subunidad $\tau 138$ reconoce a la caja B con una mayor afinidad. Posteriormente TFIIC posiciona a TFIIB en secuencias localizadas cerca del SIT para que sea reclutada la RNAP III y se lleve a cabo la transcripción. Por otro lado, el proceso transcripcional del snRNA U6 requiere en primera instancia que el PSE y la caja TATA sean reconocidos por SNAPc y TBP, respectivamente [187]. Este complejo recluta secuencialmente a Brf2, Bdp1 y a la RNAP III, lo que conlleva al inicio de la transcripción del snRNA U6 [66,71,72,74,75].

La transcripción de tRNAs, así como de snRNAs en eucariontes superiores, depende de la dirección de los elementos promotores, de tal manera que a simple vista, parecería poco probable que los tRNAs y tRNAs-like de tripanosomátidos, localizados en la hebra de DNA opuesta a sus snRNAs acompañantes, tuvieran un papel en la transcripción de los snRNAs. Sin embargo, se ha propuesto un modelo de la transcripción de snRNAs en tripanosomátidos en el que la región de DNA que separa al tRNA-like de su snRNA acompañante es doblada, permitiendo que tanto el gen de tRNA (o tRNA-like) como el snRNA sean posicionados en la misma dirección. El factor de transcripción TFIIC tiene la capacidad de doblar al DNA en su sitio de unión, por lo tanto es un buen candidato para llevar a cabo el doblamiento de DNA que conllevaría al posicionamiento de tRNAs (o tRNAs-like) y snRNAs en la misma dirección. Sin embargo, de todos los factores de transcripción que se requieren para la transcripción de tRNAs y del snRNA U6 en eucariontes, sólo se han encontrado a TFIIB y SNAPc en tripanosomátidos, es decir, TFIIC aún no se ha identificado en estos organismos. Por otro lado, los análisis de secuencia realizados en este trabajo mostraron que la secuencia GGTC está siempre presente en las cajas B de tRNAs-like asociados a genes del snRNA U2 en tripanosomátidos (Fig. 50). Se ha reportado que esta secuencia está protegida por factores proteicos en tRNAs de *T. brucei* [160], además

los homólogos de TBP y SNAPc se han encontrado asociados a los promotores de snRNAs en este organismo [188-192].

Todos estos hallazgos en conjunto nos permiten proponer un modelo de la transcripción de snRNAs para tripanosomátidos (Fig. 53), en el que el primer evento sería el reconocimiento de las cajas A y B de tRNAs (o tRNAs-like) asociados al snRNA por un factor parecido a TFIIC. Esto conllevaría al doblamiento del DNA y, en consecuencia, al posicionamiento de los genes de tRNAs y snRNAs en la misma dirección.

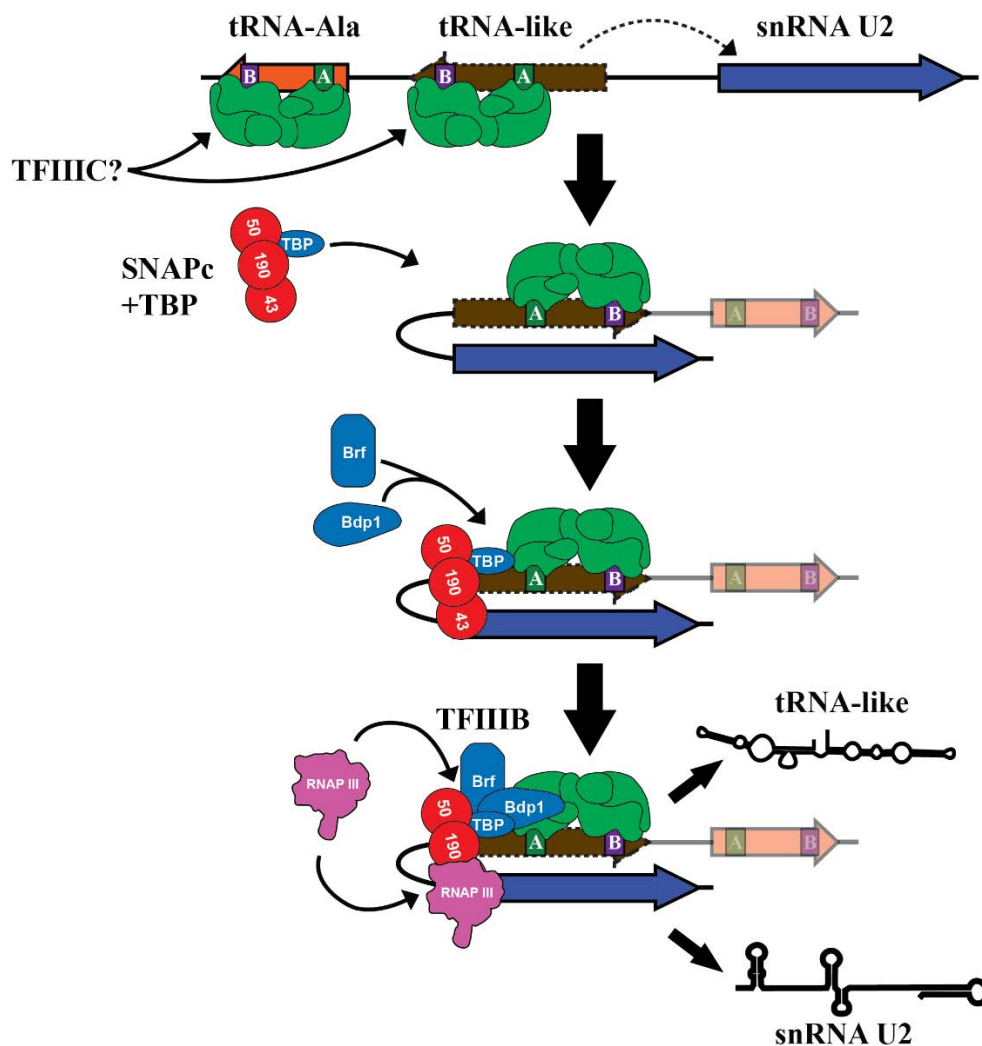


Figura 53. Modelo de la transcripción de snRNAs en tripanosomátidos. En la transcripción de snRNAs en tripanosomátidos primero un factor parecido a TFIIC (verde) se uniría a las cajas A y B de genes de tRNAs asociados a los snRNAs, este factor podría mediar el doblamiento del DNA lo que a su vez conllevaría al posicionamiento del snRNA y el tRNA o tRNA-like en la misma dirección. Posteriormente un complejo formado por TBP y SNAPc se uniría a secuencias cercanas a los SIT de los tRNAs y snRNAs. Este complejo sería entonces reconocido secuencialmente por Brf1, Bdp1 y la RNAP III, para iniciar la transcripción bidireccional de snRNAs y tRNAs o tRNAs-like. En la transcripción del snRNA U2

Por otro lado, la caja B del tRNA-Ala podría servir como un activador de la transcripción del snRNA U2 de *L. major*, de manera análoga al DSE encontrado en el promotor del snRNA U6 de humanos. Mientras que el DSE se encuentra de -200 a -220 en relación al SIT, la caja B se encuentra entre -263 y -273. En la transcripción del snRNA U6 de humanos, un nucleosoma acerca al DSE para que SNAPc interaccione con TBP (unido a la caja TATA) y el resto del complejo de pre-inicio de la transcripción. El modelo propuesto se acopla a esta posibilidad pues un nucleosoma posicionado entre el tRNA-Ala y el snRNA U2 acercaría a la caja B del tRNA-Ala y éste a su vez atraería a un TFIIC-like y a los demás factores de transcripción. De manera similar, en levadura un nucleosoma acerca a las cajas A y B del promotor del snRNA U6.

8. Conclusiones

- El snRNA U2 es un gen de copia única en la mayoría de tripanosomátidos, incluyendo a *L. major*. Las excepciones son *L. collosoma* y *T. cruzi* pues en estas especies existen tres copias del gen.
- El mapeo del SIT y extremo 3' del snRNA U2 determinó que el transcrito maduro tiene un tamaño de 143 nt en *L. major*.
- La transcripción del snRNA U2 de *L. major* termina en un tracto de Ts ubicado río abajo del gen. El transcrito primario debe procesarse para generar el transcrito maduro de 143 nt.
- La estructura secundaria del snRNA U2 de *L. major* es similar a la encontrada en eucariontes superiores. Sin embargo, carece del tallo-asa III.
- Se identificaron seis residuos de Ψ en la secuencia del snRNA U2 de *L. major*. De éstos, dos residuos (Ψ 34 y Ψ 36) se encontraron en la probable secuencia de reconocimiento del sitio de ramificación (BPRR) y uno (Ψ 93) se localizó en la secuencia reconocida por proteínas Sm (sitio Sm). Esta última región no se ha encontrado pseudouridilada en ningún otro organismo.
- Alineamientos de las secuencias del snRNA U2 entre 21 especies de tripanosomátidos revelaron la presencia de nucleótidos conservados que podrían ser pseudouridilados *in vivo*, pues la secuencia de todas las especies analizadas presentó un residuo T (U en el RNA) en cinco de las seis posiciones que se identificaron como pseudouridiladas en el snRNA U2 de *L. major*.
- La Ψ 34 del BPRR fue la única posición pseudouridilada conservada entre *L. major*, vertebrados, levaduras y *T. brucei*.
- Se localizó un tRNA-like río arriba del gen del snRNA U2, el cual se encuentra orientado en dirección contraria al snRNA y contiene cajas A y B consenso. Ensayos de Northern blot y *primer extension* indicaron que el gen del tRNA-like se transcribe *in vivo* generándose un transcrito de ~109 nt.

- El análisis de la expresión del snRNA U2 a partir de los diferentes vectores generados, sugirió que las cajas A y B del tRNA-like, además de un elemento ubicado en los primeros nucleótidos del gen, son necesarios para la expresión del snRNA *in vivo*.
- La caja B del tRNA-Ala ubicada a 263 pb del SIT del snRNA U2 podría funcionar como un elemento activador involucrado en la regulación fina de la expresión del snRNA U2 de *L. major*.

9. Bibliografía

1. **Haanstra, J. R., Gonzalez-Marcano, E. B., Gualdron-Lopez, M. & Michels, P. A.** Biogenesis, maintenance and dynamics of glycosomes in trypanosomatid parasites. *Biochimica et biophysica acta* 1863, 1038-1048, doi:10.1016/j.bbamcr.2015.09.015 (2016).
2. **Tuon, F. F., Neto, V. A. & Amato, V. S.** *Leishmania*: origin, evolution and future since the Precambrian. *FEMS immunology and medical microbiology* 54, 158-166, doi:10.1111/j.1574-695X.2008.00455.x (2008).
3. **Rioux, J. A. et al.** Taxonomy of *Leishmania*. Use of isoenzymes. Suggestions for a new classification. *Annales de parasitologie humaine et comparee* 65, 111-125 (1990).
4. **Van der Auwera, G. & Dujardin, J. C.** Species typing in dermal leishmaniasis. *Clinical microbiology reviews* 28, 265-294, doi:10.1128/CMR.00104-14 (2015).
5. **Stanley, T. G. & Wilson, I.** Multilocus enzyme electrophoresis: a practical guide. *Molecular biotechnology* 24, 203-220 (2003).
6. **Cupolillo, E., Medina-Acosta, E., Noyes, H., Momen, H. & Grimaldi, G., Jr.** A revised classification for *Leishmania* and *Endotrypanum*. *Parasitol Today* 16, 142-144 (2000).
7. **Lainson, R. & Shaw, J. J.** in *The leishmaniasis in biology AND medicine. Biology AND epidemiology* Vol. 1 (eds W. Peters & R. Killick-Kendrick) Ch. 2, 120 (Academic Press, 1987).
8. **Fraga, J., Montalvo, A. M., De Doncker, S., Dujardin, J. C. & Van der Auwera, G.** Phylogeny of *Leishmania* species based on the heat-shock protein 70 gene. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* 10, 238-245, doi:10.1016/j.meegid.2009.11.007 (2010).
9. **Leelayoova, S. et al.** Multilocus characterization and phylogenetic analysis of *Leishmania siamensis* isolated from autochthonous visceral leishmaniasis cases, southern Thailand. *BMC microbiology* 13, 60, doi:10.1186/1471-2180-13-60 (2013).
10. **Reithinger, R. et al.** Cutaneous leishmaniasis. *The Lancet. Infectious diseases* 7, 581-596, doi:10.1016/S1473-3099(07)70209-8 (2007).
11. **Goto, H. & Lauletta Lindoso, J. A.** Cutaneous and mucocutaneous leishmaniasis. *Infectious disease clinics of North America* 26, 293-307, doi:10.1016/j.idc.2012.03.001 (2012).
12. **Von Stebut, E.** Leishmaniasis. *Journal der Deutschen Dermatologischen Gesellschaft = Journal of the German Society of Dermatology : JDDG* 13, 191-200; quiz 201, doi:10.1111/ddg.12595 (2015).
13. **Lainson, R. & Shaw, J.** *Evolution, classification and geographical distribution*. Vol. 1 (Academic Press, 1987).
14. **Dutta, A. K.** Pursuit of medical knowledge: Charles Donovan (1863-1951) on kala-azar in India. *Journal of medical biography* 16, 72-76, doi:10.1258/jmb.2007.007004 (2008).
15. **Serarcangeli, C. & Pennica, A.** [Testimonies of an autochthonous illness on the anthropomorphic pottery in ancient Peru]. *Medicina nei secoli* 8, 125-141 (1996).

16. **Altamirano-Enciso, A. J., Marzochi, M. C., Moreira, J. S., Schubach, A. O. & Marzochi, K. B.** [On the origin and spread of cutaneous and mucosal leishmaniasis, based on pre- and post-colombian historical source]. *Historia, ciencias, saude--Manguinhos* 10, 852-882 (2003).
17. **Pigott, D. M. et al.** Global distribution maps of the leishmaniasis. *eLife* 3, doi:10.7554/eLife.02851 (2014).
18. **Kevric, I., Cappel, M. A. & Keeling, J. H.** New World and Old World Leishmania Infections: A Practical Review. *Dermatologic clinics* 33, 579-593, doi:10.1016/j.det.2015.03.018 (2015).
19. **Hartley, M. A., Drexler, S., Ronet, C., Beverley, S. M. & Fasel, N.** The immunological, environmental, and phylogenetic perpetrators of metastatic leishmaniasis. *Trends in parasitology* 30, 412-422, doi:10.1016/j.pt.2014.05.006 (2014).
20. **Yasri, S. & Wiwanitkit, V.** Neurological manifestation on Leishmania siamensis: existed or not? *Arquivos de neuro-psiquiatria* 73, 369, doi:10.1590/0004-282X20150019 (2015).
21. **Karunaweera, N. D.** Leishmania donovani causing cutaneous leishmaniasis in Sri Lanka: a wolf in sheep's clothing? *Trends in parasitology* 25, 458-463, doi:10.1016/j.pt.2009.07.002 (2009).
22. **Shaw, J. et al.** Characterization of Leishmania (Leishmania) waltoni n.sp. (Kinetoplastida: Trypanosomatidae), the Parasite Responsible for Diffuse Cutaneous Leishmaniasis in the Dominican Republic. *The American journal of tropical medicine and hygiene*, doi:10.4269/ajtmh.14-0774 (2015).
23. **Convit, J. et al.** The clinical and immunological spectrum of American cutaneous leishmaniasis. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 87, 444-448 (1993).
24. **Silveira, F. T., Lainson, R. & Corbett, C. E.** Clinical and immunopathological spectrum of American cutaneous leishmaniasis with special reference to the disease in Amazonian Brazil: a review. *Memorias do Instituto Oswaldo Cruz* 99, 239-251, doi:/S0074-02762004000300001 (2004).
25. **Saporito, L., Giammanco, G. M., De Grazia, S. & Colomba, C.** Visceral leishmaniasis: host-parasite interactions and clinical presentation in the immunocompetent and in the immunocompromised host. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases* 17, e572-576, doi:10.1016/j.ijid.2012.12.024 (2013).
26. **Elhassan, A. M. et al.** Post-kala-azar dermal leishmaniasis in the Sudan: peripheral neural involvement. *International journal of dermatology* 31, 400-403 (1992).
27. **Singh, S., Sharma, U. & Mishra, J.** Post-kala-azar dermal leishmaniasis: recent developments. *International journal of dermatology* 50, 1099-1108, doi:10.1111/j.1365-4632.2011.04925.x (2011).
28. **Verma, N. et al.** Clinicopathological and Immunological Changes in Indian Post Kala-Azar Dermal Leishmaniasis (PKDL) Cases in relation to Treatment: A Retrospective Study. *BioMed research international* 2015, 745062, doi:10.1155/2015/745062 (2015).
29. **Alvar, J. et al.** Leishmaniasis worldwide and global estimates of its incidence. *PloS one* 7, e35671, doi:10.1371/journal.pone.0035671 (2012).

