



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

**“Análisis transcripcional del gen de RNA de transferencia de selenocisteína en
Leishmania major”**

TESIS

QUE PARA OPTAR POR EL GRADO DE:
DOCTORA EN CIENCIAS

PRESENTA:

NORMA EDITH PADILLA MEJÍA

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

COMITÉ TUTOR: Dra. Ana María Cevallos Gaos
Instituto de Investigaciones Biomédicas

Dra. Rosaura Hernández Rivas
Posgrado en Ciencias Biológicas

MÉXICO, D.F., ABRIL, 2015



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 Comité Académico del Posgrado en Ciencias Biológicas, celebrada el día 23 de febrero de 2015, se aprobó el siguiente jurado para el examen de grado de **DOCTORA EN CIENCIAS** de la alumna **PADILLA MEJÍA NORMA EDITH**, con número de cuenta **96391673**, con la tesis titulada "**Análisis transcripcional del gen de RNA de transferencia de selenocisteína en *Leishmania major***", realizada bajo la dirección del **DR. SANTIAGO MARTÍNEZ CALVILLO**:

Presidente: DRA. YOLANDA IRASEMA CHIRINO LÓPEZ
Vocal: DRA. ROSA LAURA CAMARENA MEJÍA
Secretario: DRA. ANA MARÍA CEVALLOS GAOS
Suplente: DR. ROBERTO JOSÉ RAFAEL HERNÁNDEZ FERNÁNDEZ
Suplente: DRA. ROSAURA HERNÁNDEZ RIVAS

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

ATENTAMENTE
"POR MI RAZA HABLARA EL ESPIRITU"
Cd. Universitaria, D.F., a 20 de marzo de 2015.

M del Coro Ariz
DRA. MARÍA DEL CORO ARIZMENDI ARRIAGA
COORDINADORA DEL PROGRAMA



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

AGRADECIMIENTOS

Agradezco al Posgrado en Ciencias Biológicas, UNAM por la formación recibida durante mis estudios de Doctorado.

Agradezco al CONACyT la beca otorgada para mis estudios de Doctorado (Becario 207150).

Este trabajo fue financiado por los donativos 128461 de CONACyT, IN203909 y IN210712 de PAPIIT (DGAPA, UNAM), y proyecto 53 de PAPCA 2013 (FES Iztacala).

Asimismo, agradezco enormemente a los miembros del Comité Tutor por enriquecer mi proyecto y desarrollo académico:

Dr. Santiago Martínez Calvillo (Tutor Principal)
Dra. Ana María Cevallos Gaos
Dra. Rosaura Hernández Rivas

AGRADECIMIENTOS A TÍTULO PERSONAL

A la Universidad Nacional Autónoma de México, que ha sido mi Alma Mater desde la preparatoria hasta mis estudios de doctorado. Esta institución me ha formado profesional e intelectualmente, y estando en ella, he atravesado por momentos personales significativos. Siempre me sentiré orgullosa de egresar de ella y agradecida por todo lo que me ha brindado.

A los integrantes del jurado de examen de grado:

Dra. Yolanda Irasema Chirino López, Dra. Ana María Cevallos Gaos, Dra. Rosa Laura Camarena Mejía, Dra. Rosaura Hernández Rivas y el Dr. Roberto Hernández Fernández

por sus comentarios durante la revisión de esta tesis, los cuales ayudaron a enriquecer el presente trabajo.

A mi familia: mis padres Lucía y Carlos, y mis hermanos, Carlos Gabriel y Sandra. Gracias por su apoyo infinito, porque sin ustedes esto no sería posible.

A mis amigos del laboratorio 1 de la UBIMED: Lourdes, Fiordaliso, Maricarmen, Gaby, Juan Carlos, Daniel y Luis. Les agradezco su amistad, por las risas, por saber compartir en tiempo y espacio y hacerlo de corazón. Su amistad fue ciertamente un sostén en tiempos difíciles.

A mis amigas de otros espacios de la UBIMED: Rocio, Marissa, Marian y Carmelita. Por su invaluable amistad.

A Lillian, por tu amistad y complicidad, por estar siempre ahí.

Al Biol. Luis Enrique Florencio Martínez por su ayuda técnica.

DEDICATORIA

A mi mamá Lucía,

A mi papá Carlos,

A mi hermana Sandra,

A mi hermano Carlos Gabriel,

porque han vivido todo este proceso de cerca y han sentido junto conmigo las alegrías y las desavenencias, por su apoyo infinito (que ha sido de todo tipo y en muchos sentidos), por estar ahí incondicionales. Los quiero.

ÍNDICE

Lista de Tablas	IV
Lista de Figuras	IV
Lista de Abreviaciones	V
1. RESUMENES	
1.1 Resumen	1
1.2 Abstract	2
2. INTRODUCCIÓN	
2.1. Generalidades sobre los parásitos del género <i>Leishmania</i>	3
2.2. Ciclo de vida	7
2.3. Organización genómica	7
2.4. Expresión génica	10
2.5. Transcripción por la RNA Polimerasa III	12
2.5.1. Genes transcritos por la RNA Polimerasa III y sus elementos promotores	12
2.5.2. Factores de transcripción asociados a la RNA Polimerasa III	16
2.6. Selenocisteína y selenoproteínas	18
2.7. Elementos esenciales para la inserción de selenocisteína en las selenoproteínas	
2.7.1. Elemento SECIS	21
2.7.2. eEFSec	22
2.7.3. Proteína SBP2	23
2.7.4. Proteína ribosomal L30	23
2.8. Modelos de inserción de selenocisteína a las selenoproteínas	24
2.9. tRNA de Selenocisteína	25
2.10. Transcripción y regiones promotoras del tRNA-Sec	27
2.11. Selenoproteínas en tripanosomátidos	29
3. ANTECEDENTES	31
4. OBJETIVOS	
4.1. Objetivo General	33
4.2. Objetivos Específicos	33
5. MÉTODOLÓGIA	
5.1. Análisis <i>in silico</i>	34
5.2. Cultivos de <i>L. major</i> y <i>T. cruzi</i>	34
5.3. Análisis 5' RACE	34
5.4. Mapeo de los sitios de poliadenilación y los sitios de adición del miniexón	35

5.5. Ensayo <i>Northern blot</i>	37
5.6. Clonación de fragmentos en pGEM-T Easy	38
5.7. Transformación de células competentes <i>E. coli</i> JM109	40
5.8. Obtención de DNA de plásmido	40
5.9. Clonación en los vectores M13mp18 y M13mp19	41
5.10. Transformación de <i>E. coli</i> JM109 con construcciones M13mp18 y M13mp19	42
5.11. Inducción del DNA de cadena sencilla (ssDNA)	43
5.12. Preparación de filtros para ensayos de <i>run-on</i> nuclear	43
5.13. Ensayos de <i>run-on</i> nuclear	44
5.14. Ensayos de <i>run-on</i> nuclear con luz UV	45
5.15. Ensayos de <i>run-on</i> nuclear con inhibidores de la RNA Polimerasa	46
6. RESULTADOS	
6.1. Localización genómica del gen del tRNA-Sec en <i>L. major</i> y otros tripanosomátidos	47
6.2. En <i>L. major</i> algunos precursores del tRNA-Sec contienen el miniexón y una cola de poliadeninas	49
6.3. El tRNA-Sec en <i>T. cruzi</i> también sufre <i>trans-splicing</i> y poliadenilación	51
6.4. El tRNA-Sec de <i>L. major</i> está presente en transcritos de diferentes tamaños	52
6.5. En <i>L. major</i> , el <i>trans-splicing</i> y la poliadenilación del tRNA-Sec siguen las señales genómicas canónicas	53
6.6. Otros genes de tRNA no poseen el miniexón en su extremo 5'	54
6.7. Clonación de fragmentos en pGEM-T Easy para ensayos de <i>run-on</i> nuclear	55
6.8. Clonación de fragmentos en M13mp18 y M13mp19 para la generación de ssDNA para los ensayos de <i>run-on</i> nuclear	55
6.9. Ensayo de <i>run-on</i> nuclear, análisis del locus del tRNA-Sec	57
6.10. Ensayos de <i>run-on</i> nuclear con células irradiadas con luz UV	60
6.11. Ensayos de <i>run-on</i> nuclear con α -amanitina	62
6.12. Ensayos de <i>run-on</i> nuclear con taquetitoxina	63
6.13. Una caja tipo A y una caja B se localizan río arriba del tRNA-Sec de <i>L. major</i>	66
7. DISCUSIÓN	67
8. CONCLUSIONES	73
9. REFERENCIAS	74
10. APÉNDICES	
Apéndice I. Artículo de obtención de grado: Padilla-Mejía, N.E., <i>et al.</i> 2014. The selenocysteine tRNA gene in <i>Leishmania major</i> is transcribed by both RNA Polymerase II and RNA Polymerase III. <i>Eukaryotic Cell.</i> 14:216-227.	79
Apéndice II: Mapa del vector pGEM-T Easy	92

Apéndice III: Mapa de los vectores M13mp18 y M13mp19	93
Apéndice IV: Repetición del experimento de <i>run-on</i> con células irradiadas con luz UV	94
Apéndice V: Repetición del experimento de <i>run-on</i> con α -amanitina	95
Apéndice VI: Repetición del experimento de <i>run-on</i> con tagetitoxina	96
Apéndice VII: Artículo: Padilla-Mejía, N.E., et al. 2009. Gene organization and sequence analyses of tRNA genes in Trypanosomatid parasites. BMC Genomics 10: 232-249.	97
Apéndice VIII: Capítulo de libro: Padilla-Mejía, N.E., et al. 2013. Comparative genomics of <i>Leishmania</i> parasites. En el libro Comparative Genomics in Neglected Human Parasites. Editores: Lopez-Camarillo, C. y Marchat, L.A. Nova Science Publishers, Inc., New York, New York, USA. pp 59-80.	116

LISTA DE TABLAS

Tabla 1	Principales especies de <i>Leishmania</i> que infectan humanos	4
Tabla 2	Resumen de los principales características de los genomas de <i>Leishmania</i> spp., <i>T. brucei</i> y <i>T. cruzi</i>	9
Tabla 3	Algunos RNAs sintetizados por la RNA Polimerasa III	13
Tabla 4	Selenoproteínas caracterizadas en mamíferos	19
Tabla 5	Oligonucleótidos utilizados en los ensayos del 5' y 3' RACE y <i>Northern blot</i>	36
Tabla 6	Oligonucleótidos utilizados para la clonación de genes en el vector pGEM-T Easy	39

LISTA DE FIGURAS

Figura 1	Mapas de zonas endémicas de leishmaniasis alrededor del mundo	5
Figura 2	Incidencia y distribución en México de la Leishmaniasis.	6
Figura 3	Ciclo de vida de <i>Leishmania</i> spp.	8
Figura 4	Generación y procesamiento de los transcritos policistrónicos por <i>trans-splicing</i> .	11
Figura 5	Promotores de la RNA Polimerasa III	15
Figura 6	Factores de transcripción relacionados con la RNA Polimerasa III	17
Figura 7	Maquinaria inserción de selenoproteínas	22
Figura 8	Modelo de inserción de selenocisteína a las selenoproteínas	25
Figura 9	Particularidades estructurales del tRNA-Sec.	27
Figura 10	Particularidades del promotor del tRNA-Sec.	29
Figura 11	Particularidades del gen de tRNA-Sec en <i>L. major</i>	32
Figura 12	Sintenia del locus del tRNA-Sec en tripanosomátidos	48
Figura 13	Resultados del 5' y 3' RACE del tRNA-Sec en <i>L. major</i>	50
Figura 14	Resultados de los ensayos 5' y 3' RACE en <i>T. cruzi</i>	52
Figura 15	El tRNA-Sec de <i>L. major</i> está presente en transcritos de diferentes tamaños	53
Figura 16	Clonación de fragmentos en pGEM-T Easy	56
Figura 17	Subclonación del tRNA-Sec en los vectores M13	57
Figura 18	Ensayo de <i>run-on</i> nuclear del locus del gen de tRNA-Sec	59
Figura 19	Ensayo de <i>run-on</i> nuclear de células irradiadas con luz UV	61
Figura 20	Sensibilidad de la transcripción del tRNA-Sec a la α -amanitina	64
Figura 21	Efecto de la taquetitoxina en la transcripción del tRNA-Sec	65
Figura 22	Una caja B y una caja tipo A se localizaron río arriba del tRNA-Sec en <i>L. major</i>	66

LISTA DE ABREVIACIONES

AE	Elemento Activador, <i>Activator Element</i>
b	bases
cDNA	DNA complementario
DSE	Elemento Distal, <i>Distal Sequence Element</i>
EF	Factor de Elongación
EI	Elemento Intermedio
IC ₅₀	Concentración inhibidora 50
LC	Leishmaniasis cutánea
LMC	Leishmaniasis mucocutánea
LV	Leishmaniasis visceral
Mb	Megabases
mRNA	RNA mensajero
ncRNA	RNA no codificante
pb	pares de bases
PSE	Elemento proximal
RACE	<i>Rapid Amplification of cDNA Ends</i>
RNA Pol	RNA Polimerasa
RNAi	RNA de interferencia
rRNA	RNA ribosomal
RNP	Ribonucleoproteína
RT-PCR	Reacción de transcriptasa reversa seguida de reacción en cadena de la polimerasa
Sec	Selenocisteína
siRNA	RNA pequeño de interferencia
snRNA	RNA pequeño nuclear
ssDNA	DNA de cadena sencilla
subgen	subgénero
TF	Factor de transcripción
tRNA	RNA de transferencia
UTR	Región no traducida
UV	Ultravioleta

1. RESÚMENES

1.1 Resumen

Los tripanosomátidos son protozoarios flagelados que causan diversas enfermedades humanas que se transmiten por diferentes insectos vectores. A este grupo pertenecen los protozoarios del género *Leishmania*, agentes causales de la leishmaniasis, una enfermedad con diversas manifestaciones clínicas distribuida mundialmente en 98 países. Además de su importancia médica, estos parásitos poseen características de expresión génica únicas, que difieren de otros eucariontes, entre las que se encuentran la transcripción policistrónica por la RNA Polimerasa II (RNA Pol II) y la maduración de los mRNAs por *trans-splicing*, proceso que añade una secuencia de 39 nucleótidos (denominada miniexón) al extremo 5' de todos los mRNAs.

Nuestro grupo de trabajo está interesado en el estudio de la transcripción por la RNA Polimerasa III (RNA Pol III) en *L. major*. La RNA Pol III sintetiza diversas moléculas de RNA no codificantes esenciales para la supervivencia celular, como los RNAs de transferencia (tRNAs), el RNA ribosomal (rRNA) 5S y los RNA pequeños nucleares (snRNAs). En eucariontes, los tRNAs contienen dos elementos promotores internos, las cajas A y B. Una excepción es el tRNA de selenocisteína (tRNA-Sec), cuya transcripción en vertebrados está dirigida por elementos extragénicos y una caja interna B. El tRNA-Sec participa en la incorporación de selenocisteína, el aminoácido número 21, en un grupo de proteínas esenciales llamadas selenoproteínas. En este trabajo reportamos el análisis transcripcional del tRNA-Sec en *L. major*. En este organismo, el tRNA-Sec está codificado por un gen de copia única insertado en una unidad policistrónica del cromosoma 6, a diferencia de la mayoría de los tRNAs, que en *L. major* están agrupados en las fronteras de las unidades policistrónicas. De manera interesante, experimentos de 5'-RACE y RT-PCR mostraron la presencia de transcritos del tRNA-Sec que contienen el miniexón en el extremo 5' y una cola de poliadeninas en el 3', lo que es característico de los mRNAs, pero no de los tRNAs. Estos resultados indicaron que este tRNA es transcrito por la RNA Pol II debido a que se encuentra insertado en una unidad policistrónica y, posteriormente, es procesado por *trans-splicing* y poliadenilación. Resultados similares fueron reportados en otro tripanosomátido, el parásito *Trypanosoma brucei*. Sin embargo, ensayos de *run-on* nuclear con inhibidores de las RNA Polimerasas, así como con células irradiadas con luz UV nos indicaron que, de manera sorprendente, la RNA Pol III también está involucrada en la transcripción del tRNA-Sec. Así, este es el primer reporte de una molécula de RNA transcrita por dos RNA Polimerasas diferentes en tripanosomátidos. Nuestros resultados indican que, contrario a lo que se pensaba, la especificidad de las RNA Polimerasas no es absoluta *in vivo* en *Leishmania*.

1.2 Abstract

The trypanosomatids are flagellated protozoan that cause devastating human diseases that are transmitted by different insect vectors. This group includes the parasites of the genus *Leishmania*, causative agent of the leishmaniasis, a disease with diverse clinical manifestations that extends to 98 countries worldwide. In addition to their medical importance, these protozoan parasites show unique mechanisms of gene expression, different from those found in higher eukaryotes, such as the RNA Polymerase II (RNA Pol II) polycistronic transcription and maturation of mRNAs by *trans*-splicing, a process that attaches a 39-nucleotide miniexon to the 5' end of all the mRNAs.

Our group is interested in the study of the RNA Polymerase III (RNA Pol III) transcription in *L. major*. RNA Pol III synthesizes diverse non-coding RNAs that are essential for cell survival, such as transfer RNAs (tRNAs), 5S ribosomal RNA (rRNA) and small nuclear RNAs (snRNAs). In eukaryotes, tRNAs have two internal promoter elements which direct their transcription, the A and B boxes. The Selenocysteine tRNA (tRNA-Sec) constitutes an exception since its transcription is directed by an internal B box and three extragenic sequences in vertebrates. The tRNA-Sec carries selenocysteine, the 21st amino acid, to the biosynthesis of selenoproteins. Here we report the transcriptional analysis of the tRNA-Sec in the protozoan parasite *L. major*. In this organism, the tRNA-Sec is encoded by a single gene inserted into a RNA Pol II polycistronic unit, in contrast to most tRNAs, which are clustered at the boundaries of polycistronic units.

Interestingly, 5'-RACE and RT-PCR experiments showed that tRNA-Sec transcripts contain the miniexon at the 5' end and a poly-A tail at the 3' end, similar to what occurs to mRNAs. Thus, these results indicated that this tRNA is transcribed by RNA Pol II as part of a polycistronic unit and then processed by *trans*-splicing and polyadenylation, as it has been previously described in the related organism *Trypanosoma brucei*. However, nuclear run-on assays with RNA polymerase inhibitors, and with cells that were previously UV irradiated, showed that the tRNA-Sec gene in *L. major* is also transcribed by Pol III. To our knowledge, this is the first report of a gene transcribed by more than one RNA polymerase in trypanosomatids, namely RNA Pol II and RNA Pol III. Thus, our results indicate that RNA polymerase specificity in *Leishmania* is not absolute *in vivo*, as has been recently found in other eukaryotes.

2. INTRODUCCIÓN

Leishmania major es un parásito que pertenece a la familia *Trypanosomatidae*, la cual también incluye a otros parásitos como *Trypanosoma brucei*, *Trypanosoma cruzi* y otras especies del género *Leishmania*, todos ellos causantes de graves enfermedades humanas. Estos parásitos poseen una organización genómica y estructuras celulares similares y todos sufren cambios morfológicos durante sus ciclos de vida y son transmitidos a sus hospederos mamíferos por diferentes insectos vectores. *T. brucei* causa la enfermedad del sueño en África, *T. cruzi* la enfermedad de Chagas en el continente Americano y los parásitos del género *Leishmania* causan diferentes variantes de la leishmaniasis alrededor del mundo. Además de la importancia médica, estos parásitos muestran numerosas particularidades metabólicas y estructurales que resultan de alto interés científico (Stuart, K. et al., 2008).

2.1 Generalidades sobre los parásitos del género *Leishmania*

Los protozoarios parásitos del género *Leishmania* causan la leishmaniasis, una enfermedad que presenta diferentes manifestaciones clínicas, dependiendo de la especie de *Leishmania* infectante. A la fecha se han reportado alrededor de 20 especies diferentes de *Leishmania* patógenas. El parásito es transmitido al hombre por un insecto vector, la mosca de la arena, del género *Phlebotomus* en el Viejo Mundo o *Lutzomyia* en el continente Americano. Existen alrededor de 30 especies de insectos vectores que pueden diseminar la enfermedad. Las principales manifestaciones clínicas de la enfermedad son: leishmaniasis cutánea (LC), leishmaniasis mucocutánea (LMC) y leishmaniasis visceral (LV), según la especie de *Leishmania* infectante (Tabla 1). La leishmaniasis visceral es generalmente fatal de no seguir un tratamiento, mientras que las formas mucocutánea y cutánea generan lesiones mutilantes de tipo discapacitante que afectan el desenvolvimiento social del enfermo (Antinori, S., et al., 2012; Kaye, P. y Scott, P., 2011; www.who.int/topics/leishmaniasis/es/).

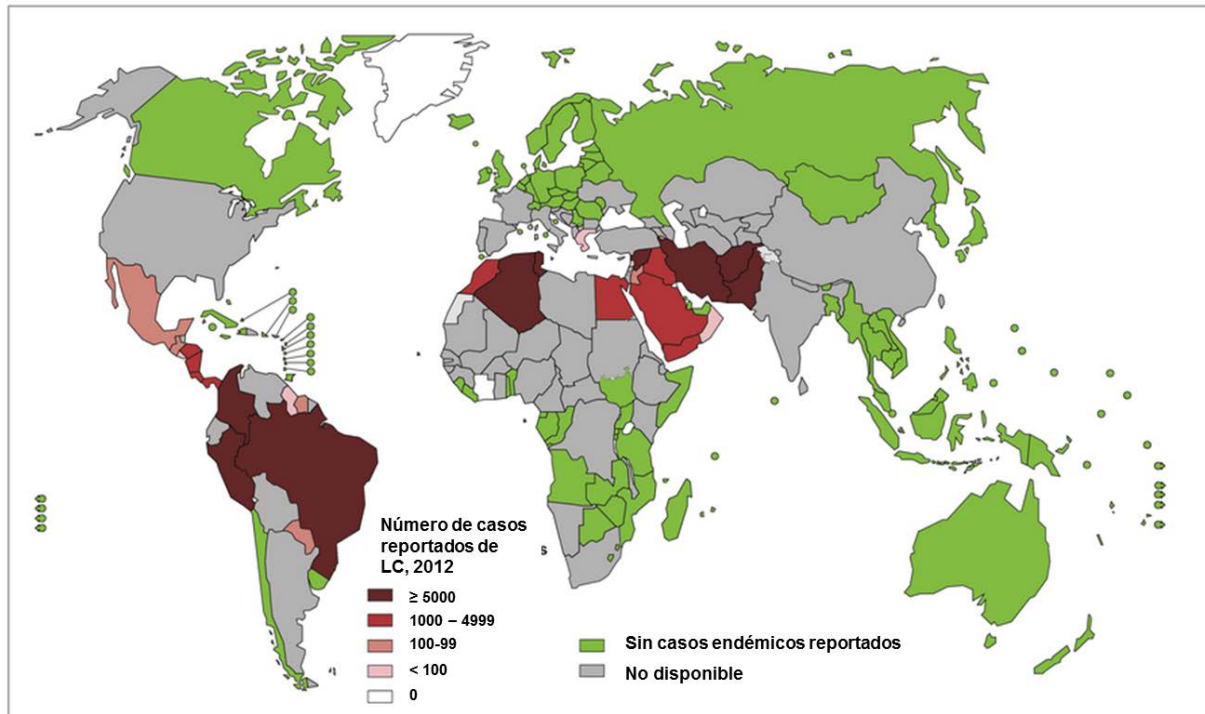
Tabla 1. Principales especies de *Leishmania* que infectan humanos
(Padilla-Mejía, N.E., et al., 2013)

Variante de la leishmaniasis	Especie infectante
Viejo Mundo subgen. Leishmania	
Leishmaniasis visceral	<i>L. donovani, Leishmania infantum</i>
Leishmaniasis cutánea	<i>L. major, L. tropica, L. aethiopica</i>
Nuevo Mundo subgen. Leishmania	
Leishmaniasis visceral	<i>L. infantum</i>
Leishmaniasis cutánea	<i>L. infantum, L. mexicana, L. pifanoi, L. amazonensis</i>
Nuevo Mundo subgen. Viannia	
Leishmaniasis cutánea	<i>L. braziliensis, L. guyanensis, L. panamensis, L. peruviana</i>
Leishmaniasis mucocutánea	<i>L. braziliensis, L. panamensis</i>

La leishmaniasis es considerada la novena enfermedad infecciosa más difundida a nivel mundial, con presencia en 98 países alrededor del mundo (Alvar, J., et al., 2012) (Figura 1). En América Latina, incluido México, se han reportado las tres formas ya mencionadas de la enfermedad. Al año, se presentan ~3700 casos de LV (7 casos/año en México), ~67,000 casos de LC (811 en México) (Alvar, J. et al., 2012).

En México, los casos de LV se han reportado principalmente en el estado de Chiapas (92% de los casos nacionales), existiendo también reportes en los estados de Guerrero, Puebla y Morelos (Figura 2A). La incidencia estimada de LV en el continente americano es de 4500 a 6800 casos anuales. En México, la forma cutánea se ha reportado en los estados de Veracruz, Tabasco, Oaxaca, Chiapas, Nayarit, Yucatán, Quintana Roo, Campeche, Coahuila, Jalisco, Sinaloa y Puebla (Figura 2B). Se ha estimado que la incidencia anual de LC en el continente americano es de 187,200 a 307,800. Sobre la LMC existen pocos reportes en México, provenientes principalmente de los estados de Chiapas y Veracruz, Tabasco y Oaxaca (Figura 2C) (Alvar, J., et al., 2012; www.indre.salud.gob.mx).

A)



B)

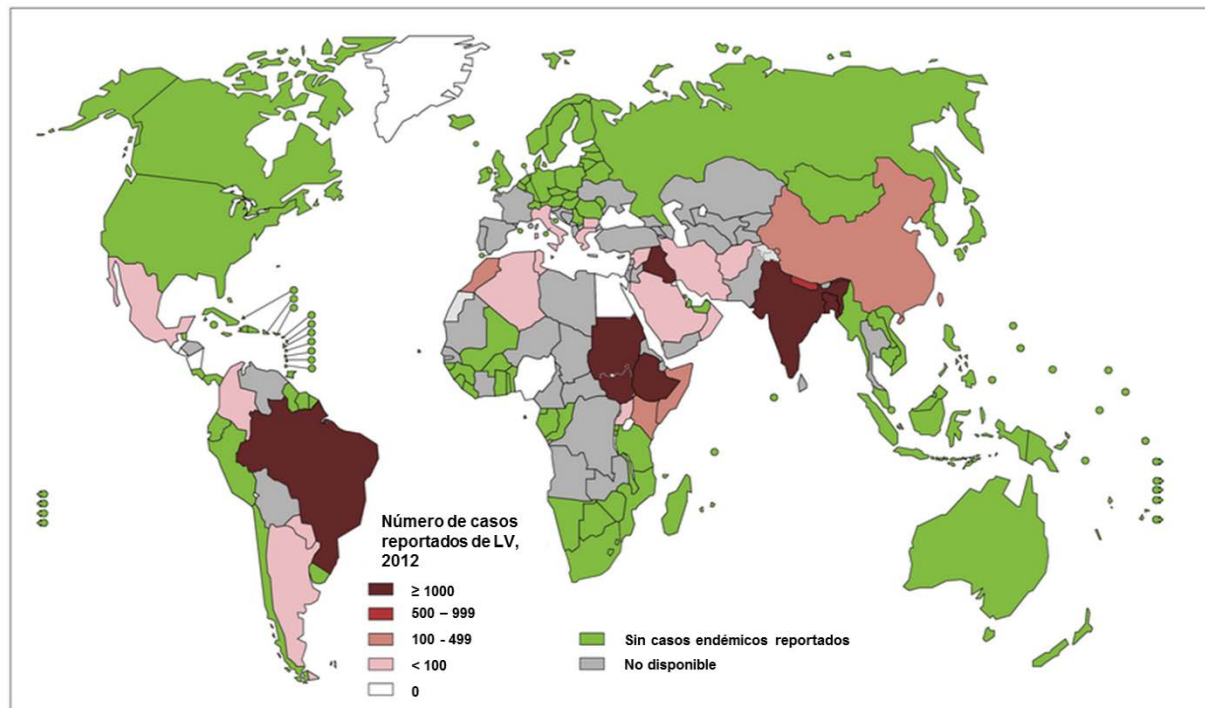
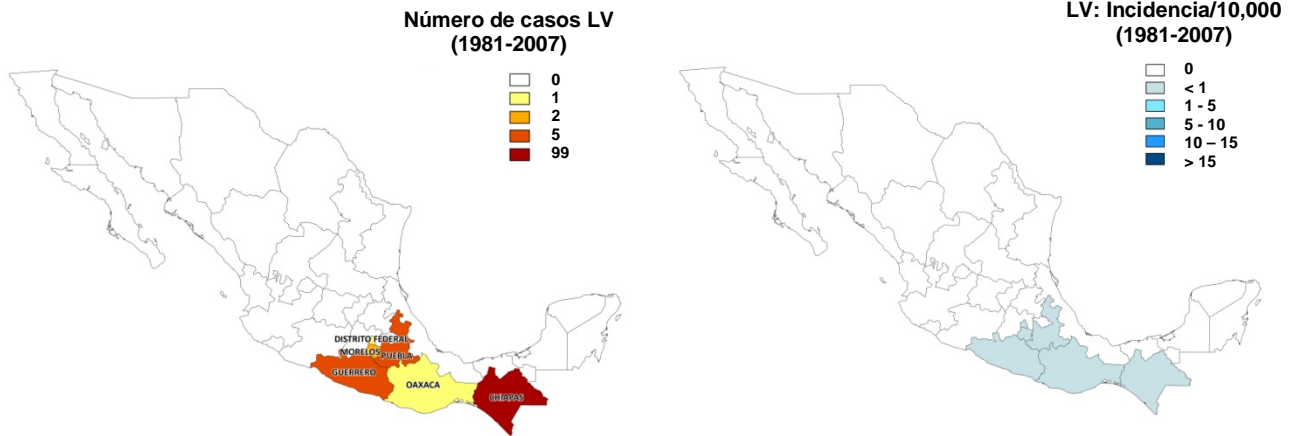
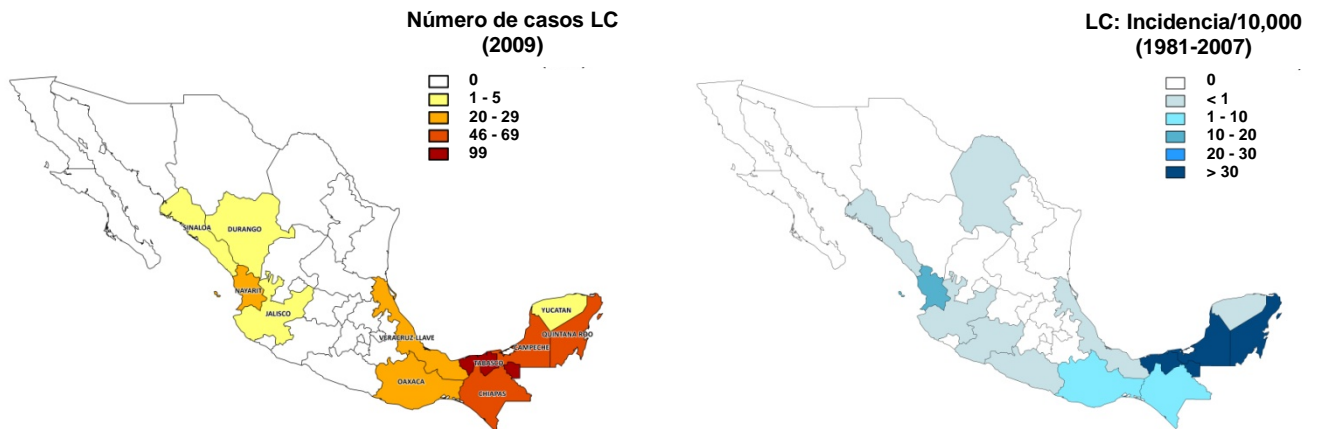


Figura 1. Mapas de zonas endémicas de leishmaniasis alrededor del mundo. Casos nuevos al 2012. A) Leishmaniasis cutánea; B) leishmaniasis visceral. (Tomado de *World Health Organization, Global Health Observatory* www.who.int/gho/neglected_diseases/leishmaniasis/en/).

A) Leishmaniasis Visceral



B) Leishmaniasis Cutánea



C) Leishmaniasis Mucocutánea

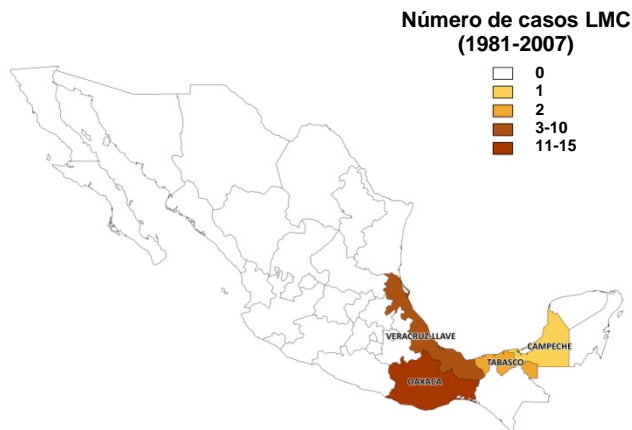


Figura 2. Incidencia y distribución en México de la Leishmaniasis. A) Visceral, B) Cutánea y C) Mucocutánea. Se muestra el número de casos y para A y B la incidencia/10,000 habitantes (Tomado de Alvar, J., et al., 2012).

2.2 Ciclo de vida

Los parásitos del género *Leishmania* son transmitidos por moscas de la arena hembras de los géneros *Lutzomyia* y *Phlebotomus*, previamente infectadas con estos protozoarios. Durante la ingesta de sangre, las moscas regurgitan en el hospedero mamífero (humano u otros reservorios) promastigotes metacíclicos, la forma infecciosa del parásito. Los promastigotes metacíclicos son fagocitados por diferentes células fagocíticas del ambiente tisular local. Una vez dentro de las células, los promastigotes metacíclicos se diferencian a amastigotes, la forma más pequeña y redondeada del parásito. Los amastigotes se replican dentro de la célula, rompiéndola y liberándose, lo que permite la reinfección de fagocitos locales y expandiendo la invasión a otros tejidos del mamífero. Otras moscas de la arena toman macrófagos infectados durante la ingesta de sangre. En el intestino de las moscas los amastigotes se transforman a promastigotes procíclicos, una forma no infecciosa y altamente replicativa del parásito. Estas formas se unen a la pared del intestino medio para luego liberarse y migrar al aparato bucal de la mosca, diferenciándose a promastigotes metacíclicos no replicativos que son transmitidos por el insecto al mamífero durante otra nueva ingesta de sangre (Figura 3) (Kaye, P. y Scott, P., 2011).

2.3 Organización genómica

El genoma de los organismos del género *Leishmania* está organizado en unidades policistrónicas, que son grupos de genes codificadores de proteínas sin una función común localizados consecutivamente en la misma cadena de DNA. Esta organización tan particular fue primeramente observada en el cromosoma 1 de *L. major* Friedlin, el cual contiene 85 genes organizados en dos unidades policistrónicas divergentes. Los primeros 32 genes se encuentran agrupados en la cadena inferior, mientras que los 53 genes restantes se encuentran en la cadena superior (Myler, P.J., et al., 1999). La publicación de la secuenciación del genoma completo de *L. major* Friedlin reveló que los genes de todos los cromosomas están organizados en unidades policistrónicas (Ivens, A.C., et al., 2005).