30. **Lievin-Le Moal, V. & Loiseau, P. M.** Leishmania hijacking of the macrophage intracellular compartments. *The FEBS journal* 283, 598-607, doi:10.1111/febs.13601 (2016).
31. **Roeder, R. G. & Rutter, W. J.** Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* 224, 234-237 (1969).
32. **Kuhn, C. D. et al.** Functional architecture of RNA polymerase I. *Cell* 131, 1260-1272, doi:10.1016/j.cell.2007.10.051 (2007).
33. **Proshkina, G. M. et al.** Ancient origin, functional conservation and fast evolution of DNA-dependent RNA polymerase III. *Nucleic acids research* 34, 3615-3624, doi:10.1093/nar/gkl421 (2006).
34. **Teichmann, M., Dieci, G., Pascali, C. & Boldina, G.** General transcription factors and subunits of RNA polymerase III: Paralogs for promoter- and cell type-specific transcription in multicellular eukaryotes. *Transcription* 1, 130-135, doi:10.4161/trns.1.3.13192 (2010).
35. **Carter, R. & Drouin, G.** The increase in the number of subunits in eukaryotic RNA polymerase III relative to RNA polymerase II is due to the permanent recruitment of general transcription factors. *Molecular biology and evolution* 27, 1035-1043, doi:10.1093/molbev/msp316 (2010).
36. **Korkhin, Y. et al.** Evolution of complex RNA polymerases: the complete archaeal RNA polymerase structure. *PLoS biology* 7, e1000102, doi:10.1371/journal.pbio.1000102 (2009).
37. **Grummt, I.** Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus. *Genes & development* 17, 1691-1702, doi:10.1101/gad.1098503R (2003).
38. **Paule, M. R. & White, R. J.** Survey and summary: transcription by RNA polymerases I and III. *Nucleic acids research* 28, 1283-1298 (2000).
39. **Smale, S. T. & Kadonaga, J. T.** The RNA polymerase II core promoter. *Annual review of biochemistry* 72, 449-479, doi:10.1146/annurev.biochem.72.121801.161520 (2003).
40. **Butler, J. E. & Kadonaga, J. T.** The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes & development* 16, 2583-2592, doi:10.1101/gad.1026202 (2002).
41. **Domitrovich, A. M. & Kunkel, G. R.** Multiple, dispersed human U6 small nuclear RNA genes with varied transcriptional efficiencies. *Nucleic acids research* 31, 2344-2352 (2003).
42. **Hernandez, N.** Small nuclear RNA genes: a model system to study fundamental mechanisms of transcription. *The Journal of biological chemistry* 276, 26733-26736, doi:10.1074/jbc.R100032200 (2001).
43. **Das, G., Henning, D., Wright, D. & Reddy, R.** Upstream regulatory elements are necessary and sufficient for transcription of a U6 RNA gene by RNA polymerase III. *The EMBO journal* 7, 503-512 (1988).
44. **Lobo, S. M. & Hernandez, N.** A 7 bp mutation converts a human RNA polymerase II snRNA promoter into an RNA polymerase III promoter. *Cell* 58, 55-67 (1989).

45. **Waibel, F. & Filipowicz, W.** RNA-polymerase specificity of transcription of Arabidopsis U snRNA genes determined by promoter element spacing. *Nature* 346, 199-202, doi:10.1038/346199a0 (1990).
46. **Jensen, R. C., Wang, Y., Hardin, S. B. & Stumph, W. E.** The proximal sequence element (PSE) plays a major role in establishing the RNA polymerase specificity of *Drosophila* U-snRNA genes. *Nucleic acids research* 26, 616-622 (1998).
47. **Barakat, N. H. & Stumph, W. E.** TBP recruitment to the U1 snRNA gene promoter is disrupted by substituting a U6 proximal sequence element A (PSEA) for the U1 PSEA. *FEBS letters* 582, 2413-2416, doi:10.1016/j.febslet.2008.06.003 (2008).
48. **Dandekar, T. & Tollervey, D.** Thirty-three nucleotides of 5' flanking sequence including the 'TATA' box are necessary and sufficient for efficient U2 snRNA transcription in *Schizosaccharomyces pombe*. *Molecular microbiology* 5, 1621-1625 (1991).
49. **Hamada, M., Huang, Y., Lowe, T. M. & Maraia, R. J.** Widespread use of TATA elements in the core promoters for RNA polymerases III, II, and I in fission yeast. *Molecular and cellular biology* 21, 6870-6881, doi:10.1128/MCB.21.20.6870-6881.2001 (2001).
50. **Zhou, D. & Lobo-Ruppert, S. M.** Transcription of the *Schizosaccharomyces pombe* U2 gene *in vivo* and *in vitro* is directed by two essential promoter elements. *Nucleic acids research* 29, 2003-2011 (2001).
51. **Frendewey, D., Barta, I., Gillespie, M. & Potashkin, J.** *Schizosaccharomyces* U6 genes have a sequence within their introns that matches the B box consensus of tRNA internal promoters. *Nucleic acids research* 18, 2025-2032 (1990).
52. **Bally, M., Hughes, J. & Cesareni, G.** SnR30: a new, essential small nuclear RNA from *Saccharomyces cerevisiae*. *Nucleic acids research* 16, 5291-5303 (1988).
53. **Eschenlauer, J. B., Kaiser, M. W., Gerlach, V. L. & Brow, D. A.** Architecture of a yeast U6 RNA gene promoter. *Molecular and cellular biology* 13, 3015-3026 (1993).
54. **Burnol, A. F. et al.** Basal promoter and enhancer element of yeast U6 snRNA gene. *Journal of molecular biology* 233, 644-658, doi:10.1006/jmbi.1993.1542 (1993).
55. **Martin, M. P., Gerlach, V. L. & Brow, D. A.** A novel upstream RNA polymerase III promoter element becomes essential when the chromatin structure of the yeast U6 RNA gene is altered. *Molecular and cellular biology* 21, 6429-6439 (2001).
56. **Brow, D. A. & Guthrie, C.** Transcription of a yeast U6 snRNA gene requires a polymerase III promoter element in a novel position. *Genes & development* 4, 1345-1356 (1990).
57. **Male, G. et al.** Architecture of TFIIC and its role in RNA polymerase III pre-initiation complex assembly. *Nature communications* 6, 7387, doi:10.1038/ncomms8387 (2015).
58. **Ducrot, C. et al.** Reconstitution of the yeast RNA polymerase III transcription system with all recombinant factors. *The Journal of biological chemistry* 281, 11685-11692, doi:10.1074/jbc.M600101200 (2006).

59. **Schramm, L. & Hernandez, N.** Recruitment of RNA polymerase III to its target promoters. *Genes & development* 16, 2593-2620, doi:10.1101/gad.1018902 (2002).
60. **Schultz, P. et al.** The two DNA-binding domains of yeast transcription factor tau as observed by scanning transmission electron microscopy. *The EMBO journal* 8, 3815-3824 (1989).
61. **Schuster, C., Krol, A. & Carbon, P.** Two distinct domains in Staf to selectively activate small nuclear RNA-type and mRNA promoters. *Molecular and cellular biology* 18, 2650-2658 (1998).
62. **Herr, W. & Cleary, M. A.** The POU domain: versatility in transcriptional regulation by a flexible two-in-one DNA-binding domain. *Genes & development* 9, 1679-1693 (1995).
63. **Yoon, J. B., Murphy, S., Bai, L., Wang, Z. & Roeder, R. G.** Proximal sequence element-binding transcription factor (PTF) is a multisubunit complex required for transcription of both RNA polymerase II- and RNA polymerase III-dependent small nuclear RNA genes. *Molecular and cellular biology* 15, 2019-2027 (1995).
64. **Kang, Y. S., Kurano, M. & Stumph, W. E.** The Myb domain of the largest subunit of SNAPc adopts different architectural configurations on U1 and U6 snRNA gene promoter sequences. *Nucleic acids research*, doi:10.1093/nar/gku905 (2014).
65. **Doherty, M. T., Kang, Y. S., Lee, C. & Stumph, W. E.** Architectural arrangement of the small nuclear RNA (snRNA)-activating protein complex 190 subunit (SNAP190) on U1 snRNA gene promoter DNA. *The Journal of biological chemistry* 287, 39369-39379, doi:10.1074/jbc.M112.407775 (2012).
66. **Jawdekar, G. W. & Henry, R. W.** Transcriptional regulation of human small nuclear RNA genes. *Biochimica et biophysica acta* 1779, 295-305, doi:10.1016/j.bbagr.2008.04.001 (2008).
67. **Glover-Cutter, K. et al.** TFIIF-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II. *Molecular and cellular biology* 29, 5455-5464, doi:10.1128/MCB.00637-09 (2009).
68. **Kuhlman, T. C., Cho, H., Reinberg, D. & Hernandez, N.** The general transcription factors IIA, IIB, IIF, and IIE are required for RNA polymerase II transcription from the human U1 small nuclear RNA promoter. *Molecular and cellular biology* 19, 2130-2141 (1999).
69. **Louder, R. K. et al.** Structure of promoter-bound TFIID and model of human pre-initiation complex assembly. *Nature*, doi:10.1038/nature17394 (2016).
70. **Zaborowska, J., Taylor, A. & Murphy, S.** A novel TBP-TAF complex on RNA polymerase II-transcribed snRNA genes. *Transcription* 3, 92-104, doi:10.4161/trns.19783 (2012).
71. **Schramm, L., Pendergrast, P. S., Sun, Y. & Hernandez, N.** Different human TFIIB activities direct RNA polymerase III transcription from TATA-containing and TATA-less promoters. *Genes & development* 14, 2650-2663 (2000).
72. **Teichmann, M., Wang, Z. & Roeder, R. G.** A stable complex of a novel transcription factor IIB-related factor, human TFIIB50, and associated proteins mediate selective transcription by RNA polymerase III of genes with upstream promoter elements. *Proceedings of the National Academy of Sciences of the United States of America* 97, 14200-14205, doi:10.1073/pnas.97.26.14200 (2000).

73. **Burnol, A. F. et al.** TFIIC relieves repression of U6 snRNA transcription by chromatin. *Nature* 362, 475-477, doi:10.1038/362475a0 (1993).
74. **Cabart, P. & Murphy, S.** BRFU, a TFIIB-like factor, is directly recruited to the TATA-box of polymerase III small nuclear RNA gene promoters through its interaction with TATA-binding protein. *The Journal of biological chemistry* 276, 43056-43064, doi:10.1074/jbc.M108515200 (2001).
75. **Cabart, P. & Murphy, S.** Assembly of human small nuclear RNA gene-specific transcription factor IIIB complex de novo on and off promoter. *The Journal of biological chemistry* 277, 26831-26838, doi:10.1074/jbc.M203119200 (2002).
76. **Jawdekar, G. W. et al.** The unorthodox SNAP50 zinc finger domain contributes to cooperative promoter recognition by human SNAPC. *The Journal of biological chemistry* 281, 31050-31060, doi:10.1074/jbc.M603810200 (2006).
77. **Hirsch, H. A., Gu, L. & Henry, R. W.** The retinoblastoma tumor suppressor protein targets distinct general transcription factors to regulate RNA polymerase III gene expression. *Molecular and cellular biology* 20, 9182-9191 (2000).
78. **Gridasova, A. A. & Henry, R. W.** The p53 tumor suppressor protein represses human snRNA gene transcription by RNA polymerases II and III independently of sequence-specific DNA binding. *Molecular and cellular biology* 25, 3247-3260, doi:10.1128/MCB.25.8.3247-3260.2005 (2005).
79. **Gomez-Roman, N., Grandori, C., Eisenman, R. N. & White, R. J.** Direct activation of RNA polymerase III transcription by c-Myc. *Nature* 421, 290-294, doi:10.1038/nature01327 (2003).
80. **Boguta, M.** Maf1, a general negative regulator of RNA polymerase III in yeast. *Biochimica et biophysica acta* 1829, 376-384, doi:10.1016/j.bbagr.2012.11.004 (2013).
81. **Orioli, A., Praz, V., Lhote, P. & Hernandez, N.** Human MAF1 targets and represses active RNA polymerase III genes by preventing recruitment rather than inducing long-term transcriptional arrest. *Genome research*, doi:10.1101/gr.201400.115 (2016).
82. **Vannini, A. et al.** Molecular basis of RNA polymerase III transcription repression by Maf1. *Cell* 143, 59-70, doi:10.1016/j.cell.2010.09.002 (2010).
83. **Das, D. & Scovell, W. M.** The binding interaction of HMG-1 with the TATA-binding protein/TATA complex. *The Journal of biological chemistry* 276, 32597-32605, doi:10.1074/jbc.M011792200 (2001).
84. **Emran, F. et al.** A role for Yin Yang-1 (YY1) in the assembly of snRNA transcription complexes. *Gene* 377, 96-108, doi:10.1016/j.gene.2006.03.012 (2006).
85. **Hu, P., Wu, S. & Hernandez, N.** A role for beta-actin in RNA polymerase III transcription. *Genes & development* 18, 3010-3015, doi:10.1101/gad.1250804 (2004).
86. **Strom, A. C., Forsberg, M., Lillhager, P. & Westin, G.** The transcription factors Sp1 and Oct-1 interact physically to regulate human U2 snRNA gene expression. *Nucleic acids research* 24, 1981-1986 (1996).

87. **Zhao, X., Pendergrast, P. S. & Hernandez, N.** A positioned nucleosome on the human U6 promoter allows recruitment of SNAPc by the Oct-1 POU domain. *Molecular cell* 7, 539-549 (2001).
88. **Boyd, D. C., Greger, I. H. & Murphy, S.** *In vivo* footprinting studies suggest a role for chromatin in transcription of the human 7SK gene. *Gene* 247, 33-44 (2000).
89. **Arimbasseri, A. G. & Bhargava, P.** Chromatin structure and expression of a gene transcribed by RNA polymerase III are independent of H2A.Z deposition. *Molecular and cellular biology* 28, 2598-2607, doi:10.1128/MCB.01953-07 (2008).
90. **Shivaswamy, S. & Bhargava, P.** Positioned nucleosomes due to sequential remodeling of the yeast U6 small nuclear RNA chromatin are essential for its transcriptional activation. *The Journal of biological chemistry* 281, 10461-10472, doi:10.1074/jbc.M512425200 (2006).
91. **Skuzeski, J. M. et al.** Synthesis of human U1 RNA. II. Identification of two regions of the promoter essential for transcription initiation at position +1. *The Journal of biological chemistry* 259, 8345-8352 (1984).
92. **Mouaikel, J. et al.** Interaction between the small-nuclear-RNA cap hypermethylase and the spinal muscular atrophy protein, survival of motor neuron. *EMBO reports* 4, 616-622, doi:10.1038/sj.embor.embor863 (2003).
93. **Singh, R. & Reddy, R.** Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. *Proceedings of the National Academy of Sciences of the United States of America* 86, 8280-8283 (1989).
94. **Singh, R., Gupta, S. & Reddy, R.** Capping of mammalian U6 small nuclear RNA in vitro is directed by a conserved stem-loop and AUAUAC sequence: conversion of a noncapped RNA into a capped RNA. *Molecular and cellular biology* 10, 939-946 (1990).
95. **Gupta, S., Singh, R. & Reddy, R.** Capping of U6 small nuclear RNA in vitro can be uncoupled from transcription. *The Journal of biological chemistry* 265, 9491-9495 (1990).
96. **Egloff, S., O'Reilly, D. & Murphy, S.** Expression of human snRNA genes from beginning to end. *Biochemical Society transactions* 36, 590-594, doi:10.1042/BST0360590 (2008).
97. **Egloff, S. et al.** Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science* 318, 1777-1779, doi:10.1126/science.1145989 (2007).
98. **Medlin, J. et al.** P-TEFb is not an essential elongation factor for the intronless human U2 snRNA and histone H2b genes. *The EMBO journal* 24, 4154-4165, doi:10.1038/sj.emboj.7600876 (2005).
99. **Medlin, J. E., Uguen, P., Taylor, A., Bentley, D. L. & Murphy, S.** The C-terminal domain of pol II and a DRB-sensitive kinase are required for 3' processing of U2 snRNA. *The EMBO journal* 22, 925-934, doi:10.1093/emboj/cdg077 (2003).
100. **Baillat, D. et al.** Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* 123, 265-276, doi:10.1016/j.cell.2005.08.019 (2005).

101. **Uguen, P. & Murphy, S.** The 3' ends of human pre-snRNAs are produced by RNA polymerase II CTD-dependent RNA processing. *The EMBO journal* 22, 4544-4554, doi:10.1093/emboj/cdg430 (2003).
102. **Stefano, J. E.** Purified lupus antigen La recognizes an oligouridylylate stretch common to the 3' termini of RNA polymerase III transcripts. *Cell* 36, 145-154 (1984).
103. **Rinke, J. & Steitz, J. A.** Association of the lupus antigen La with a subset of U6 snRNA molecules. *Nucleic acids research* 13, 2617-2629 (1985).
104. **Lund, E. & Dahlberg, J. E.** Cyclic 2',3'-phosphates and nontemplated nucleotides at the 3' end of spliceosomal U6 small nuclear RNA's. *Science* 255, 327-330 (1992).
105. **Terns, M. P., Lund, E. & Dahlberg, J. E.** 3'-end-dependent formation of U6 small nuclear ribonucleoprotein particles in *Xenopus laevis* oocyte nuclei. *Molecular and cellular biology* 12, 3032-3040 (1992).
106. **Trippe, R. et al.** Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. *Rna* 12, 1494-1504, doi:10.1261/rna.87706 (2006).
107. **Trippe, R., Richly, H. & Benecke, B. J.** Biochemical characterization of a U6 small nuclear RNA-specific terminal uridylyltransferase. *European journal of biochemistry / FEBS* 270, 971-980 (2003).
108. **Trippe, R., Sandrock, B. & Benecke, B. J.** A highly specific terminal uridylyl transferase modifies the 3'-end of U6 small nuclear RNA. *Nucleic acids research* 26, 3119-3126 (1998).
109. **Hilcenko, C. et al.** Aberrant 3' oligoadenylation of spliceosomal U6 small nuclear RNA in poikiloderma with neutropenia. *Blood* 121, 1028-1038, doi:10.1182/blood-2012-10-461491 (2013).
110. **Booth, B. L., Jr. & Pugh, B. F.** Identification and characterization of a nuclease specific for the 3' end of the U6 small nuclear RNA. *The Journal of biological chemistry* 272, 984-991 (1997).
111. **Cvitkovic, I. & Jurica, M. S.** Spliceosome database: a tool for tracking components of the spliceosome. *Nucleic acids research* 41, D132-141, doi:10.1093/nar/gks999 (2013).
112. **Chen, W. & Moore, M. J.** The spliceosome: disorder and dynamics defined. *Current opinion in structural biology* 24, 141-149, doi:10.1016/j.sbi.2014.01.009 (2014).
113. **Leung, A. K., Nagai, K. & Li, J.** Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis. *Nature* 473, 536-539, doi:10.1038/nature09956 (2011).
114. **Zhou, L. et al.** Crystal structures of the Lsm complex bound to the 3' end sequence of U6 small nuclear RNA. *Nature* 506, 116-120, doi:10.1038/nature12803 (2014).
115. **Licht, K., Medenbach, J., Luhrmann, R., Kambach, C. & Bindereif, A.** 3'-cyclic phosphorylation of U6 snRNA leads to recruitment of recycling factor p110 through LSm proteins. *RNA* 14, 1532-1538, doi:10.1261/rna.1129608 (2008).
116. **Padgett, R. A.** mRNA Splicing: Role of snRNAs. *Encyclopedia of life sciences*, 7, doi:10.1038/npg.els.0003833 (2005).

117. **Wahl, M. C., Will, C. L. & Luhrmann, R.** The spliceosome: design principles of a dynamic RNP machine. *Cell* 136, 701-718, doi:10.1016/j.cell.2009.02.009 (2009).
118. **Hoskins, A. A. et al.** Ordered and dynamic assembly of single spliceosomes. *Science* 331, 1289-1295, doi:10.1126/science.1198830 (2011).
119. **Shcherbakova, I. et al.** Alternative spliceosome assembly pathways revealed by single-molecule fluorescence microscopy. *Cell reports* 5, 151-165, doi:10.1016/j.celrep.2013.08.026 (2013).
120. **Rodgers, M. L. et al.** Conformational dynamics of stem II of the U2 snRNA. *RNA* 22, 225-236, doi:10.1261/rna.052233.115 (2016).
121. **Adachi, H. & Yu, Y. T.** Insight into the mechanisms and functions of spliceosomal snRNA pseudouridylation. *World journal of biological chemistry* 5, 398-408, doi:10.4331/wjbc.v5.i4.398 (2014).
122. **Ro-Choi, T. S. & Choi, Y. C.** Chemical approaches for structure and function of RNA in postgenomic era. *Journal of nucleic acids* 2012, doi:10.1155/2012/369058 (2012).
123. **Yu, A. T., Ge, J. & Yu, Y. T.** Pseudouridines in spliceosomal snRNAs. *Protein & cell* 2, 712-725, doi:10.1007/s13238-011-1087-1 (2011).
124. **Karjolic, J. & Yu, Y. T.** Spliceosomal snRNA modifications and their function. *RNA biology* 7, 192-204 (2010).
125. **Lafontaine, D. L., Bousquet-Antonelli, C., Henry, Y., Caizergues-Ferrer, M. & Tollervey, D.** The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. *Genes & development* 12, 527-537 (1998).
126. **Massenet, S. et al.** Pseudouridine mapping in the *Saccharomyces cerevisiae* spliceosomal U small nuclear RNAs (snRNAs) reveals that pseudouridine synthase pus1p exhibits a dual substrate specificity for U2 snRNA and tRNA. *Molecular and cellular biology* 19, 2142-2154 (1999).
127. **Ma, X. et al.** Pseudouridylation of yeast U2 snRNA is catalyzed by either an RNA-guided or RNA-independent mechanism. *The EMBO journal* 24, 2403-2413, doi:10.1038/sj.emboj.7600718 (2005).
128. **Schnauffer, A. et al.** Separate insertion and deletion subcomplexes of the *Trypanosoma brucei* RNA editing complex. *Molecular cell* 12, 307-319 (2003).
129. **Campbell, D. A., Thomas, S. & Sturm, N. R.** Transcription in kinetoplastid protozoa: why be normal? *Microbes and infection / Institut Pasteur* 5, 1231-1240 (2003).
130. **Llanes, A., Restrepo, C. M., Del Vecchio, G., Anguizola, F. J. & Leonart, R.** The genome of *Leishmania panamensis*: insights into genomics of the L. (Viannia) subgenus. *Scientific reports* 5, 8550, doi:10.1038/srep08550 (2015).
131. **Real, F. et al.** The genome sequence of *Leishmania (Leishmania) amazonensis*: functional annotation and extended analysis of gene models. *DNA research : an international journal for rapid publication of reports on genes and genomes* 20, 567-581, doi:10.1093/dnares/dst031 (2013).

132. **Rogers, M. B. et al.** Chromosome and gene copy number variation allow major structural change between species and strains of *Leishmania*. *Genome research* 21, 2129-2142, doi:10.1101/gr.122945.111 (2011).
133. **Ivens, A. C. et al.** The genome of the kinetoplastid parasite, *Leishmania major*. *Science* 309, 436-442, doi:10.1126/science.1112680 (2005).
134. **Peacock, C. S. et al.** Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. *Nature genetics* 39, 839-847, doi:10.1038/ng2053 (2007).
135. **Downing, T. et al.** Whole genome sequencing of multiple *Leishmania donovani* clinical isolates provides insights into population structure and mechanisms of drug resistance. *Genome research* 21, 2143-2156, doi:10.1101/gr.123430.111 (2011).
136. **Raymond, F. et al.** Genome sequencing of the lizard parasite *Leishmania tarentolae* reveals loss of genes associated to the intracellular stage of human pathogenic species. *Nucleic acids research* 40, 1131-1147, doi:10.1093/nar/gkr834 (2012).
137. **Wincker, P. et al.** The *Leishmania* genome comprises 36 chromosomes conserved across widely divergent human pathogenic species. *Nucleic acids research* 24, 1688-1694 (1996).
138. **Britto, C. et al.** Conserved linkage groups associated with large-scale chromosomal rearrangements between Old World and New World *Leishmania* genomes. *Gene* 222, 107-117 (1998).
139. **El-Sayed, N. M. et al.** Comparative genomics of trypanosomatid parasitic protozoa. *Science* 309, 404-409, doi:10.1126/science.1112181 (2005).
140. **Worthey, E. A. et al.** *Leishmania major* chromosome 3 contains two long convergent polycistronic gene clusters separated by a tRNA gene. *Nucleic acids research* 31, 4201-4210 (2003).
141. **Martinez-Calvillo, S. et al.** Transcription of *Leishmania major* Friedlin chromosome 1 initiates in both directions within a single region. *Molecular cell* 11, 1291-1299 (2003).
142. **Martinez-Calvillo, S., Nguyen, D., Stuart, K. & Myler, P. J.** Transcription initiation and termination on *Leishmania major* chromosome 3. *Eukaryotic cell* 3, 506-517 (2004).
143. **Bitar, M., Boroni, M., Macedo, A. M., Machado, C. R. & Franco, G. R.** The spliced leader trans-splicing mechanism in different organisms: molecular details and possible biological roles. *Frontiers in genetics* 4, 199, doi:10.3389/fgene.2013.00199 (2013).
144. **Martinez-Calvillo, S., Vizuet-de-Rueda, J. C., Florencio-Martinez, L. E., Manning-Cela, R. G. & Figueroa-Angulo, E. E.** Gene expression in trypanosomatid parasites. *Journal of biomedicine & biotechnology* 2010, doi:10.1155/2010/525241 (2010).
145. **Gunzl, A.** The pre-mRNA splicing machinery of trypanosomes: complex or simplified? *Eukaryotic cell* 9, 1159-1170, doi:10.1128/EC.00113-10 (2010).
146. **Papasaikas, P. & Valcarcel, J.** The Spliceosome: The Ultimate RNA Chaperone and Sculptor. *Trends in biochemical sciences* 41, 33-45, doi:10.1016/j.tibs.2015.11.003 (2016).

147. **Fantoni, A., Dare, A. O. & Tschudi, C.** RNA polymerase III-mediated transcription of the trypanosome U2 small nuclear RNA gene is controlled by both intragenic and extragenic regulatory elements. *Molecular and cellular biology* 14, 2021-2028 (1994).
148. **Nakaar, V., Dare, A. O., Hong, D., Ullu, E. & Tschudi, C.** Upstream tRNA genes are essential for expression of small nuclear and cytoplasmic RNA genes in trypanosomes. *Molecular and cellular biology* 14, 6736-6742 (1994).
149. **Djikeng, A. et al.** Characterization of a candidate *Trypanosoma brucei* U1 small nuclear RNA gene. *Molecular and biochemical parasitology* 113, 109-115 (2001).
150. **Nakaar, V. T., C.; Ullu, E..** An unusual liaison: small nuclear and cytoplasmic RNA genes team up with tRNA genes in trypanosomatid protozoa. *Parasitology Today* 11, 4 (1995).
151. **Tschudi, C. & Ullu, E.** Unconventional rules of small nuclear RNA transcription and cap modification in trypanosomatids. *Gene expression* 10, 3-16 (2002).
152. **Nakaar, V., Gunzl, A., Ullu, E. & Tschudi, C.** Structure of the *Trypanosoma brucei* U6 snRNA gene promoter. *Molecular and biochemical parasitology* 88, 13-23 (1997).
153. **Li, L., Otake, L. R., Xu, Y. & Michaeli, S.** The trans-spliceosomal U4 RNA from the monogenetic trypanosomatid *Leptomonas collosoma*. Cloning and identification of a transcribed tRNA-like element that controls its expression. *The Journal of biological chemistry* 275, 2259-2264 (2000).
154. **Motorin, Y., Muller, S., Behm-Ansmant, I. & Branlant, C.** Identification of modified residues in RNAs by reverse transcription-based methods. *Methods in enzymology* 425, 21-53, doi:10.1016/S0076-6879(07)25002-5 (2007).
155. **Barth, S., Hury, A., Liang, X. H. & Michaeli, S.** Elucidating the role of H/ACA-like RNAs in trans-splicing and rRNA processing via RNA interference silencing of the *Trypanosoma brucei* CBF5 pseudouridine synthase. *The Journal of biological chemistry* 280, 34558-34568, doi:10.1074/jbc.M503465200 (2005).
156. **Tschudi, C., Richards, F. F. & Ullu, E.** The U2 RNA analogue of *Trypanosoma brucei* gambiense: implications for a splicing mechanism in trypanosomes. *Nucleic acids research* 14, 8893-8903 (1986).
157. **Miller, S. I. & Wirth, D. F.** trans splicing in *Leishmania enriettii* and identification of ribonucleoprotein complexes containing the spliced leader and U2 equivalent RNAs. *Molecular and cellular biology* 8, 2597-2603 (1988).
158. **Mottram, J. et al.** Isolation and sequence of four small nuclear U RNA genes of *Trypanosoma brucei* subsp. *brucei*: identification of the U2, U4, and U6 RNA analogs. *Molecular and cellular biology* 9, 1212-1223 (1989).
159. **Hartshorne, T. & Agabian, N.** A new U2 RNA secondary structure provided by phylogenetic analysis of trypanosomatid U2 RNAs. *Genes & development* 4, 2121-2131 (1990).
160. **Tschudi, C., Williams, S. P. & Ullu, E.** Conserved sequences in the U2 snRNA-encoding genes of Kinetoplastida do not include the putative branchpoint recognition region. *Gene* 91, 71-77 (1990).

161. **He, P. & Bellofatto, V.** Structure of the *Leptomonas seymouri* trans-spliceosomal U2 snRNA-encoding gene; potential U2-U6 snRNA interactions conform to the cis-splicing counterpart. *Gene* 165, 131-135 (1995).
162. **Hitchcock, R. A., Zeiner, G. M., Sturm, N. R. & Campbell, D. A.** The 3' termini of small RNAs in *Trypanosoma brucei*. *FEMS microbiology letters* 236, 73-78, doi:10.1016/j.femsle.2004.05.020 (2004).
163. **Ambrosio, D. L., Silva, M. T. & Cicarelli, R. M.** Cloning and molecular characterization of *Trypanosoma cruzi* U2, U4, U5, and U6 small nuclear RNAs. *Memorias do Instituto Oswaldo Cruz* 102, 97-105 (2007).
164. **Broadbent, S., van der Woude, M. & Aziz, N.** Accurate and simple sizing of primer extension products using a non-radioactive approach facilitates identification of transcription initiation sites. *Journal of microbiological methods* 81, 256-258, doi:10.1016/j.mimet.2010.03.021 (2010).
165. **Wang, P. et al.** Sm core variation in spliceosomal small nuclear ribonucleoproteins from *Trypanosoma brucei*. *The EMBO journal* 25, 4513-4523, doi:10.1038/sj.emboj.7601328 (2006).
166. **Tschudi, C. & Ullut, E.** Unconventional rules of small nuclear RNA transcription and cap modification in trypanosomatids. *Gene expression* 10, 3-16 (2002).
167. **Padilla-Mejia, N. E. et al.** Gene organization and sequence analyses of transfer RNA genes in Trypanosomatid parasites. *BMC genomics* 10, 232, doi:10.1186/1471-2164-10-232 (2009).
168. **Marchetti, M. A., Tschudi, C., Silva, E. & Ullu, E.** Physical and transcriptional analysis of the *Trypanosoma brucei* genome reveals a typical eukaryotic arrangement with close interspersions of RNA polymerase II- and III-transcribed genes. *Nucleic acids research* 26, 3591-3598 (1998).
169. **Bell, M., Wohner, R. & Bindereif, A.** U4 small nuclear RNA genes of trypanosomes: cloning of the *Leptomonas seymouri* gene and mutational analysis of core snRNP assembly. *Gene* 247, 77-86 (2000).
170. **Marz, M., Kirsten, T. & Stadler, P. F.** Evolution of spliceosomal snRNA genes in metazoan animals. *Journal of molecular evolution* 67, 594-607, doi:10.1007/s00239-008-9149-6 (2008).
171. **Cross, I. & Rebordinos, L.** 5S rDNA and U2 snRNA are linked in the genome of *Crassostrea angulata* and *Crassostrea gigas* oysters: does the (CT)_n(GA)_n microsatellite stabilize this novel linkage of large tandem arrays? *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada* 48, 1116-1119, doi:10.1139/g05-075 (2005).
172. **Manchado, M. et al.** Molecular characterization and chromosomal mapping of the 5S rRNA gene in *Solea senegalensis*: a new linkage to the U1, U2, and U5 small nuclear RNA genes. *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada* 49, 79-86, doi:10.1139/g05-068 (2006).
173. **Pelliccia, F., Barzotti, R., Bucciarelli, E. & Rocchi, A.** 5S ribosomal and U1 small nuclear RNA genes: a new linkage type in the genome of a crustacean that has three different tandemly repeated units containing 5S ribosomal DNA sequences. *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada* 44, 331-335 (2001).