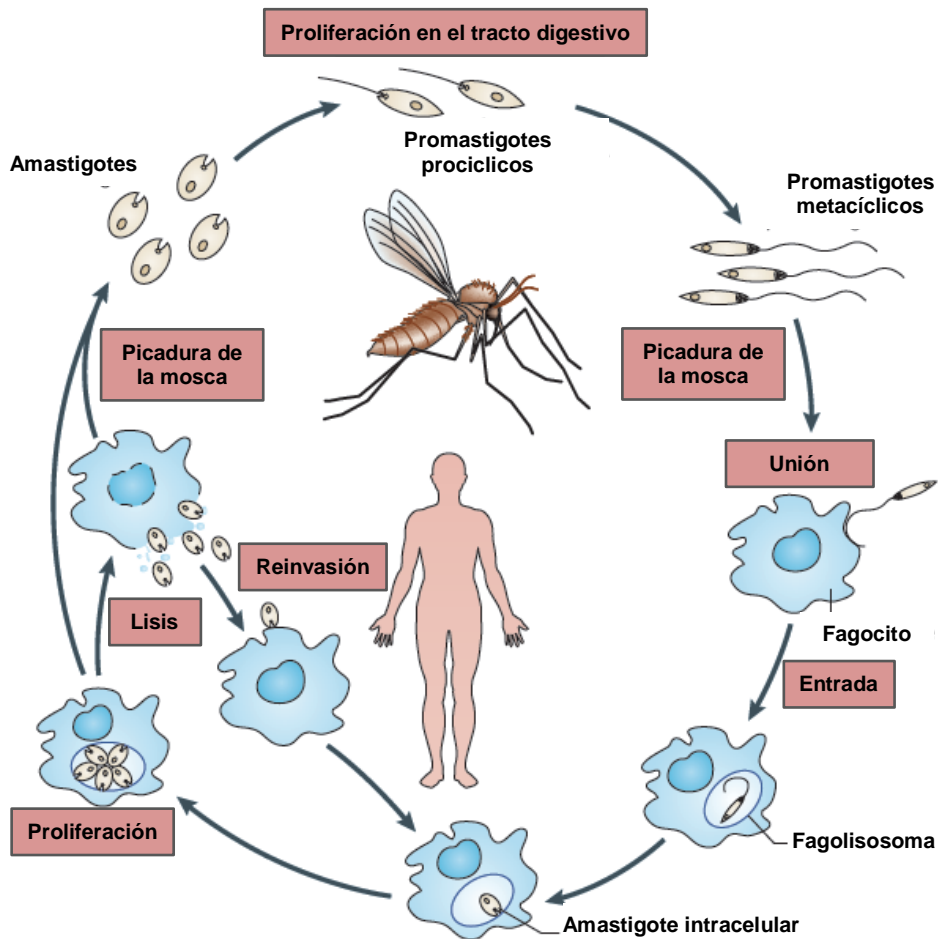


Figura 3: Ciclo de vida de *Leishmania* spp. Ver texto para detalles (Tomado de Kaye, P. y Scott, P., 2011)

Posteriormente, la secuenciación de los genomas de otras especies de *Leishmania* reveló la misma organización genómica, tal es el caso de *L. infantum* (cepa JPCM5), *L. braziliensis* (M2904), *L. mexicana* (U1103) y *L. donovani* (BPK282/0c14) (Peacock, C.S., et al., 2007; Rogers, M.B., et al., 2011; Downing, T., et al., 2011). Este tipo de organización genómica es común a otros tripanosomátidos de importancia médica, como *T. brucei* (Berriman, M., et al., 2005) y *T. cruzi* (El-Sayed, N., et al., 2005). En la Tabla 2 se enumeran las principales características de los genomas de diferentes especies de *Leishmania* spp. y *Trypanosoma* spp. (Padilla-Mejía, N.E., et al., 2013).

Tabla 2. Resumen de los principales características de los genomas de *Leishmania* spp., *T. brucei* y *T. cruzi*
(Tomado de Padilla-Mejia, N.E., et al., 2013).

	<i>L. major</i>	<i>L. infantum</i>	<i>L. braziliensis</i>	<i>L. mexicana</i>	<i>L. donovani</i>	<i>T. brucei</i>	<i>T. cruzi</i>
Número de cromosomas	36	36	35	34	36	11 ^(a)	41
Tamaño del genoma (pb)	32,855,089	32,101,728	31,997,773	32,108,741	32,44,998	26,075,396	60,372,297
Contenido general de G+C (%)	59.7	59.3	57.8	59.7	NR	46.4	51
Contenido G+C de los genes codificantes (%)	62.5	62.45	60.38	NR	NR	50.9	53.4
Genes codificantes de proteínas (predicción)	8,412	8,241	8,357	8,250	NR	9,068	21,133
Pseudogenes	97	41	161	99	13	904	3,590
Longitud media de las regions intergénicas (pb)	2,045	2,049	1,976	NR	NR	1,279	1,024
Densidad génica (genes/Mb)	252	235	258	NR	NR	317	385
Genes de tRNA ^(b)	83	77	75	84	76	66	120 ^(d)

- a) Sólo Megacromosomas
b) Estos números corresponden a lo reportado en GeneDB más una búsqueda personal utilizando tRNAScan-SE
c) NR: no reportado
d) Por genoma diploide.

Sobre los RNA de transferencia (tRNA), un análisis *in silico* del genoma de *L. major* reveló la presencia de 83 genes de tRNA distribuidos en 31 *loci*, en 19 cromosomas diferentes (Padilla-Mejia, N.E., *et al.*, 2009). La mayoría de los genes están organizados en grupos de 2-10 genes, los cuales pueden también contener otros genes transcritos por la RNA Polimerasa III, tales como el RNA ribosomal (rRNA) 5S o RNAs pequeños nucleares (snRNA). Asimismo, también se identificó un tRNA de selenocisteína (tRNA-Sec) en el genoma de *L. major* (Padilla-Mejia, N.E., *et al.* 2009). Las particularidades de este tRNA se retomaran más adelante. Otras especies de *Leishmania* también poseen una distribución y número de genes de tRNA similar a *L. major* (Tabla 2), de tal manera que la mayoría de los grupos de tRNA son altamente sinténicos entre diferentes especies de *Leishmania* (Tabla 2) (Padilla-Mejia, N.E., *et al.*, 2013).

2.4 Expresión génica

Una característica importante de la familia *Trypanosomatidae*, incluido el género *Leishmania*, es la transcripción policistrónica de los genes transcritos por la RNA Polimerasa II, es decir, la transcripción inicia únicamente río arriba del primer gen de cada unidad policistrónica, generándose transcritos policistrónicos que contienen varios pre-RNA mensajeros. Este transcrito policistrónico es procesado posteriormente para generar RNA mensajeros (mRNA) maduros. El extremo 5' de cada pre-mRNA se procesa por *trans-splicing*, proceso que consiste en la adición de una secuencia de 39 nucleótidos al extremo 5' de cada mRNA. Esta secuencia es conocida como miniexón o RNA *spliced-leader* (SL-RNA) (Van der Ploeg, L. H., 1986). Las secuencias requeridas para este procesamiento están conservadas y consisten en un dinucleótido AG en el sitio de adición del miniexón así como una región rica en pirimidinas localizada río arriba (Benz, C., *et al.*, 2005; Siegel, T.N., *et al.*, 2005). El extremo 3' de los mRNA se genera añadiendo un tracto de adeninas (Figura 4).

Todos los genes que son parte de la misma unidad policistrónica se transcriben al mismo nivel, como consecuencia de la transcripción policistrónica. Ya que los mRNA maduros de los genes adyacentes pueden diferir en sus niveles de expresión, se ha observado que la expresión génica en tripanosomátidos es regulada principalmente de manera post-transcripcional, a nivel de procesamiento y estabilidad de los mRNA. Las secuencias 3' no traducidas (3'-UTR) de los mRNA son importantes en dicha regulación post-transcripcional, un ejemplo de esto es la región 3'-UTR del mRNA de amastina en *L. infantum*, la cual tiene una región de 450 pb que es importante para incrementar la traducción del mRNA durante el estadio de amastigote (Boucher, N., *et al.*, 2002).

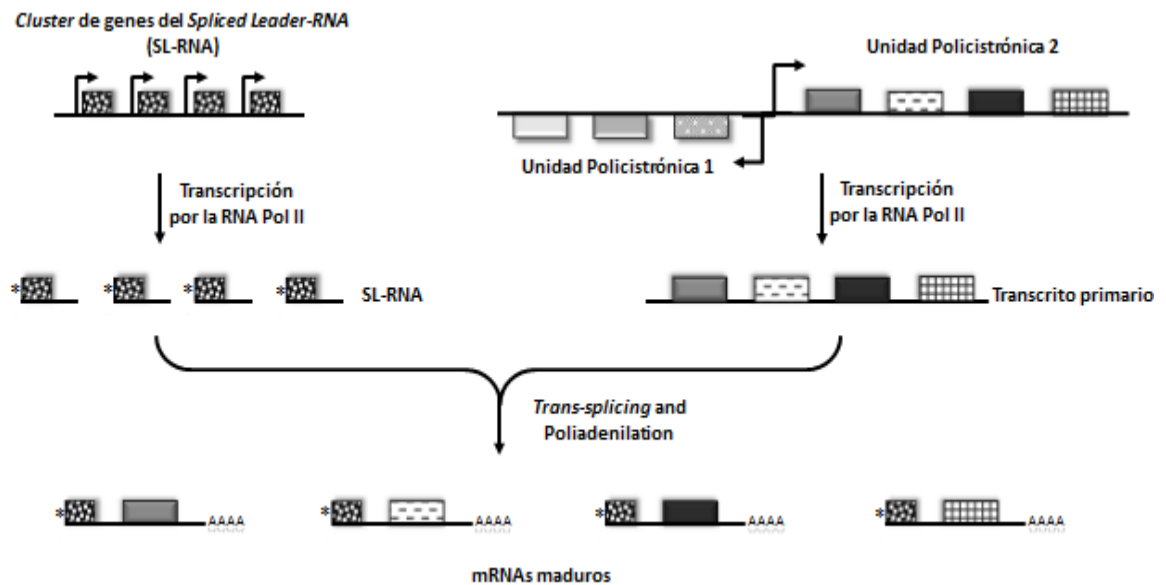


Figura 4: Generación y procesamiento de los transcritos policistrónicos por *trans-splicing*. Ver detalles en el texto

En tripanosomátidos, no se ha demostrado aún la existencia de secuencias específicas que actúen como promotores de la transcripción por la RNA Pol II, sin embargo, un análisis transcripcional del cromosoma 1 de *L. major* mostró que la transcripción inicia en la región de cambio de hebra (*strand switch región*) que se encuentra entre las dos unidades policistrónicas divergentes, de tal manera que una sola región parece dirigir la transcripción de todo el cromosoma, a diferencia de eucariontes superiores que poseen un promotor por cada gen de proteínas. Una vez iniciada la transcripción, esta procede bidireccionalmente hasta la región telomérica. En este trabajo se mapearon varios sitios de inicio de la transcripción dentro de una región <100 pb que contiene largos segmentos de G/C, sin que se hayan identificado cajas TATA u otro elemento promotor típico de la transcripción por la RNA Polimerasa II en eucariontes superiores (Martínez-Calvillo, S., et al., 2003).

2.5 Transcripción por la RNA Polimerasa III

La transcripción de genes en eucariontes está a cargo de tres RNA Polimerasas (RNA Pol): la RNA Pol I sintetiza los RNAs ribosomales (rRNA) 18S, 28S y 5.8S; RNA Pol II sintetiza los mRNA codificantes de proteínas y RNAs no codificadores (ncRNA) como los RNAs pequeños nucleares (snRNAs), RNAs pequeños nucleolares (snoRNAs) y microRNA (miRNA). La RNA Pol III transcribe una gran variedad de RNAs pequeños no codificantes, entre los que se encuentran el rRNA 5S, los RNA de transferencia (tRNA), el snRNA U6 y el RNA 7SL (Schramm, L. y Hernandez, N., 2002; Paule, M. y White, R., 2000).

En levadura, la RNA Pol II contiene 12 subunidades, mientras que la RNA Pol I contiene 14 subunidades y la RNA Pol III posee 17. La presencia de las tres RNA Polimerasas ya ha sido demostrada en tripanosomátidos: en *T. brucei*, mediante cromatografía de intercambio aniónico y experimentos de *run-on* nuclear con los inhibidores de las Polimerasas α -amanitina y O-fenantrolina (Grondal, E. J., et al., 1989). En *L. major* ya se han identificado varias subunidades de las RNA Polimerasas mediante purificación por afinidad en tándem y la subsecuente espectrometría de masas, así como por análisis *in silico* realizados con las bases de datos del genoma de este organismo (Martínez-Calvillo, S., et al., 2007).

2.5.1 Genes transcritos por la RNA Polimerasa III y sus Elementos Promotores

Como se mencionó anteriormente, la RNA Pol III transcribe una gran variedad de ncRNA, siendo los más significativos el rRNA 5S, los tRNA, el RNA 7SL. El rRNA 5S es un componente estructural ribosomal, los tRNAs descifran los codones de los mRNA y acarrean sus aminoácidos correspondientes a la cadena polipeptídica en crecimiento durante la síntesis de proteínas. Por su parte, el RNA 7SL es un componente de la partícula de reconocimiento del péptido señal para retículo endoplásmico y el snRNA U6 forma parte del complejo del splicesosoma (Dieci, G., et al. 2007). Otros RNAs transcritos por la RNA Pol III se enlistan en la Tabla 3.

Tabla 3. Algunos RNAs sintetizados por la RNA Polimerasa III (Dieci, G., et al. 2007)

Tipo de RNA	Funciones y características	Promotores
RNA H1	Componente de la RNAsa P, una endorribonucleasa que procesa el extremo 5' de los pre-tRNA	En <i>Saccharomyces cerevisiae</i> , el promotor es una secuencia tipo tRNA (tRNA-like). En vertebrados es un promotor con elementos PSE/TATA
RNA MRP	Componente de la RNAsa MRP (RNAsa mitocondrial de procesamiento de RNA) que procesa cebadores de RNA para la replicación del DNA mitocondrial. Esta RNAsa también participa en nucléolo procesando pre-rRNA	Su promotor posee elementos TATA, PSE y DSE
RNA Vault (VA RNA)	Constituyentes de ribonucleoproteínas (RNP) citoplásmicas implicados en resistencia de tumores a fármacos, ensamblaje y/o transporte macromolecular	Promotor con elementos internos y externos
RNA Y	Componente de la RNP Ro	Promotor con elementos DSE, PSE y caja TATA
RNA 7SK	Identificado sólo en vertebrados hasta la fecha, son represores de la elongación por la RNA Pol II al unirse al factor P-TEFb	Promotor con elementos DSE, PSE y caja TATA
RNA BC1 y BC200	Son específicos de ratón y primates, respectivamente, y se expresan ambos en células neuronales. Son parte de una RNP citoplasmática con una posible función en la traducción de mRNA de dendritas. Se ha demostrado que BC1 tiene relación con la espermatogénesis	Promotor con elementos internos y externos
RNAs VA-I y VA-II	Provenientes de adenovirus, inhiben la proteína-kinasa activada por RNA (PKR) que se activa en respuesta a infecciones virales, permitiendo la traducción de mRNA adenovirales	Promotor interno con elementos A y B
SINEs (<i>Short interspersed repeated DNA elements - encoded RNAs</i>)	Retrotransposones no autónomos originados por genes transcritos por RNA Pol III. Parecen regular la expresión génica a nivel post transcripcional, modular la traducción de proteínas o apagar su expresión (RNA Alu). Existen más de un millón de SINEs en el genoma humano, la mayoría inactivos	Mantienen parcialmente los elementos promotores internos (caja A y B)
microRNAs	Regulación de la expresión génica. Algunos microRNAs son transcritos por RNA Pol II	Elementos promotores internos, tipo tRNA (tRNA-like)

Los elementos promotores encargados de dirigir la transcripción de los genes por la RNA Pol III se han catalogado en tres tipos:

- a) **Tipo 1:** el rRNA 5S es el prototipo de este promotor, constituido por tres elementos internos: una caja A, una caja C y un Elemento Intermedio (EI) (Figura 5A). La caja A está localizada entre los nucleótidos +50 al +60, el EI entre el +67 al +72 y la caja C del +80 al +90. La caja C es un

elemento altamente conservado entre especies. Este promotor fue descrito primeramente en *Xenopus laevis* (Schramm, L. y Hernandez, N., 2002; Paule, M. y White, R., 2000).

- b) **Tipo 2:** se encuentran en los genes de tRNA, y se componen de las cajas A y B, las cuales son secuencias altamente conservadas (Figura 5A). La caja A se localiza cerca del sitio de inicio de transcripción, de la posición +8 a +19, mientras que la caja B se localiza en la segunda mitad del gen, frecuentemente en las posiciones +52 a +62; sin embargo, ésta tiene una localización variable debido a que algunos tRNA poseen un intrón que se localiza antes de la caja B (Schramm, L. y Hernandez, N., 2002; Paule, M. y White, R., 2000).

- c) **Tipo 3:** El ejemplo más representativo de este tipo de promotor es el de los genes del snRNA U6 en humano (Figura 5A), formado por elementos extragénicos localizados río arriba del gen. Este promotor consta de una caja TATA (localizada aproximadamente en la posición -30), una secuencia proximal (*proximal sequence element* o PSE) que va de -65 a -48 pb respecto al nucleótido +1 del gen. Otro elemento, la secuencia distal (*distal sequence element* o DSE) que se va aproximadamente de -240 a -215 ayuda a maximizar la tasa transcripcional. Este tipo de promotor se encuentra también en los genes del RNA 7SK y el RNA H1 (componente de la RNAsa P) (Schramm, L. y Hernandez, N., 2002; Paule, M. y White, R., 2000)

Los ejemplos arriba mencionados constituyen los puntos de referencia para cada tipo de promotor (Figura 5A), pero se ha observado que hay promotores diferentes que no pueden agruparse exclusivamente dentro de alguno de estos tres tipos. Por ejemplo, en los genes del rRNA 5S de *Saccharomyces cerevisiae* sólo se requiere la caja C para activar la transcripción, mientras que en *Schizosaccharomyces pombe* se encuentra una caja TATA río arriba del gen rRNA 5S, la cual es esencial para la transcripción (Hamada, M., et al., 2001). En *Acanthamoeba castellanii*, los elementos internos estándar del rRNA 5S son esenciales para la transcripción, pero además se identificó un elemento que se localiza exactamente después del sitio de inicio de la transcripción y que es también esencial para la expresión del gen (+1 a +22 pb) (Peng, Z. y Bateman, E., 2004). En *S. cerevisiae*, el gen del snRNA U6 posee una caja TATA en la región extragénica 5', una caja A interna y una caja B río abajo del gen, elementos que constituyen un promotor híbrido. Otro ejemplo de esto son los tRNA de plantas, levadura y el gusano de la seda, que además de poseer las cajas A y B

también cuentan con una caja TATA en el extremo 5' que contribuye significativamente a la eficiencia de la transcripción (Schramm, L. y Hernandez, N., 2002).

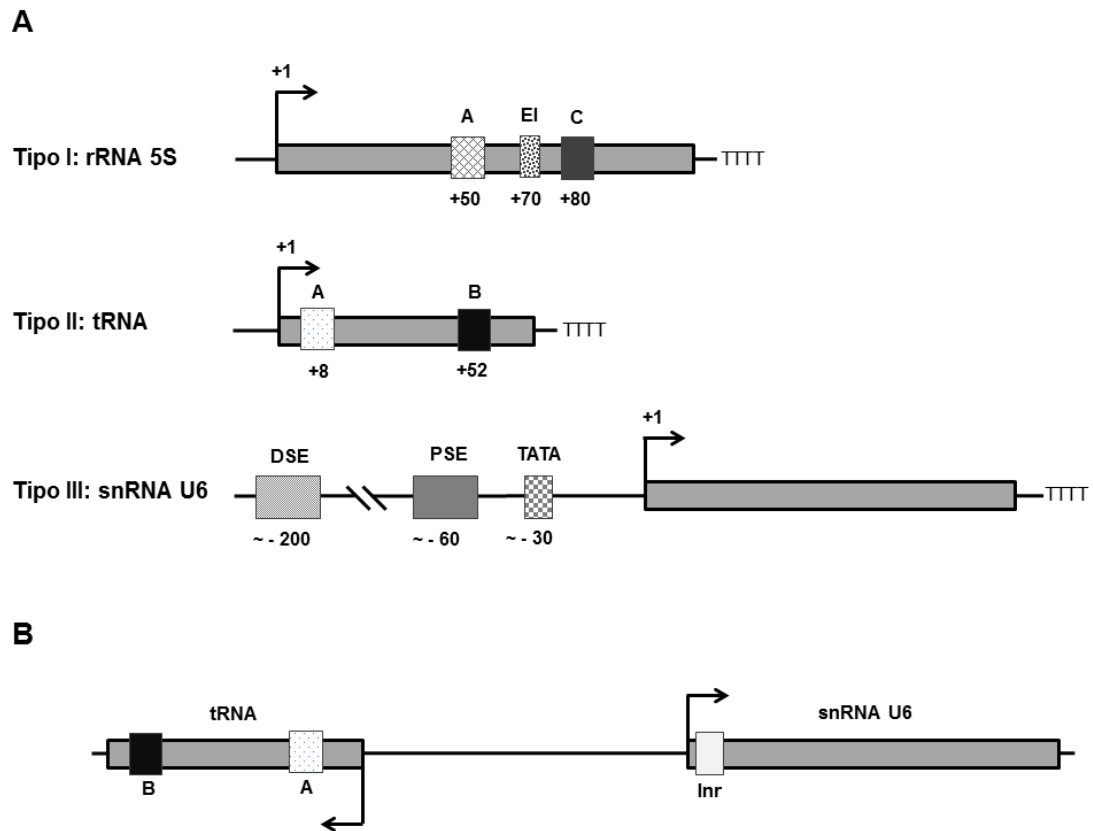


Figura 5. Promotores de la RNA Pol III. A) Constitución de los tres tipos de promotores clásicos de la RNA Polimerasa III caracterizados en eucariontes, B) Asociación entre un tRNA y el snRNA U6 en tripanosomátidos, el tRNA está localizado río arriba del snRNA y orientado divergentemente.

En tripanosomátidos, la RNA Pol III sintetiza todos los snRNA, además del rRNA 5S, el RNA 7SL y los tRNAs. Los promotores de los genes de tRNA en estos organismos contienen los elementos intragénicos clásicos caracterizados en eucariontes superiores: las cajas A y B (Padilla-Mejía, N.E., et al., 2009; Campbell, D.A., et al., 1989). Las secuencias promotoras de los snRNA U6, U3 y el del RNA 7SL ya han sido caracterizados en *T. brucei* (Nakaar, V., et al. 1994). Río arriba de estos genes (~95 pb), reside un gen de tRNA orientado en dirección contraria, cuyas cajas A y B son esenciales para la expresión de los snRNAs y del RNA 7SL (Figura 5B). El snRNA U6 además requiere de una región intragénica situada de +2 - +11 para su transcripción (elemento iniciador, Inr) (Nakaar, V., et al. 1994). Cabe mencionar que los snRNA en tripanosomátidos carecen de los elementos promotores que se han identificado en vertebrados como los responsables de dirigir su transcripción, tales como la caja

TATA, el PSE y el DSE (Figura 5B). Así, las cajas A y B del tRNA poseen un rol dual en tripanosomátidos: son elementos promotores clásicos para el propio tRNA y además se requieren para la expresión de otros genes transcritos por la RNA Pol III (Palenchar, J. y Bellofatto, V., 2006). La asociación entre los snRNAs y los tRNAs para la expresión génica de los primeros se encuentra difundida entre los organismos de la familia *Trypanosomatidae*, incluyendo especies de los géneros *Leishmania*, *Leptomonas* y *Crithidia*. Es importante recalcar que, mientras que en vertebrados los snRNA U1-U5 son transcritos por la RNA Pol II, en tripanosomátidos dichos snRNAs son transcritos por la RNA Pol III, utilizando como promotores las cajas A y B de las secuencias de tRNA (o tipo tRNA) situados río arriba. La transcripción de todos los snRNA por RNA Pol III en tripanosomátidos constituye otra de sus particularidades de expresión génica (Palenchar, J. y Bellofatto, V., 2006).

El rRNA 5S en tripanosomátidos contiene las cajas A, C y el EI (Hernández-Rivas, R., et al. 1992; Lenardo, M.J., et al. 1985), elementos promotores que se han identificado en eucariontes superiores, sin embargo, hasta la fecha dichas secuencias no han sido caracterizadas funcionalmente en estos organismos.

2.5.2 Factores de transcripción asociados a la RNA Polimerasa III

El factor TFIIIB es un factor de transcripción común para todos los promotores de la RNA Pol III. En *S. cerevisiae*, TFIIIB es un complejo proteico constituido por tres polipéptidos: TBP (*TATA Binding Protein*), B' (Bdp1 en humanos) y Brf (*TFIIIB-related factor*). En los promotores tipo 2, el factor TFIIIC reconoce las cajas A y B, esta unión DNA-proteínas permite posteriormente el reclutamiento del factor TFIIIB. Al parecer la caja B es el determinante principal para la unión de dicho factor, aunque la caja A es también contactada. A su vez, la unión de TFIIIB al promotor favorece el reclutamiento de la RNA Pol III (Figura 6) (Schramm, L. y Hernandez, N., 2002).

En los promotores tipo 1, las cajas A, C y el elemento intermedio son reconocidos por el factor TFIIIA, proteína con nueve dedos de zinc. La caja C es reconocida por tres dedos del extremo N-terminal, mientras que los dedos 7-9 contactan la caja A. El complejo TFIIIA-DNA dirige a TFIIIC hacia el promotor del rRNA 5S. Tras la unión de TFIIIC, la vía de reclutamiento de la polimerasa es similar a la

que ocurre con los promotores tipo 2, con el reclutamiento de TFIIIB y RNA Pol III (Figura 6) (Schramm, L. y Hernandez, N., 2002).

En los promotores tipo 3, el PSE es reconocido por el complejo multiproteico SNAPc (snRNA *activating protein complex*, complejo activador de snRNA). Por su parte la caja TATA es reconocida por TFIIIB. El DSE posee un octámero en su secuencia, sitio de unión para la proteína OCT1 y una secuencia SPH (*Sph1 postoctamer homology*) que recluta Staf (también conocido como ZNF o SBF). Las interacciones DNA-proteína son reforzadas por interacciones proteína-proteína entre SNAPc y TBP. La unión de SNAPc y el factor TFIIIB conducen al reclutamiento de la RNA Pol III, probablemente a través de contactos proteína-proteína con los factores mencionados (Figura 6). Los promotores tipo 3 reclutan TFIIIB, sin embargo usan una forma alternativa de la subunidad Brf1, a la que se denomina Brf2 (Schramm, L. y Hernandez, N., 2002; Paule, M. y White, R., 2000).

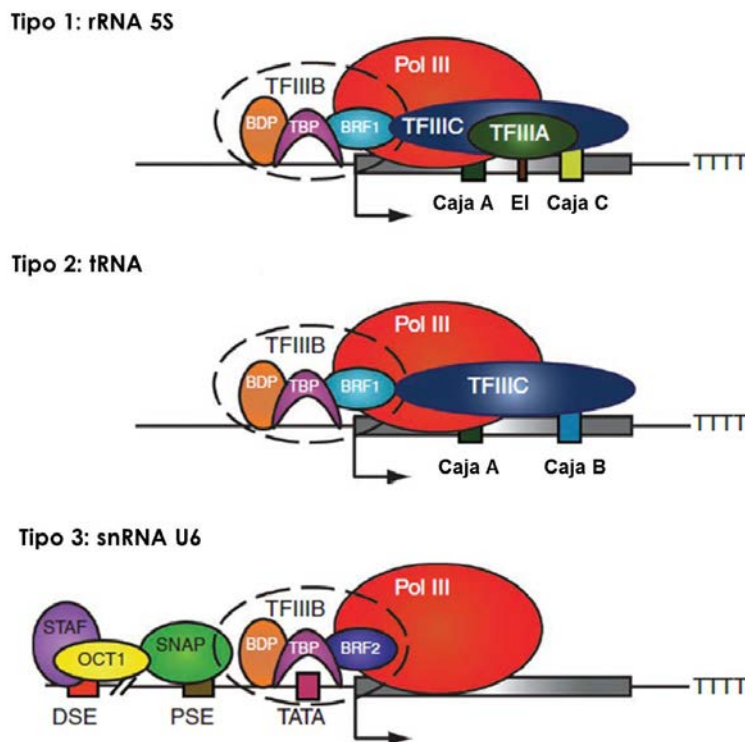


Figura 6. Factores de transcripción relacionados con la RNA Polimerasa III. Se muestran los tres tipos de promotores y los factores reclutados en cada uno de ellos (Tomado de Oler, A., et al., 2011). Ver texto para detalles.

En tripanosomátidos se han caracterizado pocos factores de transcripción. En *T. brucei* ya existen reportes sobre proteínas nucleares de unión específica a los genes de tRNA, lo cual se determinó utilizando ensayos de retardamiento de movilidad electroforética. Además, mediante ensayos de *footprinting* se mostró el posicionamiento de complejos proteicos sobre las cajas A y B (Bell, S.D. y Barry, J.D., 1995). Asimismo, en *T. brucei* se caracterizó posteriormente el factor TRF4 (TBP-related factor 4), un factor esencial para la transcripción de las tres RNA Pol, el cual es reclutado a los genes snRNA U2, U6 y el RNA 7SL, transcritos por la RNA Pol III (Ruan, J.P., et al., 2004). Por otra parte, en estos parásitos, se han detectado dos subunidades del factor de transcripción TFIIIB, Brf1 y B'' (Berriman, M., et al., 2005; El-Sayed, N., et al., 2005; Ivens, A.C., et al., 2005; Schimanski, B., et al., 2005). Estas proteínas poseen ortólogos presentes en varias especies de los géneros *Leishmania* y *Trypanosoma* (www.genedb.org y www.tritrypdb.org). Sin embargo, no se han localizado los ortólogos de TFIIIA ni TFIIIC. Mediante ensayos ChIP-on-chip se ha observado que el factor TRF4 y el factor SNAP50 se unen a grupos de genes que incluyen tRNAs, snRNAs y rRNA 5S en *L. major* (Thomas, S., et al. 2009, www.tritrypsdb.org).

2.6 Selenocisteína y Selenoproteínas

Las selenoproteínas son proteínas caracterizadas por poseer el residuo de selenocisteína (Sec), un aminoácido ya reconocido como el aminoácido número 21, el cual es un análogo de la cisteína con un átomo de selenio reemplazando el grupo tiol (Labunsky, V.M., et al. 2014). Las características nucleofílicas/catalíticas del residuo Sec lo hacen excelente para reacciones de óxido-reducción y se encuentra frecuentemente formando parte del centro activo de enzimas redox. En humano, se han identificado a la fecha 25 selenoproteínas y 24 en ratón. En este último organismo se ha demostrado que son esenciales durante el desarrollo. Además de su actividad catalítica en reacciones de óxido-reducción, las selenoproteínas están involucradas en otras funciones como defensa contra el estrés antioxidante, transporte de selenio, biosíntesis de fosfolípidos y algunas de ellas se han relacionado a desórdenes cardíacos, endócrinos y neurológicos en humanos (Reeves, M.A. y Hoffmann, P.R. 2009). En la Tabla 4 se enlistan las selenoproteínas caracterizadas hasta el momento en mamíferos.

Tabla 4. Selenoproteínas caracterizadas en mamíferos

Selenoproteínas	Características estructurales y funcionales
<p>Tiorredoxina reductasas (TrxR)</p> <p>Tres TrxR en mamíferos:</p> <p>TrxR1 (citoplásmica/nuclear)</p> <p>TrxR2 (mitocondrial, también llamada TrxR3)</p> <p>Tiorredoxina-glutatión reductasa (testicular)</p>	<p>TrxR posee dominios de unión a FAD y NADPH. El C-terminal posee una secuencia conservada G-C-U-G-OH.</p> <p>Las TrxR catalizan la reducción de la tiorredoxina, que a su vez regula el estatus redox celular. Estas selenoproteínas participan en síntesis de ácidos nucleicos, reducción de peróxidos, regulación de apoptosis, etc.</p> <p>La actividad disminuida de TrxR está relacionada con desordenes en la proliferación celular (carcinogénesis y enfermedades inmunológicas). TrxR1 es esencial para la embriogénesis y TrxR2 para la hematopoyesis y el desarrollo y función correctos del corazón ^{a, b}.</p>
<p>Glutatión peroxidasas (GPx)</p> <p>Cinco Gpx en humano:</p> <p>GPx1 (citoplásmica)</p> <p>GPx2 (gastrointestinal)</p> <p>GPx3 (plasma, extracelular)</p> <p>GPx4 (se encuentra en testículos)</p> <p>GPx6 (específica de embrión y epitelio olfatorio)</p>	<p>Principales componentes de las defensas humanas antioxidantes, actúan como mecanismos de defensa celulares. Las GPx reducen hidroperóxidos usando glutatión como cofactor. Las GPx poseen una tríada catalítica conservada U-Q-W.</p> <p>GPx1 actúa como antioxidante, pues ratones <i>knock-out</i> GPx1 son más susceptibles al daño oxidativo. Ratones doble <i>knock-out</i> GPx1/GPx2 desarrollaron progresivamente cáncer intestinal. Un alelo de GPx2 insertado de nuevo confirió protección anticancerígena, por lo que GPx2 podría servir como barrera contra hidroperóxidos intestinales.</p> <p>GPx3 es un eficiente antioxidante, importante para la protección cardiovascular.</p> <p>GPx4 reduce fosfolípidos e hidroperóxidos de colesterol; funciona en embriogénesis y es componente estructural de la capsula mitocondrial del espermatozoide maduro ^a</p>
<p>Iodotironina desiodinasas (DIO)</p> <p>DIO1</p> <p>DIO2</p> <p>DIO3</p>	<p>DIO1 y DIO3 se localizan en membrana plasmática y DIO2 en la membrana de retículo endoplásmico ^a</p> <p>Están involucradas en la activación e inactivación de la hormona tiroidea. DIO1 y DIO2 catalizan la conversión de la prohormona T4 (tiroxina) a la hormona activa 3,5,3'-triyodotironina (T3).</p> <p>DIO3 es la principal ruta de inactivación que termina la acción de T3 y previene la activación de la prohormona T4 (18), revirtiendo T4 en T3 y posteriormente T3 a 3, 3-diyodotironina. DIO1 y DIO2 también pueden modificar T3 en 3,3-diyodotironina</p>
<p>Sep15 y Sel M</p>	<p>Debido a su homología con enzimas disulfuro-isomerasas, se piensa que estas dos selenoproteínas funcionan como oxidoreductasas tiol-disulfuro. Sep15 y Sel M poseen motivos C-X-U y C-X-X-U, respectivamente y se localizan en retículo endoplásmico. Ambas proteínas se han asociado al control de calidad del plegamiento proteico en dicho organelo ^b.</p> <p>Ratones <i>knock-out</i> de SelM exhiben incremento de peso, una masa elevada del tejido adiposo, por lo que se ha sugerido que SelM juega un papel importante en la regulación del peso corporal y el metabolismo energético^f</p>
<p>Selenoproteína H (SELH)</p>	<p>Posee un motivo C-X-X-U. En <i>D. melanogaster</i> es esencial para la viabilidad y defensa antioxidante. Se localiza en núcleo y podría ser una proteína de unión a DNA que responde a cambios redox, regulando la expresión de genes involucrados en la síntesis de novo de glutatión y en la respuesta detoxificante al estatus redox. Se expresa muy activamente en etapas tempranas del desarrollo embrionario ^{b, c}</p>
<p>Selenoproteína I (SEL I, hEPT1)</p>	<p>Proteína tipo etanolamina-fosfotransferasas que actúan durante la biosíntesis de fosfolípidos. Se han encontrado 7 hélices transmembranales pero no hay evidencia experimental sobre su localización celular ^b.</p>
<p>Selenoproteína K (SelK)</p>	<p>Se localiza en membrana de retículo endoplásmico y posiblemente en membrana plasmática. Su sobreexpresión en cardiomiocitos murinos disminuye los niveles de ROS, sin embargo, en <i>D. melanogaster</i> SelK se localiza en citoplasma y no posee actividad antioxidante, por lo que su función permanece desconocida^b.</p>

Selenoproteína N (SelN, SEPN1, SepN)	<p>Proteína transmembranal de retículo endoplásmico. Está relacionada a ciertos desordenes musculares conocidos como miopatías relacionadas con SEPN-1 (clínicamente caracterizados por fuerza muscular axial, escoliosis, debilidad en el cuello y grados variables de espina rígida). Ya que se expresa en tejido fetal y células en proliferación podría participar en la formación muscular temprana. SelN podría regular el canal de Ca²⁺/receptor de rianodina(RyR) en retículo endoplásmico, que al activarse aumenta el Ca²⁺ libre ^b</p>
<p>Se conocen dos isoformas: SelN-1 (transcrito del gen completo) y SelN-2 (transcrito sin el exón 3)</p>	
Selenoproteína O (SelO)	<p>Contiene un motivo C-X-X-U. Ha sido localizada en mitocondria y expresada en células humanas HEK 293T y diversos tejidos de ratón. Se ha demostrado que SelO interacciona con una proteína blanco de manera redox, que dicha interacción es reversible y que depende de la presencia de peróxido de hidrógeno, lo cual es una propiedad típica de una oxidación-reducción tiol-dependiente ^d</p>
Selenoproteína P (SelP)	<p>Contiene 8-10 residuos Sec en mamíferos, 16-18 en peces y anfibios y 28 en erizos de mar. SelP contiene del 40-50% del selenio circulante en plasma, atribuyendo su función principal al transporte del mismo. Los ratones <i>knock-out</i> para SelP muestran problemas neurológicos y esterilidad. Otras funciones atribuidas son: quelante de metales pesados y mitigante del estrés oxidativo en cáncer ^b.</p>
Selenoproteína S (SEPS1, SELS, SELENOS, VIMP, Tanis)	<p>Proteína transmembranal localizada en retículo endoplásmico y membranas plasmáticas. Podría participar en la remoción de proteínas mal plegadas del lumen del retículo endoplásmico; en la protección celular contra el daño oxidativo; en la apoptosis inducida por el estrés de RE y parece favorecer la producción de citocinas inflamatorias. Se ha asociado a enfermedades cardíacas, apoplejía, preeclampsia, artritis reumatoide, etc. ^b</p>
Selenoproteína T (SelT)	<p>Localizada en retículo endoplásmico, junto con SelH, SelV y SelW, posee un motivo C-X-X-U. Posiblemente involucrada en fenómenos relacionados con el estrés, un estudio reciente la ha vinculado con la movilización celular de calcio. En rata, SelT se expresa en la mayoría de los tejidos a través del desarrollo embrionario y en individuos adultos ^{b, c}</p>
Selenoproteína V (SelV)	<p>Posee un motivo C-X-X-U. Parece estar restringida a testículo y, por su estructura, podría tener una función redox. SelV se expresa en rata durante las etapas de espermatogénesis tardía, la pubertad y los periodos reproductivos ^e. Su función exacta es desconocida</p>
Selenoproteína W (SEPW, SELW)	<p>Contiene un motivo C-X-X-U. Aunque su función exacta es aún desconocida, se le atribuye un papel antioxidante (por su capacidad de unión al glutatión y la protección celular contra H₂O₂) y posiblemente intervenga en la defensa celular contra el cadmio. Por su alta expresión en mioblastos proliferantes, SelW podría funcionar en el crecimiento y diferenciación musculares, protegiendo dichas células del estrés oxidativo ^{b, c}</p>
Selenoproteína R	<p>Cataliza la reducción de los residuos metionina (Met) oxidados (sulfóxidos de metionina) por ROS, dicha oxidación ocurre bajo estrés oxidativo y puede conducir a daño proteico. MsrB1 actúa sobre el epímero R del sulfóxido de Met y se encuentra en vertebrados. Ratones <i>knock-out</i> de SelR aumentan su susceptibilidad al daño oxidativo en proteínas, especialmente en hígado y riñón. Otras funciones incluyen protección contra la neurodegeneración y el daño oxidativo durante el envejecimiento ^{a, b}</p>
SelR, MsrB1 (metionin-R-sulfóxido reductasa)	
SPS2 Selenofosfato sintetasa 2	<p>Esta selenoproteína está involucrada en la vía de síntesis de Sec sobre su tRNA correspondiente ^{a, b}</p>

- a) Lu, J. y Holmgren, A., 2009.
c) Dikiy, A. *et al.*, 2007.
e) Varlamova, E.G., *et al.*, 2011

- b) Reeves, M.A. y Hoffmann, P.R. 2009.
d) Han, S.J., *et al.*, 2014
f) Pitts, M.W., *et al.*, 2013.