174. **Liao, D., Pavelitz, T., Kidd, J. R., Kidd, K. K. & Weiner, A. M.** Concerted evolution of the tandemly repeated genes encoding human U2 snRNA (the RNU2 locus) involves rapid intrachromosomal homogenization and rare interchromosomal gene conversion. *The EMBO journal* 16, 588-598, doi:10.1093/emboj/16.3.588 (1997).
175. **Ebel, C., Frantz, C., Paulus, F. & Imbault, P.** Trans-splicing and cis-splicing in the colourless Euglenoid, *Entosiphon sulcatum*. *Current genetics* 35, 542-550 (1999).
176. **Liao, D. & Weiner, A. M.** Concerted evolution of the tandemly repeated genes encoding primate U2 small nuclear RNA (the RNU2 locus) does not prevent rapid diversification of the (CT)_n(GA)_n microsatellite embedded within the U2 repeat unit. *Genomics* 30, 583-593, doi:10.1006/geno.1995.1280 (1995).
177. **Newby, M. I. & Greenbaum, N. L.** Sculpting of the spliceosomal branch site recognition motif by a conserved pseudouridine. *Nature structural biology* 9, 958-965, doi:10.1038/nsb873 (2002).
178. **Palfi, Z. et al.** SMN-assisted assembly of snRNP-specific Sm cores in trypanosomes. *Genes & development* 23, 1650-1664, doi:10.1101/gad.526109 (2009).
179. **Ben-Shlomo, H., Levitan, A., Beja, O. & Michaeli, S.** The trypanosomatid *Leptomonas collosoma* 7SL RNA gene. Analysis of elements controlling its expression. *Nucleic acids research* 25, 4977-4984 (1997).
180. **Barbezier, N. et al.** Processing of a dicistronic tRNA-snoRNA precursor: combined analysis *in vitro* and *in vivo* reveals alternate pathways and coupling to assembly of snoRNP. *Plant physiology* 150, 1598-1610, doi:10.1104/pp.109.137968 (2009).
181. **Guffanti, E. et al.** A minimal promoter for TFIIC-dependent *in vitro* transcription of snoRNA and tRNA genes by RNA polymerase III. *The Journal of biological chemistry* 281, 23945-23957, doi:10.1074/jbc.M513814200 (2006).
182. **Lee, J. Y., Evans, C. F. & Engelke, D. R.** Expression of RNase P RNA in *Saccharomyces cerevisiae* is controlled by an unusual RNA polymerase III promoter. *Proceedings of the National Academy of Sciences of the United States of America* 88, 6986-6990 (1991).
183. **Deng, W. et al.** Organization of the *Caenorhabditis elegans* small non-coding transcriptome: genomic features, biogenesis, and expression. *Genome research* 16, 20-29, doi:10.1101/gr.4139206 (2006).
184. **Pfeffer, S. et al.** Identification of microRNAs of the herpesvirus family. *Nature methods* 2, 269-276, doi:10.1038/nmeth746 (2005).
185. **Borchert, G. M., Lanier, W. & Davidson, B. L.** RNA polymerase III transcribes human microRNAs. *Nature structural & molecular biology* 13, 1097-1101, doi:10.1038/nsmb1167 (2006).
186. **Bowden, R. J., Simas, J. P., Davis, A. J. & Efstathiou, S.** Murine gammaherpesvirus 68 encodes tRNA-like sequences which are expressed during latency. *The Journal of general virology* 78 (Pt 7), 1675-1687, doi:10.1099/0022-1317-78-7-1675 (1997).
187. **Kim, M. K. et al.** Identification of SNAPc subunit domains that interact with specific nucleotide positions in the U1 and U6 gene promoters. *Molecular and cellular biology* 30, 2411-2423, doi:10.1128/MCB.01508-09 (2010).

188. **Gilinger, G., Luo, H. & Bellofatto, V.** *In vivo* transcription analysis utilizing chromatin immunoprecipitation reveals a role for trypanosome transcription factor PBP-1 in RNA polymerase III-dependent transcription. *Molecular and biochemical parasitology* 134, 169-173 (2004).
189. **Schimanski, B., Laufer, G., Gontcharova, L. & Gunzl, A.** The *Trypanosoma brucei* spliced leader RNA and rRNA gene promoters have interchangeable TbSNAP50-binding elements. *Nucleic acids research* 32, 700-709, doi:10.1093/nar/gkh231 (2004).
190. **Thomas, S., Green, A., Sturm, N. R., Campbell, D. A. & Myler, P. J.** Histone acetylations mark origins of polycistronic transcription in *Leishmania major*. *BMC genomics* 10, 152, doi:10.1186/1471-2164-10-152 (2009).
191. **Thomas, S., Yu, M. C., Sturm, N. R. & Campbell, D. A.** A non-universal transcription factor? The *Leishmania tarentolae* TATA box-binding protein LtTBP associates with a subset of promoters. *International journal for parasitology* 36, 1217-1226, doi:10.1016/j.ijpara.2006.04.004 (2006).
192. **Das, A. & Bellofatto, V.** RNA polymerase II-dependent transcription in trypanosomes is associated with a SNAP complex-like transcription factor. *Proceedings of the National Academy of Sciences of the United States of America* 100, 80-85, doi:10.1073/pnas.262609399 (2003).

10. Apéndice

10.1. Artículo de investigación derivado de los estudios de Doctorado

Referencia completa:

Rojas-Sánchez, S., Figueroa-Angulo, E., Moreno-Campos, R., Florencio-Martínez, L.E., Manning-Cela, R.G. y Martínez-Calvillo, S. Transcription of *Leishmania major* U2 small nuclear RNA gene is directed by extragenic sequences located within a tRNA-like and a tRNA-Ala gene. *Parasites & Vectors*. 2016 Jul 19;9(1):401. doi: 10.1186/s13071-016-1682-3

RESEARCH

Open Access



Transcription of *Leishmania major* U2 small nuclear RNA gene is directed by extragenic sequences located within a tRNA-like and a tRNA-Ala gene

Saúl Rojas-Sánchez¹, Elisa Figueroa-Angulo¹, Rodrigo Moreno-Campos¹, Luis E. Florencio-Martínez¹, Rebeca G. Manning-Cela² and Santiago Martínez-Calvillo^{1*}

Abstract

Background: *Leishmania* and other trypanosomatid parasites possess atypical mechanisms of gene expression, including the maturation of mRNAs by trans-splicing and the involvement of RNA Polymerase III in transcription of all snRNA molecules. Since snRNAs are essential for trans-splicing, we are interested in the study of the sequences that direct their expression. Here we report the characterization of *L. major* U2 snRNA promoter region.

Results: All species of *Leishmania* possess a single U2 snRNA gene that contains a divergently-oriented tRNA-Ala gene in the upstream region. Between these two genes we found a tRNA-like sequence that possesses conserved boxes A and B. Primer extension and RT-qPCR analyses with RNA from transiently-transfected cells showed that transcription of *L. major* U2 snRNA is almost abolished when boxes A and B from the tRNA-like are deleted or mutated. The levels of the U2 snRNA were also highly affected when base substitutions were introduced into box B from the tRNA-Ala gene and the first nucleotides of the U2 snRNA gene itself. We also demonstrate that the tRNA-like is transcribed, generating a main transcript of around 109 bases. As pseudouridines in snRNAs are required for splicing in other organisms, we searched for this modified nucleotide in the *L. major* U2 snRNA. Our results show the presence of six pseudouridines in the U2 snRNA, including one in the Sm site that has not been reported in other organisms.

Conclusions: Four different regions control the transcription of the U2 snRNA gene in *L. major*: boxes A and B from the neighbor tRNA-like, box B from the upstream tRNA-Ala gene and the first nucleotides of the U2 snRNA. Thus, the promoter region of *L. major* U2 snRNA is different from any other promoter reported for snRNAs. Pseudouridines could play important roles in *L. major* U2 snRNA, since they were found in functionally important regions, including the branch point recognition region and the Sm binding site.

Keywords: *Leishmania*, U2 snRNA, Promoter region, tRNA-like, Pseudouridine

Background

The information in eukaryotic protein-coding genes is frequently interrupted by introns, which have to be removed from the precursors of the mRNAs before they can be translated. The process in which introns are removed and exons are joined together is called splicing. This procedure is carried out in a stepwise coordinated fashion by a large

ribonucleoprotein complex, named spliceosome, that consists of around 200 proteins [1] and five major small nuclear RNAs (snRNAs): U1, U2, U4, U5 and U6 [2, 3]. Each snRNA binds several specific proteins, as well as common proteins, to form a small nuclear ribonucleoprotein particle (snRNP). Among these proteins, the Sm proteins form a donut-shaped heptamer that interacts with a single-stranded region, the Sm site, within the U1, U2, U4 and U5 snRNAs [4]. On the other hand, the U6 snRNA binds Sm-like proteins that recognize a U-tract sequence located at the 3' end of the RNA [5].

* Correspondence: scalv@campus.iztacala.unam.mx

¹Unidad de Biomedicina, Facultad de Estudios Superiores Iztacala, Universidad Nacional Autónoma de México, Av. de los Barrios 1, Col. Los Reyes Iztacala, Tlalnepantla, Edo. de México CP 54090, Mexico
Full list of author information is available at the end of the article

The sequence and secondary structure of snRNAs are highly conserved, even between phylogenetically distant organisms [6]. All snRNAs are extensively post-transcriptionally modified and pseudouridine is the most abundant modified nucleotide in these RNA molecules [7]. Notably, most pseudouridines are located in functionally important regions of the snRNAs. In fact, several studies have demonstrated that pseudouridines play important roles in both snRNP assembly and splicing [8, 9].

Two different RNA polymerases (RNAP) are involved in the synthesis of snRNAs: U1 to U5 are transcribed by RNAP II, whereas U6 is transcribed by RNAP III [10]. In humans, the snRNA core promoter recognized by both RNAP II and III is composed of two elements: a proximal sequence element that is located between nucleotides -50 to -65, and a distal sequence element lying between positions -200 to -220, in relation to the transcription start site (TSS). Additionally, the RNAP III promoter contains a TATA box element situated around position -25 [11, 12].

Although splicing is widely conserved among eukaryotes, some organisms possess special forms of this mechanism to process mRNAs. For instance, protein-coding genes in *Leishmania* and other protozoan parasites that belong to the family Trypanosomatidae are organized into polycistronic gene clusters that are transcribed by RNAP II to generate polycistronic transcripts; and mature mRNAs are produced from these precursors by trans-splicing and polyadenylation. Trans-splicing is a process that adjoins a capped 39-nucleotide minixenon or spliced leader to the 5' termini of all the mRNAs [13]. Like cis-splicing, trans-splicing occurs via two transesterification reactions and requires the participation of snRNAs, but it involves the formation of a Y structure instead of a lariat intermediate [14, 15]. Although every snRNA has an essential participation in splicing, it has been shown that U2 and U6 snRNAs are key role players in the process, as they base-pair with each other to form the splicing catalytic core [16].

Another peculiarity of gene expression in trypanosomatids is that all snRNAs are transcribed by RNAP III [17]. Moreover, snRNA genes in *Trypanosoma brucei* have a divergently oriented tRNA gene, or a tRNA-like sequence, in their 5'-flanking region, and boxes A and B from the neighboring tRNA gene are essential for expression of the snRNAs [17–19]. Notably, the distance between box A from the tRNA gene and the snRNA gene is conserved, as it is usually ~104 bp in length [20, 21]. A third element located within the first 21 nucleotides of the snRNA coding region is also required for proper transcription initiation [17, 22]. However, not every snRNA gene needs all three sequence elements for transcription, as box B is dispensable for the *in vitro*

synthesis of the U6 snRNA [22]. Also, *in vivo* expression of the U1 snRNA does not require an element located within the snRNA itself [19]. Moreover, box A of the associated tRNA gene is dispensable for U4 snRNA transcription in the related trypanosomatid *Leptomonas collosoma* [23].

Little attention has been paid to the study of snRNAs in *Leishmania*. In the present study we have analyzed the sequences that direct transcription of the U2 snRNA gene in *L. major*, the causative agent of cutaneous leishmaniasis in the Old World. As reported in other trypanosomatids, analysis of the genomic context of the U2 snRNA gene allowed us to identify a tRNA-like sequence located in the 5' flanking region that contains typical boxes A and B. By performing RT-qPCR analysis we found that these two elements, together with an intragenic sequence, are required for transcription of the U2 snRNA. Unlike other snRNAs in trypanosomatids, our results showed that an additional box B that resides on a tRNA-Ala gene located 211 bp upstream the U2 snRNA gene, is also important for the optimal transcription of this gene in *L. major*.

Methods

Leishmania major culture and transfection

Promastigotes from *L. major* MHOM/IL/81/Friedlin (LSB-132.1) were grown in BM medium (1× M199 medium pH 7.2, containing 10 % heat-inactivated fetal bovine serum, 0.25× brain heart infusion, 40 mM HEPES, 0.01 mg/ml hemin, 0.0002 % biotin, 100 IU/ml penicillin, 100 µg/ml streptomycin and 1× L-glutamine) at 26 °C and harvested in the mid-log phase. Electroporations were performed following the high-voltage protocol previously described [24]. Usually, 25–50 µg of test plasmid and cotransfection plasmid pLMRIB [25] were aliquoted into 4-mm gap cuvettes, and 500 µl of cells (2×10^8 cells/ml) were added to the cuvette and mixed. The cells were electroporated twice at 1500 V and 25 µF (ECM 630 Electroporation System, BTX, Holliston, USA), pausing 10 s between pulses. Following electroporation, cells were transferred to 10 ml of BM medium and incubated at 26 °C, and total RNA was isolated 24 h post-transfection using TRI reagent (Sigma, St. Louis, USA).

Bioinformatic analyses

Sequence information from different trypanosomatids was obtained from the NCBI (<http://www.ncbi.nlm.nih.gov/>) and TriTrypDB (Release 27) (<http://tritrypdb.org/tritrypdb/>) databases. BLAST searches were performed to obtain information for synteny maps and to search for U2-associated tRNA-like sequences. Alignments were generated using the MUSCLE (<http://www.ebi.ac.uk/Tools/msa/muscle/>) [26] and DNAMAN (version 6,

Lynnon Corporation, Quebec, Canada) programs and corrected manually. Secondary structure analysis was performed using the RNAfold Web Server (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>) [27] with default parameters. Graphical representations of patterns within multiple sequence alignments were generated with the WebLogo application (<http://weblogo.threeplusone.com/>).

Primer extension analysis

Total RNA was analyzed by primer extension with 5' end-labelled synthetic DNA oligonucleotides as described previously [28]. The following oligonucleotides were used: 3'endLmjU2 (5'-GAA AAA AGG AGT TGC TCC CCT GGA AAC GTG-3'), complementary to U2 nt 122 to 151; LmjU2tag-Rev (5'-ATT AGA GTC GAG GTC AGA CC-3'), complementary to a 15-nt tag sequence inserted between U2 nt 83 and 84 (tag sequence is underlined); and 5'LmjU2DcajaB (5'-ATC CGG CCC TGG TCT CCA AA-3'), complementary to tRNA-like nt 58 to 77. In a typical experiment, 20 to 50 µg of total RNA were hybridized to 1×10^5 cpm of γ - 32 P-labelled oligonucleotide at 55 °C for 90 min. The reaction volume was then increased to 50 µl by the addition of a solution containing 39.5 µl of RT buffer and 0.5 µl of SuperScript III Reverse Transcriptase (Invitrogen, Carlsbad, USA). The reaction was incubated at 55 °C for 60 min. Nucleic acids were precipitated and an aliquot of the products was separated on 8 % polyacrylamide-8 M urea gels and exposed to phosphor imaging plates. Autoradiograms were scanned on a Fujifilm FLA-5000 system (Fuji Photo Film Co., Ltd., Tokyo, Japan). When the U2 promoter region was analyzed, bands corresponding to a non-specific product were used to normalize samples of the U2 tagged-specific primer extension. Sequence ladders of U2, U2-tagged and tRNA-like sequences were obtained using a Sequenase 7-deaza-dGTP Sequencing Kit (USB, Cleveland, USA) as specified by the manufacturer. Constructs ptRNAU2 and pComp were used as templates in the sequencing reactions.

5'-RACE analysis

5' Rapid Amplification of cDNA Ends (5'-RACE) experiments were performed with 5 µg of total RNA from *L. major* with a kit from Thermo Scientific (Waltham, USA). First strand cDNA was synthesized with primer 3'U2GST1 (5'-AAC GTG GAA CTC CAA GGA AA-3'), and PCR amplifications were performed with nested primer 3'U2GST2 (5'-CCT TGA GTT GTA ATT TCT AT-3') and nested Abridged Anchor Primer, AAP (5'-GGC CAC GCG TCG ACT AGT ACG GGI IGG GII GGG IIG-3'). The nested PCR products were cloned into the pGEM-T Easy vector (Promega, Fitchburg, USA) and sequenced.