La maquinaria de síntesis de selenoproteínas contiene varios aspectos inusuales: a) el aminoácido Sec se sintetiza sobre su tRNA específico (tRNA-Sec) a partir del residuo de serina, b) el tRNA-Sec es el tRNA más largo conocido y posee particularidades estructurales que difieren de los tRNA canónicos, c) el tRNA-Sec posee elementos promotores que difieren de los tRNA clásicos, d) la incorporación de Sec a las proteínas se lleva a cabo en respuesta a un codón UGA, cuya función normal es como un codón de paro, e) la presencia del elemento SECIS (*Selenocysteine Insertion Sequence*) en la región 3'-UTR de los mRNA, el cual es esencial para la traducción, f) la existencia del eEFSec, un factor de traducción especializado en la síntesis de selenoproteínas (Labunskyy, V.M., et al. 2014). A continuación se describen las moléculas participantes más relevantes en la incorporación de Sec a las selenoproteínas. Las características del tRNA-Sec se discuten más adelante en la sección 2.9.

2.7 Elementos esenciales para la inserción de Sec en las selenoproteínas

2.7.1 Elemento SECIS (*Selenocysteine Insertion Sequence*)

El elemento SECIS es una estructura de tallo-asa que se localiza en la región 3'-UTR de los mRNA de las selenoproteínas. Los elementos SECIS están altamente conservados a nivel de estructura secundaria (aunque no a nivel de secuencia), la cual consiste en un asa apical y dos hélices separadas por un asa interna (Figura 7A). En la hélice II, junto al asa interna se encuentra un dinucleótido AG:GA conservado, esencial para el funcionamiento del SECIS y su interacción con la proteína SBP2 (descrita más adelante). En algunos elementos SECIS el asa apical forma un minitallo-asa adicional, el cual es utilizado para clasificar a los elementos SECIS en dos tipos: aquellos cuya asa apical carece del minitallo (Tipo 1) y aquellos que sí lo poseen (Tipo 2) (Figura 7A). En el asa apical se encuentra el trinucleótido RAA, también requerido para la incorporación de Sec a las selenoproteínas (Walczak, R., et al., 1998). El elemento SECIS posee un motivo de RNA conocido como "*kink-turn*", que es un sitio importante para la interacción y reconocimiento entre RNA y proteínas. En el caso de SECIS, este motivo podría ser un puente molecular entre el ribosoma y la maquinaria de síntesis de selenoproteínas, favoreciendo la unión del eEFSec al ribosoma o el acomodo del tRNA-Sec (Allmang, C. y Krol, A., 2006b).

Los elementos SECIS son particulares para cada mRNA en cada especie, es decir, homólogos de la misma selenoproteína en diferentes organismos pueden poseer diferentes elementos SECIS. Por ejemplo, el mRNA de la selenoproteína SelM en mamíferos posee un RNA SECIS tipo 2, mientras que el

mRNA de SeIM en el pez cebra posee un elemento tipo 1 (Allmang, C. y Krol, A., 2006b). El elemento SECIS interacciona con diversas proteínas incluyendo eEFSec, SBP2 y L30 para regular la adición de selenocisteína a las selenoproteínas (Labunskyy, V.M., *et al.* 2014).

2.7.2 eEFSec

eEFSec es un factor de elongación de la traducción especializado en la síntesis de selenoproteínas y su función es posicionar al tRNA-Sec en el sitio A ribosomal para la incorporación de Sec a las proteínas nacientes (Figura 7B). Su dominio N-terminal tiene alta similitud con el factor de elongación canónico EF1A y, al igual que éste, posee actividad de GTPasa. A diferencia de EF1A, eEFSec posee alta afinidad por el tRNA-Sec aminoacilado, pero no se une a otros tRNAs. Por ensayos de co-inmunoprecipitación se ha demostrado que eEFSec, a través de su extremo C-terminal, interactúa con la proteína SBP2, requiriéndose al tRNA-Sec para la formación de este complejo (Allmang, C. 2009 y 2006a).

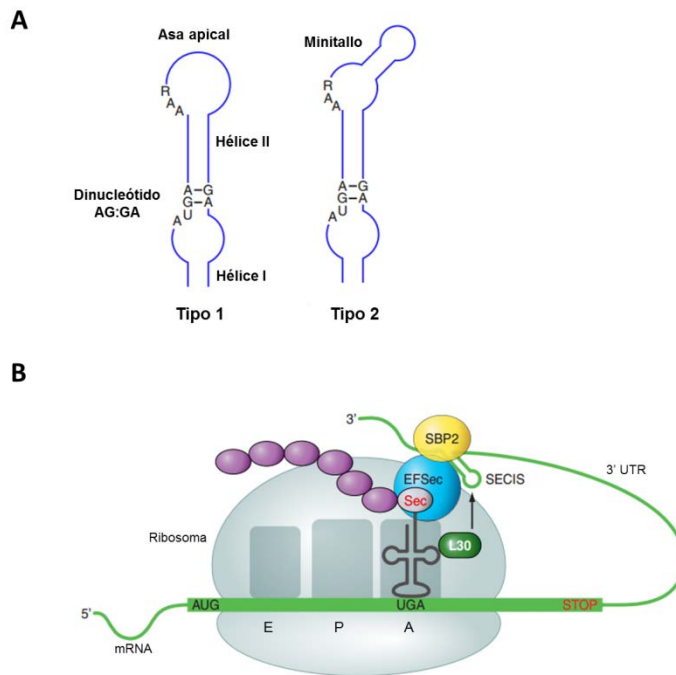


Figura 7. Maquinaria de inserción de selenocisteína. A) Estructura secundaria del elemento SECIS. Se muestran los dos tipos de elementos SECIS identificados, B) Maquinaria esencial para la inserción de Sec en las selenoproteínas. El elemento SECIS se localiza en la región 3' no traducida (3'-UTR) del mRNA. Sec se inserta en respuesta al codón UGA. Se muestran los factores requeridos para la incorporación de Sec: tRNA-Sec, el factor eEFSec, y las proteínas SBP2 y L30 (Tomado de Labunskyy, V.M., *et al.*, 2014). E, sitio de salida; P, sitio peptidil; A, sitio aminoacil.

2.7.3 Proteína SBP2 (SECIS binding protein 2)

SBP2 posee tres dominios: el dominio N-terminal, cuya función a la fecha se desconoce, un dominio de incorporación a Sec, y un dominio C-terminal de unión a RNA (conocido como RBD). El RBD pertenece a una familia de proteínas de unión a RNAs con motivos *kink-turn* (como el rRNA 28S y el elemento SECIS). A través de este dominio, se ha detectado una interacción con el ribosoma, específicamente con la subunidad 60S a través del rRNA 28S. La función de SBP2 es la unión al elemento SECIS y estimular la incorporación de Sec a las selenoproteínas al reclutar al complejo eEFSec/tRNA-Sec al ribosoma, a través de su dominio de unión a eEFSec (Figura 7B). Aunque SBP2 puede unirse al rRNA 28S y al elemento SECIS, no puede hacerlo al mismo tiempo, por lo que se ha propuesto un modelo competitivo en el que SBP2 se une sólo a uno de ellos en un momento determinado (Labunskyy, V.M., et al., 2014).

SBP2 es también parte de los mecanismos de regulación de la síntesis de selenoproteínas, pues se une al elemento SECIS de los mRNA de selenoproteínas de manera diferencial, lo cual repercute en una traducción diferencial de las mismas (Allmang, C., et al., 2009).

2.7.4 Proteína ribosomal L30

La proteína ribosomal L30 es un componente de la subunidad mayor ribosomal (60S), e interacciona con el elemento SECIS (Figura 7B). L30 compite con SBP2 por dicho elemento, es decir, el elemento SECIS puede unirse a L30 o a SBP2 pero no a las dos al mismo tiempo. Se ha observado que las sales de magnesio (las cuales inducen la conformación *kink-turn* en SECIS) favorecen la estabilidad del complejo L30/SECIS y disminuyen la de SBP2/SECIS, lo que sugiere que estas proteínas tienen diferentes especificidades de unión a SECIS (Gonzalez-Flores, J., et al., 2013). L30 interactúa con SECIS con mucha mayor afinidad cuando está asociada a los ribosomas, en comparación con la proteína L30 libre, lo cual sugiere que L30 adopta una conformación más favorable para la unión a SECIS cuando está formando parte del ribosoma. Se ha sugerido que L30 desplaza transitoriamente a SBP2 para llevar al elemento SECIS a la proximidad del sitio A del ribosoma (Allmang, C., et al., 2009).

2.8 Modelos de inserción de Selenocisteína a las Selenoproteínas

Se ha propuesto un par de modelos para explicar la inserción de Sec en respuesta al codón UGA. En el primer modelo, la proteína SBP2 está unida permanentemente a un grupo de ribosomas que participarán en la biosíntesis de selenoproteínas. Estos ribosomas se posicionan sobre el codón UGA al tiempo que el elemento SECIS se repliega hacia el ribosoma, lo que provoca la separación de SBP2 del ribosoma y su unión al elemento SECIS. Estos movimientos favorecen un cambio conformacional en el sitio A del ribosoma, permitiendo la entrada del complejo eEFSec/tRNA-Sec. Posteriormente, la proteína ribosomal L30 desplaza a SBP2 del elemento SECIS, lo que permite el regreso de SBP2 al ribosoma (Figura 8) (Allmang, C. y Krol, A., 2006a; Kossinova, O., *et al.*, 2013).

En el segundo modelo, SBP2 se une directamente al elemento SECIS, lo cual sirve como plataforma para reclutar al complejo eEFSec/tRNA-Sec, el cual ingresa al sitio A del ribosoma. Posteriormente, la proteína L30 desplaza a SBP2 del SECIS, causando cambios conformacionales en este último elemento que inducen la liberación del complejo eEFSec/tRNA-Sec/GTP del ribosoma (Figura 8). La función de la proteína L30 aún no está del todo clara, pero podría funcionar como un puente entre el ribosoma y el elemento SECIS, o bien, L30 podría desplazar a SBP2 y permanecer unida al elemento SECIS. Este último movimiento favorecería la liberación del tRNA-Sec y la hidrólisis de GTP (Allmang, C. y Krol, A., 2006a; Kossinova, O., *et al.* 2013).

Recientemente se ha observado que, en reticulocitos de conejo, el elemento SECIS ya se encuentra asociado a SBP2 antes de la formación del complejo de inicio de la traducción 48S, evidencia que apoyaría el segundo modelo (Kossinova, O., *et al.*, 2013). Asimismo, se observó que SBP2 permanece unida al ribosoma tras el acomodo del tRNA-Sec al sitio A ribosomal y se disocia durante el paso de transpeptidación; y finalmente, que el elemento SECIS contacta a la subunidad ribosomal 60S después del acomodo del tRNA en el sitio A. Se propone que la incorporación del residuo Sec en la cadena polipeptídica está gobernada por rearrreglos conformacionales tanto en el SECIS como en el rRNA 28S (Kossinova, O., *et al.*, 2013).

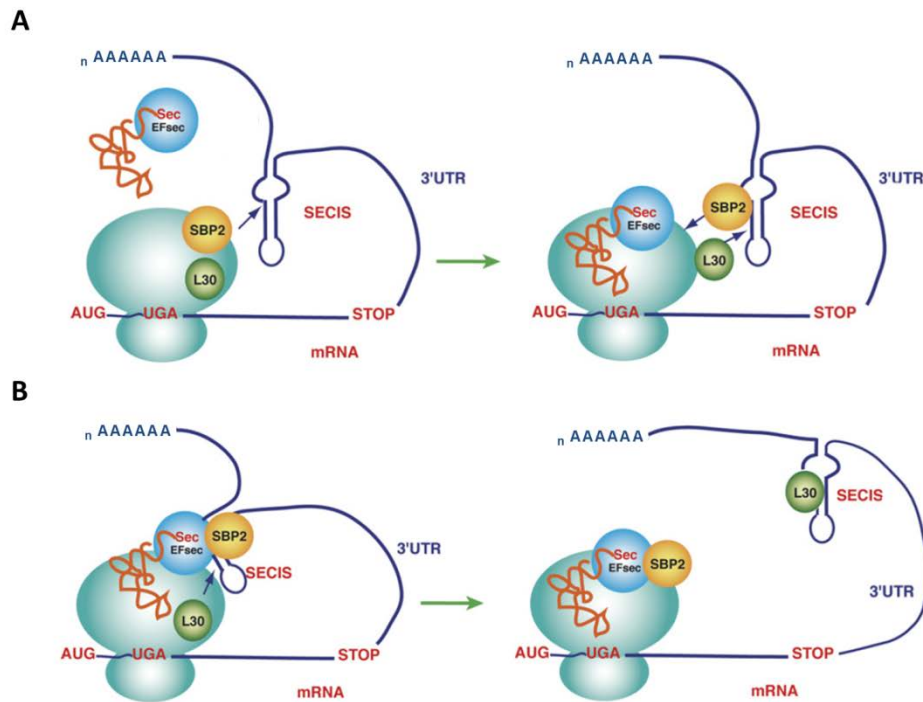


Figura 8. Modelos de inserción de selenocisteína a las selenoproteínas. En el primer modelo (A), SBP2 está unida a ribosomas y posteriormente se une al elemento SECIS (pero no a ambos). En el segundo modelo (B) SBP2 se une directamente al elemento SECIS desde el inicio. Ver texto para detalles (Tomada de Allmang, C. y Krol, A., 2006a).

2.9 tRNA de Selenocisteína

El tRNA-Sec es el encargado de transportar el residuo Sec hasta el ribosoma durante la síntesis de proteínas. Este tRNA difiere en algunos aspectos respecto a los tRNA convencionales: a nivel de secuencia, el tRNA-Sec es más largo (88-100 nucleótidos) que los tRNA clásicos (73-83 nucleótidos), pues posee los brazos D, T y variable más extensos (Figura 9). Asimismo, el brazo variable es más largo que en el resto de los tRNA. Como resultado, el tRNA-Sec posee una estructura secundaria denominada 9/4, ya que el brazo aceptor tiene 9 nt y el brazo T tiene 4 nt. En contraste, los tRNA convencionales tienen una estructura 7/5. El brazo D del tRNA-Sec está compuesto por un tallo de 6 nt y un asa de 4 nt, mientras que otros tRNA, el tallo tiene 4 nt y el asa 8 nt. El tRNA-Sec posee el anticodón UCA, complementario al codón UGA, que codifica para selenocisteína (Figura 9) (Hatfield, D. y Gladishev, V., 2002). Otra característica inusual del tRNA-Sec es que su transcripción comienza en

el primer nucleótido de su secuencia codificante (Lee, B.J., *et al.*, 1987) mientras que en los tRNA canónicos, ésta empieza varios nucleótidos río arriba, generando una secuencia 5' que posteriormente será procesada y eliminada (Orioli, A., *et al.*, 2012). En la mayoría de los eucariontes en los que se ha identificado el tRNA-Sec se encuentra una sola copia del gen, excepto en el pez cebra, organismos que posee dos copias (Labunsky, V.M., *et al.*, 2014).

Los tRNA poseen diversas modificaciones post-transcripcionales a lo largo de su estructura, las cuales ya han sido ampliamente estudiadas en una gran variedad de organismos. En eucariontes, un tRNA convencional tiene entre 7 y 17 nucleótidos modificados (Phizicky, E.M. y Alfonzo, J.D., 2010). Por su parte, el tRNA-Sec contiene sólo 4 bases modificadas: una 1-metiladenosina en la posición 58 (m¹A58), una pseudouridina en la posición 55 (ψ55), una N⁶-isopenteniladenosina en posición 37 (i⁶A37) y una 5-metilcarboximetiluridina (mcm⁵U) en posición 34 (Figura 9) (Itoh, Y., *et al.*, 2009; Kim, L.K., *et al.*, 2000). Este número bajo de modificaciones post-transcripcionales en comparación con los tRNA convencionales es otra característica atípica del tRNA-Sec (Allmang, C. y Krol, A., 2006a).

En células humanas y de ratón, se han detectado una isoforma del tRNA-Sec que posee un grupo 2'-O-metil unido a la ribosa de la mcm⁵U en la posición 34 (5-metilcarboximetiluridina-2'-O-metilribosa, mcm⁵Um). De manera interesante, la formación de mcm⁵Um está en función de la ingesta de selenio en ratas (Diamond, A.M., *et al.*, 1993). Asimismo, se ha demostrado que esta metilación de la ribosa en la mcm⁵U influye en la estabilidad de las estructuras secundaria y terciaria del tRNA-Sec. Además, aunque las formas metilada y no metilada de la mcm⁵U coexisten en la célula, sus niveles intracelulares varían entre diferentes tejidos y líneas celulares (Hatfield, D. y Gladishev, V., 2002; Kim, L.K., *et al.*, 2000). De hecho, ya se ha demostrado que estas isoformas modulan la expresión de diferentes selenoproteínas (Carlson, B.A., *et al.*, 2005). Aquellas selenoproteínas relacionadas con el estrés celular, como la glutatión peroxidasa (GPx) 1 y 3, SelR, SelT y SelW se sintetizan por el tRNA-Sec que posee mcm⁵Um. Por su parte, la presencia de mcm⁵U favorece la expresión de selenoproteínas cuyas funciones son esenciales para la supervivencia celular, por ejemplo, la tioredoxina reductasa (TrxR) 1 y 3. Sin embargo, existen selenoproteínas, como GPx4 y SelP, que parecen ser sintetizadas por ambas isoformas del tRNA-Sec (Carlson, B.A., *et al.*, 2007 y 2005).

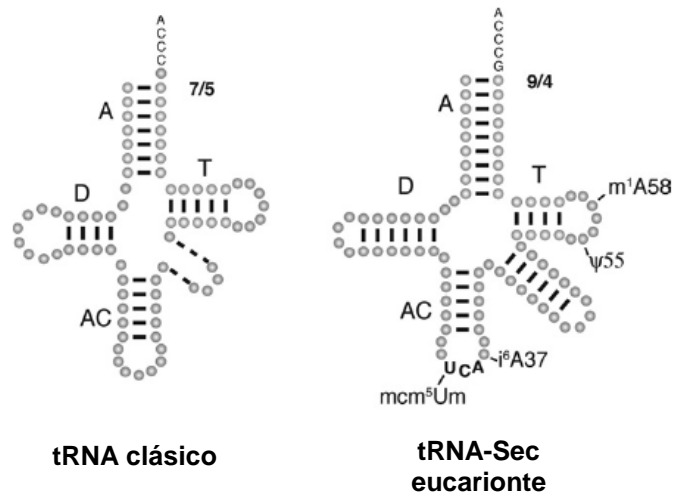


Figura 9. Particularidades estructurales del tRNA-Sec. Comparación de estructura secundaria de un tRNA convencional y el tRNA-Sec. A, brazo del aminoácido; D, brazo D; AC, brazo del anticodón; T, brazo T. 7/5 y 9/4 indican el número de pares de bases que van del brazo del aminoácido al brazo T. El brazo V en el tRNA clásico (línea punteada) es de longitud variable (3-4 nt), mientras que en el tRNA-Sec siempre es largo (~ 7 nt). En el tRNA-Sec se señalan las modificaciones post-transcripcionales posibles (Tomado de Allmang, C. y Krol, A., 2006).

2.10 Transcripción y regiones promotoras del tRNA-Sec

En *Xenopus laevis* se observó que el tRNA-Sec es transcrito por la RNA Pol III, lo cual se determinó mediante ensayos de transcripción en presencia de α -amanitina (Park, J.M., *et al.*, 1996; Lee, B.J., *et al.*, 1989). Asimismo, en células HeLa (células humanas de adenocarcinoma de cérvix) se ha comprobado mediante ensayos de inmunoprecipitación de la cromatina, la presencia de la RNA Polimerasa III y el factor Brf2 en el *locus* del tRNA-Sec (Faresse, N.J., *et al.*, 2012).

La región promotora del tRNA-Sec posee elementos intragénicos y extragénicos, por lo que es considerado como un promotor híbrido entre los promotores tipo 2 (que poseen sólo elementos intragénicos) y los tipo 3 (que poseen únicamente elementos extragénicos) (Figura 10). Así, río arriba del gen de tRNA-Sec se localizan 3 elementos: un Elemento Proximal (PSE, *Proximal Sequence Element*), una caja TATA y un Elemento Activador (AE, *Activator Element*). Dentro de la secuencia codificante, la caja B parece ser la única que participa en la transcripción (Hatfield, D. y Gladishev, V., 2002).

Por experimentos de transcripción *in vitro* e *in vivo* realizados en *X. laevis* se demostró que las cajas A y B por sí mismas no activan la transcripción y que son los elementos PSE y TATA los que más contribuyen a esta actividad, pues al mutar una de estas secuencias la transcripción disminuye considerablemente y si son eliminadas, la transcripción se abate por completo. El elemento PSE se encontró en la posición -72 a -48 y la caja TATA de -32 a -25 (Figura 10). Asimismo, se observó que la caja B tiene un papel activo en la transcripción del tRNA-Sec, pues la sustituciones en la secuencia conlleva a una reducción de ~5 veces en el nivel de transcripción, resultado que sugiere que la caja B coopera con los elementos PSE y TATA para maximizar la transcripción. La eliminación completa de la caja B resulta en la obtención de transcritos de diferentes tamaños. La caja B del tRNA-Sec de *X. laevis* si tiene una secuencia consenso (Carbon, P. y Krol, A. 1991; Lee, B.J., et al., 1989).

Como se ha mencionado anteriormente, la caja A es uno de los elementos promotores intragénicos que dirigen la transcripción en los tRNA clásicos. En el tRNA-Sec de *X. laevis* se encontró que la secuencia de la caja A está modificada respecto a la secuencia consenso (altamente conservada) por la inserción de dos nucleótidos. Mediante ensayos de mutagénesis y transcripción *in vitro*, se observó que la eliminación completa de la caja A no altera el nivel de transcripción, el cual permanece similar a la tasa de transcripción con las construcciones que poseen la secuencia silvestre del tRNA-Sec (Lee, B. J., et al., 1989).

El elemento AE fue mapeado por mutagénesis y transcripción *in vivo* en *X. laevis* y fue localizado de -209 a -195 respecto al sitio de inicio de la transcripción. En el modelo del snRNA U6, el elemento equivalente es el DSE (Figura 10) el cual funciona como un elemento activador que posee convencionalmente una secuencia octamérica ATGCAAAT. Sin embargo, una característica del AE es la ausencia de dicho octámero, lo que lo hace diferente a los DSE clásicos. Este elemento AE contiene un motivo conocido como SPH (*SphI* Postoctamer Homology) con la secuencia GCATGC, la cual también se encuentra típicamente en los DSE de los snRNA. Al introducir mutaciones en el AE la transcripción disminuye considerablemente, aunque no se abate del todo, por lo que se ha considerado que el AE sirve para maximizar la expresión del tRNA-Sec (Myslinski, E., et al., 1991).

En ratón, este elemento sí se ha encontrado como un DSE clásico, el cual posee el motivo SPH y el octámero. Por estudios de transfección en células NIH 3T3 (fibroblastos embrionarios de ratón) se observó que la expresión del tRNA-Sec disminuía hasta en un 38% al mutar o eliminar la secuencia del

DSE (Kelly, V.P., et al., 2005). Por otra parte, en dicho estudio se reportó también que la introducción de un fragmento de 3.2 kb entre los elementos DSE y PSE en células madre embrionarias es letal en organismos homocigotos debido a la pérdida de la transcripción del tRNA-Sec. En individuos heterocigotos, los niveles del tRNA-Sec disminuyeron diferencialmente en varios tejidos: su expresión en células de riñón bajo 50% respecto a los ratones silvestres, mientras que en células de corazón no hubo disminución. Estas observaciones sugirieron que el DSE regula la transcripción del tRNA-Sec de manera tejido-específica y que su actividad es vital para la producción de dicho tRNA durante la embriogénesis de ratón (Kelly, V.P., et al., 2005).

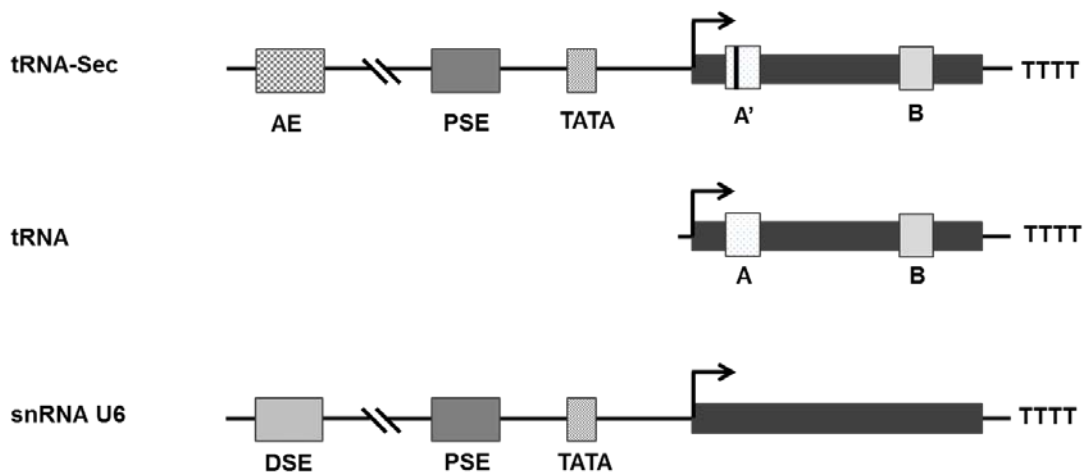


Figura 10. Particularidades del promotor del tRNA-Sec. Elementos promotores de tRNA-Sec (elementos intragénicos y extragénicos) caracterizados en vertebrados en comparación con los tRNA canónicos (elementos intragénicos) y el snRNA U6 (elementos extragénicos).

2.11 Selenoproteínas en Tripanosomátidos

A la fecha ya se reportó la presencia de selenoproteínas y su maquinaria de biosíntesis en tripanosomátidos (Cassago, A., et al., 2006; Lobanov, A., et al., 2006). El gen de tRNA-Sec en estos organismos se identificó originalmente por búsquedas en bases de datos, y posteriormente se confirmó su expresión por ensayos de RT-PCR en *T. brucei*, lo que fue el primer indicio de que la maquinaria de biosíntesis de selenoproteínas está activa en estos parásitos (Lobanov, A., et al., 2006). En *T. brucei*, el tRNA-Sec se expresa en parásitos procíclicos y sanguíneos, lo cual sugiere que este tRNA está activo durante todo el ciclo de vida (Cassago, A., et al., 2006). En varias especies de

Leishmania, incluyendo *L. major*, existe una copia génica del tRNA-Sec en el cromosoma 6, mientras que *T. brucei* y *T. cruzi* poseen 2 copias en los cromosomas 9 y 33, respectivamente (Padilla-Mejía, N.E., et al, 2009). Esto es similar a otros eucariontes, que poseen generalmente 1 copia de este tRNA en su genoma.

De las selenoproteínas identificadas en vertebrados (Tabla 4), los ortólogos de SelK y SelT se identificaron en tripanosomátidos. En los genomas de varias especies de *Leishmania* y *Trypanosoma* se identificó una selenoproteína de probable localización mitocondrial, SelTryp. Por experimentos de incorporación *in vivo* de ⁷⁵Se a proteínas, se observaron tres diferentes proteínas con selenio incorporado en *T. cruzi*, presumiblemente SelK, SelT y SelTryp (Lobanov, A., et al., 2006). Según predicciones computacionales, en esas tres selenoproteínas los residuos Sec se localizan dentro de motivos redox. También mediante herramientas bioinformáticas, se identificaron las selenoproteínas LmSel1 y LinSel1, en *L. major* y *L. infantum*, respectivamente, cuya expresión se comprobó por RT-PCR (Cassago, A., et al., 2006), así como SelX, selenoproteína putativa exclusiva de *T. brucei* (Lobanov, A., et al., 2006). En los mRNA de estas selenoproteínas se han identificado los elementos SECIS respectivos, los cuales poseen características similares a las reportadas en otros eucariontes: una estructura en tallo-asa, la presencia del dinucleótido AG:GA y un par AA en el asa apical (Figura 7A) (Cassago, A., et al., 2006; Lobanov, A., et al., 2006).

Otros componentes de la maquinaria de síntesis de selenoproteínas han sido identificados en estos parásitos: SPS, PSTK, SerRS (proteínas involucradas en la síntesis de Sec sobre el tRNA-Sec), Secp43 (proteína involucrada en la traducción de las selenoproteínas) y el factor de elongación específico eEFSec (Cassago, A. et al., 2006; Lobanov, A., et al., 2006). La participación de eEFSec como factor activo en la síntesis de selenoproteínas en tripanosomátidos se demostró por ensayos de RNA de interferencia (RNAi) en *T. brucei*: al activar el sistema de RNAi para eEFSec, las selenoproteínas SelK, SelT y SelTryp dejaron de expresarse gradualmente. Este efecto fue también observado al inducir el sistema de RNAi de SPS2 (Aeby, E., et al., 2009). Mediante experimentos de RNAi también se demostró que SerRS es esencial para el crecimiento normal de *T. brucei*. Esta enzima posee un dominio específico de tripanosomátidos que no se encuentra en SerRS de otros eucariontes (Geslain, R., et al., 2006). En *L. major*, la proteína SPS ya ha sido clonada y caracterizada. Se comprobó que ésta se expresa a lo largo de todo el ciclo de vida y que participa en la biosíntesis de Sec a partir de fosfatos inorgánicos y de selenio, al igual que en otros eucariontes (Sculaccio, S.A., et al., 2008).

3. ANTECEDENTES

La transcripción en parásitos tripanosomátidos muestra variaciones respecto a lo que ocurre en la mayoría de los eucariontes superiores. Entre estos procesos atípicos se encuentran la transcripción policistrónica; la carencia de promotores clásicos de la RNA Pol II y la función de genes de tRNA como promotores de los snRNA, entre otros (Martínez-Calvillo, S., et al. 2010).

En un trabajo previo publicado por nuestro grupo de trabajo (Padilla-Mejía, N.E., et al., 2009) se reportó que existe una copia única del gen del tRNA-Sec en el cromosoma 6 de *L. major* y que este gen se localiza dentro de una unidad policistrónica, flanqueado por los genes de proteínas LmjF.06.0210 y LmjF.06.0200 (Figura 11A), cuyos productos están identificados como proteínas hipotéticas en GeneDB. El análisis *in silico* de las secuencias promotoras del tRNA-Sec, las cajas típicas de los tRNA A y B, reveló que la caja A contiene una base extra en comparación a la secuencia consenso de tripanosomátidos (Figura 11B y C) (Padilla-Mejía, N.E., et al., 2009). En otros eucariontes se ha reportado la inserción de dos bases respecto a la secuencia consenso de vertebrados (Lee, B. J., et al., 1989). Por su parte, la caja B, la cual no muestra variaciones respecto a la secuencia consenso en otros eucariontes, en *L. major* sí difiere en la primera y última posición de la secuencia (Figura 11D y E). Ya que ambas cajas promotoras poseen cambios respecto a la secuencia consenso típica de los tRNA, se consideró que secuencias externas al gen podrían participar en dirigir su transcripción (Padilla-Mejía, N.E., et al., 2009), tal y como ocurre para los snRNAs o el RNA 7SL en tripanosomátidos. En *T. brucei*, ha sido reportado que el tRNA-Sec es transcrito por la RNA Polimerasa II (Aeby, E., et al., 2010). En dicho trabajo se reportó que el tRNA-Sec sufre la adición del miniexón (o RNA *Spliced-Leader*, SL-RNA) en la región 5', lo cual se determinó a partir de bibliotecas de cDNA poliadenilados. Esto sugirió que el tRNA-Sec es transcrito como parte de un precursor policistrónico y que, posteriormente, es procesado por *trans-splicing* y poliadenilación, tal como ocurre con los mRNA.

Nuestro laboratorio está interesado en el estudio de los mecanismos de transcripción por la RNA Pol III en *L. major*. A la fecha no se sabe cuál es la RNA Pol que transcribe el tRNA-Sec en este organismo, ¿este tRNA es transcrito por la RNA Pol II como en *T. brucei*?, ¿o es transcrito por la RNA Pol III como ocurre con los tRNA-Sec en eucariontes superiores?, ¿existen mecanismos de transcripción atípicos relacionados con el tRNA-Sec en *L. major*? Estos cuestionamientos constituyen las principales preguntas del presente trabajo.