RT-PCR analysis

U2 snRNA transcription termination sites were mapped by poly(A) tailing of total RNA, as previously reported [29]. The cDNA was prepared with oligonucleotide Nested(dT) (5'-CCT CTG AAG GTT CAC GGA TCC ACA TCT AGA TTT TTT TTT TTT TTT TTT VN-3'). The first PCR was performed with primers U2Sa-5' (5'-ATA TCT TCT CGG CTA TTT AGC-3') and B1 (5'-CCT CTG AAG GTT CAC GGA T-3'), and the second PCR was done with primers U2qFw2 (5'-CTA TTT AGC TAA GAT CAT GTT TAT A-3') and B2 (5'-CAC GGA TCC ACA TCT AGA T-3'). The final PCR products were cloned into the pGEM-T Easy vector (Promega) and sequenced.

Mapping of pseudouridines

A primer extension-based pseudouridylation assay was carried out as previously described [30, 31]. Briefly, RNA from 5×10^8 cells was treated with *N*-cyclohexyl-*N'*-β-(4-methylmorpholinium)ethyl-carbodiimide *p*-tosylate (CMCT) at 37 °C for 20 min in 50 µl of CMCT buffer (50 mM bicine, pH 8.0, 7 M urea and 4 mM EDTA) containing 170 mM CMCT reagent. To remove CMCT groups, the CMCT-treated RNA was subjected to alkali hydrolysis with 80 µl of Na₂CO₃ (50 mM, pH 10.4) at 37 °C for 2 h. One control experiment was done without CMCT treatment, and another one without alkali hydrolysis. The treated RNA (25 to 50 µg) was used as template in primer extension analysis with the 3'end-LmjU2 oligonucleotide. In this method, the reverse transcriptase stops 1 nt before the modified base.

Plasmid constructs

To obtain the plasmids used in this work, first a 429-bp fragment containing the U2 snRNA gene and 292 nt from its 5'-flanking sequence, was amplified by PCR with primers 5'LmjtRNAaala (5'-TGA AAA AGT TGG AGA AGT TG-3') and 3'U2Lmjend (5'-TCC CCT GGA AAC GTG GAA CT-3'). The PCR product was cloned into the pGEM-T vector (Promega) to obtain plasmid ptRNAU2. This construct was used as template in a site-directed mutagenesis assay in which a 15-bp tag sequence was inserted between positions 83 and 84 of the U2 snRNA gene with oligonucleotides 5'U2oligoTag (5'-GCC TTC GGG CAA AGG TCT GAC CTC GAC TCT AAT AGA AAT TAC AAC-3') and 3'U2oligoTag (5'-GTT GTA ATT TCT ATT AGA GTC GAG GTC AGA CCT TTG CCC GAA GGC-3'). The resultant plasmid was called ptRNAU2M. This construct lacked the T tract located downstream of the U2 gene. Thus, to add the transcription termination signal, a 473-bp DNA fragment from ptRNAU2M was amplified with oligonucleotides LmjtRNAaala-XmaI-F (5'-ACC CGG GTG AAA AAG TTG GAG AAG TTG-3') and endLmjU2-XbaI-R (5'-ATC TAG AGA AAA AAG

GAG TTG CTC CCC TGG AAA CGT G-3') and cloned into the pGEM-T Easy vector (Promega), generating construct pComp. To obtain the vectors that contain sequential deletions of the U2 snRNA 5'-flanking sequence, several DNA fragments were amplified from pComp. The same reverse oligonucleotide (endLmjU2-XbaI-R) was used in all PCR reactions, together with the following forward oligonucleotides: U2tRNA-like-XmaI-F (5'-ACC CGG GAG GAA AAG ATG CTT TCG ACG AG-3') was used for pΔtRNA-Ala; LmjU2dcajaB-XmaI-F (5'-ACC CGG GAT CCC GCC CTG GTC TCC AAA-3') was used for pΔBox-B; LmjU2dcajaA-XmaI-F (5'-ACC CGG GGT AAG CGG CAC GGC AGT GAG-3') was employed for pΔBox-A; and primer U2Sa-XmaI-F (5'-ACC CGG GAT ATC TTC TCG GCT ATT TAG C-3') was used for pΔ-293/-1. All amplicons were cloned into the pGEM-T Easy vector (Promega). To obtain plasmids with base substitutions in the coding and 5'-flanking regions of the U2 snRNA, construct pComp was used as template in mutagenesis reactions performed with the following oligonucleotides: 5U2LmjmutFw (5'-TGG TAC TAA CAT ATC CAT GGA ACT ATT TAG CTA AGA T-3') and 5U2LmjmutRv (5'-ATC TTA GCT AAA TAG TTC CAT GGA TAT GTT AGT ACC A-3') for pBS + 6/+12; primers TLBTWU2LmjmutFw (5'-CAG GTA CCG TTT GGA TAG GAT CGC CCG ATT CCT TCG GGG TT-3') and TLBTWU2LmjmutRv (5'-AAC CCC GAA GGA ATC GGG CGA TCC TAT CCA AAC GGT ACC TG-3') for pBS-128/-138; oligonucleotides BalaU2LmjmutFw (5'-ATG CGG TAG GTA TTG TTG GAC GTA CCC AAC TTC TCC A-3') and BalaU2LmjmutRv (5'-TGG AGA AGT TGG GTA CGT CCA ACA ATA CCT ACC GCA T-3') for pBS-263/-269; primers AalaU2LmjmutFw (5'-ACC GCG TCG GGG ATG GAT ATC ATA GCC TAG AGC GAC CGC TTA-3') and AalaU2LmjmutRv (5'-TAA GCG GTC GCT CTA GGC TAT GAT ATC CAT CCC CGA CGC GGT-3') for pBS-219/-230; oligonucleotides BlikeU2LmjmutFw (5'-GCG GGA TTC CTT CGG TTG GAC GAC CTC GTC GAA AGC A-3') and BlikeU2LmjmutRv (5'-TGC TTT CGA CGA GGT CGT CCA ACC GAA GGA ATC CCG C-3') for pBS-150/-156; and primers AlikeU2LmjmutFw (5'-CGC GGC GTT TTT TTT GAT ACC ATC GCC TAC CGT TTG GAG ACC-3') and AlikeU2LmjmutRv (5'-GGT CTC CAA ACG GTA GGC GAT GGT ATC AAA AAA AAC GCC GCG-3') for pBS-105/-116. Finally, to generate pDBS that possesses a double mutation, plasmid pBS-105/-116 was used as template in a mutagenesis reaction with oligonucleotides BlikeU2LmjmutFw and BlikeU2LmjmutRv. All mutagenesis reactions were performed with a QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, USA). In boxes A and B, the mutated nucleotides correspond to the most conserved positions, which have

been shown to be important for Pol III transcription in other organisms [32, 33]. The identity of each insert was confirmed by sequencing using the T7 and SP6 oligonucleotides.

Real-time quantitative PCR

Around 1 μg of DNase I-treated RNA from transiently transfected cells (from three independent transfections) was used as template for first strand cDNA synthesis using SuperScript™ III Reverse Transcriptase (Invitrogen) and 2 pmol of oligonucleotides 3'U2Lmjend or Alfa-tub-3' (5'-GTA GTT GAT GCC GCA CTT GAA G-3'). The cDNA was analyzed by real-time quantitative PCR (RT-qPCR) assays using the Platinum SYBR Green qPCR SuperMix-UDG kit (Invitrogen) in a Rotor-Gene 3000 cycler (Corbett Research, Mortlake, Australia) according to the manufacturer's recommendations. All RT-qPCR reactions were performed at least in duplicate, using primers and conditions that were optimized to produce a single amplicon of the correct size. Each amplification product was analyzed for specificity by both agarose gel electrophoresis and melt curve analysis. Standard curves for primer pairs were derived from genomic and plasmid DNA dilution series and ranged in their r^2 value from 0.98 to 1.0. PCR efficiencies were near to 100 % for the used genes, so the data were analyzed by the $2^{-\Delta\Delta C_q}$ method. For normalization of the data we used α -tubulin (*LmjF.13.0280*) as a reference gene, and all values were represented relative to data from cells transfected with plasmid pComp. The tagged-U2 snRNA was amplified with primers U2Sa-5' and LmjU2tag-Rev. The α -tubulin transcript was amplified with oligonucleotides Alfa-tub-5' (5'-AGA AGT CCA AGC TCG GCT ACA C-3') and a-tubLmjrvs (5'-GGT CGT AGA TGG CCT CAT TG-3'). To analyze RNA from cells transfected with pBS + 6/+12, the same conditions indicated above were performed, but using U2qFw2 as the sense primer in the PCR assays.

Northern blot analysis

Total RNA (10–30 μg) was separated on 10 % polyacrylamide-8 M urea gels. After electrophoresis, RNA was transferred to Hybond N+ membranes (GE Healthcare, Little Chalfont, England) by electro-blotting using a Trans-Blot SD Semi-Dry Transfer Cell (Bio-Rad, Hercules, USA). The tRNA-Ala probe corresponds to the Ala-Rvs (5'-TCG ATC CCA ATA CCT ACC GC-3') oligonucleotide, the tRNA-like probe corresponds to the 5'LmjU2dcajaB oligonucleotide, the U2 snRNA probe to the 3'endLmjU2 oligonucleotide, and the tag probe to the LmjU2tag-Rev oligonucleotide. Primers were labelled with [γ - 32 P]ATP using T4 kinase. Hybridizations were performed in 6× SSPE (60 mM Na₂HPO₄, 0.9 M NaCl and 6 mM EDTA), 5× Denhardt's reagent and 1 % SDS

at 42 °C. Washing was carried out at 55 °C in 0.2× SSC and 0.1 % SDS.

Results

Genomic context of the U2 snRNA in *L. major* and other trypanosomatids

The *L. major* genome contains a single copy of the U2 snRNA gene (*LmjF.31.snRNA.01*), which is embedded into an RNAP II polycistronic unit on chromosome 31 [34]. The U2 snRNA gene is flanked by two different amino acid transporter genes: *LmjF.31.0580* and *LmjF.31.0570* (Fig. 1a). A divergently-oriented tRNA-Ala gene (*LmjF.31.*

TRNAALA.01) is located 211 bp upstream of the U2 snRNA gene (Fig. 1a). To determine the genomic context of U2 snRNA genes in other species of *Leishmania*, we examined the genome databases of *L. donovani* (strain BPK282A1) [35], *L. infantum* (JPCM5), *L. braziliensis* (M2904) [36, 37], *L. mexicana* (U1103) [37] and *L. tarentolae* (Parrot-TarII) [38]. Similarly to *L. major*, in all of the *Leishmania* species analyzed the U2 snRNA is encoded by a single-copy gene inserted into the same polycistronic unit (Fig. 1a), which is located on chromosome 31 in *L. donovani*, *L. infantum*, *L. braziliensis* and *L. tarentolae*, but on chromosome 30 in *L. mexicana* (as chromosomes

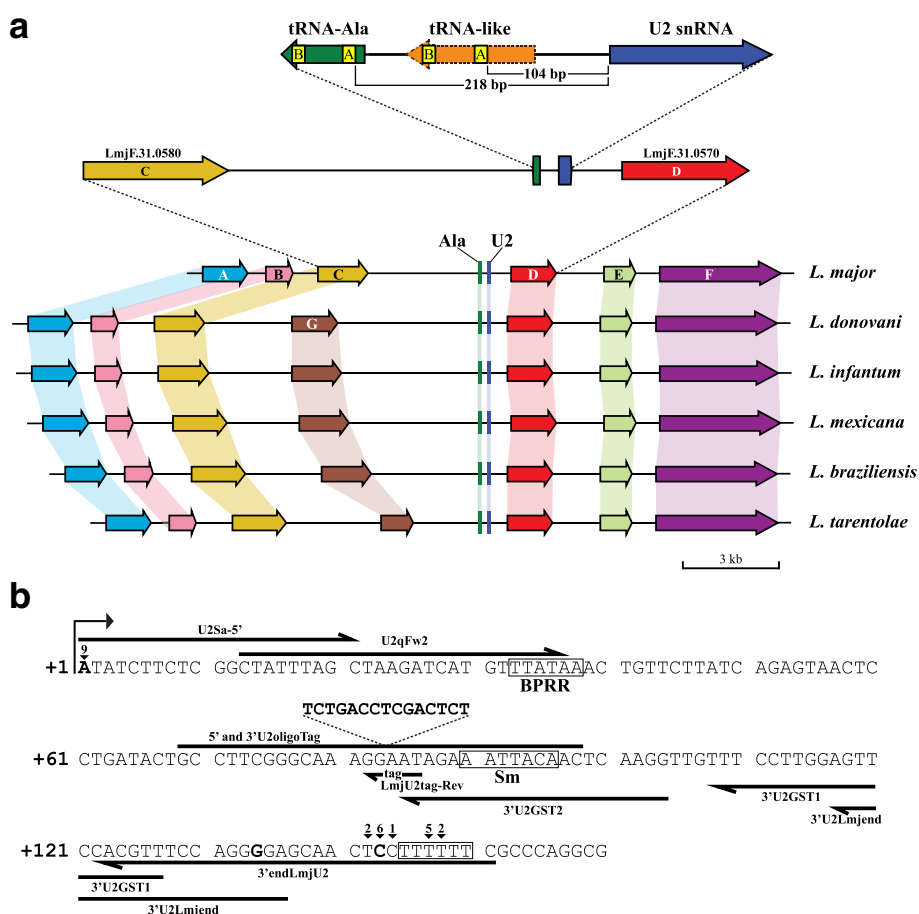


Fig. 1 Genomic context and sequence of the U2 snRNA gene in *L. major*. **a** Synteny of U2 snRNA genes in different species of *Leishmania*. Orthologues genes are joined by colored lines. In *L. major*, gene A (light blue) corresponds to *LmjF.31.0600*, gene B (pink) to *LmjF.31.0590*, gene C (mustard) to *LmjF.31.0580*, gene D (red) to *LmjF.31.0570*, gene E (light green) to *LmjF.31.0560* and gene F (purple) to *LmjF.31.0550*. Gene G (brown) is absent in *L. major*, but corresponds to *LdBPK_310600.1* in *L. donovani*. Two enlargements of the *L. major* U2 snRNA gene are shown. On the top scheme are presented the U2 snRNA (blue), the tRNA-like (orange) and tRNA-Ala (green) genes. Boxes A and B within the tRNA-like and tRNA-Ala gene are shown in yellow. Distances between the U2 snRNA gene and boxes A from the tRNA-like (104 bp) and the tRNA-Ala gene (218 bp) are indicated. Figure is drawn to scale. **b** Sequence of the *L. major* U2 snRNA gene and 3'-flanking region. The A residue shown in bold type (position +1) corresponds to the TSS mapped by primer extension and 5'-RACE analyses (the arrowhead specify the number of clones found at that position). At the 3' end, the cluster of T nucleotides is shown in a box. The 3' termini mapped by RT-PCR are denoted with arrowheads, specifying the number of clones found at each position. The mature 3' terminus of the U2 snRNA (C at position +143) is shown in bold type. The missannotated 3' end of U2 at position 134 (G) is also shown in bold type. The locations of putative BPRR and Sm site are shown. The sequence of the tag, inserted between positions 83 (G) and 84 (A), is also indicated. The sequence and location of some of the oligonucleotides used in this study are shown. To facilitate the analysis of the results, the sequence is shown as the reverse-complement of the annotated sequence of chromosome 31 from the TriTrypDB database

8 and 29 are fused in this species). Moreover, the divergently oriented tRNA-Ala gene is present in all of the species analyzed. Thus, synteny of the U2 snRNA locus is observed among *Leishmania* species. Interestingly, the genome of *L. major* lacks an extra amino acid transporter gene (Fig. 1a, gene G) that is present in all other species of *Leishmania*.

To extend the analysis of gene copy number of the U2 snRNA to other trypanosomatids, the genomic databases of members of the genera *Trypanosoma*, *Crithidia*, *Endotrypanum* and *Leptomonas* were examined. A single copy of the U2 snRNA gene is present in *T. brucei* (strains Lister 427, DAL972 and TREU927), *T. evansi* (STIB 805), *T. vivax* (Y486), *C. fasciculata* (Cf-Cl), *E. monterogeii* (LV88), *L. pyrrocoris* (H10) and *L. seymouri* (ATCC 30220). Interestingly, two different strains of *T. cruzi* (CL Brener and Dm28_c) have three copies of the U2 snRNA gene. Thus, trypanosomatids possess only one copy of the U2 snRNA gene, with the exception of *T. cruzi* that contains three copies of this gene.

Mapping of transcription initiation and termination sites in the U2 snRNA

According to the TriTrypDB databases, the length of the *L. major* U2 snRNA gene is 134 bp (Fig. 1b). However, sequence comparisons (Additional file 1: Figure S1) suggest that the 3' end of the gene is missannotated. Thus, to determine the correct size of the U2 snRNA we mapped both the 5' and 3' termini of this snRNA. First, to locate the TSS, a primer extension assay was carried out. As shown in Fig. 2a, the TSS of U2 snRNA corresponds to an A residue (T in the complementary sequence), which is equivalent to the TSS previously identified in *T. cruzi*, *T. brucei* and *L. collosoma* [39, 40]. To corroborate the location of the TSS in the *L. major* U2 snRNA, a 5'-RACE assay was conducted. Analysis of nine cloned sequences confirmed that transcription of the U2 snRNA initiates at the A nucleotide (Fig. 1b). Thus, the TSS of the U2 snRNA is conserved in trypanosomatids.

To determine transcription termination sites of U2 snRNA in *L. major*, RT-PCR was performed with total RNA that was poly(A)-tailed *in vitro*. Transcription of U2 snRNA was found to end within a tract of six Ts located from 145 to 150 bp downstream of the TSS (Fig. 1b). In two of the clones analyzed, the poly(A) tail was added after the fourth T, and in five clones it was added after the third T residue. Trimming of these transcripts generate shorter RNA molecules, as one clone ended at nucleotide +144 (C), six clones ended at position +143 (C), and two clones at base +142 (U). Thus, it appears that during the processing of the 3' end of the U2 snRNA there are several populations of RNA with different 3' ends. Since most clones ended at the C residue located at position +143, we considered this

nucleotide as the mature 3' terminus of the U2 snRNA in *L. major* (Fig. 1b). Consequently, by mapping both 5' and 3' ends of U2 snRNA in *L. major*, we have determined that the mature transcript is 143 nt long.

Localization of pseudouridine residues in the U2 snRNA

Pseudouridylation of the U2 snRNA is essential for snRNP assembly and splicing in vertebrates and yeast [9]. To determine the presence of pseudouridine residues in the *L. major* U2 snRNA, total RNA was treated with CMCT followed by alkaline treatment and primer extension analysis. Considering that in this procedure the reverse transcriptase extension stops one base downstream of the CMCT-modified nucleotide, we found six pseudouridines in the U2 snRNA sequence at positions 14, 34, 36, 55, 86 and 93 (Fig. 2b, Lane 6). A control experiment was done without CMCT treatment to detect natural pauses of the reverse transcriptase, probably as a result of RNA degradation or the presence of stable RNA secondary structures (Fig. 2b, Lane 5). Another control without alkali hydrolysis was carried out to identify all the positions where the CMCT reagent was attached (Fig. 2b, Lane 7). In the experiment shown in Fig. 2b, the band that corresponds to pseudouridine at position 14 is very faint. However, in other experiments this band was clearly observed (Fig. 2c, Lane 6).

To determine if the pseudouridines are located in important functional domains of the U2 snRNA, the predicted secondary structure of the RNA molecule was generated and putative branch point recognition region (BPRR) and Sm binding sites were identified (Fig. 3a). Regarding the BPRR, the well-conserved metazoan sequence GUAGUA is replaced by UUAUAA in the *L. major* U2 snRNA (Fig. 3a). Also, the Sm site in the U2 snRNA, previously identified as AAUUACA (from nucleotides 90 to 96) (Fig. 3a), differs from the metazoan consensus sequence RAU₃₋₆GR [41]. Interestingly, pseudouridines at positions 34 and 36 are located within the BPRR, and pseudouridine at base 93 is situated in the Sm site (Fig. 3a). It is worth noting that a pseudouridine in the Sm site has not been reported in any other organism.

As expected, the sequence of the U2 snRNA is conserved across trypanosomatids, as shown by a multiple sequence alignment (Additional file 1: Figure S1) and a sequence logo created with the WebLogo application [42] (Fig. 3b). Homology between the *L. major* U2 snRNA and its orthologues in other species of *Leishmania* is very high, ranging from 95.1 % identity with *L. braziliensis* to 98.6 % identity with *L. tarentolae*. Also, *L. major* U2 snRNA shows 69.4 % and 76.2 % identity to its orthologues in *T. cruzi* and *T. brucei*, respectively. Notably, we found that the U2 snRNA genes from all of the 21 species of trypanosomatids that we analyzed

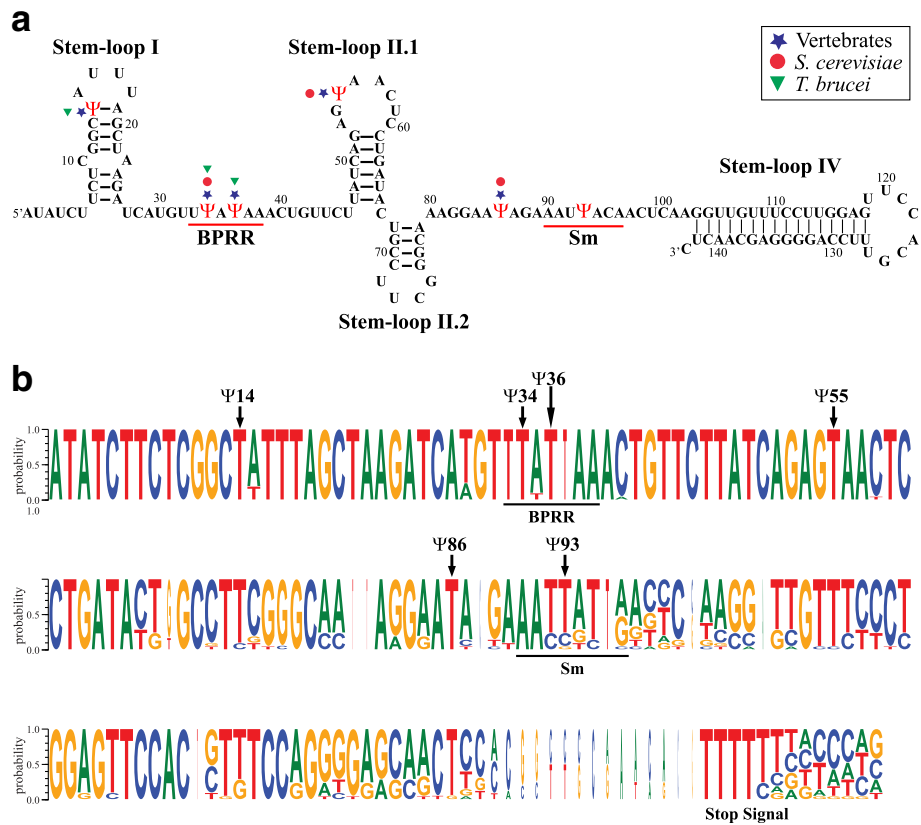
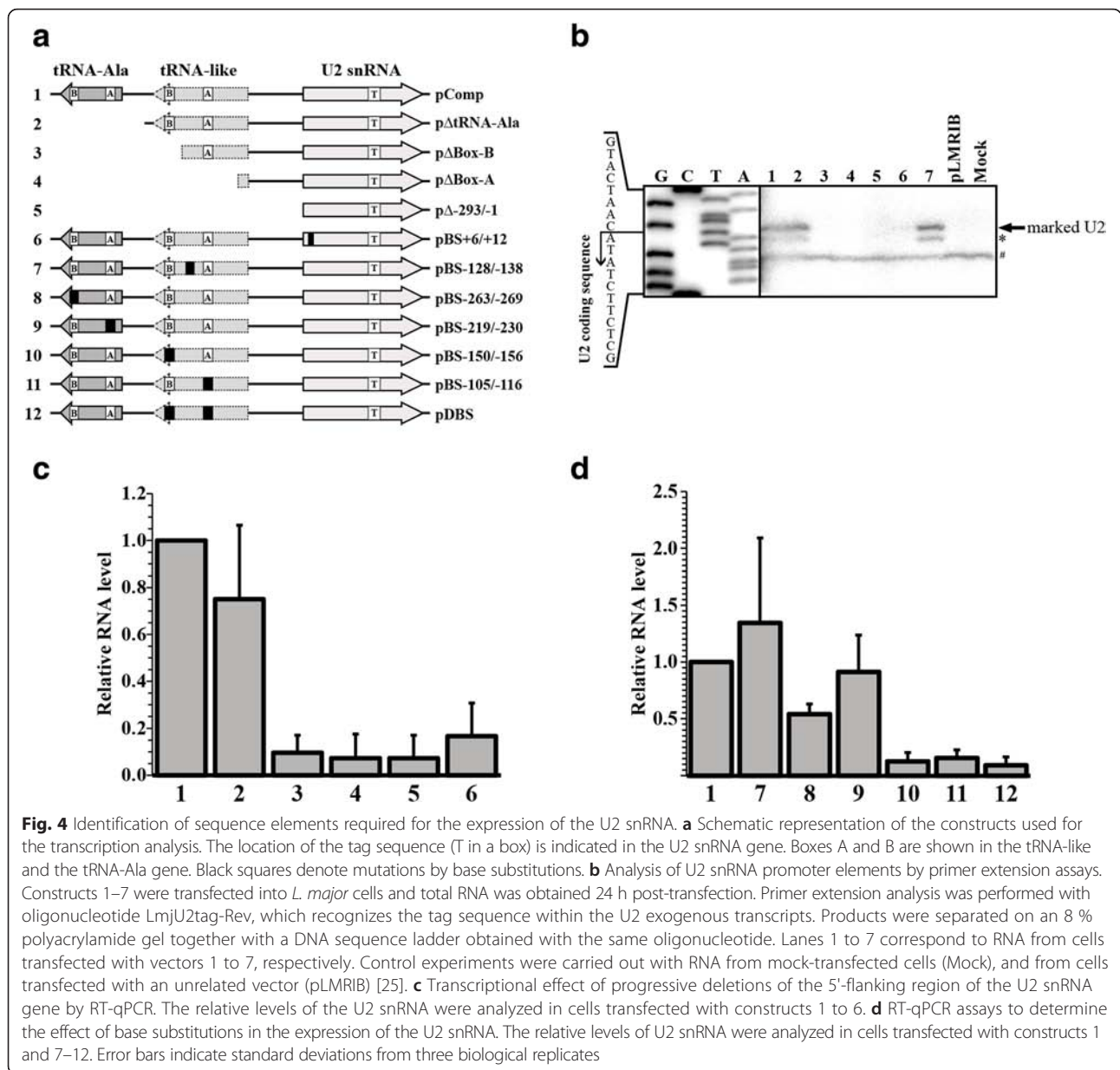


Fig. 3 Sequence conservation and secondary structure of the U2 snRNA. **a** Predicted secondary structure of the *L. major* U2 snRNA. Stem-loops I, II.1, II.2 and IV are indicated. Also, BPRR and Sm site are shown. Pseudouridine residues (red Ψ) identified in this study are denoted. Pseudouridines conserved between *L. major* and vertebrates (blue star), *S. cerevisiae* (red circle) and *T. brucei* (green triangle) are also indicated. **b** Consensus sequence of the U2 snRNA gene in trypanosomatids. The sequences of the U2 snRNA genes and 3'-flanking regions from 21 species of trypanosomatids (see Additional file 1: Figure S1) were aligned with the DNAMAN software. Using the WebLogo application, the generated pattern was represented as a sequence logo by residue probabilities. BPRR, Sm site and stop signal are shown. The positions of the pseudouridine residues (Ψ) mapped in *L. major* are indicated with arrows

external U2 snRNA gene by insertion of a 15-nt tag sequence between positions +83 and +84 (Figs. 1b and 4a). Since no promoter sequences have been reported in this region, it is unlikely that the tag sequence could compromise the activity of a potential internal promoter element. Indeed, expression of tagged U2 snRNA was demonstrated by Northern blot (Additional file 2: Figure S2).

Thus, total RNA was extracted 24 h post-transfection and primer extension analyses were performed using an oligonucleotide complementary to the tag. As shown in Fig. 4b, with pComp (plasmid 1), which contains the complete 5' upstream region (tRNA-Ala and tRNA-like sequences) two primer extension products were detected (Fig. 4b, Lane 1). The longest product corresponds to the U2 snRNA TSS, while the other product was mapped to an A located at position +3 (marked with an asterisk in Fig. 4b) that might represent a reverse transcriptase stop. As expected, these bands were not present in the mock-transfected cells and cells transfected with a control vector (Fig. 4b, Mock and pLMRIB). Deletion of

the tRNA-Ala gene (pΔtRNA-Ala) did not affect the abundance of the U2 snRNA bands (Fig. 4b, Lane 2). In contrast, elimination of box B from the tRNA-like (pΔBox-B, Lane 3), deletion of box A from the tRNA-like (pΔBox-A, Lane 4) and removal of the entire 5' flanking region (pΔ-293/-1, Lane 5) reduced the U2 snRNA transcripts to undetectable levels (Fig. 4b). Therefore, these results show that sequences within the tRNA-like are needed for transcription of the *L. major* U2 snRNA. To evaluate the presence of a possible internal promoter element in the snRNA gene, the sequence of the first nucleotides of the gene (from +6 to +12) were mutated (Additional file 3: Figure S3). In the primer extension analysis with this vector (pBS +6/+12, Lane 6) the snRNA transcripts were strongly reduced, showing that the first bases of the U2 snRNA gene are required for expression of the gene. In contrast, mutation of the sequence located between boxes A and B in the tRNA-like did not affect the abundance of the U2 snRNA (pBS-128/-138, Lane 7) (Fig. 4b). A non-specific product that was observed in all the lanes was



used as a control for RNA recovery (marked with a hash character in Fig. 4b).

To obtain a more precise determination of the abundance of the U2 snRNA transcripts, RNA from transiently-transfected cells was analyzed by RT-qPCR. As in the primer extension analysis, to discriminate between endogenous U2 and U2-tagged transcripts, an oligonucleotide complementary to the tag sequence was used. As shown in Fig. 4c, the U2 snRNA relative levels were severely reduced (by 90–93 %) when box B of the tRNA-like and sequences downstream were deleted (compare bar 1 to 3–5). Thus, these results are in agreement with the primer extension results (Fig. 4b, Lanes 3–5), and indicate that the tRNA-like is needed for the *in vivo* expression of

the U2 snRNA. To analyze independently the transcriptional roles of boxes A and B from the tRNA-like, base substitutions were produced in these two elements (Additional file 3: Figure S3). We observed that mutations in box B (Fig. 4d, bar 10), box A (Fig. 4d, bar 11) and both boxes (Fig. 4d, bar 12) drastically diminished the U2 snRNA levels to 13, 16 and 9 %, respectively. By contrast, a slight increase in the levels of U2 snRNA was found with the control vector that bears a mutation in the region located between boxes A and B in the tRNA-like, where no promoter element is found (Fig. 4d, bar 7). In agreement with the primer extension analysis, mutation of the first nucleotides of the U2 snRNA gene diminished the RNA levels to 17 % of the value of the control (Fig. 4c, bar 6).

An unexpected result was that deletion of the tRNA-Ala gene, located upstream of the tRNA-like, produced a 25 % reduction in the levels of the U2 snRNA (Fig. 4c, bar 2). To further explore this effect, base substitutions in boxes A and B of the tRNA-Ala gene were generated and analyzed. Interestingly, when box B was mutated, the expression of the U2 snRNA was reduced to 54 % of the value for the control (Fig. 4d, bar 8). In contrast, mutations in box A from the tRNA-Ala gene have a slight impact on the abundance of the U2 snRNA (Fig. 4d, bar 9). Therefore, our results show that box B from the tRNA-Ala gene located upstream of the tRNA-like is important for transcription of the U2 snRNA in *L. major*.

The tRNA-like sequence associated to the U2 snRNA is transcribed

To determine whether the tRNA-like sequence linked to the *L. major* U2 snRNA gene is transcribed, a Northern blot experiment was performed with total RNA isolated

from wild type *L. major* promastigotes. A band of around 109 nt was detected, indicating that the tRNA-like is transcribed in *L. major* (Fig. 5a, middle panel). Several larger and weaker bands were also observed, which may represent longer tRNA-like transcripts or unspecific bands. Control experiments to detect the U2 snRNA and the tRNA-Ala showed the expected bands of around 143 and 73 nt, respectively (Fig. 5a). Densitometry analysis of the observed bands showed that the tRNA-like is at least 10-fold less abundant than the U2 snRNA.