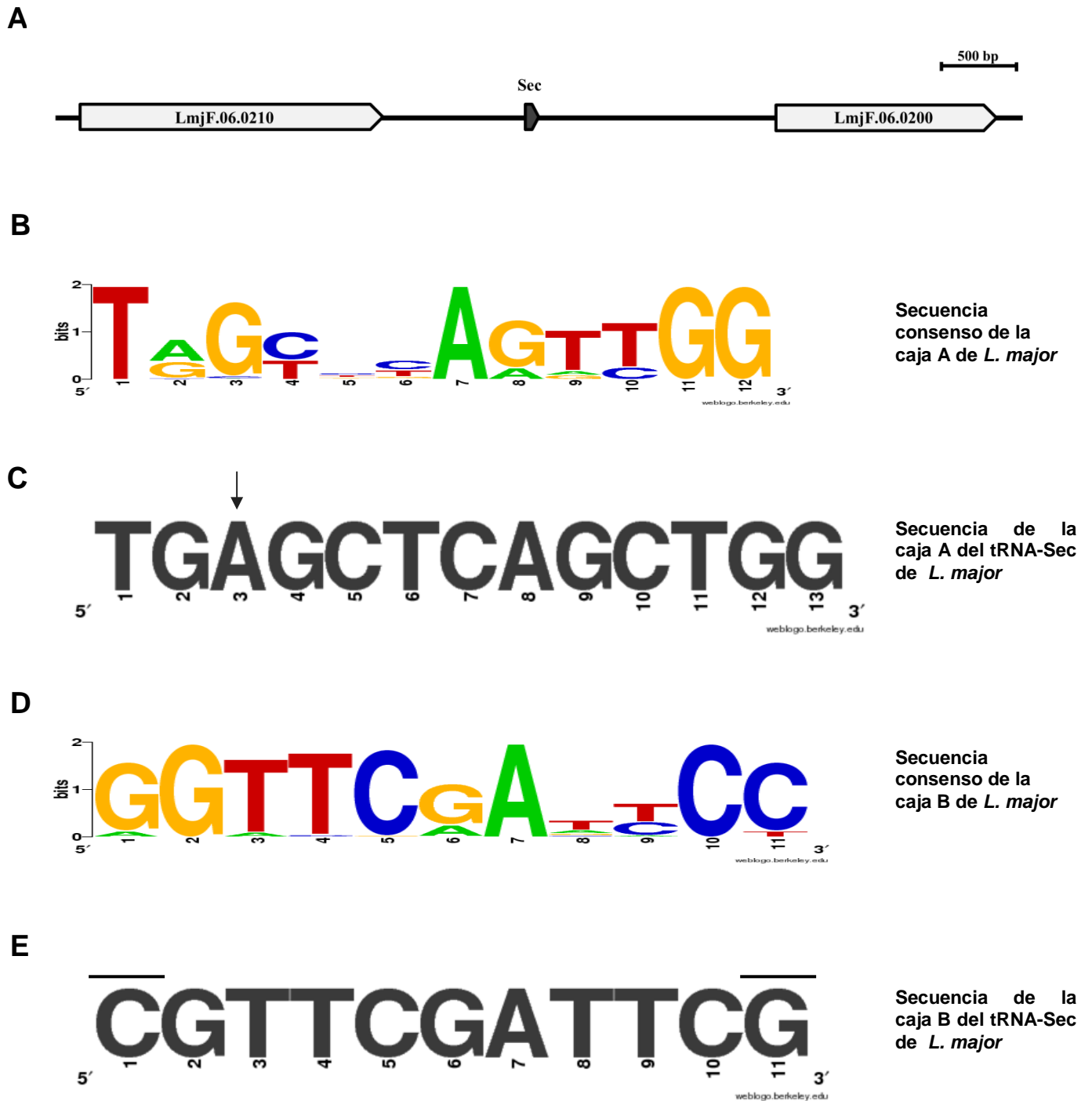


Figura 11. Particularidades del gen de tRNA-Sec en *L. major*. A) Localización genómica del tRNA-Sec, se ubica en el cromosoma 6 de *L. major*, entre dos genes codificadores de proteínas: LmjF.06.0210 y LmjF.06.0200; B) Secuencia consenso de la caja A de los tRNA en *L. major*; C) Secuencia de la caja A del tRNA-Sec, se señala con una flecha la base extra respecto a la secuencia consenso; D) Secuencia consenso de la caja B de los tRNA en *L. major*; E) Secuencia de la caja B del tRNA-Sec, las bases que varían respecto a la secuencia consenso se señalan con una barra superpuesta (Padilla-Mejía, N.E., et al., 2009). En B y D, la altura de la base indica que tan conservado está esa base, cuando dos bases ocupan una misma posición, su altura representa la frecuencia dentro de esa posición (Crooks, G.E., et al. 2004).

4. OBJETIVOS

4.1 Objetivo General

Identificar cuál es la RNA Polimerasa que transcribe el gen de tRNA de selenocisteína en *L. major*

4.2 Objetivos Particulares

- Conocer el contexto genómico del tRNA-Sec y determinar si este gen se encuentra conservado entre especies de *Leishmania* y *Trypanosoma*.
- Identificar los transcritos del tRNA-Sec y determinar sus sitios de procesamiento en *L. major* y *T. cruzi*.
- Evaluar si el gen de tRNA-Sec de *L. major* es transcrito como parte de un policistrón.
- Evaluar si la tasa transcripcional del tRNA-Sec de *L. major* es similar a la de genes transcritos por la RNA Pol II o a la de los transcritos por la RNA Pol III.
- Evaluar si el gen de tRNA-Sec de *L. major* posee su propia región promotora
- Determinar si la transcripción del gen de tRNA-Sec es sensible a la α -amanitina y a la tagetitoxina
- Determinar si la RNA Polimerasa II o la RNA Polimerasa III sintetizan el tRNA-Sec en *L. major*

5. METODOLOGÍA

5.1. Análisis *in silico*

Los análisis bioinformáticos se realizaron utilizando las bases de datos de los genomas de las especies de *Leishmania* (*L. major*, *L. donovani*, *L. infantum*, *L. braziliensis*, *L. mexicana* y *L. tarentolae*) y *Trypanosoma* (*T. cruzi*, *T. brucei*, *T. vivax* y *T. congolense*) disponibles en línea en los sitios de www.genedb.org y www.tritrypdb.org (versión 8.1). Las comparaciones de secuencias se llevaron a cabo con el programa ClustalW2 (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>).

5.2. Cultivos de *L. major* y *T. cruzi*

Los promastigotes de *L. major* MHOM/IL/81/Friedlin (LSB-132.1) se crecieron a 26°C en medio BM (medio M199 pH 7.2 suplementado con suero fetal bovino 10%, infusión de cerebro - corazón 0.25x, HEPES 40 mM, hemina 0.01 mg/ml, biotina 0.0002 %, penicilina 100 IU/ml / estreptomicina 100 µg/ml y L-glutamina 1x). Los cultivos se recuperaban en fase media logarítmica.

Epimastigotes de *T. cruzi* CL Brener se mantuvieron a 28°C en medio LIT (*liver infusion-tryptose*) suplementado con suero fetal bovino 10%, penicilina 50 IU/ml / estreptomicina 50 µg/ml y hemina 0.025 mg/ml (Florencio-Martínez, L.E., et al., 2010)

5.3. Análisis 5´-RACE (5´ *Rapid Amplification of cDNA Ends*)

Para el ensayo 5´-RACE se utilizó un kit de *Life Technologies* (18374-058) con 5 µg de RNA total de *L. major* o *T. cruzi*. Las secuencias de los oligonucleótidos empleados se enlistan en la Tabla 5.

El cDNA fue amplificado utilizando los oligonucleótidos tRNA-SECgsp1 (*L. major*) y Tb-tRNA^{Sec}-GSP1 (*T. cruzi*). Posteriormente, el cDNA proveniente de los genes de *L. major* se sometió a una amplificación por PCR, utilizando los oligonucleótidos tRNA-SECgsp2 y el *Abridged Anchor Primer* (AAP). Para el gen de tRNA-^{Sec} de *T. cruzi*, se realizaron dos amplificaciones por PCR. La primera amplificación se realizó con los oligonucleótidos ME23 y Tb-tRNA^{Sec}-GSP1, mientras que la segunda se realizó utilizando el producto de la primer PCR como molde y utilizando los oligonucleótidos Tc-tRNA^{Sec}-GSP2 y ME23.

Para el gen de tRNA-^{Asp} de *L. major*, el cDNA se sintetizó con el oligonucleótido Lm24Asp-GSP1 y las amplificaciones por PCR se realizaron con el Lm24Asp-GSP2 y el AAP. Para el gen tRNA-^{Pro} de *L. major*, el cDNA se sintetizó con el oligonucleótido Lmc24-ProGSP1 y las amplificaciones por PCR se realizaron con los oligonucleótidos Lmc24-ProGSP2 y el AAP.

En todos los casos los productos de PCR se clonaron en el vector pGEM-T Easy (Promega) y se secuenciaron.

5.4. Mapeo de los sitios de poliadenilación y los sitios de adición del miniexón

Se realizaron ensayos de RT-PCR para mapear los sitios de poliadenilación para los tRNA-^{Sec} de *L. major* y *T. cruzi*, así como del gen LmjF.06.0210 de *L. major*, utilizando un cDNA preparado con el oligonucleótido Nested(dT) (Tabla 5). Para el gen de tRNA-^{Sec} de *L. major* y *T. cruzi* se realizó primero una PCR con los oligonucleótidos Lm-^{Sec}1 y B1. La segunda PCR se realizó con los oligonucleótidos Lm-^{Sec}2 y B2. Para el gen LmjF.06.0210, la primer PCR se realizó con los oligonucleótidos Lmj06.0210PA1 y B1, y la segunda PCR con los oligonucleótidos Lmj06.0210PA2 y B2.

El sitio de adición del miniexón para el gen LmjF.06.0200 de *L. major* se localizó también por RT-PCR. El cDNA se sintetizó con el oligonucleótido Lmj06.0200.ME1. Para la reacción de PCR se utilizaron los oligonucleótidos Lmj06.0200.ME2 y Miniexón. En todos los casos los productos de PCR se clonaron en el vector pGEM-T Easy (Promega) y se secuenciaron.

Tabla 5. Oligonucleótidos utilizados en los ensayos de 5'-, 3' – RACE y Northern blot

Nombre del oligonucleótido	Secuencia del oligonucleótido
AAP (Abridged Anchor Primer)	5'-GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG
B1	5'-CCTCTGAAGGTTACGGAT
B2	5'-CACGGATCCACATCTAGAT
Lm24Asp-GSP1	5'-CCGGCCGGGAATTGAAC
Lm24Asp-GSP2	5'-GGGTCACCCGCGTGACAGGC
Lmc24-ProGSP1	5'-GGGCCGCTAGGGGAATTGAA
Lmc24-ProGSP2	5'-TGACCTCCCGCACCCGAAG
Lm-Sec1	5'-AGCCGCGATGAGCTCAGCT
Lm-Sec2	5'-TGGGTGCGGGCTTCAA
Lmj06.0200.ME1	5'- AGAGCGACACCCGTGACTTC
Lmj06.0200ME2	5'- ACGGAACCCAGAACGCAGGA
Lmj06.0210PA1	5'- AGCCGACTCATACTGCGGCT
Lmj06.0210PA2	5'- CTCATGCACITTAAGCTGTA
Miniexón	5'-AACGCTATATAAGTATCAGTT
ME23	5'-CGCTATTATTGATACAGTTTCTG
Nested dT	5'CCTCTGAAGGTTACGGATCCACATCTAGATTTTTTTTTTTTTTTTTVN
Tb-tRNASec-GSP1	5'-CACCACAAAGGCCGA
Tc-tRNASec-GSP2	5'-AACGGCTGCGAGTCCAAC
tRNA-SECgsp1	5'-TGGCACGCCACGAAG
tRNA-SECgsp2	5'-ATCGAACGGCTGTGAGAGCA
TRNA-Lysc03-GSP1	5'-GCGCACTCCGTGGGG

V = G, C, A

N = A, C, G, T

I = Inosina

5.5. Ensayo de Northern-blot

El RNA total se aisló utilizando el reactivo TRI (Sigma), bajo las instrucciones del fabricante. Después, 12 µg de RNA se corrieron en un gel de poliacrilamida 10%/urea 8 M en TBE 1x a 120 V. Tras la electroforesis, los ácidos nucleicos se transfirieron a membranas Hybond N⁺ (Amersham) utilizando el sistema *Trans-Blot Semy Dry System* (Biorad) a 25V por 1 hora. Tras la transferencia se fijó el RNA con luz UV en el equipo Stratalinker.

Para el tRNA-Sec: las membranas se hibridaron con fragmentos de DNA marcados con [α -³²P]dCTP con el kit *High Prime DNA Labeling* (Roche). La sonda correspondió a un fragmento de 190 pb proveniente del vector pRACE225, el cual incluye la secuencia genómica del tRNA-Sec de 88 pb más 86 pb de la región 5'. El fragmento que contiene el tRNA-Sec se liberó del vector pGEM-T Easy utilizando la enzima de restricción *EcoRI*, y la banda correspondiente se purificó por el kit *NucleoSpin Gel Extraction* (Macherey-Nagel) siguiendo las instrucciones del fabricante. Para la reacción de marcaje se utilizaron 100 ng de DNA, el cual se desnaturalizó calentándolo a ebullición por 10 minutos (min) y se enfrió rápidamente en hielo. Se centrifugó la muestra y se agregaron 4 µl de mezcla de High Prime 5x (que contiene 4U de Polimerasa Klenow y Buffer de reacción 5x), 1 µl de dATP, 1 µl de dGTP, 1 µl de dTTP (0.5 mM) y 5 µl [α -³²P]dATP (3000 Ci/mmol, 10 µCi/µl). El volumen total de la reacción fue de 20 µl. La reacción se incubó a 37°C, 10 min y se detuvo agregando 2 µl de EDTA 0.2 M (pH 8.0) y calentando a 65°C por 10 min. La sonda se purificó con las columnas de sefadex G-50 (GE Healthcare) siguiendo las instrucciones del fabricante.

Para el tRNA-Lys: como sonda se utilizó el oligonucleótido TRNA-Lysc03-GSP1 (Tabla 5), el cual se marcó en su extremo 5' con [γ -³²P]dATP con la enzima T4 cinasa. Para la reacción de marcaje se utilizaron 50 pmoles del oligonucleótido mencionado, buffer de cinasa 1x, 5 µl de [γ -³²P]dATP (3000 Ci/mmol, 10 µCi/µl), 20 U de T4 cinasa (Promega). La reacción se incubó a 37°C por 10 min y se detuvo agregando 2 µl de EDTA 0.5 M. Se recuperó la fase acuosa y la sonda se purificó utilizando columnas de sefadex G-50 (GE Healthcare) siguiendo las instrucciones del fabricante.

Las hibridaciones se realizaron en solución SSPE 6x (Na_2HPO_4 60 mM, NaCl 0.9 M y EDTA 6 mM), solución Denhardt 5x y SDS 1% a 42°C por 18 horas con 5×10^6 cpm de la sonda ya mencionada. Los lavados se llevaron a cabo en SSPE 2x/SDS 0.5% a 55°C. Las imágenes se revelaron en un equipo Fuji-Phosphoimager FLA 5000. Composición de la solución Denhardt 50x: Ficoll 1%, Polivinilpirrolidona (PVP) 1%, Albúmina Sérica Bovina (BSA)1%.

5.6. Clonación de fragmentos en pGEM-T Easy

Varios fragmentos de DNA del cromosoma 6 de *L. major* se amplificaron por PCR y se clonaron en el vector pGEM-T Easy (Promega) (Apéndice II). La secuencia de cada oligonucleótido empleado se encuentra en la Tabla 6. El gen LmjF.06.0370 se amplificó con los oligonucleótidos Lm06-0370-5' y Lm06-0370-3'; el gen LmjF.06.0360 con los oligonucleótidos Lm06-0360-5' y Lm06-0360-3'. El gen LmjF.06.0350 se amplificó con los oligonucleótidos Lm06-0350-5' y Lm06-0350-3'; y el gen LmjF.06.0340 con Lm06-0340-5' y Lm06-0340-3'. El gen LmjF.06.0260 fue amplificado con los oligonucleótidos Lm06-0260-5' y Lm06-0260-3' y el gen LmjF.06.0210 con los oligonucleótidos Lm06-0210-5' y Lm06-0210-3'. El gen LmjF.06.0200 fue amplificado con Lm06-0200-5' y Lm06-0200-3'; y el gen LmjF.06.0110 con Lm06-0110-5' y Lm06-0110-3'. El gen de tRNA-Sec fue amplificado con los oligonucleótidos Lm-TRNASEC524-5' y Lm-TRNASEC524-3' y el gen de tRNA-Asp con los Lm24-TRNAASP-5' y Lm24-TRNAASP-3'. El tRNA-Phe fue amplificado con los oligonucleótidos Lm09-TRNAPHE-5' y Lm09-TRNAPHE-3'; el gen de tRNA-Pro con Lm24-TRNAPRO-5' y Lm24-TRNAPRO-3'. El gen de tRNA-Tyr fue amplificado con los oligonucleótidos Lm36-TRNATYR-5' y Lm36-TRNATYR-3'; y el gen de rRNA 5S con Lm15-rRNA5S-5' y Lm15-rRNA5S-3'. El gen de rRNA 18S fue amplificado con los oligonucleótidos Lm-rRNA18S-5' y Lm-rRNA18S-3'.

Una vez amplificados y purificados los fragmentos, estos se ligaron en el vector pGEM-T Easy utilizando T4 DNA Ligasa (Promega). En la reacción de ligación se agregaron 6 μl de buffer 2x, 50 ng de pGEM-T Easy (1 μl), los productos amplificados de cada gen mencionado en el párrafo anterior, y 1 μl de T4 Ligasa (3 Unidades Weiss/ μl). El volumen final de las reacciones fue de 12 μl , ajustando las relaciones inserto:vector a cada caso. Las reacciones de ligación se incubaron de 12-16 horas a 4°C. Posteriormente, las construcciones fueron utilizadas para transformar células competentes *Escherichia coli* JM109.

Tabla 6. Oligonucleótidos utilizados para la clonación de genes en el vector pGEM-T Easy

Nombre del oligonucleótido	Secuencia del oligonucleótido
Lm06-0370-5'	5'- GAAGCGATGGACTGTTCTGG
Lm06-0370-3'	5'-CGGTCCTTGCTGCGAATATC
Lm06-0360-5'	5'-CTCCTCTTCTGGACATTTGCT
Lm06-0360-3'	5'- TTCCTCCACTTGCAACATAG
Lm06-0350-5'	5'-CGGTTCAACGGTAGTCTCTTC
Lm06-0350-3'	5'-TAGAGGAACCAGAACGGGTAG
Lm06-0340-5'	5'- CCTCGCATACACCCTTTCGG
Lm06-0340-3'	5'- CGCGAATGTACACCACACGG
Lm06-0260-5'	5'- CCTGCTTGCTGTCGATGGTG
Lm06-0260-3'	5'-TCGCCTCATCCTCCTTTCG
Lm06-0210-5'	5'- GCCGGAGACATTTGCGTAC
Lm06-0210-3'	5'- CTATGGCGACGGGATCATC
Lm06-0200-5'	5'- CCATCCCATGACAAGAGC
Lm06-0200-3'	5'-TGTAGTCGCTGTACTCGC
Lm06-0110-5'	5'-TTCACTACCGCGATAGGGTTG
Lm06-0110-3'	5'- CCATATCCAGATCCTGCATCC
Lm-TRNA ^{SEC524} -5'	5'-CCGGCTGCCTTCATCAACTC
Lm-TRNA ^{SEC524} -3'	5'-GCGCATACGTTTCGGAGTCC
Lm24-TRNA ^{ASP} -5'	5'-GAATGCGCTGCTGAGTCTCT
Lm24-TRNA ^{ASP} -3'	5'-GCGGTATGCGTGTGGTGTA
Lm09-TRNA ^{PHE} -5'	5'-TTCATCCGCGCAAAGAGG
Lm09-TRNA ^{PHE} -3'	5'-GGCCTCCACGTATTTCC
Lm24-TRNA ^{PRO} -5'	5'-GCGATCTCGTGGCTCTGGAG
Lm24-TRNA ^{PRO} -3'	5'- ACAGCTCATCCAACGGGCGC
Lm36-TRNA ^{TYR} -5'	5'-AGTGCCGAGAAGTTCGACG
Lm36-TRNA ^{TYR} -3'	5'-TCGTCTCCGTTCTGTTC
Lm15-rRNA ^{5S} -5'	5'-GAAAGCATCTCTGTGGGTTCTGA
Lm15-rRNA ^{5S} -3'	5'-CCCGGGGTCCTGCAAATG

Lm-rRNA18S-5'	5'-CGGCCTCTAGGAATGAAGG
Lm-rRNA18S-3'	5'-CCCCTGAGACTGTAACCTC

5.7. Transformación de células competentes *E. coli* JM109

Las células competentes *E. coli* JM109 fueron transformadas con el DNA de plásmido de las construcciones mencionadas en la sección anterior. Las reacciones de ligación se colocaron en tubos de polipropileno y se agregaron 200 μ l de *E. coli* JM109 competentes caseras (o 50 μ l de células competentes de alta eficiencia). Se incubó en hielo por 30 min y se dio un choque térmico a 42°C e inmediatamente se regresó el tubo a hielo por 2 min más. Se añadieron 900 μ l de medio SOC a los tubos y se incubaron por 1.5 horas a 37°C en agitación (200 rpm). Al término de la incubación se transfirió el medio a un tubo de 1.5 ml estéril y se centrifugó 1 min a 11,000 rpm, se retiró el medio y se resuspendió la pastilla de bacterias en 100 μ l de medio SOC, los cuales se plaquearon en cajas Petri con LB agar suplementado con ampicilina, IPTG (isopropil β -D-1-tiogalactopiranosido) y X-gal (5-bromo-4-cloro-3-indolil- β -D-galactósido). Las cajas Petri se incubaron de 16-18 horas a 37°C y se observó el crecimiento de colonias azules y blancas.

Composición del medio SOC: triptona 2%, extracto de levadura 0.5%, NaCl 10 mM, KCl 2.5 mM. Esterilizar por autoclave y posteriormente agregar MgCl₂ 10 mM y glucosa 20 mM. Composición del medio LB agar - ampicilina: triptona 1%, extracto de levadura 0.5%, NaCl 0.5%, agar bacteriológico 1.5%. Esterilizar por autoclave, dejar enfriar a ~50°C y agregar ampicilina a concentración final de 100 μ g/ml. Verter en cajas Petri y dejar solidificar. Para medio suplementado con IPTG y X-gal, agregar al medio solidificado 100 μ l de IPTG 100 mM y 20 μ l de X-gal 50 mg/ml y esparcirlos sobre la superficie.

5.8. Obtención de DNA de plásmido

Las colonias que hayan incorporado el vector pGEM-T Easy adquirirán la resistencia a ampicilina. Este vector también permite el desarrollo de colonias azules y blancas, según se haya insertado el fragmento de interés o no. El vector pGEM-T Easy codifica para la enzima β -

galactosidasa, la cual hidroliza el compuesto X-gal para producir un pigmento azul, por lo que las colonias con plásmido no recombinante aparecerán de color azul. El sitio de clonación múltiple del vector se localiza dentro de la secuencia de β -galactosidasa, por lo que si se inserta un fragmento de DNA en este sitio no se produce una enzima funcional y por lo tanto, la colonia bacteriana aparecerá de color blanco (www.sigmaaldrich.com).

Así, una vez transformadas las células competentes *E. coli* JM109 con el DNA de las construcciones de interés, se seleccionaron las colonias blancas (que potencialmente habían incorporado el vector ligado al inserto de interés) y se aisló el DNA de plásmido mediante el kit *NucleoSpin Plasmid* (Macherey-Nagel) siguiendo las instrucciones del fabricante. Para identificar clonas que se ligaron con los insertos se realizaron digestiones con enzimas de restricción. Posteriormente, la identidad de cada inserto fue confirmada secuenciándolos con los oligonucleótidos T7 y/o SP6.

5.9. Clonación en los vectores M13mp18 y M13mp19

Los vectores M13mp18 y M13mp19 (Apéndice III) fueron donados amablemente por el Dr. Roberto Hernández Fernández (Instituto de Investigaciones Biomédicas, UNAM). Varios genes fueron clonados en los vectores M13mp18 y mp19 para generar DNA de cadena sencilla (ssDNA): tRNA-Sec, tRNA-Tyr, tRNA-Pro, rRNA 5S, LmjF.06.0200, LmjF.06.0210, rRNA 18S y α -tubulina. Para esto, los insertos fueron liberados de sus respectivas construcciones en pGEM-T Easy por enzimas de restricción y clonados en los mismos sitios en la forma replicativa de los vectores M13. Las secuencias del tRNA-Sec y de α -tubulina fueron clonadas en los sitios *Sal*I y *Sph*I de los vectores M13. El resto de las secuencias mencionadas fueron clonadas en los sitios *Sac*I y *Sph*I.

Una vez liberados y purificados, los fragmentos se ligaron en los vectores M13mp18 y M13mp19 utilizando una T4 DNA Ligasa (Promega). En la reacción de ligación se utilizaron de 150 – 300 ng del vector linearizado, 1 μ l de Buffer 10x, los productos amplificados de cada gen mencionado en el párrafo anterior, y 1 μ l de T4 Ligasa (3 Unidades Weiss/ μ l). El volumen final de las reacciones fue de 10 μ l, ajustando las relaciones inserto:vector a cada caso. Las reacciones de ligación se incubaron de 12-16 horas a 4°C. Posteriormente, las construcciones fueron utilizadas para transformar células competentes *E. coli* JM109.

5.10. Transformación de *E. coli* JM109 con construcciones M13mp18 y M13mp19

Los métodos para la transformación de células competentes con los vectores M13, así como la generación de ssDNA fueron amablemente proporcionados por la Dra. Rosaura Hernández Rivas (Departamento de Biomedicina Molecular, CINVESTAV).

Se colocaron 5 ml de medio LB en un tubo de vidrio estéril y se inocularon con células *E. coli* JM109 previamente crecidas en placas de agar LB. Se incubó a 37°C toda la noche (Tubo 1). Al siguiente día, se colocaron 150 ng de la construcción en un tubo estéril (Tubo 2) previamente enfriado en hielo y se añadieron 50 µl de células competentes *E. coli* JM109 (Promega). Se incubaron los tubos en hielo por 30 min. Mientras transcurrió este tiempo, se prepararon tubos con 3 ml de top agar (LB, suplementado con MgCl₂ 5 mM) y se mantuvieron en baño maría a 52°C.

Una vez terminada la incubación de 30 min de las células competentes con la construcción (Tubo 2), se dio un choque térmico en baño maría a 42°C por 45 seg y rápidamente se transfirió a hielo por 2 min. A este tubo se agregaron: 3 ml de top agar, 200 µl del cultivo de toda la noche (Tubo 1), 24 µl de X-gal (50 mg/ml) y 34 µl de IPTG (100 mM). Esta mezcla se vortexeó rápidamente y se vertió en placas de agar YT 2x previamente calentadas a 37°C. Se dejaron reposar por 15 min para que el top agar solidificara y se incubaron toda la noche a 37°C. Se debe observar el crecimiento de placas líticas incoloras si la ligación inserto – M13 ha sido exitosa, un vector vacío M13 desarrollaría una placa color azul (Sambrook, J. y Russell, D., 2001).

Composición del top agar LB: triptona 1%, extracto de levadura 0.5%, NaCl 0.5%, agar bacteriológico 0.6%. El medio se esterilizó por autoclave. Composición del medio YT 2x: triptona 1.6%, extracto de levadura 1%, NaCl 0.5%. Para YT 2x agregar agar bacteriológico al 1.5%. El medio se esterilizó por autoclave y se suplementó con MgCl₂ a concentración final de 5 mM.

5.11. Inducción del DNA de cadena sencilla (ssDNA)

Para la producción de ssDNA, se colocó LB (5 ml) en un tubo de vidrio estéril y se inoculó con células *E. coli* JM109 previamente crecidas en placas de agar LB. Se incubó a 37°C toda la noche. Al día siguiente, se tomaron 100 µl de cultivo y se colocaron en 5 ml de medio YT 2x (diluir 1:50) fresco suplementado con MgCl₂ 5 mM (Tubo 1). Este tubo se incubó 2 horas a 37°C con agitación vigorosa (250 – 275 rpm). Mientras transcurrió este tiempo, se tomó 1 ml de medio YT 2x fresco y se colocó en tubos de 1.5 ml, se tomó una placa lítica incolora y colocó en este medio. Los tubos de 1.5 ml se incubaron a temperatura ambiente para difundir el fago. Por otra parte, se preparó un matraz con 8 ml de medio YT 2x (suplementado con MgCl₂ a concentración final 5 mM) por cada placa que se quería difundir. Se agregó 1 ml del cultivo de 2 horas (Tubo 1) más 1 ml de medio que contiene la placa lítica difundiéndose. Se incubó 5 horas a 37°C con agitación 200 - 220 rpm (Sambrook, J. y Russell, D., 2001). Transcurrida la incubación, se centrifugó el medio a 5000 rpm, 15 min a 4°C. Se recuperó el sobrenadante en un tubo nuevo y se extrajo el ssDNA utilizando columnas QIAprep Spin M13 (Qiagen) siguiendo las instrucciones del fabricante.

5.12. Preparación de filtros para ensayos de *run-on* nuclear

Los filtros se prepararon utilizando el aparato Bio-Dot (Biorad) y membranas de nylon Hybond N⁺ (GE Healthcare). La membrana y tres hojas de papel Whatman se prehumedecieron en agua destilada. Se ensambla el aparato Bio-Dot colocando las hojas de papel Whatman y sobre ellas la membrana de nylon. Una vez ensamblado el aparato Bio-Dot, éste se conectó al vacío. Los pozos donde se colocaron las muestras de DNA se hidrataron con 500 µl de agua destilada. Se aplicó el DNA desnaturalizado (ver abajo la preparación de las muestras) en un volumen de 200 µl. La muestra se filtró utilizando vacío suave. Una vez filtrada la muestra, los pozos se enjuagan con 500 µl de NaOH 0.4 M (para DNA de plásmido) o con 500 µl de NaOH 10 mM / EDTA 1 mM (para ssDNA) y se aplicó nuevamente vacío para pasar la solución a través del filtro. Posteriormente, se desensambló el aparato Bio-Dot y se enjuagaron los filtros en solución SSC 2x por 10 min. Se dejaron secar al aire y se realizó un *cross-linking* con luz UV utilizando un aparato Stratalinker (Stratagene). Para tal propósito, el aparato se programó en el modo *Autocrosslink*, procedimiento durante el cual se aplicó a los filtros una energía de 120 mJoules/cm² para fijar el DNA.

Para los filtros con DNA de doble cadena, se utilizaron 2 μg de DNA, se colocaron en un tubo de 1.5 ml y se agregaron NaOH y EDTA a concentraciones finales de 0.4 M y 10 mM, respectivamente. La muestra se calentó a 100°C por 10 minutos y se pasó rápidamente a hielo por 5 *min*. Tras este tiempo, el DNA se aplicó al aparato Bio-Dot ensamblado con las membranas de nylon.

Para los filtros con ssDNA, se utilizaron 2 μg de DNA, se colocaron en un tubo de 1.5 ml y se agregaron NaOH y EDTA a concentraciones finales de 10 mM y 1 mM, respectivamente. La muestra se calentó a 55°C por 5 min y se pasó rápidamente a hielo por 2 min. Tras este tiempo, el DNA se aplicó al aparato Bio-Dot ensamblado con las membranas de nylon.

Composición de la solución SSC 20x: NaCl 3 M, citrato de sodio 0.3 M. Se ajustó a pH de 7.0 y se esterilizó por autoclave.

5.13. Ensayos de *run-on* nuclear

Para el aislamiento de núcleos de los ensayos de *run-on* nuclear se utilizaron 2.5×10^8 células, las cuales se lavaron dos veces en PBS-G (NaCl 137 mM, KCl 2.7 mM, Na_2HPO_4 8 mM, KH_2PO_4 1.5 mM más glucosa 2%) y se resuspendieron en 4 mL de buffer de lisis (Tris-HCl pH 7.5 10 mM, CaCl_2 3mM, MgCl_2 2mM) pre-enfriado en hielo y se agregó NP-40 a una concentración final de 0.5%. Las células se transfirieron a un homogenizador *Dounce* y se rompieron manualmente. Una alícuota de estas células se verificó al microscopio para constatar la ruptura de las células y la liberación de los núcleos. A continuación, los núcleos se colectaron por centrifugación (1,400 g) y se lavaron una vez más con buffer de lisis. Los núcleos se resuspendieron en 100 μl de mezcla de *run-on* (*run-on mix*): Tris-HCl (pH 7.5) 100 mM, glicerol 25%, espermina 0.15 mM, espermidina 0.5 mM, DTT 2 mM, RNasin 40U (Promega), MgCl_2 2 mM, MnCl_2 4 mM, NaCl 50 mM, KCl 50 mM, rATP 2 mM, rGTP 2 mM, rUTP 2 mM, rCTP 10 μM y 250 μCi de [α - ^{32}P]-CTP (3000 Ci/mmol, Amersham).

Los núcleos resuspendidos en la mezcla de *run-on* se incubaron por 6 min a 26°C para permitir la elongación de los transcritos nacientes. Tras esta incubación se agregaron 10 U de DNAsa I y se continuó la incubación por 5 min a 37°C, y posteriormente se detuvo la reacción con la adición de 100 μl de buffer de paro (Tris-HCl 10 mM, pH 7.5, EDTA 10 mM, SDS 1% y proteinasa K 100 $\mu\text{g}/\text{ml}$),

incubando 15 min a 37°C. Posteriormente, el RNA se obtuvo mediante una extracción con 200 µl de fenol:cloroformo:isoamílico, una vez extraída la fase acuosa, se agregaron 100 µl de Tris 10 mM a la fase orgánica, recuperándose la segunda fase acuosa. Para separar los nucleótidos radiactivos libres, se utilizaron columnas de sefadex G-50 (GE Healthcare).

El RNA naciente marcado se hibridó con filtros (Amersham) que contenían puntos (*dots*) con 2 µg de DNA de plásmido (clonado en pGEM-T Easy) o de DNA de cadena sencilla (clonado en vectores M13). Las hibridaciones se realizaron durante 48 horas a 50°C en solución de hibridación (formamida 50%, SSC 5x, SDS 0.2%, solución Denhardt 4x y DNA de esperma de salmón 100 µg/ml). Los lavados se llevaron a cabo en SSC 0.1x y SDS 0.1% a 65°C. Las membranas se expusieron en *Imaging Plates* y se escanearon en un equipo Fuji FLA-3000.

Las densitometrías de todos los ensayos de *run-on* se realizaron utilizando el programa Multi-Gauge 3.0 (Fujifilm). El análisis estadístico de los datos se realizó utilizando Microsoft Excel y para la realización de la prueba de ANOVA de una vía se utilizó el programa Sigma Stat incluido en el programa SigmaPlot (12.0).

5.14. Ensayos de *run-on* nuclear con luz UV

Para los ensayos en presencia de luz UV, los promastigotes se colocaron en una caja Petri (en un volumen de 15 ml) y se irradiaron con luz UV, en agitación, en un aparato Stratalinker UV *cross-linker* (Stratagene). Las dosis de luz UV utilizadas fueron 1.25, 2.5 y 5.0 kJ/m². Tras la irradiación, las células se incubaron por una hora y media a 28°C para permitir la elongación por las RNA Polimerasas reclutadas o posicionadas antes de la irradiación. El resto del ensayo de *run-on* nuclear se realizó tal como se describe en el apartado anterior. La hibridación se realizó utilizando los filtros con DNA plasmídico fijado como se describió previamente.

5.15. Ensayos de *run-on nuclear* con inhibidores de la RNA Polimerasa

La elongación del RNA naciente en la presencia de inhibidores de la transcripción se realizó incubando los núcleos (tras la lisis de células) por 15 min en 100 μ l de Buffer de Lisis con diferentes concentraciones de α -amanitina (Roche Molecular Biochemicals) o tagetitoxina (Epicentre). Los núcleos se centrifugaron a 4000 rpm por 10 min y se resuspendieron en mezcla de *run-on* más α -amanitina (100, 200 y 400 μ g/ml) o tagetitoxina (160 μ M). A continuación se permitió la elongación de RNA naciente incubando la reacción por 6 min a 26°C. El resto del ensayo de *run-on nuclear* se realizó como se describió en la sección 5.13. Para el *run-on* con α -amanitina, se utilizaron membranas con DNA de cadena sencilla (ssDNA).