The TSS of the tRNA-like was mapped by primer extension (Fig. 5b). The most abundant band corresponded to a C residue (G in the complementary sequence) (labelled with a red asterisk in Fig. 5b) located 65 bases upstream of the U2 snRNA (Fig. 5c). If we assume that transcription of the tRNA-like ends in the T-tract located downstream of the gene (Fig. 5c), the tRNA-like transcript would be 109 nt long, which is in agreement with the size observed in the Northern blot (Fig. 5a). The smaller bands detected in the primer extension

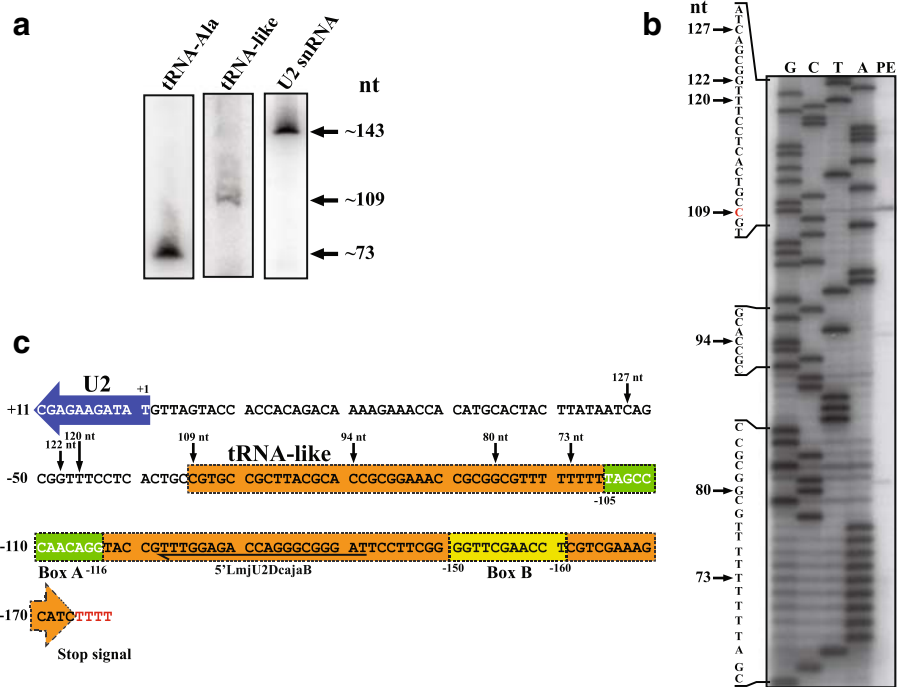


Fig. 5 Transcription of the tRNA-like sequence associated to the U2 snRNA gene. **a** Northern blot analyses with specific probes to detect the tRNA-Ala (left image), tRNA-like (middle image) and U2 snRNA (right image). **b** Mapping of the 5' ends of tRNA-like molecules in wild type promastigotes of *L. major*. Total RNA was subjected to primer extension analysis with oligonucleotide 5'LmjU2DcajaB, and the products were separated on an 8 % polyacrylamide denaturing gel (Lane PE) together with a DNA sequencing ladder obtained with the same oligonucleotide (Lanes G, C, T and A). The primer extension products are marked with asterisks; the red asterisk denotes the most abundant product. Partial sequences of the tRNA-like are shown on the left, indicating with arrows the positions of the 5' termini and the expected size of tRNA-like molecules. **c** Sequence of the tRNA-like, which is boxed with dotted lines (orange box with an arrow indicating the direction of transcription). Primer extension products are marked with arrows, indicating the length of the transcript. The most abundant primer extension product was considered as the TSS (C, 109 nt). Boxes A and B are colored in green and yellow, respectively. The tRNA-like transcription stop signal is shown in red. The location of the 5'LmjU2DcajaB oligonucleotide is indicated. The first eleven nucleotides of the U2 snRNA gene are also shown. Sequence numbers are relative to the TSS (+1) from the U2 snRNA. To facilitate the analysis of the results, the sequence is shown as presented in the TriTrypDB database (which is the reverse-complement of the sequence shown in Fig. 1)

analysis (73, 80 and 94 nt) might represent reverse transcriptase stops, while the larger bands (120, 122 and 127 nt) could denote alternative TSS (black asterisks in Fig. 5b, and black arrows in Fig. 5c).

Sequence analysis of boxes A and B from tRNA-like sequences associated to U2 snRNA genes

Sequence analysis of boxes A (TAG CCC AAC AGG) and B (GGT TCG AAC CT) from the tRNA-like associated to the *L. major* U2 snRNA revealed that they are similar to boxes A (TAG CTC AGA TGG) and B (GGA TCG ATA CC) from the tRNA-Ala gene. To compare these sequence elements across trypanosomatids, we first searched for boxes A and B in tRNA-like sequences associated to U2 snRNA genes in the genome databases of 18 additional species of trypanosomatids (Fig. 6a). A tRNA-like sequence with boxes A and B was found in all the species analyzed (Fig. 6a); and sequence logos were obtained for both sequence elements (Fig. 6b, upper logos). For comparison, sequences of boxes A and B from all of the 258 classic tRNAs found in TriTryps [43] were aligned and logo sequences were generated (Fig. 6b, lower logos). Regarding tRNA-like sequences, analysis of the sequence logo from box A showed a high conservation of most bases, with the exception of the position labelled with an asterisk (which is absent in most species analyzed) and base 9 (Fig. 6b, upper logo). Box B is even more conserved, as positions 1 to 5 (GGTTC) are present in all the sequences, and the rest of the bases are conserved, with the exception of positions 8 and 11 (Fig. 6b, upper logo). In general, high sequence conservation was observed when comparing boxes B from tRNA-like sequences and classic tRNA genes, since position 11 is the only one that is not conserved (Fig. 6b, right logos). On the other hand, boxes A are not as highly conserved as boxes B, since positions 5, 6, 7, 9, 11 and (*) differ between tRNA-like and classic tRNA genes (Fig. 6b, left logos).

It is interesting to note that a distance of 104 bp separates the U2 snRNA gene and box A from the tRNA-like sequence in most trypanosomatids (Fig. 6a). Moreover, the distance that separates box A from box B in the tRNA-like sequences is conserved in *Leishmania*, fluctuating between 33 bases (in *L. major* and *L. braziliensis*) and 36 bases (in *L. mexicana*). In contrast, this distance is not conserved in tRNA-like sequences from trypanosomes, as it varies from 30 bases (*T. vivax*) to 44 bases (*T. cruzi*) (Fig. 6a). Thus, our analysis shows that, while boxes A are more divergent in sequence, their position in relation to the U2 snRNA gene is almost invariable. On the other hand, the sequences of boxes B are highly conserved but their distance to box A (or to the snRNA gene) is variable across trypanosomatids.

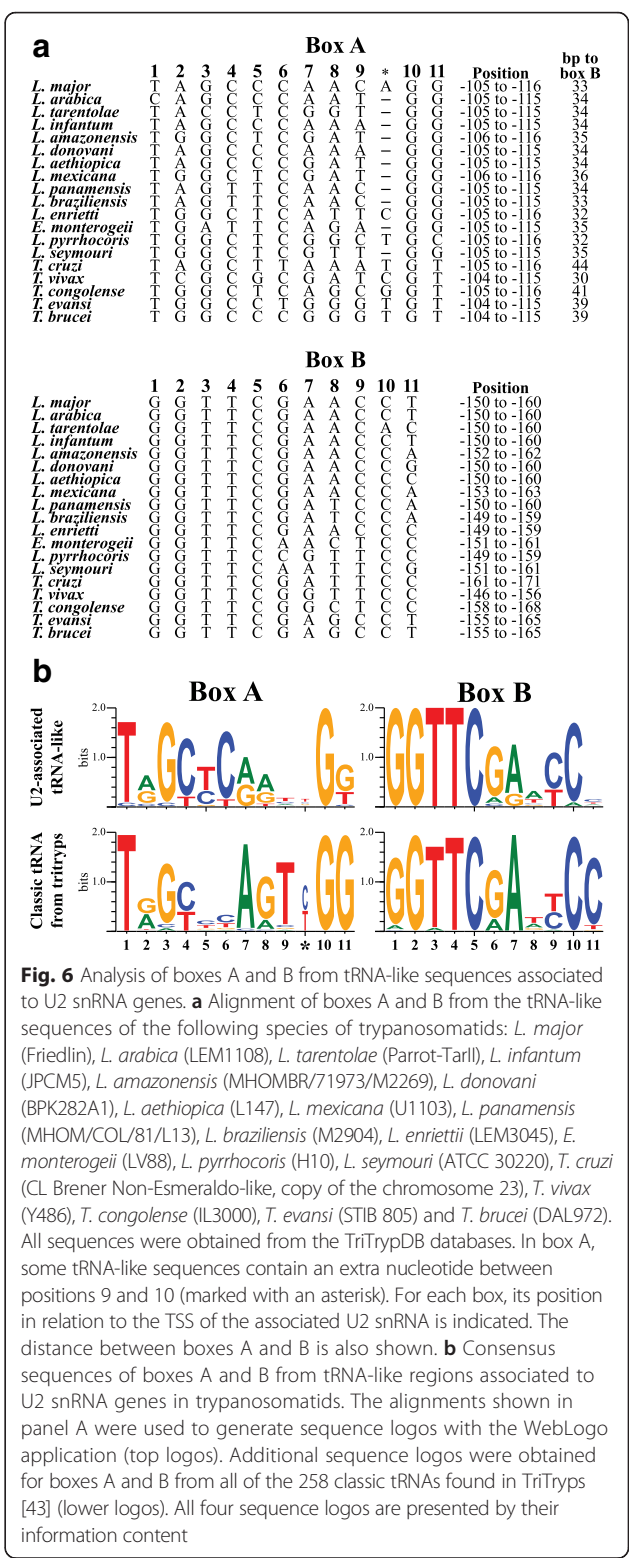


Fig. 6 Analysis of boxes A and B from tRNA-like sequences associated to U2 snRNA genes. **a** Alignment of boxes A and B from the tRNA-like sequences of the following species of trypanosomatids: *L. major* (Friedlin), *L. arabica* (LEM1108), *L. tarentolae* (Parrot-Tarll), *L. infantum* (JPCM5), *L. amazonensis* (MHOMBR/71973/M2269), *L. donovani* (BPK282A1), *L. aethiopica* (L147), *L. mexicana* (U1103), *L. panamensis* (MHOM/COL/81/L13), *L. braziliensis* (M2904), *L. enrietti* (LEM3045), *E. monterogeii* (LV88), *L. pyrrocoris* (H10), *L. seymouri* (ATCC 30220), *T. cruzi* (CL Brener Non-Esmeraldo-like, copy of the chromosome 23), *T. vivax* (Y486), *T. congolense* (IL3000), *T. evansi* (STIB 805) and *T. brucei* (DAL972). All sequences were obtained from the TriTrypDB databases. In box A, some tRNA-like sequences contain an extra nucleotide between positions 9 and 10 (marked with an asterisk). For each box, its position in relation to the TSS of the associated U2 snRNA is indicated. The distance between boxes A and B is also shown. **b** Consensus sequences of boxes A and B from tRNA-like regions associated to U2 snRNA genes in trypanosomatids. The alignments shown in panel A were used to generate sequence logos with the WebLogo application (top logos). Additional sequence logos were obtained for boxes A and B from all of the 258 classic tRNAs found in TriTryps [43] (lower logos). All four sequence logos are presented by their information content

Discussion
 In this work we have studied the sequence, structure and genomic context of the U2 snRNA gene in *Leishmania* and other trypanosomatids, and have identified

the sequences that drive transcription of the U2 snRNA in *L. major*. Although in general the sequence of the U2 snRNA is highly conserved in trypanosomatids, the sequence and length of its 3' end are variable (Fig. 3b and Additional file 1: Figure S1). Thus, we mapped both 5' and 3' ends of U2 snRNA in *L. major* and determined that the mature transcript is 143 nt long (Fig. 1b), which is similar to the 143 and 145 nt reported for the U2 snRNA in *T. cruzi* and *T. brucei*, respectively [44, 45].

The predicted secondary structure of the U2 snRNA in *L. major* (Fig. 3a) resembles the structure reported for *T. brucei*, as they both lack stem-loop III (Fig. 3a). Out of the six pseudouridine residues that we found in the *L. major* U2 snRNA (Fig. 2b), five are also present in vertebrates (positions 14, 34, 36, 55 and 86 in *L. major*), three have been found in *S. cerevisiae* (positions 34, 55 and 86) [8, 9], and three are present in *T. brucei* (positions 14, 34 and 36) [31] (Fig. 3a). It is worth noting that the only pseudouridine that is conserved in all of the organisms analyzed is the one located at position 34, which is part of the motif involved in the branch site recognition (Fig. 3a). Pseudouridine at position 36 is also located in the BPRR. Interestingly, the *L. major* U2 snRNA contains one pseudouridine that has not been reported in other organisms, which is located within the Sm site at position 93 (Fig. 3a). Thus, it is likely that the pseudouridines we mapped in the *L. major* U2 snRNA are important for snRNP assembly and trans-splicing.

A tRNA-like was found between the U2 snRNA and the tRNA-Ala genes in *L. major* (Fig. 1a). Primer extension and RT-qPCR analyses showed that transcription of the U2 snRNA gene is dependent on boxes A and B located inside the tRNA-like sequence, since deletion or mutation of these elements strongly reduced the level of the snRNA. Also, an intragenic promoter element was found close to the 5' end of the U2 snRNA gene (Fig. 4). Similar promoter elements drive the expression of the U2 snRNA gene in *T. brucei* [17]. Moreover, boxes A and/or B located in tRNA-like or tRNA genes have been shown to control the expression of the U1, U3 and U6 snRNA genes in *T. brucei*, the U4 snRNA gene in *L. collosoma* and 7SL RNA genes in *T. brucei* and *L. collosoma* [18, 19, 22, 23, 46]. However, unlike any other snRNA or 7SL RNA genes in trypanosomatids, our data show that box B from the upstream tRNA-Ala also participates in the regulation of transcription of the U2 snRNA gene in *L. major*. While mutation of box B from the tRNA-like reduced the U2 snRNA expression by 87 %, base substitutions in box B from the tRNA-Ala decreased the levels of the U2 snRNA by 46 % (Fig. 4d, compare bars 8 and 10). Thus, it is possible that box B from the tRNA-Ala participates in the fine-tuning of transcription of the *L. major* U2 snRNA.

At present it is not known how boxes A and B from the associated tRNA or tRNA-like control transcription of the snRNA. It has been proposed that transcription factors binding to boxes A and B would be shared by the snRNA and tRNA (or tRNA-like) genes. Alternatively, boxes A and B could regulate snRNA expression by an indirect mechanism involving the binding of transcription factors to the tRNA gene, which in turn could generate changes in chromatin structure that would allow transcription of the snRNA [21]. In this regard, very little is known about RNAP III transcription factors in trypanosomatids. From the two transcription factors required for the expression of tRNA genes, TFIIC and TFIIB, only the latter has been identified in trypanosomatids [47]. Thus, it is possible that the highly-transcribed tRNA-Ala gene, and the tRNA-like sequence, might bind abundant transcription factors and produce an open chromatin conformation that could facilitate transcription of the U2 snRNA in *L. major*.

Our results showed that the tRNA-like associated to the U2 snRNA is transcribed, generating a major transcript of around 109 nt (Fig. 5a, middle panel). Transcription of tRNA-like sequences associated with the *T. brucei* U2 snRNA and the *L. collosoma* U4 snRNA have also been demonstrated, producing transcripts of 89 and 90 bases, respectively [20, 23]. Densitometric analysis indicated that the tRNA-like transcript is at least 10-fold less abundant than the U2 snRNA. Thus, if the U2 snRNA and the tRNA-like are transcribed together and at the same rate, the tRNA-like should be less stable than the U2 snRNA. The predicted secondary structure of the tRNA-like was very different from the expected for a tRNA, as it does not fold into the classical cloverleaf structure (Additional file 4: Figure S4). Consequently, it is possible that this atypical secondary structure might contribute to the apparent instability of the tRNA-like transcript. Other possibility is that the transcription rate in the two directions is different and, accordingly, lower amounts of tRNA-like are produced.

Our data show that all of the *Leishmania* species analyzed contain a single U2 snRNA gene, which is highly syntenic (Fig. 1a). Similarly, in other trypanosomatids the U2 snRNA is encoded by a single-copy gene, with the exception of *T. cruzi* and *L. collosoma* that contain three genes [39]. In *T. cruzi*, only one of the U2 snRNA genes is transcribed [39]. Interestingly, sequence analysis allowed us to determine that the two *T. cruzi* U2 snRNA genes that are not expressed possess a box A in the associated tRNA-like sequence, but do not contain a box B (Additional file 5: Figure S5). Thus, the absence of box B in the neighbor tRNA-like sequence might contribute to the silencing of two U2 snRNA genes in *T. cruzi*. Remarkably, from all the trypanosomatids with genome sequences available, *C. fasciculata* is the only organism

that does not have a tRNA-like associated with the U2 snRNA, as it contains a tRNA-Ala gene whose box A is located 105 bp upstream of the U2 snRNA gene (Additional file 6: Figure S6).