6. RESULTADOS

6.1. Localización genómica del gen de tRNA-Sec en *L. major* y otros tripanosomátidos

La mayoría de los genes de tRNA en *L. major* están agrupados junto con otros genes de tRNA u otro tipo de genes transcritos por la RNA Pol III (Ivens, A.C., et al., 2005; Padilla-Mejía, N.E., et al., 2009). Estos agrupamientos se localizan frecuentemente en los límites de las unidades policistrónicas, las cuales son transcritas por la RNA Pol II. El gen de tRNA-Sec se encuentra insertado en la unidad policistrónica más larga del cromosoma 6 de *L. major* (Figura 12A) sin la compañía de otros tRNA (Padilla-Mejía, N.E., et al., 2009).

Se analizó el contexto genómico de los genes de tRNA-Sec en varias especies de *Leishmania*: *L. infantum* (strain JPCM5), *L. braziliensis* (M2904), *L. mexicana* (U1103), *L. donovani* (BPK282/0cl4) y *L. tarentolae* (Parrot-Tarll) (Peacock, C., et al., 2007; Rogers, M.B., et al., 2011; Downing, T., et al., 2011). Se encontró que en todas estas especies de *Leishmania* el tRNA-Sec es un gen de copia única que está dentro de la misma unidad policistrónica en el cromosoma 6 (Figura 12A). Así, el *locus* de tRNA-Sec es sinténico entre especies de *Leishmania*.

En *T. brucei* existen dos copias del tRNA-Sec localizadas en el cromosoma 9 (Padilla-Mejía, N.E., et al., 2009). El análisis de las bases de datos de los genomas de *T. cruzi* (CL Brener)(El-Sayed, N.M., et al., 2005), *T. vivax* (Y486) y *T. congolense* (IL3000), mostró que la presencia de dos genes de tRNA-Sec está conservada en este género (Figura 12B). Aunque el *locus* del tRNA-Sec es sinténico en estas especies, el gen de una Ser/Thr protein - cinasa (gen C en la Figura 12B) está duplicado en *T. cruzi*.

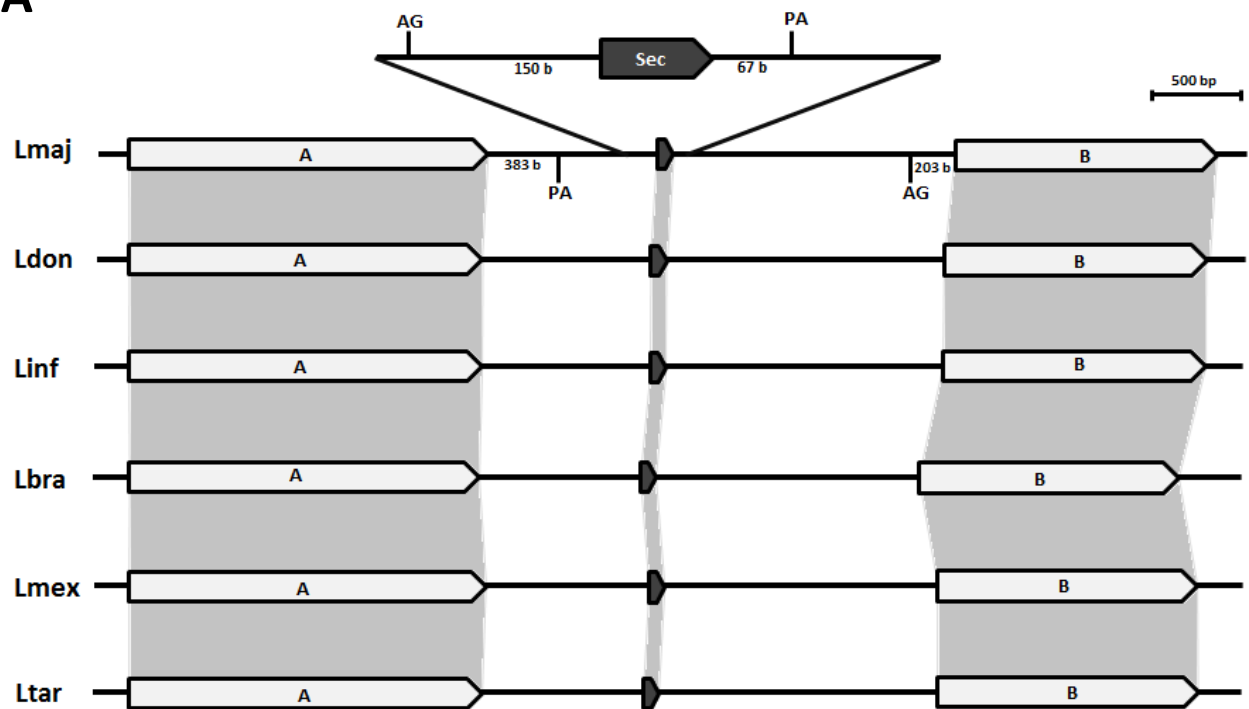
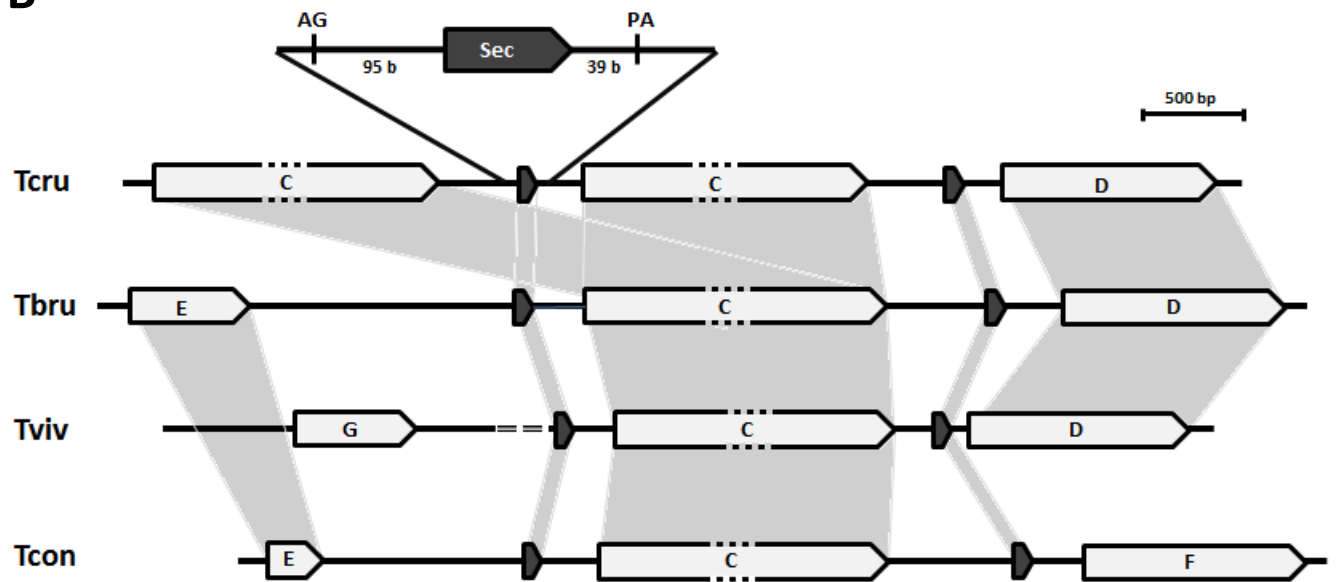
A**B**

Figura 12. Sintenia del locus del tRNA-Sec en Tripanosomátidos. Pie de figura en página siguiente

Figura 12. Sintenia del locus del tRNA-Sec en Tripanosomátidos. Se muestra el contexto genómico del tRNA-Sec en los genomas de *L. major* (Lmaj), *L. donovani* (Ldon), *L. infantum* (Linf), *L. braziliensis* (Lbra), *L. mexicana* (Lmex) y *L. tarentolae* (**panel A**); y para *T. cruzi* (Tcru), *T. brucei* (Tbru), *T. vivax* (Tviv) y *T. congolense* (Tcon) (**panel B**). Los genes ortólogos se unen mediante barras grises. Se muestra una ampliación en la escala de los tRNA-Sec respectivos en *L. major* y *T. cruzi*, indicando los sitios de adición del miniexón (AG) y el de poliadenilación (PA). En *T. cruzi*, la secuencia codificante del tRNA-Sec así como de sus regiones flanqueantes son idénticas. Se muestra la localización de los AG y de PA de sólo uno de los genes del esquema. En *L. major*, se indican la posición de PA para el gen A, y el sitio AG para el gen B. Los genes A corresponden al *LmjF.06.0210* (Lmaj), *LdBPK_060210.1* (Ldon), *LinJ.06.0210* (Linf), *LbrM.06.0190* (Lbra), *LmxM.06.0210* (Lmex) y *LtaP06.0180* (Ltar), todos ortólogos. Los genes B corresponden a *LmjF.06.0200*, *LdBPK_06200.1*, *LinJ.06.0200*, *LbrM.06.0180*, *LmxM.06.0200* y *LtaP06.0170*. Los genes C corresponden a *TcCLB.506467.50* y *TcCLB.506467.40* (Tcru), *Tb927.9.2350* (Tbru), *TvY486_0900740* (Tviv) and *TcIL3000_0_10490* (Tcon). Los genes D corresponden a *TcCLB.506467.29*, *Tb927.9.2320* y *TvY486_0900730*. El gen E corresponde a *Tb927.9.2390* y *TcIL3000_0_10500*. El gen F corresponde a *TcIL3000_0_10480* y el gen G a *TvY486_0900750*.

6.2. En *L. major*, algunos precursores del tRNA-Sec contienen el miniexón y una cola de poliadeninas

La presencia de cajas A y B atípicas en el gen de tRNA-Sec en *L. major*, así como su localización dentro de una unidad policistónica, nos sugirió que los mecanismos que regulan la transcripción de este gen podrían ser diferentes a los que regulan otros genes de tRNA. Para analizar esta posibilidad, se procedió a mapear los sitios de inicio de transcripción del gen de tRNA-Sec mediante experimentos de 5'-RACE. De las amplificaciones de estos ensayos se obtuvieron tres bandas de DNA de 300, 225 y 150 pares de bases (pb) aproximadamente (Figura 13A), las cuales fueron clonadas en el vector pGEM-T Easy (Promega) y secuenciadas.

El análisis de varias de las clonas de la banda más grande (~300 pb) evidenció la presencia del miniexón en el extremo 5', a 150 pb río arriba de la secuencia del tRNA-Sec (Figura 13C). Este resultado fue inesperado ya que la adición del miniexón se creía exclusivo de los mRNA, lo cual nos sugirió que el gen de tRNA-Sec podría ser transcrito por la RNA Pol II, tal y como ha sido demostrado en *T. brucei* (Aeby, E., et al., 2010).

El análisis de la secuencia de las clonas de la banda de ~225 pb mostró transcritos sin el miniexón conteniendo 68 y 86 pb río arriba del gen de tRNA-Sec. Finalmente, las clonas de la banda de ~150 pb contenían secuencias que se extendían 5 y 9 pb hacia el extremo 5' del gen del tRNA-Sec, así como transcritos que iniciaban en el nucleótido +1 del gen (Figura 13C). Ya se ha descrito que los transcritos primarios de los tRNA clásicos, transcritos por la RNA Pol III, inician entre 5-12 nucleótidos río arriba del gen (Orioli, A., et al., 2012), aunque también se ha reportado que en

Xenopus, el tRNA-Sec, a diferencia de los tRNA canónicos, comienza su transcripción en el primer nucleótido del gen (Lee, B.J., *et al.*, 1987). Así, de manera interesante, estos últimos transcritos podrían ser sintetizados por la RNA Polimerasa III. Sin embargo, también existe la posibilidad de que los transcritos que comienzan en el nucleótido +1 correspondan a tRNAs maduros a los que ya se les ha eliminado toda secuencia de su extremo 5', independientemente de la RNA Polimerasa involucrada en su transcripción.

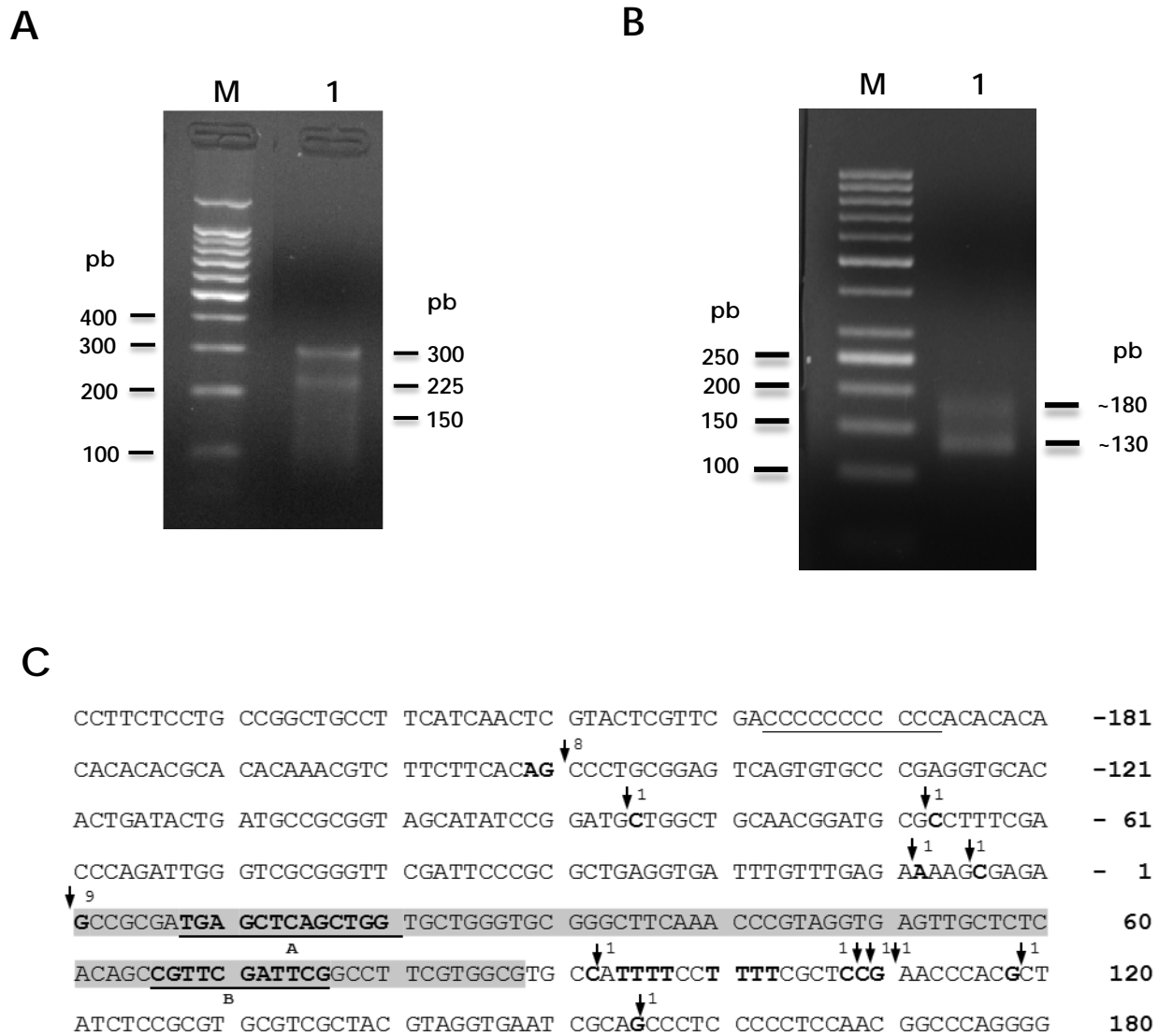


Figura 13. Resultados del 5' y 3'-RACE del tRNA-Sec en *L. major*. (A) Gel de agarosa al 1.5% donde se muestran los productos finales del experimento de 5'-RACE del gen de tRNA-Sec (carril 1). M, marcador de 1 kb (Invitrogen). (B) Gel de agarosa al 2% donde se muestran los productos finales del experimento de 3'-RACE del gen de tRNA-Sec (carril 1). M, marcador de 100 pb. (C) Secuencia del tRNA-Sec y sus secuencias flanqueantes. Con sombra gris se identifica la secuencia codificadora del tRNA-Sec, río arriba (150 pb) se encuentra el dinucleótido AG (en negritas), sitio de adición de miniexón. También se señalan los extremos 5' de los productos más pequeños del 5'-RACE, así como los sitios de poliadenilación río abajo del tRNA-Sec (nucleótidos en negritas). El número de clones de cada uno de los productos del 5' y 3'-RACE se marca con una flecha. La región rica en pirimidinas localizada río arriba del tRNA-Sec se muestra subrayada.

Como se ha mencionado, las secuencias requeridas para el *trans-splicing* constan de un dinucleótido AG que constituye el sitio de adición de miniexón así como una región rica en pirimidinas localizada río arriba (Benz, C., *et al.*, 2005; Siegel, T.N., *et al.*, 2005). En el caso del tRNA-Sec, se localizó el dinucleótido AG así como un tracto de pirimidinas río arriba (Figura 13C), por lo que el *trans-splicing* del tRNA-Sec cumple con las señales canónicas caracterizadas en tripanosomátidos.

Ya que se observó que los transcritos del tRNA-Sec poseían la secuencia del miniexón, nos preguntamos si también estaban poliadenilados en el extremo 3'. Para determinar esto, se realizó un ensayo de 3'-RACE. Las bandas resultantes, de ~130 y ~180 pb se clonaron igualmente en pGEM-T Easy para ser secuenciadas (Figura 13B). El análisis de la secuencia de ambas bandas demostró que el tRNA-Sec efectivamente posee una cola de poliadeninas en el extremo 3' (Figura 13C). Se encontraron clonas con diferentes sitios de poliadenilación: de 22 a 67 bases río abajo del transcrito de tRNA-Sec (Figura 13C). Así, estos resultados muestran que los precursores del tRNA-Sec en *L. major* contienen el miniexón en el extremo 5' y un tracto de poliadeninas en el extremo 3'. Estos datos sugieren que el tRNA-Sec es transcrito por la RNA Pol II como parte de un policistrón y procesado por *trans-splicing* y poliadenilación. Posteriormente, este transcrito necesitaría ser adicionalmente procesado para generar un tRNA maduro.

6.3. El tRNA-Sec en *T. cruzi* sufre *trans-splicing* y poliadenilación

Los resultados anteriores sugieren fuertemente la participación de la RNA Pol II en la síntesis del tRNA-Sec en *L. major*, de manera similar a lo que se reportó para *T. brucei* (Aeby, E., *et al.*, 2010), por lo que este fenómeno podría representar una característica general de la expresión génica en tripanosomátidos. Así, se quiso determinar si en *T. cruzi*, el tRNA-Sec también se somete a los procesos de adición del miniexón y poliadenilación.

Para esto, se realizaron ensayos de 5' y 3'-RACE, por los cuales se demostró la presencia del miniexón 95 bases río arriba del tRNA-Sec, así como la presencia del tracto de poliadeninas a 36-39 bases río abajo del tRNA-Sec (Figura 12B y 14). Así, la participación de RNA Pol II en la transcripción del gen de tRNA-Sec parece una característica de la expresión génica en tripanosomátidos.

De manera similar a lo previamente reportado para *T. brucei* (Aeby, E., et al., 2010) y a nuestros hallazgos en *L. major*, para el tRNA-Sec de *T. cruzi* se encontraron los elementos canónicos requeridos para el trans-splicing: un dinucleótido AG como sitio de adición del miniexón y una región de polipirimidinas río arriba (Figura 14). Cabe mencionar que las secuencias codificantes de las dos copias del tRNA-Sec en *T. cruzi* son idénticas, al igual que las regiones intergénicas 5' y 3' (Figura 12B), por lo que en los resultados del 5' y 3' – RACE es imposible distinguir de cuál de las dos copias del tRNA-Sec se trata.

```

TTGGATCCCT CACCATGTTC GAACGCCCC AC GCGTCATG CCGCAACCAT TCCCTTTGTT -121
GTTTCTATTC TTTGCTTGTC GCTAGAGTGA TAGGGCTTAG TGTAGCCAAT GATGGTGCTG - 61
GAGCGCCCTG GGCCCATCTT TGC GGCGCTG AGAAGGAGGC GGAGGTGTGG AGTGGAGTGT - 1
GCCACGATGA GCTCAGCTGG TGCTGGGTGC GGGCTTCAA CCCGTAGGCG AGTTGGACTC 60
GCAGCCGTTG GATTCG GCCT TTGTGGTGT TTTTTTTTTT GC GGA ACTGC AGTCCGTCTT 120
TTTTATTCTT TGTCCGTTGT GTTGTGCTGT AAGGCGTATG TATTTTTTAT TTTTATTTTT 180

```

↓⁹
 A
 B
 ↓ ↓ ↓ ↓
 1 1 1 1

Figura 14. Resultados de los ensayos 5' y 3' RACE en *T. cruzi*. Secuencia del tRNA-Sec y sus regiones flanqueantes. Las secuencias correspondientes al tRNA-Sec se marcan en sombreado gris. Las cajas A y B intragénicas se muestran en negritas y subrayadas. El sitio aceptor de miniexón (AG, río arriba del tRNA-Sec) y las regiones de poliadenilación (río abajo del tRNA-Sec) se muestran en negritas. El número de clones de cada uno de los productos del se marca con una flecha. La región rica en pirimidinas localizada río arriba del tRNA-Sec se muestra subrayada.

6.4. El tRNA-Sec de *L. major* está presente en transcritos de diferentes tamaños

Para demostrar la presencia de los diferentes transcritos que contienen el tRNA-Sec se realizó un ensayo de *Northern blot*. Para este experimento, RNA total de *L. major* se hibridó con una sonda que contiene el tRNA-Sec. Se identificaron 4 bandas de diferentes tamaños: 80, 400, 600 y 1000 bases (Figura 15A). La banda de ~80 nucleótidos corresponde al tRNA-Sec maduro (el gen es de 88 pb), mientras que la banda de ~400 nucleótidos podría corresponder al transcrito que posee el miniexón y la cola de poliadenilación. Las bandas más grandes representarían RNAs precursores antes de ser procesados.

Se realizó otro experimento de *Northern blot*, hibridando esta vez RNA total con una sonda que correspondía a la secuencia del tRNA de lisina (tRNA-Lys), y se observó únicamente una banda

con un tamaño correspondiente a un tRNA maduro (Figura 15B). Esto nos indicó que la presencia de precursores largos de tRNA no es una característica general de los transcritos de dichos RNAs en *L. major*.

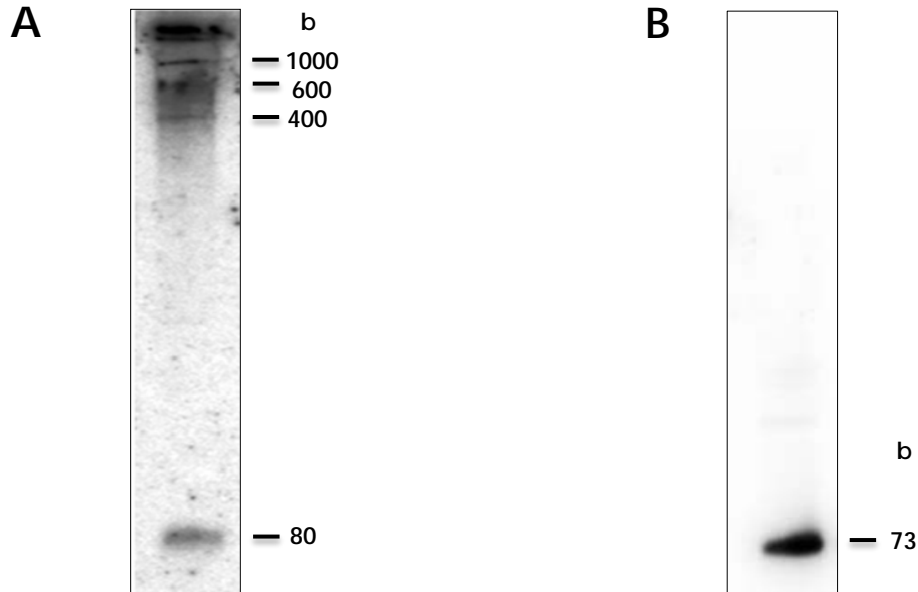


Figura 15. El tRNA-Sec de *L. major* está presente en transcritos de diferentes tamaños. (A) Ensayo de *Northern blot* del tRNA-Sec. Se observan 4 bandas de diferentes tamaños: ~80, 400, 600 y 1000 bases. (B) Ensayo de *Northern blot* del tRNA-Lys. Se observa una única banda de ~73 bases.

6.5. En *L. major*, el *trans-splicing* y la poliadenilación del tRNA-Sec ocurren en las señales genómicas canónicas para estos mecanismos

Como se mencionó anteriormente, se localizó un dinucleótido AG como sitio de adición del miniexón del tRNA-Sec de *L. major* (Figura 13C). Para estudiar si existían otras señales de procesamiento canónicas en el transcrito policistrónico, se realizaron experimentos de RT-PCR para identificar el sitio de poliadenilación del gen LmjF.06.0210, localizado río arriba del tRNA-Sec, así como el sitio de adición del miniexón del gen LmjF.06.0200, localizado río abajo (Figura 12A).

El tracto de poliadeninas del mRNA del gen LmjF.06.0210 se encontró a 383 bases río abajo del codón de paro. Por otra parte, el sitio de adición de miniexón del mRNA del gen LmjF.06.0200 se localizó 203 bases río arriba del codón de inicio (Figura 12A). Asimismo, también se localizaron regiones ricas de pirimidinas en las regiones intergénicas, las cuales se requieren para el *trans-splicing*

del gen río abajo así como para la poliadenilación del gen río arriba. Una de estas regiones ricas en pirimidinas se localizó entre el sitio de poliadenilación del LmjF.06.0210 y el sitio de adición del miniexón del tRNA-Sec, así como entre la región de poliadenilación del tRNA-Sec y el sitio de adición del miniexón del LmjF.06.0200. Asimismo, 35 pares de bases río arriba del dinucleótido AG se localizó un tracto de polipirimidinas (11 citosinas consecutivas) (Figura 13C). Tanto el dinucleótido AG como el tracto de polipirimidinas constituyen las señales canónicas del *trans-splicing* en tripanosomátidos. Por lo tanto, nuestros resultados indican que las señales canónicas regulan el procesamiento del transcrito policistrónico del tRNA-Sec en *L. major*.

6.6. Otros genes de tRNA no poseen el miniexón en su extremo 5'

Como se mencionó antes, el tRNA-Sec está situado dentro de una unidad policistrónica de genes codificantes de proteínas. Sin embargo, este no es el único gen de tRNA que está insertado en grupos de genes transcritos por RNA Pol II, pues el gen de tRNA-Asp (LmjF.24.TRNAASP.01) y el tRNA-Pro (LmjF.24.TRNAPRO.01) se localizan dentro de dos diferentes unidades policistrónicas en el cromosoma 24, como genes únicos de la RNA Pol III (Padilla-Mejia, N.E., *et al.*, 2009).

Nosotros nos preguntamos si estos tRNA también son transcritos por la RNA Pol II y sufren la adición del miniexón en su extremo 5'. Para determinar esto se realizaron ensayos de 5'-RACE. Se secuenciaron aproximadamente dos decenas de clonas de ambos tRNA sin que se encontrara el miniexón río arriba de estos tRNA. La mayoría de las clonas comenzaron en el extremo 5' del tRNA maduro. De manera interesante, varias de ellas comenzaron 4-5 bases río arriba del gen de tRNA, posiciones que podrían corresponder a sitios de inicio de la transcripción de la RNA Pol III.

Así estos resultados sugieren que los tRNA-Asp y tRNA-Pro residentes en el cromosoma 24 no son transcritos por la RNA Pol II, por lo que la localización dentro de una unidad policistrónica de genes codificadores de proteínas no parece ser el único requisito para que un tRNA sea transcrito como parte del policistrón por la RNA Pol II.

6.7. Clonación de fragmentos en pGEM-T Easy para ensayos de *run-on* nuclear

Para seguir con el análisis de la transcripción del gen de tRNA-Sec en *L. major*, se decidió realizar ensayos de *run-on* nuclear. Para realizar los ensayos de *run-on* se utilizaron diversas secuencias de DNA de genes transcritos por las tres RNA Polimerasas: los genes codificadores de proteínas del cromosoma 6 (Figura 18A) transcritos por la RNA Pol II, diversos tRNA y el rRNA 5S (transcritos por la RNA Pol III), así como el rRNA 18S (transcrito por la RNA Pol I). Los fragmentos de interés para estos ensayos fueron amplificados por PCR a partir de DNA genómico de *L. major*, y posteriormente clonados en el vector pGEM-T Easy mediante los procedimientos descritos en la sección de Materiales y Métodos. Para comprobar que, efectivamente, las construcciones contenían los insertos del tamaño apropiado, éstos se liberaron mediante enzimas de restricción y se visualizaron en geles de agarosa.

En la Figura 16 se muestran los geles de agarosa de las digestiones realizadas a las construcciones. En la mayoría de los casos las construcciones fueron digeridas con la enzima *EcoR* I, ya que pGEM-T Easy posee dos sitios *EcoR* I que flanquean el sitio de inserción del fragmento en el vector (Apéndice II). Para la construcción que contiene el fragmento 0350 se realizó una digestión doble *Sph* I / *Sac* I, debido a que este inserto posee un sitio *EcoR* I dentro de su secuencia. La identidad de todas las clonas se verificó por análisis con enzimas de restricción y mediante secuenciación.

6.8. Clonación de fragmentos en M13mp18 y M13mp19 para la generación de ssDNA para los ensayos de *run-on* nuclear

A partir de las construcciones en pGEM-T Easy, algunos de los fragmentos se subclonaron en los vectores M13mp18 y M13mp19 para la generación ssDNA, entre ellos tRNA-Sec, tRNA-Tyr, tRNA-Pro, rRNA 5S, LmjF.06.0200, LmjF.06.0210, rRNA 18S y α -tubulina. Como se mencionó anteriormente, las secuencias del tRNA-Sec y de α -tubulina fueron clonadas en los sitios *Sal* I y *Sph* I de los vectores M13 (Apéndice III). El resto de las secuencias mencionadas fueron clonadas en los sitios *Sac* I y *Sph* I. Para comprobar la inserción de los fragmentos en los vectores M13, se realizaron reacciones de digestión con diferentes enzimas de restricción y/o PCR y los resultados se visualizaron en geles de agarosa. Las clonas positivas se mandaron secuenciar para verificar la identidad de los insertos. Como ejemplo de estas subclonaciones se muestran los resultados de la inserción del tRNA-Sec en los vectores M13mp18 (M18-Sec) y M13mp19 (M19-Sec) (Figura 17).

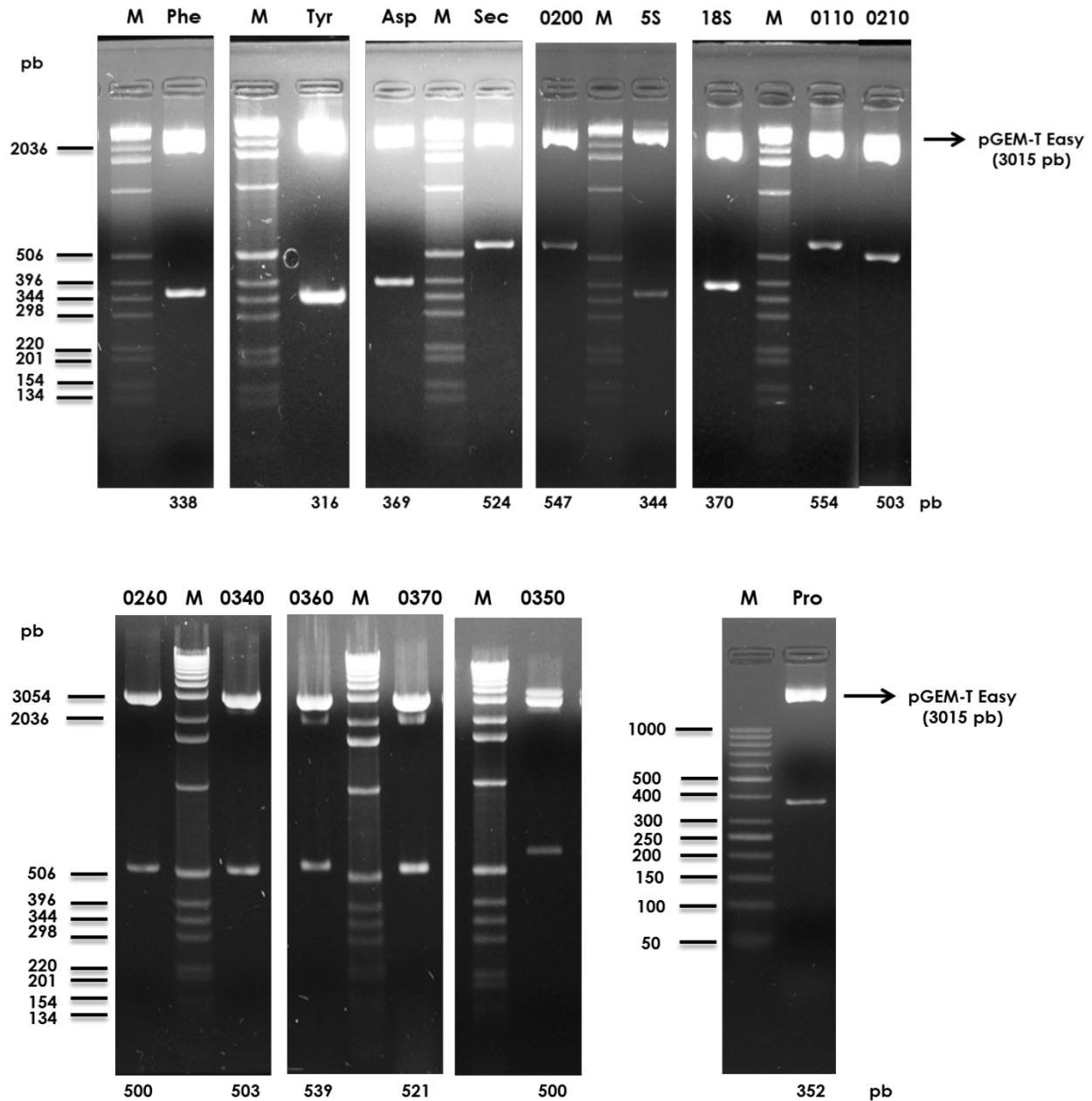


Figura 16. Clonación de fragmentos en pGEM-T Easy. Geles de agarosa al 2%, donde se muestran las reacciones de digestión de las construcciones pGEM-T Easy con los insertos de interés para los ensayos de *run-on* nuclear. M, marcador de 1 kb (Invitrogen), los tamaños de las bandas del marcador se muestran del lado izquierdo, en el caso del tRNA-Pro, el marcador corresponde al de 50 pb (Fermentas). En la parte superior se identifica el inserto correspondiente: tRNA-Phe (Phe), tRNA-Tyr (Tyr), tRNA-Asp (Asp), tRNA-Sec (Sec), tRNA-Pro (Pro), rRNA 5S (5S), rRNA 18S (18S), las secuencias de los genes codificadores de proteínas LmjF.06.0200 (0200), LmjF.06.0210 (0210), LmjF.06.0110 (0110), LmjF.06.0260 (0260), LmjF.06.0340 (0340), LmjF.06.0360 (0360), LmjF.06.0370 (0370), LmjF.06.0350 (0350). El tamaño de cada inserto se indica en la parte inferior de cada carril. Es importante mencionar que, tras la digestión los insertos acarrean al menos 20 pb de la secuencia del vector. Con una flecha se indica la posición de pGEM-T Easy (3015 pb).

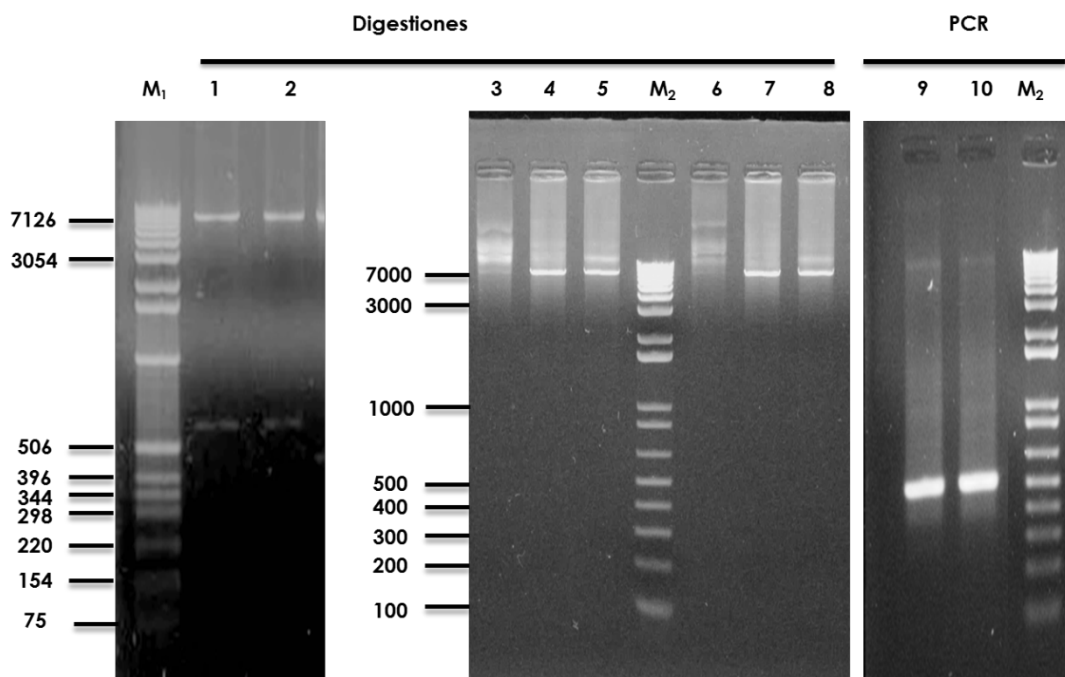


Figura 17. Subclonación del tRNA-Sec en los vectores M13. Geles de agarosa al 1.2%, donde se muestran las reacciones de digestión de las construcciones M18-Sec y M19-Sec, así como PCR realizados a partir de las mismas para verificar la inserción de tRNA-Sec. M₁, marcador de 1 kb (Invitrogen), los tamaños de las bandas del marcador se muestran del lado izquierdo, M₂, marcador 1kb Plus (Invitrogen). Carril 1) M18-Sec, 2) M19-Sec, digestiones dobles con las enzimas *Sph* I / *Sac* I, que liberan el inserto tRNA-Sec (524 pb), la banda de ~ 7000 pb corresponde a los vectores M13. Carriles 3) M18-Sec y 6) M19-Sec no digeridos, 4) M18-Sec y 7) M19-Sec digeridos con la enzima *Nco* I; 5) M18-Sec y 8) M19-Sec digeridos con la enzima *Spe* I. Las digestiones con *Nco* I y *Spe* I linearizarían la construcción sólo si el inserto está presente, pues los vectores no poseen estos sitios. En los carriles 4, 5, 7 y 8 se aprecia la construcción linearizada (7778 pb). La presencia del tRNA-Sec también se verificó por PCR, amplificando dicho fragmento a partir del 9) M18-Sec y 10) M19-Sec. Se aprecia la banda de 524 pb. La presencia del resto de los insertos en los vectores M13 se comprobó mediante estrategias similares: análisis por enzimas de restricción y amplificación por PCR.

6.9. Ensayo de *run-on* nuclear, análisis del locus del tRNA-Sec

Para seguir con el análisis de la transcripción del gen de tRNA-Sec se realizaron ensayos de *run-on* nuclear. Además de este gen, también se analizaron varios genes codificadores de proteínas (transcritos por la RNA Pol II) de la misma unidad policistónica que el tRNA-Sec: LmjF.06.0360, LmjF.06.0350, LmjF.06.0340, LmjF.06.0260, LmjF.06.0210, LmjF.06.0200 and LmjF.06.0110. También se incluyó el primer gen de la unidad policistónica adyacente: LmjF.06.0370 (Figura 18A).

Se incluyeron también varios genes transcritos por la RNA Pol III: tRNA-Asp (LmjF.24.TRNAASP.01), tRNA-Phe (LmjF.09.TRNAPHE.01), tRNA-Pro (LmjF.24.TRNAPRO.01), tRNA-Tyr

(LmjF.36.TRNATYR.01) y el rRNA 5S (LmjF.15.5SrRNA.01). El rRNA 18S (LmjF.27.rRNA.01) se incluyó como gen representante de la transcripción por RNA Pol I. Para este ensayo, los núcleos se aislaron y se incubaron en presencia de [α - 32 P]-CTP, de tal manera que el nucleótido radiactivo se incorporó en los transcritos nacientes. Posteriormente, el RNA naciente marcado radiactivamente se purificó y se hibridó con los genes de interés fijados previamente en membranas de nylon. Como se muestra en la Figura 18B, la intensidad de la señal obtenida por el tRNA-Sec es mayor en comparación con los genes codificadores de proteínas de la misma unidad policistónica, siendo similar a la señal observada con otros tRNA y el rRNA 5S, que son genes transcritos por la RNA Pol III.

El *run-on* nuclear se repitió 5 veces, obteniendo resultados muy similares (Figura 18C). Se cuantificaron por densitometría las señales del tRNA-Sec y la de los genes codificadores de proteínas que lo flanquean, LmjF.06.0210 y LmjF.06.0200, los cuales forman parte de la misma unidad policistónica. Para graficar estos datos la señal de LmjF.06.0210 se consideró como el 100%. En promedio, la señal del tRNA-Sec fue ~3.5 veces más alta que la señal obtenida con los otros dos genes (Figura 18D), y una prueba estadística de ANOVA de una vía ($F_2=13.7$, $p<0.0001$) nos indicó que el comportamiento transcripcional del tRNA-Sec es diferente al de los genes flanqueantes de proteínas, por lo que este ensayo de *run-on* nuclear nos sugirió que el tRNA-Sec tiene un aporte transcripcional adicional además del de la RNA Pol II, muy probablemente, por la RNA Pol III.

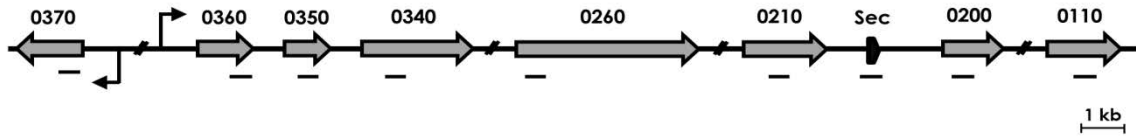
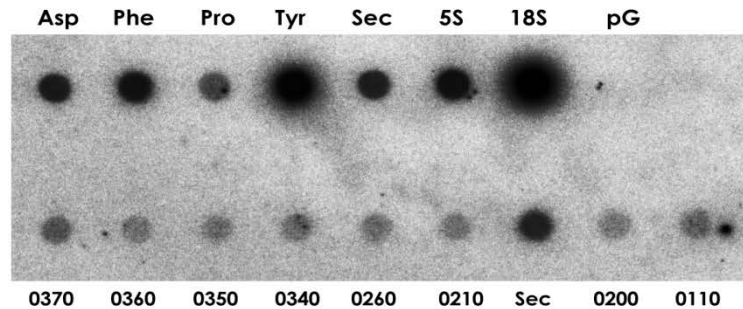
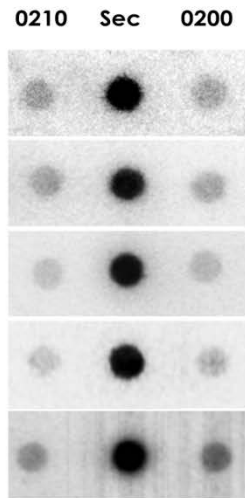
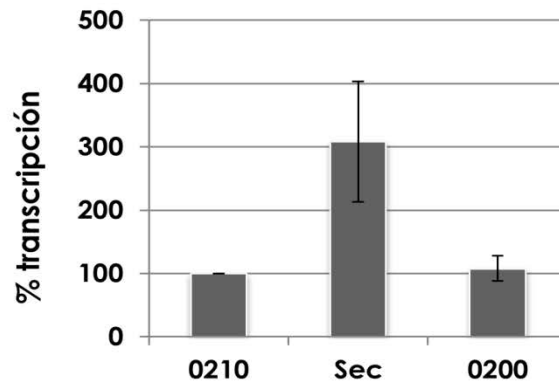
A**B****C****D**

Figura 18. Ensayo de *run-on* nuclear del locus del gen tRNA-Sec. (A) Mapa genómico de la unidad policistónica donde reside el tRNA-Sec en el cromosoma 6 de *L. major*. Los genes mostrados en el mapa fueron utilizados para el ensayo de *run-on* nuclear: LmjF.06.0370 (0370), LmjF.06.0360 (0360), LmjF.06.0350 (0350), LmjF.06.0340 (0340), LmjF.06.0260 (0260), LmjF.06.0210 (0210), LmjF.06.0200 (0200) y LmjF.06.0110 (0110). (B) Ensayo de *run-on* nuclear, en el que el RNA naciente marcado de *L. major* fue hibridado con DNA de doble cadena (2 μ g) clonado en pGEM-T Easy. Además de los genes señalados en el panel en A, también se analizaron varios genes transcritos por RNA Pol III (tRNA-Asp, tRNA-Phe, tRNA-Pro, tRNA-Tyr y rRNA 5S) así como el gen rRNA 18S, transcrito por RNA Pol I. Como control se incluyó el vector pGEM-T Easy vacío (pG). (C) Señal de hibridación del LmjF.06.0210, tRNA-Sec y LmjF.06.0200 proveniente de cinco experimentos independientes. (D) Promedio de los porcentajes de transcripción de la figura C obtenidos mediante densitometría. La señal del LmjF.06.0210 se consideró como 100%. Se realizó una ANOVA de una vía ($F_2=13.7$, $p<0.0001$; comparaciones múltiples por el método de Holm-Sidak nos indicaron que los valores son significativamente diferentes entre el tRNA-Sec y LmjF.06.0200 ($t=4.46$, $P=0.002$) y entre el tRNA-Sec y LmjF.06.0210 ($t=4.60$ y $P=0.002$).

6.10. Ensayos de *run-on* con células irradiadas con luz UV

Anteriormente se había determinado que el gen del tRNA^{Sec} de *L. major* posee secuencias promotoras internas similares a las encontradas en los tRNA^{Sec} de otras especies (Figura 10, Padilla-Mejia, N.E., *et al.*, 2009), las cuales dirigen la transcripción por la RNA Pol III. La transcripción dirigida por esta polimerasa comienza muy cerca del sitio de inicio de la transcripción, por lo que se quiso evaluar si la transcripción del tRNA^{Sec} inicia cerca de su secuencia codificante. Para este propósito, se realizó un ensayo de *run-on* nuclear con parásitos irradiados con luz UV. En este ensayo, la irradiación de células con luz UV detiene la elongación del RNA durante la transcripción al producir dímeros de pirimidinas en el DNA. La inactivación por la radiación UV de la transcripción de un gen está en función de a) la distancia entre dicha secuencia y su región promotora y b) la dosis de luz UV, pues la irradiación de células con dosis crecientes de luz UV resulta en una disminución progresiva de los transcritos nacientes. En tripanosomátidos, la irradiación con dosis bajas con luz UV inhibe el procesamiento de RNA, causando una acumulación de transcritos de las secuencias que están situadas adyacentes o muy cerca de sus regiones promotoras (Martínez-Calvillo, S., *et al.*, 2003 y 2004).

Para este experimento, se utilizaron los mismos fragmentos de DNA que para el ensayo de *run-on* nuclear anterior (Figura 18 A y B), las señales de hibridación fueron cuantificadas por densitometría y se graficó la disminución de la señal de hibridación de cada gen respecto a la dosis de luz UV, para lo cual se consideraron los valores de la membrana control (sin luz UV) como 100%. Este experimento fue realizado por duplicado, obteniendo resultados reproducibles. El duplicado de este experimento se muestra en el Apéndice IV.

Tal y como se esperaba, la señal de los genes transcritos por la RNA Pol II, así como la del rRNA 18S se redujo considerablemente a medida que se incrementó la dosis de luz UV. Por otra parte, la señal de los genes transcritos por la RNA Pol III se observó más resistente a las irradiaciones crecientes con luz UV. El patrón de reducción en la intensidad de la señal del tRNA^{Sec} fue similar a la de otros tRNA y al rRNA 5S y difirió del de los genes codificantes de proteínas localizados en la misma unidad policistónica (Figura 19A). Por ejemplo, a la dosis de luz UV de 2.5 kJ/m², la señal de hibridación del LmjF.06.0210 y LmjF.06.0200 se redujo en un 54% y 66%, respectivamente (en base a la membrana control). Por su parte, la señal del tRNA^{Sec} mostró una reducción del 15% a la misma dosis, similar a los genes transcritos por la RNA Pol III, cuyas disminuciones en la señal van del 10%-34% (Figura 19B).

Este resultado sugiere que la secuencia promotora que dirige la transcripción del tRNA-Sec se localiza cerca o dentro de la secuencia codificante.

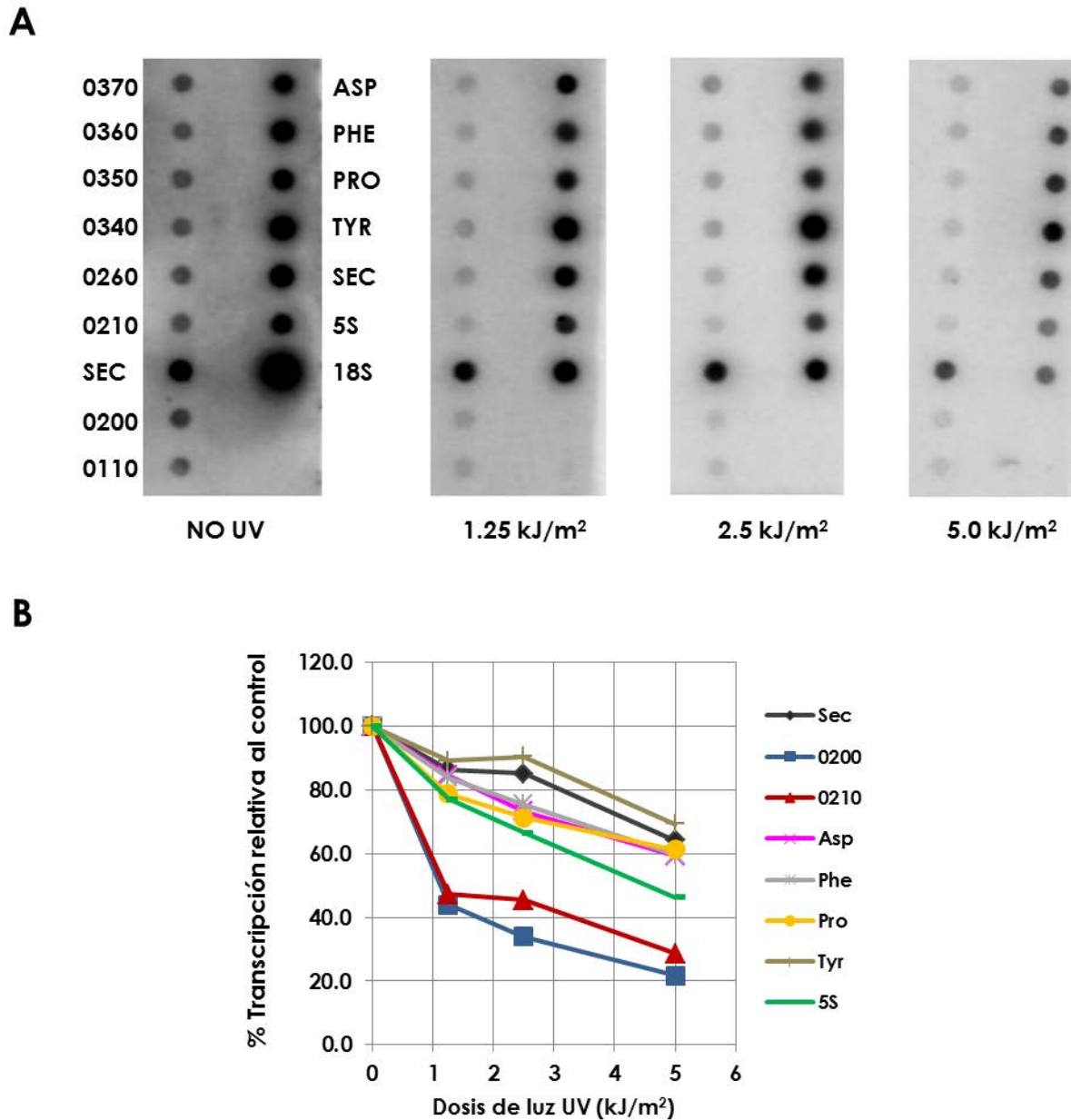


Figura 19. Ensayo de *run-on* nuclear con células irradiadas con luz UV. (A) Ensayo de *run-on* nuclear utilizando núcleos aislados de promastigotes irradiados con dosis crecientes de luz UV (1.25, 2.5 y 5 kJ/m²). Los genes analizados son los mismos que en la Figura 16 (B) Resultados del panel A cuantificados por densitometría. Las señales de cada gen obtenidas en la membrana sin luz UV se consideraron como 100%, las cuales se grafican en función de la dosis de luz UV. Los valores de la gráfica corresponden al promedio de dos experimentos independientes.

6.11. Ensayos de *run-on* nuclear con α -amanitina

La transcripción por las RNA Pol que existen en eucariontes puede ser inhibida al exponer a la célula a diferentes drogas, entre ellas la α -amanitina, la cual actúa como un inhibidor selectivo pues las tres RNA Pol (I, II y III) poseen diferentes sensibilidades a dicho compuesto. Típicamente, la RNA Pol I es insensible a la α -amanitina, la RNA Pol II es altamente sensible y la RNA Pol III muestra una sensibilidad intermedia entre las dos primeras, es decir, bajas concentraciones de la droga inhiben la transcripción por la RNA Pol II, pero tienen poco efecto en la transcripción de RNA Pol III (Nguyen, V.T., *et al.*, 1996). En *T. brucei* la RNA Pol I es resistente a 1 mg/ml de α -amanitina, la RNA Pol II tiene una ID_{50} = 2-10 μ g/ml, y la RNA Pol III muestra un nivel intermedio de resistencia entre las dos primeras con ID_{50} = 150 μ g/ml (Campbell, D.A., *et al.*, 2003). Esta diferencia en la sensibilidad de las RNA Pol a la α -amanitina es aprovechada para estudiar la transcripción de los genes por las diferentes polimerasas (Nguyen, V.T., *et al.*, 1996).

Para evaluar si la inhibición de la transcripción de la RNA Polimerasa II disminuye la transcripción del tRNA-Sec, se evaluó el efecto de diferentes concentraciones de α -amanitina en la transcripción del tRNA-Sec. Dicho efecto se determinó mediante un ensayo de *run-on* nuclear. Ya que en tripanosomátidos existe cierto nivel de transcripción inespecífica de la cadena no codificante, para el ensayo de *run-on* nuclear se utilizó DNA de cadena sencilla clonado y producido en los vectores M13. Similar al experimento de *run-on* de la sección anterior, las señales de hibridación fueron cuantificadas por densitometría y se graficaron en función de la dosis de α -amanitina. Los valores de la membrana control (sin droga) se consideraron el 100%. Este experimento fue realizado por duplicado, obteniendo resultados reproducibles. El duplicado de este experimento se muestra en el Apéndice V.

La concentración más baja de α -amanitina (100 μ g/ml) redujo la transcripción mediada por la RNA Pol II de LmjF.06.0210 en un 63% y de LmjF.06.0200 en un 65% respecto al control (sin α -amanitina) (Figura 20 A y B). Sin embargo, la señal de hibridación del tRNA-Sec se redujo sólo en un 25%. La transcripción de los genes transcritos por la RNA Pol III, tRNA-Tyr, tRNA-Pro y rRNA 5S, se redujo en un 31, 27 y 28%, respectivamente.

En general, la curva de inhibición del tRNA-Sec indica que la transcripción de este gen es más resistente a la α -amanitina que los genes flanqueantes codificantes de proteínas (LmjF.06.0210 y

LmjF.06.0200) y que posee una sensibilidad similar a la de otros genes transcritos por la RNA Pol III. Estos resultados aportan evidencia adicional de que la RNA Pol III participa en la transcripción del tRNA-Sec en *L. major*. Como se esperaba, la señal de hibridación de todos los genes se vio altamente reducida con la concentración más alta de α -amanitina (400 μ g/ml), con excepción del rRNA 18S (transcrito por la RNA Pol I) (Figura 20 A y B). Similar a lo que ya ha sido reportado en *L. major* (Martínez-Calvillo, S., et al., 2004), el gen de α -tubulina, con 12 copias en el cromosoma 13, es más resistente a la α -amanitina en comparación con genes codificantes de proteínas de copia única.

Asimismo, con este experimento se detectó una fuerte señal de la transcripción antisentido (Figura 20A), lo cual fue consistente con lo que ya se ha reportado previamente para los genes transcritos por la RNA Pol II y III (Martínez-Calvillo, S., et al., 2004 y 2003).

6.12. Ensayos de *run-on* nuclear con tagetitoxina

La tagetitoxina es un inhibidor selectivo de la RNA Pol III, el cual no posee efecto sobre las otras dos RNA Pol. Estos efectos han sido previamente observados en diversos eucariontes incluyendo vertebrados, insectos, levaduras (Steinberg, T.H., et al., 1990) y tripanosomátidos (Günzl, A., et al., 1997; Fantoni, A., et al., 1994).

Para confirmar si la RNA Pol III está involucrada en la transcripción del tRNA-Sec, se realizaron experimentos de *run-on* nuclear con tagetitoxina. La señal de hibridación del tRNA-Sec se redujo en un 70% (Figura 21). Esta reducción es de nuevo similar a la reducción de la señal de otros genes transcritos por la RNA Pol III: los tRNA de prolina y aspartato así como el rRNA 5S, los cuales se redujeron en un 66, 76 y 85%, respectivamente. Los genes transcritos por la RNA Pol II, LmjF.06.0200 y LmjF.06.0210, redujeron su señal con la tagetitoxina en un 31 y 17%. Este experimento fue realizado por duplicado, obteniendo resultados reproducibles. La repetición del mismo se muestra en el Apéndice VI. Estos resultados nos aportan una evidencia experimental adicional de que la RNA Pol III está involucrada en la transcripción del tRNA-Sec en *L. major*.

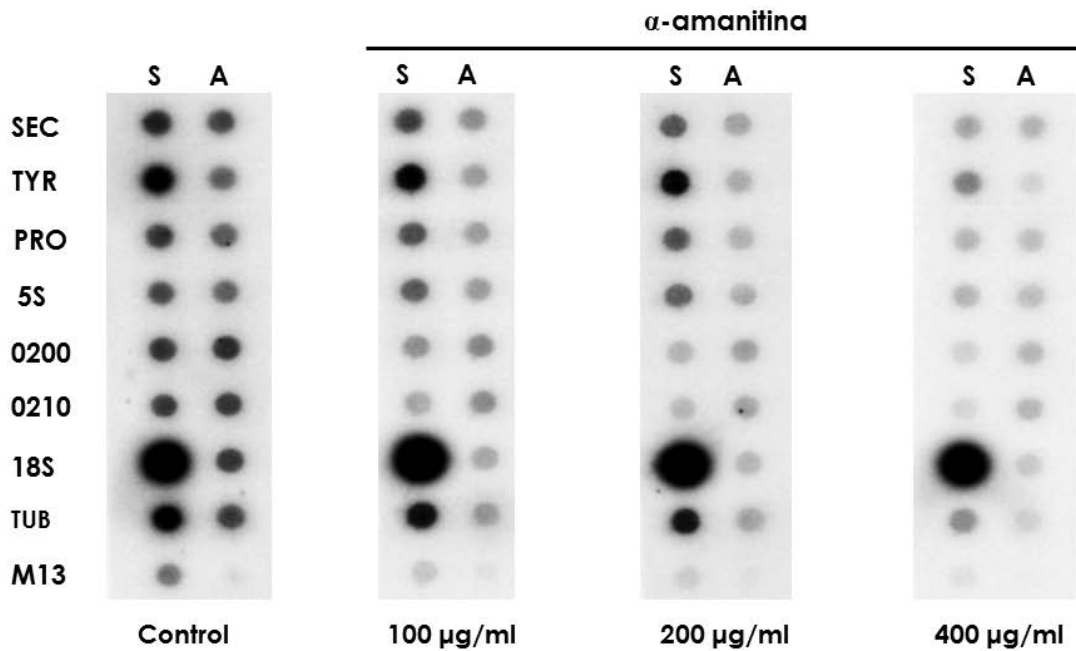
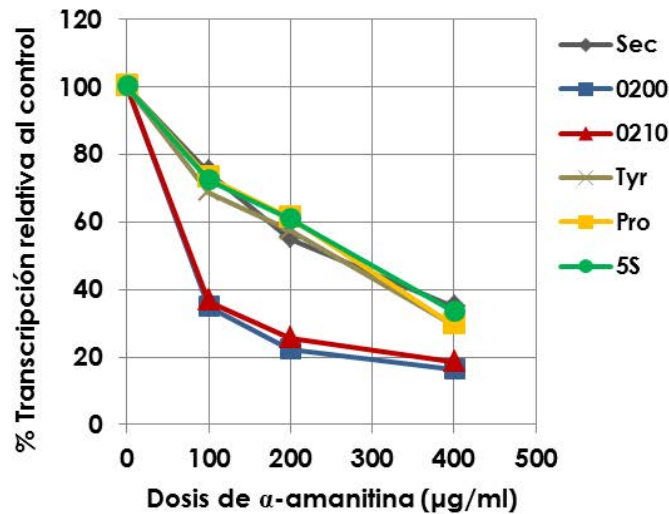
A**B**

Figura 20. Sensibilidad de la transcripción del tRNA-Sec a la α -amanitina. (A) Experimento de *run-on* nuclear. Se permitió la elongación de los transcritos nacientes en presencia de diferentes dosis de α -amanitina (0, 100, 200 y 400 $\mu\text{g/ml}$) y se hibridaron con membranas en las que se fijó previamente DNA de cadena sencilla clonado en los vectores M13. Se incluyeron los DNAs correspondientes a: tRNA-Sec, LmjF.06.0210 y LmjF.06.0200. Los genes usados como control fueron: tRNA-Tyr, tRNA-Pro, rRNA 5S (transcritos por la RNA Pol III); α -tubulina (RNA Pol II); rRNA 18S (Pol I). Se incluyó también el vector M13mp18 vacío sin inserto (M13). S: DNA complementario a la cadena sentido, A: DNA complementario a la cadena antisentido. B) Cuantificación por densitometría de las señales de hibridación con la cadena sentido. Las señales de cada gen obtenidas en la membrana sin α -amanitina se consideraron como 100%. Este porcentaje se grafica contra la dosis del inhibidor. Todos las señales están normalizadas respecto a la del RNAr 18S. Los valores de la gráfica corresponden al promedio de dos experimentos independientes.

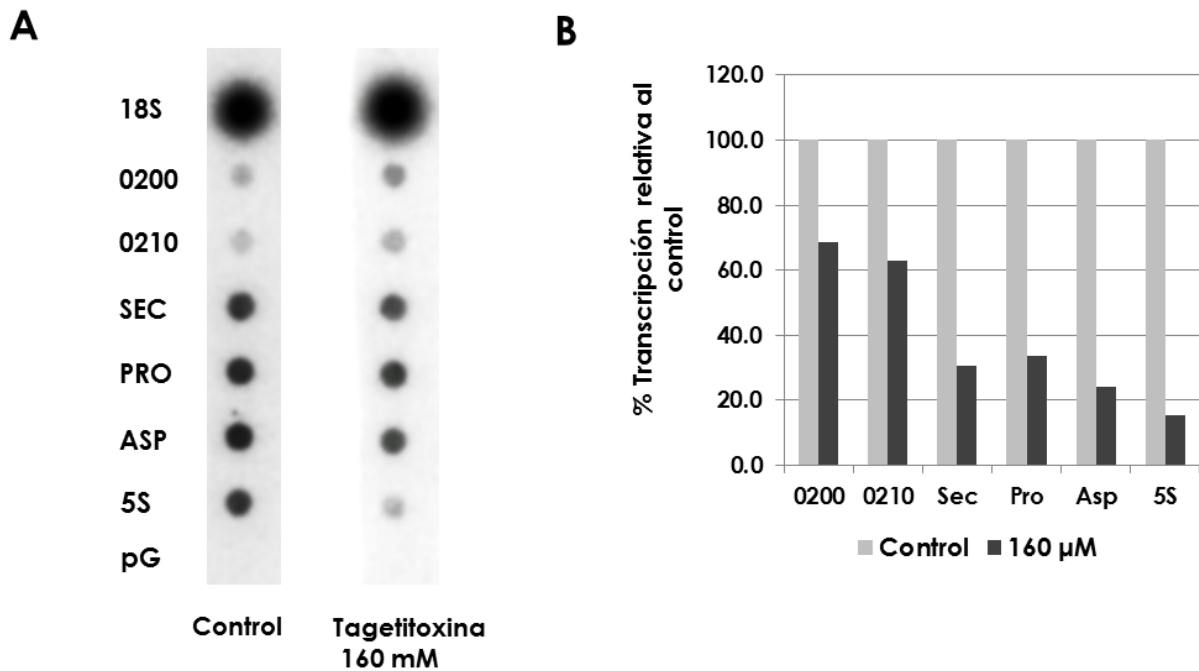


Figura 21. Efecto de la tagetitoxina en la transcripción del tRNA-Sec. (A) Experimento de *run-on* nuclear. Se permitió la elongación de los transcritos nacientes en presencia de Tagetitoxina (160 μM) y se hibridaron con membranas en las que se fijaron 2 μg de DNA clonado en pGEM-T Easy. Se incluyeron los DNAs correspondientes al tRNA-Sec, LmjF.06.0210 y LmjF.06.0200, además de varios genes transcritos por RNA Pol III (tRNA-Pro, tRNA-Asp y rRNA 5S) y por la RNA Pol I (rRNA 18S). Se incluyó también el vector pGEM-T Easy sin inserto (pG) como control. (B) Resultados del panel A cuantificados por densitometría. Las señales de cada gen obtenidas en la membrana control sin tagetitoxina se consideraron como 100%, las barras más oscuras muestran la reducción de la transcripción para cada gen. Todas las señales están normalizadas con la señal del rRNA 18S. Los valores mostrados corresponden al promedio de dos experimentos independientes.

6.13. Una caja tipo A y una caja B se localizan río arriba del tRNA-Sec en *L. major*

En tripanosomátidos, ciertos genes transcritos por la RNA Pol III (los snRNA o el RNA 7SL) utilizan como secuencias promotoras las cajas A y B intragénicas de un tRNA o tRNA-like situado río arriba (Palenchar, J. y Bellofatto, V., 2006). Se realizó una búsqueda informática para determinar si existían secuencias similares río arriba del tRNA-Sec de *L. major*. De manera interesante, se localizó una caja B clásica y una secuencia tipo caja A (A-like) río arriba del tRNA-Sec (Figura 22C).

La caja B clásica (GGTTCGATTCC) se localizó de -44 a -34 respecto a la secuencia codificante del tRNA-Sec, mientras que la caja tipo A (TGGCTGCAACGG) se localizó de -85 a -74. Estas cajas son similares a las que se encuentran río arriba de los snRNA. Estos hallazgos constituyen un indicio más de la participación de la RNA Pol III en la transcripción del tRNA-Sec. Cabe mencionar que secuencias similares no fueron localizadas río arriba del tRNA-Sec en *T. cruzi* (Figura 14) ni en *T. brucei*.

```

ACTGATACTG ATGCCGCGGT AGCATATCCG GATGCTGGCT GCAACGGATG CGCCTTTTCGA - 61
                                     A-like
CCCAGATTGG GTCGCGGGTT CGATTCCCGC GCTGAGGTGA TTTGTTTGAG AAAAGCGAGA - 1
                                     B
GCCGCGATGA GCTCAGCTGG TGCTGGGTGC GGGCTTCAA AAAAGCGAGA - 60
                                     A
ACAGCCGTT GATTCCGGCCT TCGTGGCGTG CCATTTTCCT TTTCGCTCCG AACCCACGCT 120
                                     B

```

Figura 22. Una caja B y una caja tipo A se localizaron río arriba del tRNA-Sec de *L. major*. Secuencia del tRNA-Sec y sus secuencias flanqueantes. Con sombra gris se identifica la secuencia codificadora del tRNA-Sec, río arriba se encontraron una caja B (B) y una caja tipo A (A-like), las cuales se muestran en negritas y subrayadas. También se señalan las cajas internas A y B, subrayadas y en negritas.

7. DISCUSIÓN

En este trabajo hemos analizado la transcripción del gen de tRNA-Sec en *L. major*. En eucariontes superiores, como *Xenopus*, ratón y humano, este gen es transcrito por la RNA Pol III, aunque las secuencias promotoras que dirigen su transcripción son inusuales en comparación con el resto de los tRNA (Kelly, V.P., *et al.*, 2005; Carbon, P. y Krol, A., 1991; Myslinski, E., *et al.*, 1991; Lee, B.J., *et al.*, 1989). En *L. major*, este gen también posee características atípicas en sus elementos promotores y nuestros resultados indican que tanto la RNA Pol II como la RNA Pol III están involucradas en su síntesis.

Nosotros hemos encontrado que el tRNA-Sec de *L. major*, localizado dentro de una unidad policistrónica del cromosoma 6 (Figura 11A, 18A), es transcrito junto con los genes codificadores de proteínas de ese policistrón y procesado mediante la adición del miniexón en su extremo 5' y de una cola de poliadeninas en el extremo 3', a semejanza de los mRNA en estos parásitos. Asimismo, identificamos las señales de procesamiento de dicho transcrito relacionadas con el *trans-splicing*, las cuales consisten típicamente en un dinucleótido AG en el sitio aceptor del miniexón precedido por un tracto de polipirimidinas (Figura 13C). Estos datos nos sugieren que la RNA Pol II está involucrada en la transcripción del tRNA-Sec, tal y como ya se reportó en *T. brucei* (Aeby, E., *et al.*, 2010). También demostramos que el tRNA-Sec en *T. cruzi* sufre la adición del miniexón en el extremo 5' y la poliadenilación en 3', por lo que el involucramiento de la RNA Pol II en la transcripción del tRNA-Sec parece ser una característica de la expresión génica en tripanosomátidos.

Durante mucho tiempo se creyó que en tripanosomátidos la adición del miniexón por *trans-splicing* y la poliadenilación eran fenómenos exclusivos del procesamiento de los pre-mRNA. Sin embargo, recientemente se ha encontrado que otros tipos de RNA no codificantes sufren este tipo de modificaciones en estos organismos. Por ejemplo, en *T. cruzi*, organismo en el que se ha demostrado que el precursor del rRNA 18S y su región 5'-ETS (5'-*External Transcribed Spacer*) forman un transcrito que sufre la adición del miniexón en la región 5' y la poliadenilación en la 3' (Mayer, M.G., *et al.*, 2012). Esto ocurre antes de que el rRNA 18S sea separado de la secuencia 5'-ETS para generar un rRNA maduro. El *trans-splicing* de este precursor ribosomal también se observó en *T. brucei*, *L. amazonensis* y *Crithidia fasciculata* (Mayer, M.G., *et al.*, 2012). Otro ejemplo de RNAs procesados por *trans-splicing* y poliadenilación, lo constituyen un tipo de ncRNA estadio-específicos en amastigotes de *L. infantum* (Dumas, C., *et al.*, 2006). Estos ncRNAs largos de entre 300 y 600 nucleótidos sufren la adición del miniexón en 5' y la poliadenilación en 3', aunque no se sabe si sufren

procesamientos adicionales. Estos ncRNAs largos se encontraron formando parte de un complejo ribonucleoproteico en el citosol de los parásitos, sin embargo, se desconoce su función. Estos hallazgos indican que el fenómeno de *trans-splicing* y la poliadenilación son modificaciones que cubren un rango amplio de RNAs en tripanosomátidos y que no es exclusivo de los mensajeros codificadores de proteínas. Aún no se sabe si estos procesamientos son importantes para la estabilidad y/o el transporte de RNAs no codificantes tal como lo son para los mRNA. En el alga *Chlamydomonas reinhardtii* (Komine, Y., et al., 2000) y en la bacteria *E. coli* (Mohanty, B.K., et al., 2012) se ha reportado la existencia de tRNAs poliadenilados en su extremo 3', modificación que afecta la estabilidad del transcrito. Tomando en cuenta estos hallazgos, se requerirían más investigaciones para determinar si el miniexón y la tracto de poliadeninas regulan la estabilidad del tRNA-Sec en tripanosomátidos, o si en *L. major* existe una estabilidad diferencial entre el tRNA-Sec poliadenilado y con miniexón y el tRNA-Sec que carece de estas modificaciones.