Conclusions

In the present study we have demonstrated that transcription of the *L. major* U2 snRNA gene is dependent on boxes A and B from the associated tRNA-like and a sequence located at the 5' end of the U2 snRNA. Notably, unlike any other snRNA or 7SL RNA genes in trypanosomatids, box B from the upstream tRNA-Ala is also important for the expression of the *L. major* U2 snRNA gene. Thus, the U2 snRNA gene promoter shows unique features in *L. major*. Also, we have mapped several pseudouridines in the *L. major* U2 snRNA gene, including one located in the Sm binding site that has not been reported in any other organism. Future studies will explore the role that pseudouridines might play in splicing and snRNP assembly in *Leishmania*.

Additional files

Additional file 1: Figure S1. Sequence alignment of U2 snRNA genes from trypanosomatids. Sequences were obtained from the TriTrypDB databases of the following species of trypanosomatids: *L. major* Friedlin (Lmj), *L. arabica* LEM1108 (Lar), *L. tarentolae* Parrot-Tarll (Ltr), *L. infantum* JPCM5 (Lin), *L. amazonensis* MHOMBR/71973/M2269 (Laz), *L. donovani* BPK282A1 (Ldn), *L. aethiopia* L147 (Lae), *L. mexicana* U1103 (Lmx), *L. panamensis* MHOM/COL/81/L13 (Lpn), *L. braziliensis* M2904 (Lbr), *L. enriettii* LEM3045 (Len), *E. monterogei* LV88 (Emg), *C. fasciculata* CF-CI (Cfs), *L. pyrrocoris* H10 (Lpy), *L. seymouri* ATCC 30220 (Lsy), *L. collosoma* (unspecified strain) [39] (Lcl), *T. cruzi* CL Brener Non-Esmeraldo-like, copy of the chromosome 23 (Tcr), *T. vivax* Y486 (Tvx), *T. congolense* IL3000 (Tcn), *T. evansi* STIB 805 (Tev) and *T. brucei* DAL972 (Tbr). The alignment was obtained with the DNAMAN software, and manually corrected. Arrows indicate the identified pseudouridines (Ψ) in the *L. major* U2 snRNA (this work). BPRR and Sm sites are also indicated. The 3' ends (C residues) that have been experimentally mapped are shown in red font for *L. major* (this work), *T. cruzi* [45] and *T. brucei* [44]. The T-tract length is indicated for each species. Sequence numbers are relative to the TSS (+1) from the U2 snRNA. This alignment was used to generate the sequence logo depicted in Fig. 3b. (PDF 2137 kb)

Additional file 2: Figure S2. Northern blot analysis of the tagged U2 snRNA. The experiment was performed with total RNA from cells transfected with construct 1 (pComp), construct 6 (pBS + 6/+12), the unrelated vector pLMRIB or mock-transfected cells. The probe was the oligonucleotide LmjU2tag-Rev, which specifically recognizes the tag sequence. (TIF 807 kb)

Additional file 3: Figure S3. Mutations introduced in the tRNA-Ala/U2 snRNA locus. For each vector, sequence elements are indicated inside colored boxes: tRNA-Ala in green, tRNA-like in orange and U2 snRNA in blue. Boxes A and B are also indicated. The tag sequence, inserted within the U2 snRNA gene, is also shown. Base substitutions are indicated in red font. (PDF 3122 kb)

Additional file 4: Figure S4. Predicted secondary structure of the tRNA-Ala (panel a) and the tRNA-like associated to the *L. major* U2 snRNA gene (panel b). The minimum free energy structures are shown. The color scale indicates low (blue) to high (red) probabilities of base pairing. (PDF 355 kb)

Additional file 5: Figure S5. Sequence comparisons of U2 snRNA genes and flanking regions from *T. cruzi* (CL Brener Non-Esmeraldo-like).

Sequences from the genes located on chromosomes 23, 37 and 6 are shown. The U2 snRNA gene from chromosome 23 is presented in blue font. The position of boxes A and B is indicated. Sequence numbers are relative to the TSS (+1) from the U2 snRNA. (PDF 1404 kb)

Additional file 6: Figure S6. Genomic context of U2 snRNA genes in trypanosomatids. Schematic representations of the U2 snRNA loci from *Leishmania* spp., *E. monterogei* (LV88), *L. seymouri* (ATCC 30220), *C. fasciculata* (CF-CI), *T. cruzi* (CL Brener Non-Esmeraldo-like, copy of the chromosome 23) and *T. brucei* (DAL972). All of the U2 snRNA genes contain divergently-oriented boxes A and B at conserved distances (around 104 bases), which are contained in tRNA-like sequences. The exception is *C. fasciculata*, where boxes A and B are located inside a tRNA-Ala gene. In *T. cruzi*, a tRNA-Arg gene is located downstream of the U2 snRNA gene. Figure is drawn to scale. (TIF 268 kb)

Abbreviations

5'-RACE, 5' Rapid Amplification of cDNA Ends; BPRR, branch point recognition region; CMCT, N-cyclohexyl-N'-β-(4-methylmorpholinium)ethyl-carbodiimide p-tosylate; mRNA, messenger RNA; RNAP, RNA polymerase; RT-qPCR, real-time quantitative polymerase chain reaction; snRNA, small nuclear RNA; tRNA, transfer RNA; TSS, transcription start site

Acknowledgments

This work is one of the requirements to obtain the PhD degree in Posgrado en Ciencias Biológicas (UNAM) for Saúl Rojas-Sánchez, who was the recipient of doctoral fellowships from CONACyT (Fellowship 256179, CVU 384986), and CEP/UNAM (303180353).

Funding

This work was supported by grant 251831 from CONACyT, grants IN210712 and IN214715 from PAPIIT (UNAM), and grant FESI-DIP-PAPCA-2014-24 (PAPCA, FES Iztacala) to S. Martínez-Calvillo; and by grants 132312 and 139898 from CONACyT to R. Manning-Cela. The funders had no role in the design of the study, data collection and analysis, or writing of the manuscript.

Availability of data and material

The datasets supporting the conclusions of this article are included within the article and its additional files.

Authors' contributions

SRS performed most of the experiments, analyzed data and drafted the manuscript. EFA generated some vectors and analyzed data. RMC performed some *in silico* analyses and helped perform RT-PCR analysis. LEFM helped perform transfections, Northern blot and RT-qPCR analyses. RGMC participated in the writing and editing of the manuscript. SMC conceived the study, analyzed data and wrote the manuscript. All authors read and approved the manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Author details

¹Unidad de Biomedicina, Facultad de Estudios Superiores Iztacala, Universidad Nacional Autónoma de México, Av. de los Barrios 1, Col. Los Reyes Iztacala, Tlalnepantla, Edo. de México CP 54090, Mexico.

²Departamento de Biomedicina Molecular, Centro de Investigación y de Estudios Avanzados del IPN, Av. IPN 2508, México, DF CP 07360, Mexico.

Received: 24 March 2016 Accepted: 5 July 2016

Published online: 19 July 2016

References

- Cvitkovic I, Jurica MS. Spliceosome database: a tool for tracking components of the spliceosome. *Nucleic Acids Res.* 2013;41(Database issue):D132–41.

2. Chen W, Moore MJ. The spliceosome: disorder and dynamics defined. *Curr Opin Struct Biol.* 2014;24:141–9.
3. Wahl MC, Will CL, Luhrmann R. The spliceosome: design principles of a dynamic RNP machine. *Cell.* 2009;136(4):701–18.
4. Leung AK, Nagai K, Li J. Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis. *Nature.* 2011;473(7348):536–9.
5. Zhou L, Hang J, Zhou Y, Wan R, Lu G, Yin P, et al. Crystal structures of the Lsm complex bound to the 3' end sequence of U6 small nuclear RNA. *Nature.* 2014;506(7486):116–20.
6. Rodgers ML, Tretbar US, Dehaven A, Alwan AA, Luo G, Mast HM, et al. Conformational dynamics of stem II of the U2 snRNA. *RNA.* 2016;22(2):225–36.
7. Ro-Choi TS, Choi YC. Chemical approaches for structure and function of RNA in postgenomic era. *J Nucleic Acids.* 2012;2012:369058.
8. Adachi H, Yu YT. Insight into the mechanisms and functions of spliceosomal snRNA pseudouridylation. *World J Biol Chem.* 2014;5(4):398–408.
9. Yu AT, Ge J, Yu YT. Pseudouridines in spliceosomal snRNAs. *Protein Cell.* 2011;2(9):712–25.
10. Eglhoff S, O'Reilly D, Murphy S. Expression of human snRNA genes from beginning to end. *Biochem Soc Trans.* 2008;36(Pt 4):590–4.
11. Mroczek S, Dziembowski A. U6 RNA biogenesis and disease association. *Wiley Interdiscip Rev RNA.* 2013;4(5):581–92.
12. Orioli A, Pascali C, Pagano A, Teichmann M, Dieci G. RNA polymerase III transcription control elements: themes and variations. *Gene.* 2012;493(2):185–94.
13. Martinez-Calvillo S, Vizuet-de-Rueda JC, Florencio-Martinez LE, Manning-Cela RG, Figueroa-Angulo EE. Gene expression in trypanosomatid parasites. *J Biomed Biotechnol.* 2010;2010:525241.
14. Bitar M, Boroni M, Macedo AM, Machado CR, Franco GR. The spliced leader trans-splicing mechanism in different organisms: molecular details and possible biological roles. *Front Genet.* 2013;4:199.
15. Gunzl A. The pre-mRNA splicing machinery of trypanosomes: complex or simplified? *Eukaryot Cell.* 2010;9(8):1159–70.
16. Papasaikas P, Valcarcel J. The Spliceosome: The Ultimate RNA Chaperone and Sculptor. *Trends Biochem Sci.* 2016;41(1):33–45.
17. Fantoni A, Dare AO, Tschudi C. RNA polymerase III-mediated transcription of the trypanosome U2 small nuclear RNA gene is controlled by both intragenic and extragenic regulatory elements. *Mol Cell Biol.* 1994;14(3):2021–8.
18. Nakaar V, Dare AO, Hong D, Ullu E, Tschudi C. Upstream tRNA genes are essential for expression of small nuclear and cytoplasmic RNA genes in trypanosomes. *Mol Cell Biol.* 1994;14(10):6736–42.
19. Djikeng A, Ferreira L, D'Angelo M, Dolezal P, Lamb T, Murta S, et al. Characterization of a candidate *Trypanosoma brucei* U1 small nuclear RNA gene. *Mol Biochem Parasitol.* 2001;113(1):109–15.
20. Tschudi C, Ullu E. Unconventional rules of small nuclear RNA transcription and cap modification in trypanosomatids. *Gene Expr.* 2002;10(1–2):3–16.
21. Nakaar VT, Tschudi C, Ullu E. An unusual liaison: small nuclear and cytoplasmic RNA genes team up with tRNA genes in trypanosomatid protozoa. *Parasitol Today.* 1995;11:4.
22. Nakaar V, Gunzl A, Ullu E, Tschudi C. Structure of the *Trypanosoma brucei* U6 snRNA gene promoter. *Mol Biochem Parasitol.* 1997;88(1–2):13–23.
23. Li L, Otake LR, Xu Y, Michaeli S. The trans-spliceosomal U4 RNA from the monogenetic trypanosomatid *Leptomonas collosoma*. Cloning and identification of a transcribed tRNA-like element that controls its expression. *J Biol Chem.* 2000;275(4):2259–64.
24. Robinson KA, Beverley SM. Improvements in transfection efficiency and tests of RNA interference (RNAi) approaches in the protozoan parasite *Leishmania*. *Mol Biochem Parasitol.* 2003;128(2):217–28.
25. Martinez-Calvillo S, Sunkin SM, Yan S, Fox M, Stuart K, Myler PJ. Genomic organization and functional characterization of the *Leishmania major* Friedlin ribosomal RNA gene locus. *Mol Biochem Parasitol.* 2001;116(2):147–57.
26. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics.* 2004;5:113.
27. Lorenz R, Bernhart SH, Honer Zu Siederdisen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA Package 2.0. *Algorithms Mol Biol.* 2011;6:26.
28. Carey MF, Peterson CL, Smale ST. The primer extension assay. *Cold Spring Harb Protoc.* 2013;2013(2):164–73.
29. Martinez-Calvillo S, Nguyen D, Stuart K, Myler PJ. Transcription initiation and termination on *Leishmania major* chromosome 3. *Eukaryot Cell.* 2004;3(2):506–17.
30. Motorin Y, Muller S, Behm-Ansmant I, Branlant C. Identification of modified residues in RNAs by reverse transcription-based methods. *Methods Enzymol.* 2007;425:21–53.
31. Barth S, Hury A, Liang XH, Michaeli S. Elucidating the role of H/ACA-like RNAs in trans-splicing and rRNA processing via RNA interference silencing of the *Trypanosoma brucei* CBF5 pseudouridine synthase. *J Biol Chem.* 2005;280(41):34558–68.
32. Hofstetter H, Kressman A, Birnstiel ML. A split promoter for a eucaryotic tRNA gene. *Cell.* 1981;24(2):573–85.
33. Baker RE, Gabrielsen O, Hall BD. Effects of tRNATyr point mutations on the binding of yeast RNA polymerase III transcription factor C. *J Biol Chem.* 1986;261(12):5275–82.
34. Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, et al. The genome of the kinetoplastid parasite, *Leishmania major*. *Science.* 2005;309(5733):436–42.
35. Downing T, Imamura H, Decuypere S, Clark TG, Coombs GH, Cotton JA, et al. Whole genome sequencing of multiple *Leishmania donovani* clinical isolates provides insights into population structure and mechanisms of drug resistance. *Genome Res.* 2011;21(12):2143–56.
36. Peacock CS, Seeger K, Harris D, Murphy L, Ruiz JC, Quail MA, et al. Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. *Nat Genet.* 2007;39(7):839–47.
37. Rogers MB, Hilley JD, Dickens NJ, Wilkes J, Bates PA, Depledge DP, et al. Chromosome and gene copy number variation allow major structural change between species and strains of *Leishmania*. *Genome Res.* 2011;21(12):2129–42.
38. Raymond F, Boisvert S, Roy G, Ritt JF, Legare D, Isnard A, et al. Genome sequencing of the lizard parasite *Leishmania tarentolae* reveals loss of genes associated to the intracellular stage of human pathogenic species. *Nucleic Acids Res.* 2012;40(3):1131–47.
39. Hartshorne T, Agabian N. A new U2 RNA secondary structure provided by phylogenetic analysis of trypanosomatid U2 RNAs. *Genes Dev.* 1990;4(12A):2121–31.
40. Tschudi C, Richards FF, Ullu E. The U2 RNA analogue of *Trypanosoma brucei* gambiense: implications for a splicing mechanism in trypanosomes. *Nucleic Acids Res.* 1986;14(22):8893–903.
41. Wang P, Palfi Z, Preusser C, Lücke S, Lane WS, Kambach C, et al. Sm core variation in spliceosomal small nuclear ribonucleoproteins from *Trypanosoma brucei*. *EMBO J.* 2006;25(19):4513–23.
42. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. *Genome Res.* 2004;14(6):1188–90.
43. Padilla-Mejia NE, Florencio-Martinez LE, Figueroa-Angulo EE, Manning-Cela RG, Hernandez-Rivas R, Myler PJ, et al. Gene organization and sequence analyses of transfer RNA genes in trypanosomatid parasites. *BMC Genomics.* 2009;10:232.
44. Hitchcock RA, Zeiner GM, Sturm NR, Campbell DA. The 3' termini of small RNAs in *Trypanosoma brucei*. *FEMS Microbiol Lett.* 2004;236(1):73–8.
45. Ambrosio DL, Silva MT, Cicarelli RM. Cloning and molecular characterization of *Trypanosoma cruzi* U2, U4, U5, and U6 small nuclear RNAs. *Mem Inst Oswaldo Cruz.* 2007;102(1):97–105.
46. Ben-Shlomo H, Levitan A, Beja O, Michaeli S. The trypanosomatid *Leptomonas collosoma* 7SL RNA gene. Analysis of elements controlling its expression. *Nucleic Acids Res.* 1997;25(24):4977–84.
47. Velez-Ramirez DE, Florencio-Martinez LE, Romero-Meza G, Rojas-Sanchez S, Moreno-Campos R, Arroyo R, et al. BRF1, a subunit of RNA polymerase III transcription factor TFIIB, is essential for cell growth of *Trypanosoma brucei*. *Parasitology.* 2015;142(13):1563–73.
48. Broadbent S, van der Woude M, Aziz N. Accurate and simple sizing of primer extension products using a non-radioactive approach facilitates identification of transcription initiation sites. *J Microbiol Methods.* 2010;81(3):256–8.