Por otra parte, dado que el tRNA-Sec se encuentra localizado dentro de una unidad policistrónica, nos preguntamos si la localización genómica de este tRNA determinaba su transcripción por la RNA Pol II. Para responder esto, se seleccionaron dos tRNA, el tRNA de asparto (tRNA-Asp, LmjF.24.TRNAASP.01) y el de prolina (tRNA-Pro, LmjF.24.TRNAPRO.01) que, de manera similar al tRNA-Sec, se encuentran ubicados dentro de sus unidades policistrónicas respectivas en el cromosoma 24 de *L. major* y se investigó si poseían el miniexón y si estaban poliadenilados. Se encontró que los transcritos del tRNA-Asp y el tRNA-Pro no sufren ni la adición del miniexón ni la poliadenilación, lo que sugiere que la ubicación genómica del tRNA-Sec dentro de una unidad policistrónica no es el único requisito para que éste sea transcrito por la RNA Pol II. A la fecha, no se ha reportado otro tRNA en tripanosomátidos que sufra modificaciones similares a las descritas aquí para el tRNA-Sec, sin embargo, esta posibilidad no puede descartarse.

Los tRNA canónicos son transcritos por la RNA Pol III y sus transcritos iniciales comienzan entre 5-12 nucleótidos río arriba del gen (Orioli, A., et al. 2012). Como ya se ha mencionado, se identificaron transcritos del tRNA-Sec que comenzaban 5-9 nucleótidos río arriba del gen, así como varios que comenzaban en el nucleótido +1 (Figura 13C), por lo que pensamos que éstos pudieran corresponder a transcritos sintetizados por la RNA Pol III. Para investigar si la RNA Pol III tenía alguna participación en la transcripción del tRNA-Sec se realizaron ensayos de *run-on* nuclear. Si todos los transcritos del tRNA-Sec eran sintetizados por la RNA Pol II, el nivel de transcripción de este tRNA sería similar al nivel de sus genes flanqueantes de proteínas (LmjF.06.0200 y LmjF.06.0210), mientras que, de haber un aporte transcripcional por la RNA Pol III a la síntesis del tRNA-Sec sería posible identificar un

nivel mayor de transcritos del tRNA-Sec reflejado en una señal de hibridación más alta en comparación con los genes de proteínas de la misma unidad policistónica. El primer ensayo de *run-on* nuclear (Figura 18) nos sugirió que el tRNA-Sec poseía un aporte transcripcional extra en comparación con los genes de proteínas flanqueantes, muy posiblemente por la RNA Pol III. Además, en este ensayo, la señal de hibridación del tRNA-Sec fue muy similar a la de otros genes transcritos por la RNA Pol III (por ejemplo, otros tRNAs y el rRNA 5S) (Figura 18).

Como se ha comentado, la transcripción dirigida por la RNA Pol III comienza muy cerca del sitio de inicio de transcripción y, en genes con regiones promotoras internas, cerca de las mismas. Por otra parte, la irradiación con dosis bajas de luz UV causa una acumulación de transcritos de los genes que poseen promotores internos como los tRNA (Martínez-Calvillo, S., *et al.*, 2004), por lo que se decidió realizar ensayos de *run-on* con células irradiadas con luz UV para evaluar si la transcripción del tRNA-Sec comienza cerca de su región codificante, de ser así, esto aportaría otro indicio de la participación de la RNA Pol III en la transcripción de dicho gen. De manera interesante, la señal del tRNA-Sec se mantiene alta ante las dosis crecientes de luz UV, tal cual ocurre con otros genes de tRNA y el rRNA 5S, lo que indica que la región promotora y el sitio de inicio de la transcripción del tRNA-Sec se encuentran cerca de su región codificante, una característica de las secuencias transcritas por la RNA Pol III. En contraste, la señal de los genes de proteínas flanqueantes disminuyen considerablemente con las dosis crecientes de luz UV, tal y como se esperaba para regiones alejadas de su sitio de inicio de la transcripción (Figura 19A).

Para obtener más evidencia de la participación de la RNA Pol III en la transcripción del tRNA-Sec se realizaron ensayos de *run-on* con inhibidores específicos de la transcripción por las RNA Polimerasas: α -amanitina y tagetitoxina. La RNA Pol II es altamente sensible a la α -amanitina, mientras que la RNA Pol III muestra una sensibilidad intermedia. Así, en los ensayos de *run-on* con α -amanitina la señal de la transcripción de un gen transcrito por la RNA Pol III es más resistente a la droga que los genes transcritos por la RNA Pol II (Martínez-Calvillo, S., *et al.*, 2004 y 2003). Los ensayos de *run-on* con α -amanitina nos mostraron que, a las concentraciones de 100 y 200 $\mu\text{g/ml}$, la transcripción del tRNA-Sec fue más resistente a la droga que con los genes flanqueantes (LmjF.06.0200 y LmjF.06.0210), los cuales son transcritos por la RNA Pol II. La resistencia del tRNA-Sec a esta droga fue muy similar a la de otros tRNA y a la del rRNA 5S. Esto representó una evidencia más de que en *L. major*, el tRNA-Sec es transcrito por la RNA Pol III. La resistencia de la transcripción del tRNA-Sec a la α -amanitina fue reportada originalmente en *Xenopus* (Lee, B.J., *et al.*, 1989). Sin embargo, en *T. brucei* el tRNA-Sec es

una secuencia altamente sensible a dicha droga, incluso más que el SL-RNA, indicando que en dicho organismo, la RNA Pol II dirige la transcripción del tRNA-Sec (Aeby, E., *et al.*, 2010).

Los resultados de los ensayos de *run-on* con tagetitoxina son congruentes con los resultados encontrados en los ensayos de α -amanitina. Cabe recordar que la tagetitoxina es un inhibidor específico de la RNA Pol III sin efecto sobre las otras dos RNA polimerasas (Steinberg, T.H., *et al.*, 1990). Así, en los ensayos de *run-on* con tagetitoxina la señal de la transcripción de un gen transcrito por la RNA Pol III disminuye ante la presencia de la droga, mientras que los genes transcritos por la RNA Pol II no ven afectada su transcripción. Los resultados de los *run-on* con esta droga nos mostraron que la transcripción del tRNA-Sec fue más sensible a esta droga que sus genes flanqueantes de proteínas, los cuales se vieron menos afectados por la tagetitoxina. La sensibilidad del tRNA-Sec a esta droga fue muy similar a la de otros tRNA y a la del rRNA 5S, indicando que el tRNA-Sec es transcrito por la RNA Pol III. Como se esperaba para un gen transcrito por la RNA Pol I, el rRNA 18S fue resistente a la tagetitoxina. La sensibilidad de los tRNAs y el rRNA 5S ante la tagetitoxina ya ha sido ampliamente estudiada en tripanosomátidos (Martínez-Calvillo, S., *et al.*, 2004; Crenshaw-Williams, K. y Bellofatto, V., 1999; Fantoni, A., *et al.*, 1994).

En eucariontes superiores, el tRNA-Sec posee un promotor híbrido pues las secuencias encargadas de dirigir la transcripción consisten en un elemento intragénico (caja B típica de los tRNA) y tres elementos extragénicos (típicos de los snRNA) (Hatfield, D. y Gladishev, V., 2002). En tripanosomátidos, los snRNA utilizan como secuencias promotoras las cajas A y B intragénicas de un tRNA o tRNA-like situado río arriba (Palenchar, J. y Bellofatto, V., 2006). Un análisis más profundo de las secuencias flanqueantes del tRNA-Sec nos indicó la existencia de una secuencia tipo caja A y de una caja B clásica río arriba del tRNA-Sec (Figura 22), hallazgos que sugerirían que secuencias tipo tRNA (*tRNA-like*) podrían estar involucradas en la transcripción del tRNA-Sec por la RNA Pol III, tal y como ocurre con los snRNAs y el RNA 7SL en estos organismos, aspecto que sería interesante abordar experimentalmente. Estas cajas no se localizaron en las secuencias flanqueantes del tRNA-Sec en *T. brucei* (Aeby, E., *et al.*, 2010), organismo en el que el tRNA-Sec es transcrito sólo por la RNA Pol II. En *T. cruzi* tampoco localizamos ninguna caja A o B río arriba del tRNA-Sec. Es posible que la ausencia de la caja tipo A y la caja B extragénicas situadas río arriba del tRNA-Sec de *T. brucei* explique por qué este tRNA no es transcrito por la RNA Pol III en este organismo.

Como se ha mencionado, la mayoría de los genes de tRNA en tripanosomátidos están agrupados junto con otros genes transcritos por la RNA Pol III, principalmente otros tRNA (Padilla-Mejía, N.E., *et al.* 2009). En el género *Leishmania* el *locus* del tRNA-Sec es completamente sinténico entre varias especies (Figura 12A), incluyendo las regiones intergénicas entre el tRNA-Sec y sus genes flanqueantes, por lo que es probable que la transcripción del tRNA-Sec ocurra también por ambas RNA Polimerasas en otras especies de *Leishmania*. En las bases de datos del genoma de cinetoplastidos (TriTrypDB, www.tritrypdb.org) se encuentran disponibles los datos de microarreglos hibridados con DNA proveniente de ensayos de inmunoprecipitación de la cromatina (ChIP-on-chip) para la proteína TBP (TATA-Binding Protein), subunidad del factor de la transcripción TFIIB (parte de la maquinaria de la RNA Pol III) (Oler, A., *et al.*, 2011). Estos datos muestran que en el *locus* del tRNA-Sec en *L. major* hay un pico de enriquecimiento de TBP que no está presente en el resto de los genes codificadores de proteínas en la misma unidad policistónica, lo cual nuevamente sugiere la presencia de la RNA Pol III en el tRNA-Sec. La unión de TBP también se observó en los *loci* de otros genes de tRNA.

El presente trabajo contribuye a ampliar el conocimiento sobre la transcripción del tRNA-Sec en tripanosomátidos. Como ya se ha mencionado, un trabajo previo con *T. brucei* demostró la participación de la RNA Pol II en la transcripción del tRNA-Sec (Aeby, E., *et al.*, 2010). Estos autores observan la adición de miniexón al extremo 5' del tRNA-Sec y la poliadenilación del 3' y demuestran que el tRNA-Sec forma parte de un policistón. Nuestros hallazgos indican que estos procesamientos ocurren también para el tRNA-Sec en *L. major*, indicándonos la participación de la RNA Pol II en su transcripción. En *T. brucei* se observó que la transcripción del tRNA-Sec es sensible a bajas concentraciones de α -amanitina y que la reducción de la actividad transcripcional de la RNA Pol II mediante RNAi disminuye la presencia del transcrito de tRNA-Sec, lo que no ocurre al bloquear la actividad de la RNA Pol III por RNAi, evidencias que sugieren que la RNA Pol II es la única encargada de la transcripción del tRNA-Sec en dicho organismo (Aeby, E., *et al.*, 2010). Las evidencias presentadas en este trabajo muestran que, a diferencia de lo que ocurre en *T. brucei*, la RNA Pol III sí participa en la transcripción del tRNA-Sec en *L. major*. Así, en este organismo, ambas RNA Pol, II y III, transcriben el tRNA-Sec.

Varias preguntas surgen a partir de nuestros resultados; por ejemplo, si en *L. major* el tRNA-Sec que posee el miniexón y el tracto de poliadeninas, ¿es procesado para generar un transcrito maduro o no?; ¿cuál es la RNA Pol que transcribe el tRNA-Sec en *T. cruzi*?, ¿es transcrito por la RNA Pol II como en *T. brucei*? ¿o por ambas RNA Pol, II y III, como en *L. major*?; o si en *L. major*, ¿las cajas extragénicas

situadas río arriba del tRNA-Sec son funcionales y, efectivamente, dirigen la transcripción por la RNA Pol III?. También sería interesante evaluar cuánto del tRNA-Sec celular es transcrito por la RNA Pol II y cuánto por la RNA Pol III.

El presente trabajo representa el primer reporte de un gen transcrito por dos RNA Pol en *L. major* y otros tripanosomátidos. De hecho, hay muy pocos reportes que indiquen la participación de más de una RNA Pol en la síntesis de algún RNA en otros organismos. Por ejemplo, se ha reportado que el gen del RNA H1 (componente de la RNAsa P), típicamente un gen transcrito por la RNA Pol III, puede ser transcrito tanto por la RNA Pol II como por la RNA Pol III en células de humano (Faresse, N.J., *et al.*, 2012). Esto fue demostrado por ensayos de ChIP-seq, mediante los cuales se observó la presencia de la maquinaria de ambas polimerasas, de la RNA Pol III (Brf2 y una subunidad de la polimerasa) y de la RNA Pol II (el factor de transcripción GTF2B y una subunidad de la polimerasa) en el *locus* del tRNA-Sec. El transcrito de la RNA H1 termina en los residuos de T localizados río abajo del gen, como ocurre con otros genes de RNA Pol III, sin embargo, también se identificó una caja 3', secuencia relacionada con la terminación de la transcripción en genes de snRNA, los cuales son transcritos por la RNA Pol II en vertebrados. Un segundo ejemplo se encuentra en levadura, organismo en el que la RNA Pol II transcribe dos secuencias localizadas en los extremos 5' y 3' del rRNA 35S. Esta transcripción se superpone con la de la RNA Pol I, la cual transcribe el precursor del rRNA 35S. Además, la inhibición de la transcripción de la RNA Pol II disminuye la del rRNA 5S, transcrito por la RNA Pol III (Mayan, M., 2013). Por otra parte, en *Arabidopsis* se ha reportado que la RNA Pol II coordina el reclutamiento de otras polimerasas necesarias para la biogénesis de siRNAs reguladores de la expresión génica (Zheng, B., *et al.*, 2009). Los ejemplos anteriores indican que la interacción entre las diferentes RNA Pol es más compleja de lo que se creía anteriormente. De hecho, en años recientes, se han realizado varios trabajos en células de humano que muestran la presencia de la RNA Pol II en *loci* de genes transcritos por la RNA Pol III, lo cual sugiere que la RNA Pol II podría influenciar la expresión de los genes de RNA Pol III (Oler, A., *et al.*, 2011; Raha, D., *et al.*, 2010; Listerman, I., *et al.*, 2007). Estos hallazgos sugieren que, las RNA Pol, más que actuar de manera individual, parecen estar coordinadas para regular la expresión génica y es posible que este tipo de coordinación exista también en *L. major* y en otros tripanosomátidos.

8. CONCLUSIONES

- El *locus* del tRNA-Sec es sinténico entre las diferentes especies de *Leishmania* y parcialmente sinténico entre las diferentes especies de *Trypanosoma*.
- En las secuencias río arriba del gen del tRNA-Sec en *L. major* fueron identificadas una caja B canónica y una secuencia similar a una caja A, las cuales podrían participar en el reclutamiento de la RNA Pol III.
- En *L. major* y *T. cruzi*, el tRNA-Sec sufre la adición del miniexón en su extremo 5' y la poliadenilación de su extremo 3'. Estos hallazgos sugieren que la RNA Pol II participa en la transcripción del tRNA-Sec en estos organismos.
- En *L. major*, existen transcritos del tRNA-Sec que comienzan 5-9 bases río arriba del gen o en el primer nucleótido del mismo. Estos transcritos podrían ser sintetizados por la RNA Pol III.
- En *L. major*, el tRNA-Sec está presente en transcritos de diferente peso molecular, lo que sugiere que este tRNA se transcribe como parte de la unidad policistrónica más grande del cromosoma 6 de este organismo. Esto nuevamente sugiere la participación de la RNA Pol II en la transcripción del tRNA-Sec.
- En *L. major*, el tRNA-Sec se transcribe aproximadamente 3.5 veces más que los genes de proteínas que lo flanquean. Dicha transcripción adicional comienza en una región promotora localizada muy cerca o dentro de su secuencia codificante, tal y como ocurre con los genes transcritos por la RNA Pol III.
- La transcripción del tRNA-Sec es resistente a la α -amanitina y sensible a la tagetitoxina en una manera similar a la de otros tRNA y el rRNA 5S. Estas evidencias sugieren fuertemente que la RNA Pol III está involucrada también en la transcripción del gen de tRNA-Sec.
- Dos RNA Polimerasas, RNA Pol II y RNA Pol III, son capaces de transcribir el tRNA-Sec en *L. major*.

9. REFERENCIAS

- Aeby, E. et al. 2010. tRNA^{Sec} is transcribed by RNA Polymerase II in *Trypanosoma brucei* but not in humans. *Nucleic Acids Res.* 38: 5833-5843.
- Aeby, E. et al. 2009. The canonical pathway for selenocysteine insertion is dispensable in Trypanosomes. *Proc Natl Acad Sci U S A.* 106: 5088-5092.
- Alvar, J., et al. 2012. Leishmaniasis worldwide and global estimates of its incidence. *PLoS One.* 7: e35671.
- Allmang, C. y Krol, A. 2006a. Selenoprotein synthesis: UGA does not end the story. *Biochimie.* 88: 1561-1571.
- Allmang, C. y Krol, A. 2006b. SECIS RNAs and K-turn binding proteins. A survey of evolutionary conserved RNA and protein motifs. En *Selenium: its molecular biology and role in human health.* Segunda edición. Editores: Hatfield, D., Berry, M., y Gladishev, V. Springer Publishers, USA. pp 51-61.
- Allmang, C., et al. 2009. The selenium and the selenoprotein pathway in eukaryotes: more molecular partners than anticipated. *Biochim Biophys Acta* 1790: 1415-1423.
- Antinori, S., et al. 2012. Leishmaniasis: new insights from old and neglected disease. *Eur J Clin Microbiol Infect Dis.* 31: 109-118.
- Bell, S.D. y Barry, J.D. 1995. Trypanosome nuclear factors which bind to internal promoter elements of tRNA genes. *Nucleic Acids Res.* 23: 3103-3110.
- Benz, C., et al. 2005. Messenger RNA processing sites in *Trypanosoma brucei*. *Mol Biochem Parasitol.* 143: 125-134.
- Berriman, M., et al. 2005. The genome of the African trypanosome *Trypanosoma brucei*. *Science.* 5733:416-422.
- Boucher, N., et al. 2002. A common mechanism of stage-regulated gene expression in *Leishmania* mediated by a conserved 3'-untranslated region element. *J Biol Chem.* 277: 19511-19520.
- Carbon, P. y Krol, A. 1991. Transcription of the *Xenopus laevis* selenocysteine tRNA(Ser)^{Sec} gene: a system that combines an internal B box and upstream elements also found in U6 snRNA genes. *EMBO J.* 10: 599-606.
- Carlson, B.A., et al. 2007. Selective restoration of the selenoprotein population in a mouse hepatocyte selenoproteinless background with different mutant selenocysteine tRNAs lacking Um34. *J Biol Chem.* 282: 32591-32602.
- Carlson, B.A., et al. 2005. Selective rescue of selenoprotein expression in mice lacking a highly specialized methyl group in selenocysteine tRNA. *J Biol Chem.* 280: 5542-5548.
- Campbell, D.A., et al. 2003. Transcription in kinetoplastid protozoa: why be normal?. *Microbes Infect.* 5: 1231-1240.
- Campbell, D.A., et al. 1989. Genomic organization of nuclear tRNA^{Gly} and tRNA^{Leu} genes in *Trypanosoma brucei*. *Mol Biochem Parasitol.* 37: 257-262.
- Cassago, A., et al. 2006. Identification of *Leishmania* selenoproteins and SECIS element. *Mol Biochem Parasitol.* 149: 128-134.
- Crenshaw-Williams, K. y Bellofatto, V. 1999. *In vivo* transcriptional analysis of the spliced leader RNA gene in the trypanosomatid *Leptomonas seymouri*. *Parasitol Res.* 85: 700-706.
- Crooks, G.E., et al. 2004. WebLog: a sequence logo generator. *Genome Res.* 14: 1188-1190.

- Diamond, A.M., et al. 1993. Dietary selenium affects methylation of the wobble nucleoside in the anticodon of selenocysteine tRNA[Ser]^{Sec}. *J Biol Chem.* 268:14215-14223.
- Dieci, G., et al. 2007. The expanding RNA polymerase III transcriptome. *Trends Genet.* 23: 614-622.
- Dikiy, A., et al. 2007. SelT, SelW, SelH and Rdx12: genomics and molecular insights into the functions of selenoproteins of a novel thioredoxin-like family. *Biochemistry.* 46: 6871-6882.
- Downing, T., et al. 2011. Whole genome sequencing of multiple *Leishmania donovani* clinical isolates provides insights into population structure and mechanism of drug resistance. *Genome Res.* 21: 2143-2156.
- Dumas, C., et al. 2006. A novel class of developmentally regulated noncoding RNAs in *Leishmania*. *Eukaryot Cell.* 5: 2033-2046.
- El-Sayed, N.M., et al. 2005. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science.* 309: 409-415.
- Fantoni, A., et al. 1994. RNA Polymerase III – mediated transcription of the trypanosoma U2 small nuclear RNA gene is controlled by both intragenic and extragenic regulatory elements. *Mol Cell Biol.* 14: 2021-2028.
- Faresse, N.J., et al. 2012. Genomic study of RNA Polymerase II and III SNAPc-bound promoters reveals a gene transcribed by both enzymes and a broad use of common activators. *PLoS Genet.* 8: e1003028.
- Florencio-Martinez, L.E., et al. 2010. Cellular analysis of host cell infection by different developmental stages of *Trypanosoma cruzi*. *Exp Parasitol.* 126: 332-336.
- Geslain, R. et al. *Trypanosoma* seryl-tRNA synthetase is a metazoan-like enzyme with high affinity for tRNA^{Sec}. *J Biol Chem.* 281: 38217-38225.
- Gonzalez-Flores, J., et al. 2013. The molecular biology of selenocysteine. *Biomol Concepts.* 4: 349-365.
- Grondal, E.J., et al. 1989. Characterization of the RNA Polymerases of *Trypanosoma brucei*: trypanosomal mRNAs are composed of transcripts derived from both RNA polymerase II and III. *EMBO J.* 8: 3383-3389.
- Günzl, A., et al. 1997. Transcription of the *Trypanosoma brucei* spliced leader RNA gene is dependent only on the presence of upstream regulatory elements. *Mol Biochem Parasitol.* 85: 67-76.
- Hamada, M., et al. 2001. Widespread use of TATA elements in the core promoters for RNA polymerases III, II and I in fission yeast. *Mol Cell Biol.* 21: 6870-6881.
- Han, S.J., et al. 2014. Characterization of Mammalian Selenoprotein O: a redox-active mitochondrial protein. *PLoS One.* 9: e95518.
- Hatfield, D. y Gladishev, V. 2002. How selenium has altered our understanding of the genetic code. *Mol Cell Biol.* 22: 3565-3576.
- Hernández-Rivas, R., et al. 1992. *Trypanosoma cruzi* 5S rRNA genes: molecular cloning, structure and chromosomal organization. *FEMS Microbiol Lett.* 71: 63-67.
- Ivens, A.C., et al. 2005. The genome of the kinetoplastid parasite, *Leishmania major*. *Science.* 309: 436-442.
- Ito, Y., et al. 2009. Crystal structure of human selenocysteine. *Nucleic Acids Res.* 37: 6259-6268.
- Kaye, P. y Scott, P. 2011. Leishmaniasis: complexity at the host-pathogen interface. *Nat Rev Microbiol.* 9: 604-615.
- Kelly, V.P., et al. 2005. The Distal Sequence Element of the Selenocysteine tRNA gene is a tissue-dependent enhancer essential for mouse embryogenesis. *Mol Cell Biol.* 25: 3658-3669.

- Kim, L.K., *et al.* 2000. Methylation of the ribosyl moiety at position 34 of selenocysteine tRNA^{[Ser]Sec} is governed by both primary and tertiary structure. *RNA*. 6: 1306-1315.
- Komine, Y., *et al.* 2000. Polyadenylation of three classes of chloroplast RNA in *Chlamydomonas reinhardtii*. *RNA*. 6: 598-607.
- Kossinova, O., *et al.* 2013. A novel insight into the mechanism of mammalian selenoprotein synthesis. *RNA*. 19: 1147-1158.
- Labunsky, V.M., *et al.* 2014. Selenoproteins: molecular pathways and physiological roles. *Physiol Rev*. 94: 739-777.
- Lee, B.J., *et al.* 1989. Transcription of *Xenopus* Selenocysteine tRNA^{Ser} gene is directed by multiple 5' extragenic regulatory elements. *J Biol Chem*. 264: 9696-9702.
- Lee, B.J., *et al.* 1987. Unique pathway of expression of an opal suppressor phosphoserine tRNA. *Proc Natl Acad Sci U S A*. 84: 6384-6388.
- Lenardo, M.J., *et al.* 1985. Characterization of the *Trypanosoma brucei* 5S ribosomal RNA gene and transcript: the 5S rRNA is a spliced-leader independent species. *Gene*. 35: 131-141.
- Listerman, I., *et al.* 2007. Extragenic accumulation of RNA Polymerase II enhances transcription by RNA Polymerase III. *PLoS Genet*. 3: e212.
- Lobanov, A., *et al.* 2006. Selenium metabolism in *Trypanosoma*: characterization of selenoproteomes and identification of a Kinetoplastid-specific selenoprotein. *Nucleic Acids Res*. 34: 4012-4024.
- Lu, J. y Holmgren, A. 2009. Selenoproteins. *J Biol Chem*. 284: 723-727
- Martinez-Calvillo, S., *et al.* 2010. Gene expression in Trypanosomatid parasites. *J Biomed Biotechnol*. 2010:525241. doi: 10.1155/2010/525241.
- Martinez-Calvillo, S., *et al.* 2007. Characterization of the RNA polymerases II and III complexes in *Leishmania major*. *Int J Parasitol*. 37: 491-502
- Martinez-Calvillo, S., *et al.* 2004. Transcription initiation and termination on *Leishmania major* chromosome 3. *Eukaryot Cell*. 3: 506-517.
- Martinez-Calvillo, S., *et al.* 2003. Transcription of *Leishmania major* chromosome 1 initiates in both directions within a single region. *Mol Cell*. 11: 1291-1299.
- Mayan, M. 2013. RNAP-II transcribes two small RNAs at the promoter and terminator regions of the RNAP-I gene in *Saccharomyces cerevisiae*. *Yeast*. 30: 25-32.
- Mayer, M.G., *et al.* 2012. Footprints of a trypanosomatid RNA world: pre-small subunit rRNA processing by spliced leader addition trans-splicing. *Mem Inst Oswaldo Cruz*. 107: 522-531.
- Mohanty, B.K., *et al.* 2012. Polyadenylation helps regulate functional tRNA levels in *Escherichia coli*. *Nucleic Acids Res*. 40: 4589-4603.
- Myler, P.J., *et al.* 1999. *Leishmania major* Friedlin Chromosome 1 has an unusual distribution of protein-coding genes. *Proc Natl Acad Sci U S A*. 96: 2902-2906.
- Myslinski, E., *et al.* 1991. Optimal tRNA^{(Ser)Sec} gene activity requires an upstream SPH motif. *Nucleic Acids Res*. 20: 203-209.
- Nakaar, V., *et al.* 1994. Upstream tRNAgenes are essential for expression of small nuclear and cytoplasmic RNA genes in trypanosomes. *Mol Cell Biol*. 14: 6736-6742.

- Nguyen, V.T., *et al.* 1996. *In vivo* degradation of RNA Polymerase II largest subunit triggered by α -amanitina. *Nucleic Acids Res.* 24: 2924-2929.
- Oler, A.J., *et al.* 2011. Human RNA Polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. *Nat Struct Mol Biol.* 17:620-629.
- Palenchar, J. y Bellofatto, V. 2006. Gene transcription in trypanosomes. *Mol Biochem Parasitol.* 146: 135-141.
- Padilla-Mejia, N.E., *et al.* 2013. Comparative genomics of *Leishmania* parasites. In *Comparative Genomics in Neglected Human Parasites*. Editores: Lopez-Camarillo, C. y Marchat, L.A. Nova Science Publishers, Inc., New York, New York, USA.
- Padilla-Mejia, N.E., *et al.* 2009. Gene organization and sequence analyses of transfer RNA genes in Trypanosomatid parasites. *BMC Genomics.* 10: 232-249.
- Park, J.M., *et al.* 1996. Analysis of the Selenocysteine tRNA(Ser)^{Sec} gene transcription in vitro using *Xenopus* oocyte extracts. *Biochem Biophys Res Commun.* 226: 231-236.
- Paule, M. y White, R. 2000. Transcription by RNA polymerases I and III. *Nucleic Acids Res.* 6: 1283-1298.
- Peacock, C.S. *et al.* 2007. Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. *Nat Genet.* 39: 839-847.
- Peng, Z. y Bateman, E. 2004. Analysis of the 5S rRNA gene promoter from *Acanthamoeba castellanii*. *Mol Microbiol.* 52: 1123-1132.
- Phizicky, E.M. y Alfonzo, J.D. 2010. Do all modifications benefit all tRNAs? *FEBS Lett.* 584: 265-271.
- Pitts, M.W., *et al.* 2013. Deletion of selenoprotein M leads to obesity without cognitive deficits. *J Biol Chem.* 288: 26121-26134.
- Raha, D., *et al.* 2010. Close association of RNA polymerase II and many transcription factors with Pol III genes. *Proc Natl Acad Sci U S A.* 107: 3639-3644.
- Reeves, M.A. y Hoffmann, P.R. 2009. The human selenoproteome: recent insights into functions and regulation. *Cell Mol Life Sci.* 66: 2457-2478.
- Rogers, M.B., *et al.* 2011. Chromosome and gene copy number variation allow major structural change between strain and species of *Leishmania*. *Genome Res.* 21: 2129-2142.
- Ruan, J.P., *et al.* 2004. Functional characterization of a *Trypanosoma brucei* TATA-binding protein-related factor points to a universal regulator of transcription in trypanosomes. *Mol Cell Biol.* 24:9610-9618.
- Sambrook, J y Russell, D. 2001. *Molecular Cloning. A laboratory manual*. Tercera edición. Cold Spring Harbor, New York, USA. Editorial Cold Spring Harbor Laboratory Press.
- Sculaccio, S.A., *et al.* 2008. Selenocysteine incorporation in kinetoplastid: selenophosphate synthetase (SELD) from *Leishmania major* and *Trypanosoma brucei*. *Mol Biochem Parasitol.* 162: 165-171
- Schramm, L. y Hernandez, N. 2002. Recruitment of RNA Polymerase III to its target promoters. *Genes Dev.* 16: 2593-2620.
- Schimanski, B., *et al.* 2005. Characterization of a multisubunit transcription factor complex essential for spliced-leader RNA gene transcription in *Trypanosoma brucei*. *Mol Cell Biol.* 25: 7303-7313.
- Siegel, T.N., *et al.* 2005. Systematic study of sequence motifs for RNA *trans*-splicing in *Trypanosoma brucei*. *Mol Cell Biol.* 25: 9586-9594.
- Steinberg, T.H., *et al.* 1990. Tagetitoxin: a new inhibitor of eukaryotic transcription by RNA Polymerase III. *J Biol Chem.* 265: 499-505.

- Stuart, K., *et al.* 2008. Kinetoplastids: related protozoan pathogens, different diseases. *J Clin Invest.* 118: 1301-1310.
- Thomas, S., *et al.* 2009. Histone acetylation mark origins of polycistronic transcription in *Leishmania major*. *BMC Genomics.* 10: doi:10.1186/1471-2164-10-152
- Van der Ploeg, L.H. 1986. Discontinuous transcription and splicing in trypanosomes. *Cell.* 47: 479-480.
- Varlamova, E.G., *et al.* 2011. New mammalian selenium-containing protein V: the search for protein partners. *Dokl Biochem Biophys.* 441: 255-257.
- Walczak, R. *et al.* 1998. An essential non-Watson-Crick base pair motif in 3'-UTR to mediate selenoprotein translation. *RNA.* 4: 74-84.
- Zheng, B., *et al.* 2009. Intergenic transcription by RNA Polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in *Arabidopsis*. *Genes Dev.* 23: 2850-2860.

APÉNDICE I: ARTÍCULO DE GRADO

Padilla-Mejía, N. E., Florencio-Martínez, L.E., Vizuet-de-Rueda, J.C., Cevallos, A.M., Hernández-Rivas, R., Manning-Cela, R. y Martínez-Calvillo, S. **2014**. The selenocysteine tRNA gene in *Leishmania major* is transcribed by both RNA Polymerase II and RNA Polymerase III. *Eukaryotic Cell*. 14:216-227

The Selenocysteine tRNA Gene in *Leishmania major* Is Transcribed by both RNA Polymerase II and RNA Polymerase III

Norma E. Padilla-Mejía,^{a,b} Luis E. Florencio-Martínez,^a Rodrigo Moreno-Campos,^{a,b} Juan C. Vizuet-de-Rueda,^{a,b} Ana M. Cevallos,^c Rosaura Hernández-Rivas,^d Rebeca Manning-Cela,^d Santiago Martínez-Calvillo^{a,b}

Unidad de Biomedicina, Facultad de Estudios Superiores Iztacala, Universidad Nacional Autónoma de México, Tlalnepantla, Estado de México, Mexico^a; Posgrado en Ciencias Biológicas, Facultad de Estudios Superiores Iztacala, Universidad Nacional Autónoma de México, Tlalnepantla, Estado de México, Mexico^b; Departamento de Biología Molecular y Biotecnología, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Mexico City, Mexico^c; Departamento de Biomedicina Molecular, Centro de Investigación y de Estudios Avanzados del IPN, Mexico City, Mexico^d

Eukaryotic tRNAs, transcribed by RNA polymerase III (Pol III), contain boxes A and B as internal promoter elements. One exception is the selenocysteine (Sec) tRNA (tRNA-Sec), whose transcription is directed by an internal box B and three extragenic sequences in vertebrates. Here we report on the transcriptional analysis of the tRNA-Sec gene in the protozoan parasite *Leishmania major*. This organism has unusual mechanisms of gene expression, including Pol II polycistronic transcription and maturation of mRNAs by *trans* splicing, a process that attaches a 39-nucleotide miniexon to the 5' end of all the mRNAs. In *L. major*, tRNA-Sec is encoded by a single gene inserted into a Pol II polycistronic unit, in contrast to most tRNAs, which are clustered at the boundaries of polycistronic units. 5' rapid amplification of cDNA ends and reverse transcription-PCR experiments showed that some tRNA-Sec transcripts contain the miniexon at the 5' end and a poly(A) tail at the 3' end, indicating that the tRNA-Sec gene is polycistronically transcribed by Pol II and processed by *trans* splicing and polyadenylation, as was recently reported for the tRNA-Sec genes in the related parasite *Trypanosoma brucei*. However, nuclear run-on assays with RNA polymerase inhibitors and with cells that were previously UV irradiated showed that the tRNA-Sec gene in *L. major* is also transcribed by Pol III. Thus, our results indicate that RNA polymerase specificity in *Leishmania* is not absolute *in vivo*, as has recently been found in other eukaryotes.

Eukaryotic cells use three different RNA polymerases (designated Pol I, Pol II, and Pol III) to transcribe their nuclear genome. Each of the three RNA polymerases transcribes specific classes of genes. Pol I synthesizes the large rRNA precursor, while Pol II produces mRNAs, most snRNAs, and snoRNAs (1, 2). Pol III synthesizes tRNAs, 5S rRNA, U6 snRNA, and several other small RNA molecules (3).

Leishmania and other trypanosomatids, such as *Trypanosoma cruzi* and *Trypanosoma brucei*, possess the three typical RNA polymerases (4, 5). However, these protozoan parasites present unusual mechanisms of gene expression, including Pol II polycistronic transcription (6, 7). The genomes of these organisms are organized into large polycistronic gene clusters (PGCs), i.e., 10s to 100s of protein-coding genes arranged sequentially on the same strand of DNA. Most chromosomes contain at least two PGCs, which can be either divergently transcribed (toward the telomeres) or convergently transcribed (away from the telomeres). Pol II transcription of an entire PGC initiates at a single region located upstream of the first gene of the cluster (8–10), and mature nuclear mRNAs are generated from primary transcripts by *trans* splicing and polyadenylation (11). *trans* Splicing is a process that adjoins a capped 39-nucleotide miniexon or spliced leader (SL) to the 5' termini of all the mRNAs (12, 13). The most conserved sequences needed for this process are an AG dinucleotide at the 3' splice site and an upstream pyrimidine-rich region (14–17). The *trans* splicing and polyadenylation of contiguous genes are linked, as the selection of a splice site for a gene influences the choice of a polyadenylation site for the upstream gene (18).

Pol III transcription in trypanosomatids is also atypical, as this enzyme transcribes all snRNA genes (not only U6 snRNA) in these organisms (19). In *Leishmania major*, most tRNA genes are orga-

nized into clusters of 2 to 10 genes that may contain other Pol III-transcribed genes, and the majority of the clusters are located at the boundaries of PGCs (20, 21). However, some tRNA genes are single genes that are not part of a cluster, and some of them are located inside PGCs. One such gene is the selenocysteine (Sec) tRNA (tRNA-Sec) gene, which is embedded into a large cluster of protein-coding genes on chromosome 6.

Selenocysteine, the 21st amino acid, is present in a group of proteins known as selenoproteins in bacteria, archaea, and eukarya (22). The presence of selenoproteins and all the machinery required for its synthesis has been demonstrated in *L. major* and other trypanosomatids (23, 24). Sec insertion is directed by a UGA codon, which is usually a stop codon, assisted by a specific structural signal located in the 3' untranslated region of the mRNA (25). A special tRNA species, tRNA-Sec, inserts Sec into nascent selenoproteins. Like other eukaryotic tRNA genes, tRNA-Sec is transcribed by Pol III in vertebrates. One of the typical characteristics of most tRNA genes is that their promoter sequences are

Received 21 October 2014 Accepted 23 December 2014

Accepted manuscript posted online 29 December 2014

Citation Padilla-Mejía NE, Florencio-Martínez LE, Moreno-Campos R, Vizuet-de-Rueda JC, Cevallos AM, Hernández-Rivas R, Manning-Cela R, Martínez-Calvillo S. 2015. The selenocysteine tRNA gene in *Leishmania major* is transcribed by both RNA polymerase II and RNA polymerase III. *Eukaryot Cell* 14:216–227. doi:10.1128/EC.00239-14.

Address correspondence to Santiago Martínez-Calvillo, scalv@campus.iztacala.unam.mx.

Copyright © 2015, American Society for Microbiology. All Rights Reserved.

doi:10.1128/EC.00239-14

internal and consist of two conserved elements: boxes A and B (26). Nevertheless, in *Xenopus laevis* and other vertebrates, transcription of tRNA-Sec genes is directed by an internal box B and three extragenic domains: a TATA box, a proximal sequence element, and an activator element (27, 28).

The consensus sequences of trypanosomatid tRNA promoter elements were determined by analyzing the sequences of all tRNA genes in *L. major*, *T. brucei*, and *T. cruzi* and comparing them to the sequences of boxes A and B from *Saccharomyces cerevisiae* (21, 29). Analysis of the promoter sequences from tRNA-Sec genes in trypanosomatids indicated that box A contains an additional A residue between bases 2 and 3 (TGAGCTCAGCTGG, in which the additional A residue is underlined) compared with the consensus sequence (TGGCTCAGCTGG) (21). A similar insertion was previously reported in tRNA-Sec genes from other organisms (30). Regarding box B, tRNA-Sec genes from trypanosomatids present two changes (CGTTCGATTCG, in which the two changes are underlined) compared to the highly conserved consensus sequence (GGTTCGANTCC): a C (instead of a G) at position 1 and a G (in place of a C) at position 11. In other species, the sequence of box B from tRNA-Sec is identical to the corresponding consensus sequence. Since the sequences of both internal control elements from tRNA-Sec genes in trypanosomatids differ from the corresponding consensus sequences, it was hypothesized that the synthesis of tRNA-Sec is regulated by external elements in these organisms (21). In fact, it was demonstrated that the tRNA-Sec gene in *T. brucei* is transcribed by Pol II (31).

In the work described here, we performed a transcriptional analysis of the tRNA-Sec gene in *L. major*. Our data show that, similar to the findings for *T. brucei*, the *L. major* tRNA-Sec gene is polycistronically transcribed by Pol II, generating transcripts that contain the minixon and a poly(A) tail. The same result was observed in *T. cruzi*, indicating that the participation of Pol II in the transcription of the tRNA-Sec gene might represent a hallmark of trypanosomatids. Interestingly, nuclear run-on data show that Pol III also transcribes the tRNA-Sec gene in *L. major*. Thus, our results indicate that RNA polymerase specificity in *Leishmania* is not absolute *in vivo* and reveal that the relationship between Pol II and Pol III is more complex than was previously believed.

MATERIALS AND METHODS

In silico analysis. Information for sequence analysis and synteny maps of species of *Leishmania* (*L. major*, *L. donovani*, *L. infantum*, *L. braziliensis*, *L. mexicana*, and *L. tarentolae*) and *Trypanosoma* (*T. cruzi*, *T. brucei*, *T. vivax* and *T. congolense*) was obtained from the TriTrypDB databases (version 8.1; <http://tritrypdb.org/tritrypdb/>). Sequence comparisons were performed using the ClustalW2 program (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>).

Culture of *L. major* and *T. cruzi*. 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, 1× L-glutamine) at 26°C and harvested in the mid-log phase. Epimastigotes of *T. cruzi* CL Brener were maintained in liver infusion-tryptose (LIT) medium containing 10% heat-inactivated fetal bovine serum, 50 IU/ml penicillin, 50 µg/ml streptomycin, and 0.025 mg/ml hemin at 28°C, as previously described (32).

5'-RACE analysis. 5' rapid amplification of cDNA ends (5'-RACE) experiments were performed with 5 µg of total RNA from *L. major* or *T. cruzi* with a kit from Life Technologies, Inc. For the *L. major* tRNA-Sec gene, the first-strand cDNA was synthesized with primer tRNA-SECgsp1

(5'-TGGCAGCCACGAAG) and the PCR amplifications were performed with the nested primer tRNA-SECgsp2 (5'-ATCGAACGGCTGTGAGAGCA) and the nested abridged anchor primer (AAP; 5'-GGCCACGCGTCTAGTACGGGIIIGGGIIIGGGIIIG). For the *L. major* tRNA-Asp gene (*LmjF.24.TRNAASP.01*), the first-strand cDNA was synthesized with primer Lm24Asp-GSP1 (5'-CCGGCCGGGAATTGAAC) and the PCR amplifications were carried out with primer Lm24Asp-GSP2 (5'-GGTCACCCGCGTGACAGGC) and the nested AAP. For the tRNA-Pro gene (*LmjF.24.TRNAPRO.01*) from *L. major*, the cDNA was produced with primer Lmc24-ProGSP1 (5'-GGGCCGCTAGGGGAATTGAA) and the PCR amplifications were performed with primer Lmc24-ProGSP2 (5'-TGACCTCCCGCACCCGAAG) and AAP. For the *T. cruzi* tRNA-Sec gene, the first-strand cDNA was synthesized with primer Nested(dT) (5'-CCTCTGAAGGTTACGGATCCACATCTAGATTTTTTTTTTTTTTTTTTTTTV) and two PCR amplifications were performed. The first PCR was performed with primers ME23 (5'-CGCTATTATTGATACAGTTCTG) and Tb-tRNA-Sec-GSP1 (5'-CACCACAAAGGCCGA). The second PCR amplification was performed using the first PCR product as the template and primers Tc-tRNA-Sec-GSP2 (5'-AACGGCTGCGAGTCCAAC) and ME23. The nested PCR products were cloned into the pGEM-T Easy vector (Promega) and sequenced.

RT-PCR assays. To map polyadenylation sites for the tRNA-Sec genes from *L. major* and *T. cruzi* and for the *LmjF.06.0210* gene from *L. major*, reverse transcription-PCR (RT-PCR) experiments were performed using cDNA prepared with oligonucleotide Nested(dT). For the *L. major* and *T. cruzi* tRNA-Sec genes, the first PCR was carried out with primers Lm-Sec1 (5'-AGCCGCGATGAGCTCAGCT) and B1 (5'-CCTCTGAAGGTTCCACGGAT) and the second PCR was done with primers Lm-Sec2 (5'-TGGGTGGGGCTTCAAAA) and B2 (5'-CACGGATCCACATCTAGAT). For the *LmjF.06.0210* gene, the first PCR was performed with primers Lmj06.0210PA1 (5'-AGCCGACTCATACTGCGGCT) and B1 and the second PCR was done with primers Lmj06.0210PA2 (5'-CTCATGCACTTTAAGCTGTA) and B2. The minixon addition site for the *LmjF.06.0210* gene from *L. major* was also located by RT-PCR. The cDNA was prepared with oligonucleotide Lmj06.0200.ME1 (5'-AGAGCGACCCCGTACTC) and PCR was performed with primers Lmj06.0200.ME2 (5'-ACGGAACCCAGAACGCAGGA) and minixon (5'-AACGCTATATAAGTATCAGTT). The final PCR products were cloned into the pGEM-T Easy vector (Promega) and sequenced. Pol III transcription termination sites were mapped by poly(A) tailing of total RNA. For this purpose, 2 µg of total RNA was mixed with 1 µl of 25 mM ATP, 2 µl of 5× poly(A) polymerase reaction buffer (USB), and 1,200 units of *S. cerevisiae* yeast poly(A) polymerase (USB) in a final volume of 20 µl. The mixture was incubated for 20 min at 37°C, and the reaction was terminated by heating at 65°C for 10 min. The cDNA was prepared with oligonucleotide Nested(dT). The first PCR was done with primers LmSec1 (5'-AGCCGCGATGAGCTCAGCT) and B1, and the second PCR was done with primers LmSec2 (5'-TGGGTGGGGCTTCAAAA) and B2.

Northern blot analysis. Total RNA was isolated using the TRI Reagent (Sigma), as specified by the supplier. RNA (12 µg) was separated on 10% polyacrylamide–8% urea gels. After electrophoresis, nucleic acids were transferred to Hybond N+ membranes (Amersham) by electroblotting using a Trans-Blot semidry system. The tRNA-Sec probe corresponds to a 190-bp fragment that includes the 88 bp of the tRNA-Sec gene plus 86 bp of the 5' upstream region, and that was labeled with [α -³²P]dCTP using a High Prime labeling system (Amersham). For tRNA-Lys, primer TRNA-Lysc03-GSP1 (5'-GCGCACTCCGTGGGG) was labeled with [γ -³²P]ATP by T4 kinase. Hybridizations were performed in 6× SSPE (60 mM Na₂HPO₄, 0.9 M NaCl, 6 mM EDTA), 5× Denhardt's reagent, and 1% SDS at 42°C. Washing was carried out at 55°C in 0.2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) and 0.1% SDS.

Molecular cloning into pGEM-T Easy. DNA fragments from *L. major* chromosome 6 were amplified by PCR and cloned into the pGEM-T Easy vector. *LmjF.06.0370* (521 bp) was amplified with oligonucleotides Lm06-0370-5' (5'-GAAGCGATGGACTGTTCTGG) and Lm06-0370-3' (5'-CGG

TCCTTGCTGCGAATATC), and *LmjF.06.0360* (539 bp) was amplified with primers Lm06-0360-5' (5'-CTCCTCTTCTGGACATTTGCT) and Lm06-0360-3' (5'-TTCCCTCCACTTGCAACATAG). *LmjF.06.0350* (500 bp) was amplified with oligonucleotides Lm06-0350-5' (5'-CGGTTCAACGG TAGTCTCTTC) and Lm06-0350-3' (5'-TAGAGGAACCAGAACGG GTAG), and *LmjF.06.0340* (503 bp) was amplified with oligonucleotides Lm06-0340-5' (5'-CCTCGCATACACCCTTTCGG) and Lm06-0340-3' (5'-CGCGAATGTACACCACACGG). *LmjF.06.0260* (500 bp) was amplified with oligonucleotides Lm06-0260-5' (5'-CCTGCTTGC TGTCGATGGTG) and Lm06-0260-3' (5'-TCGCCTCATCTCCTCT TGC), and *LmjF.06.0210* (503 bp) was amplified with oligonucleotides Lm06-0210-5' (5'-GCCGGAGACATTTGCGTAC) and Lm06-0210-3' (5'-CTATGGCGACGGGATCATC). *LmjF.06.0200* (547 bp) was amplified with oligonucleotides Lm06-0200-5' (5'-CCATCCCATGAC AAGAGC) and Lm06-0200-3' (5'-TGATAGTCGCTGTACTCGC), and *LmjF.06.0110* (554 bp) was amplified with oligonucleotides Lm06-0110-5' (5'-TTCCTACTACCGCATAGGGTTG) and Lm06-0110-3' (5'-C CATATCCAGATCCTGCATCC). The tRNA-Sec gene (524 bp) was amplified with oligonucleotides Lm-TRNASEC524-5' (5'-CCGGTGCCTT CATCAACTC) and Lm-TRNASEC524-3' (5'-GCGCATACGTTTCGGA GTCC), and the tRNA-Asp gene (369 bp) was amplified with oligonucleotides Lm24-TRNAASP-5' (5'-GAATGCGCTGCTGAGTC TCT) and Lm24-TRNAASP-3' (5'-GCGGTATGCGTGTGGTGTA). The tRNA-Phe gene (*LmjF.09.TRNAPHE.01*; 338 bp) was amplified with oligonucleotides Lm09-TRNAPHE-5' (5'-TTCATCCGCGCAAAGAGG) and Lm09-TRNAPHE-3' (5'-GGCCTCCACGTATTTCG), and the tRNA-Pro gene (352 bp) was amplified with oligonucleotides Lm24-TRNAPRO-5' (5'-GCGATCTCGTGGCTCTGGAG) and Lm24-TRNAPRO-3' (5'-ACAG CTCATCCAACGGGCGC). The tRNA-Tyr gene (*LmjF.36.TRNATYR.01*; 316 bp) was amplified with oligonucleotides Lm36-TRNATYR-5' (5'-AGT GCCGAGAAGTTCGACG) and Lm36-TRNATYR-3' (5'-TCGTCTCCGTT CCTGTTGC), and the 5S rRNA gene (*LmjF.15.5SrRNA.01*; 344 bp) was amplified with oligonucleotides Lm15-rRNA5S-5' (5'-GAAAGCATCTCTGTG GGTTGCA) and Lm15-rRNA5S-3' (5'-CCCGGGTCTCTGCAAATG). The 18S rRNA gene (*LmjF.27.rRNA.01*; 370 bp) was amplified with oligonucleotides Lm-rRNA18S-5' (5'-CGGCCTCTAGGAATGAAGG) and Lm-rRNA18S-3' (5'-CCCCTGAGACTGTAACCTC). The α -tubulin gene (338 bp) was amplified with primers alfa-tub-5' (5'-AGAAGTCCAAGCTCGGC TACAC) and alfa-tub-3' (5'-GTAGTTGATGCCGCACTTGAAG). After transformation of JM109 competent cells, plasmid DNA was purified from colorless colonies with NucleoSpin plasmid columns (Macherey-Nagel) as specified by the supplier. The identity of each insert was confirmed by sequencing using the T7 and SP6 primers.

Molecular cloning into M13 and preparation of single-stranded DNA. DNA fragments from tRNA-Sec, tRNA-Tyr, tRNA-Pro, 5S rRNA, *LmjF.06.0200*, *LmjF.06.0210*, 18S rRNA, and α -tubulin were excised from their respective constructs in pGEM-T Easy and cloned into M13mp18 and M13mp19 replicative-form DNA. The tRNA-Tyr, tRNA-Pro, 5S rRNA, *LmjF.06.0200*, *LmjF.06.0210*, and 18S rRNA sequences were cloned into M13 SacI and SphI sites. The tRNA-Sec and α -tubulin sequences were cloned into M13 Sall and SphI sites. After transformation of *Escherichia coli* JM109 cells (Promega), single-stranded DNA was purified from colorless plaques with QIAprep Spin M13 columns (Qiagen) as specified by the supplier.

Nuclear run-on assays. Nuclei were isolated from 2.5×10^8 *L. major* promastigotes by washing twice in phosphate-buffered saline (PBS), resuspending the cells in 4 ml ice-cold lysis buffer (10 mM Tris-HCl, pH 7.5, 3 mM CaCl₂, 2 mM MgCl₂), and adding NP-40 to a final concentration of 0.5%. Cells were transferred to a Dounce homogenizer and broken with 40 strokes; the nuclei were collected by centrifugation (1,400 \times g) and washed once with lysis buffer. RNA elongation was performed as described elsewhere (8, 33). Briefly, nuclei were resuspended in 100 μ l of run-on mix: 100 mM Tris-HCl (pH 7.5), 25% glycerol, 0.15 mM spermine, 0.5 mM spermidine, 2 mM dithiothreitol, 40 U RNasin (Promega), 2 mM MgCl₂, 4 mM MnCl₂, 50 mM NaCl, 50 mM KCl, 2 mM ATP, 2 mM

GTP, 2 mM UTP, 10 μ M CTP, and 250 μ Ci of [α -³²P]CTP (3,000 Ci/mmol; Amersham). The incubation was carried out for 6 min at 26°C, after which DNase I (10 U) was added. Incubation was continued for 5 min at 37°C and then stopped by the addition of 100 μ l of 10 mM Tris-HCl (pH 7.5), 10 mM EDTA, 1% SDS, and 100 μ g/ml proteinase K. After 15 min incubation at 37°C, RNA was extracted with phenol-chloroform and separated from free nucleotides by G-50 Sephadex chromatography. Labeled nascent RNA was hybridized to Hybond filters (Amersham) containing dots of 2 μ g of plasmid DNA (for the experiments whose results are shown in Fig. 3, 4, and 6) or single-stranded M13 DNA (for the experiments whose results are shown in Fig. 5). Hybridization was performed for 48 h at 50°C in 50% formamide, 5 \times SSC, 0.2% SDS, 4 \times Denhardt's reagent, and 100 μ g/ml salmon sperm DNA. Posthybridization washes were carried out in 0.1 \times SSC and 0.1% SDS at 65°C. In the assays carried out in the presence of UV light, promastigotes (in a total volume of 15 ml) were irradiated in petri dishes, with agitation, in a Stratilinker UV cross-linker (Stratagene). After irradiation, cells were incubated for at least 1.5 h at 28°C to allow the clearing of RNA polymerases engaged prior to irradiation. Elongation of nascent RNA in the presence of transcription inhibitors was performed by preincubating the nuclei with α -amanitin (Roche Molecular Biochemicals) or tagetitoxin (Tagetin; Epicentre Biotechnologies) for 15 min on ice in lysis buffer. The nuclei were next pelleted and resuspended in elongation buffer in the presence of the drugs.

RESULTS

Genomic location of the tRNA-Sec gene in *L. major* and other trypanosomatids. The majority of tRNA genes in *L. major* are clustered with other tRNA genes or other genes transcribed by Pol III (20, 21). These clusters are frequently located at the boundaries of PGCs, which are transcribed by Pol II. However, the tRNA-Sec gene is a single gene that is inserted into the largest PGC of chromosome 6 (Fig. 1A). In order to examine the genomic context of the tRNA-Sec genes in other species of *Leishmania*, we analyzed the genomic databases of several species recently sequenced: *L. infantum* (strain JPCM5), *L. braziliensis* (M2904) (34, 35), *L. mexicana* (U1103) (35), *L. donovani* (BPK282/0cl4) (36), and *L. tarentolae* (Parrot-TarII) (37). We found that in all the *Leishmania* species the tRNA-Sec gene is a single-copy gene embedded into the same PGC of chromosome 6 (data not shown). Thus, synteny of the tRNA-Sec locus is observed among *Leishmania* species.

In *T. brucei* there are two copies of the tRNA-Sec gene located on chromosome 9 (21, 38). Analyses of the genomic databases of *T. cruzi* (CL Brener) (39), *T. vivax* (Y486), and *T. congolense* (IL-3000) (40) showed that the presence of two tRNA-Sec genes is conserved across the genus *Trypanosoma* (Fig. 1B). Although the tRNA-Sec locus is syntenic in these species, several differences were observed; for instance, the Ser/Thr protein kinase gene (gene A in Fig. 1B) is duplicated in *T. cruzi*.

tRNA-Sec transcripts contain the minixon and a poly(A) tail in *L. major* and *T. cruzi*. The presence of atypical boxes A and B in the tRNA-Sec gene in *L. major* and its location inside a cluster of protein-coding genes suggest that the mechanisms that regulate the transcription of this gene might be different from those that regulate the transcription of other tRNA-Sec genes. To explore this possibility, we performed 5'-RACE experiments in order to map the transcription start sites of the tRNA-Sec gene. From this assay, we obtained two DNA bands of approximately 300 and 225 bp and a smaller and fainter band of about 150 bp contained within a smear (Fig. 2A). These bands were cloned and sequenced. Interestingly, analysis of several clones from the largest DNA band showed the presence of the minixon sequence at the 5' end, located 150 bp upstream of the tRNA-Sec gene (Fig. 2B), which

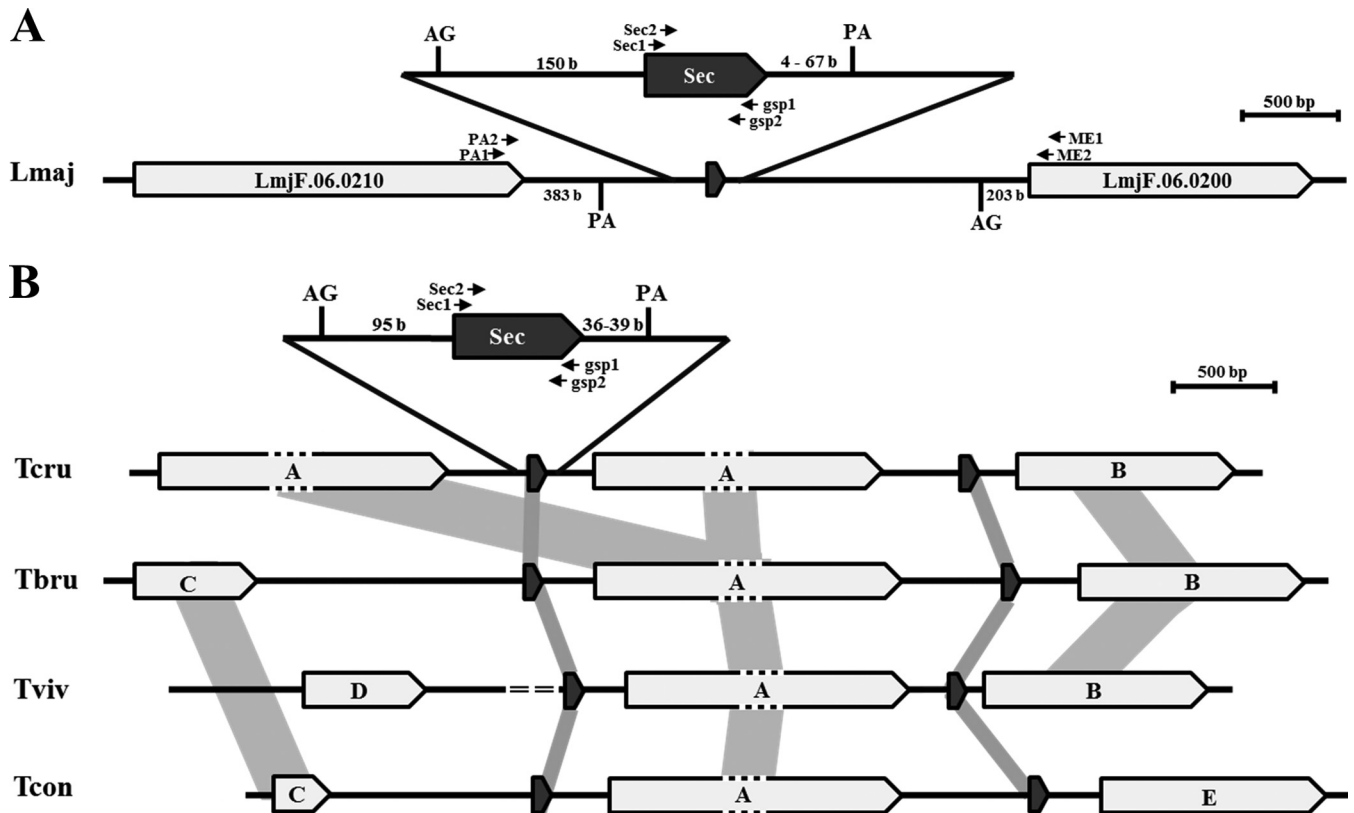


FIG 1 Synteny of the tRNA-Sec loci in trypanosomatids. The genomic context of the tRNA-Sec genes is shown for *L. major* (Lmaj) (A) and for *T. cruzi* (Tbru), *T. brucei* (Tbru), *T. vivax* (Tviv), and *T. congolense* (Tcon) (B). Orthologous genes are joined by gray lines. Enlargements of the tRNA-Sec gene in *L. major* and one of the tRNA-Sec genes in *T. cruzi* are shown, indicating the miniexon addition site (AG) and the poly(A) region (PA) of the tRNA-Sec genes. In *L. major*, the positions of the poly(A) region for *LmjF.06.0210* and AG for *LmjF.06.0200* are indicated. The positions of the primers used to map processing signals are also denoted with arrows. In panel B, genes A correspond to *TcCLB.506467.50* and *TcCLB.506467.40* (*T. cruzi*), *Tb927.9.2350* (*T. brucei*), *TvY486_0900740* (*T. vivax*), and *TcIL3000_0_10490* (*T. congolense*). Genes B correspond to *TcCLB.506467.29*, *Tb927.9.2320*, and *TvY486_0900730*. Genes C correspond to *Tb927.9.2390* and *TcIL3000_0_10500*. Gene D corresponds to *TvY486_0900750*, and gene E corresponds to *TcIL3000_0_10480*. b, number of bases.

suggested that the tRNA-Sec gene might be transcribed by Pol II, as if it were a protein-coding gene. Indeed, it was reported that the tRNA-Sec gene in the related parasite *T. brucei* is transcribed by Pol II (31).

Sequence analysis of clones from the 225-bp band showed non-*trans*-spliced transcripts whose sequences extended 68 or 86 bases upstream of the tRNA-Sec gene. These clones might represent intermediates between the miniexon-containing transcripts and the mature tRNA-Sec. Finally, sequence analysis of clones obtained from the 150-bp band revealed RNA molecules with a 5' end that corresponded to the mature tRNA-Sec, as well as some smaller molecules. Interestingly, the sequences of two of the clones extended 5 and 9 bp upstream of the mature tRNA's 5' end (see below) (Fig. 2B).

Since the tRNA-Sec transcripts were found to bear the miniexon sequence, we wondered if they also possessed a poly(A) tail. To determine if the tRNA-Sec transcripts in *L. major* are polyadenylated, an RT-PCR assay was performed. Only one band was identified and cloned. Analysis of the cloned sequences demonstrated that some tRNA-Sec molecules indeed contain a poly(A) tail, which was added at different positions between 4 and 67 bases downstream of the tRNA-Sec gene (Fig. 2B). Thus, our results show that some of the tRNA-Sec transcripts in *L. major* contain the miniexon at the 5' end and a poly(A) tail at the 3' end. Con-

sequently, the data suggest that the tRNA-Sec gene is polycistronically transcribed by Pol II and processed by *trans* splicing and polyadenylation.

If the tRNA-Sec is indeed polycistronically transcribed, it should be possible to demonstrate the presence of precursor transcripts. To examine this possibility, total RNA was hybridized with a tRNA-Sec probe. As shown in Fig. 2C, RNAs of approximately 80, 400, 650, and 1,000 bases were observed. The ~80-base band represents the mature tRNA-Sec gene, while the ~400-base RNA most likely corresponds to the transcript that contains the miniexon and a poly(A) tail of about 130 nucleotides. The larger bands represent longer RNA precursors. To demonstrate that the presence of large precursors was not a general feature of tRNA transcripts in *Leishmania*, a similar experiment was performed to detect transcripts using a tRNA-Lys-specific probe (transcribed by Pol III), finding only the mature tRNA and not any larger bands (Fig. 2D).

A typical AG sequence was identified as the miniexon attachment site in the transcript of the *L. major* tRNA-Sec gene (Fig. 2B). To gain insight into other processing signals in the polycistronic transcript, RT-PCR experiments were performed to localize the polyadenylation site of the gene (*LmjF.06.0210*) located upstream of the tRNA-Sec gene and the miniexon addition site of the downstream gene (*LmjF.06.0200*). The poly(A) tail of *LmjF.06.0210* was found 383 bases downstream of its stop codon, and the miniexon

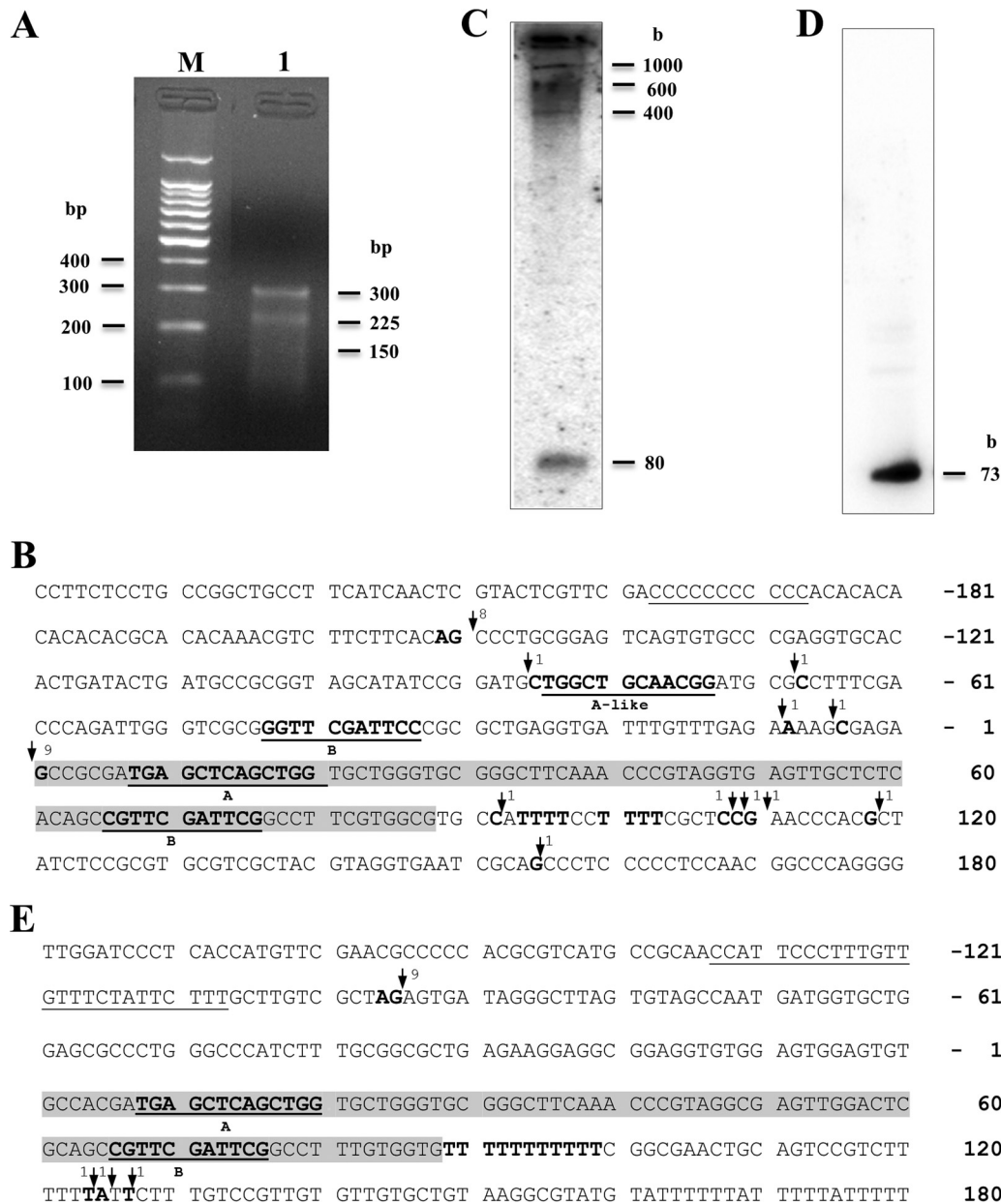


FIG 2 Mapping of processing sites and Northern blot analysis of the tRNA-Sec gene in *L. major*. (A) The final products of a 5'-RACE experiment of the tRNA-Sec gene were analyzed on a 1.5% agarose gel (lane 1). The size marker corresponds to a 1-kb ladder (Invitrogen) (lane M). The sequences of the tRNA-Sec gene and flanking regions for *L. major* (B) and *T. cruzi* (E) are shown. In both cases, the tRNA-Sec gene is highlighted in gray. Upstream of the genes, the minixion acceptor sites (AG) and the position of all clones found in the 5'-RACE analyses are shown in bold type and marked with an arrow with a number that indicates the total number of clones with the indicated sequences found at that position. Pyrimidine-rich regions found upstream of the minixion acceptor sites are underlined. Polyadenylation sites found downstream of the tRNA-Sec genes are shown in bold type and marked with an arrow with a number that indicates the total number of clones with the indicated sequences found at that position. The clusters of T residues located downstream from the tRNA-Sec genes are shown in bold type. Internal boxes A and B are labeled and shown in bold type and underlined. For *L. major*, upstream boxes B and A-like are also labeled and shown in bold type and underlined. (C and D) Hybridization of RNA with a tRNA-Sec-specific probe (C) and a tRNA-Lys-specific probe (D).

addition site of *LmjF.06.0200* was located 203 bases upstream of the start codon of the gene (Fig. 1A). Both sites correspond to processing sites reported on the TriTrypDB web page. However, the *LmjF.06.0210* poly(A) site that we found is not reported to be the dominant one, which was located 1,942 bases downstream of the stop codon. This clone might actually represent a dicistronic precursor that contains both *LmjF.06.0210* and the tRNA-Sec gene, taking into consideration the fact that the tRNA-Sec gene is

located 920 bases downstream of the *LmjF.06.0210* stop codon. Pyrimidine-rich regions, which are required for both *trans* splicing and polyadenylation, were located between the polyadenylation region of *LmjF.06.0210* and the minixion addition site of the tRNA-Sec gene and between the polyadenylation region of the tRNA-Sec gene and the minixion addition site of *LmjF.06.0200*. Therefore, canonical signals seem to regulate the processing of the polycistronic transcript of the tRNA-Sec gene in *L. major*.

Our results showed the participation of Pol II in the synthesis of the tRNA-Sec gene in *L. major*, similar to what has been reported in *T. brucei* (31). Therefore, it is likely that this might represent a more generalized feature of gene expression in trypanosomatid species. We thus determined if the tRNA-Sec gene of *T. cruzi* is also *trans* spliced and polyadenylated. As predicted, sequence analysis of several clones obtained by 5'-RACE and RT-PCR assays demonstrated the presence of the minixon 95 bases upstream of the tRNA-Sec gene and the presence of a poly(A) tail between 36 and 39 bases downstream of the tRNA-Sec gene (Fig. 1B and 2E). Thus, the participation of Pol II in transcription of the tRNA-Sec genes seems to be a hallmark of trypanosomatids.

The results of the experiments described above demonstrate the participation of Pol II in transcription of the tRNA-Sec gene in *L. major*. However, they do not exclude the possibility that the tRNA-Sec gene can be also transcribed by Pol III, considering that putative Pol III internal promoter elements are present in this gene. As mentioned above, two clones obtained in the 5'-RACE assay contained sequences that extended 5 and 9 bases upstream of the tRNA-Sec; since Pol III transcription of tRNAs usually starts a few nucleotides upstream of the tRNA gene, it is possible that these clones represent Pol III transcription start sites. Moreover, sequence analysis of the downstream region of tRNA-Sec showed the presence of two runs of 4 T residues separated by a CC sequence (Fig. 2B), a typical sequence that is associated with the termination of transcription by Pol III. Thus, to determine the presence of clones that end within the runs of T residues located downstream of the tRNA-Sec gene, RT-PCR was performed with total RNA that was poly(A) tailed *in vitro*. Interestingly, some clones were found to end within the first run of T residues (data not shown), suggesting the participation of Pol III in transcription of the tRNA-Sec gene in *L. major*. Transcription of tRNA-Sec genes in vertebrates is regulated by three sequence elements located upstream of the gene: a TATA box at about position -30, a proximal sequence element located at about position -70, and an activator element located at position -200 (28). These promoter elements are not contained in the upstream region of the tRNA-Sec gene in *L. major*. However, we found a consensus box B (GG TTCGATTCC) located 34 bases upstream of the tRNA-Sec gene and a box A-like sequence (TGGCTGCAACGG) located 74 bases upstream of the gene (Fig. 2B). These sequence elements are not present in the tRNA-Sec genes from *T. brucei* and *T. cruzi* (Fig. 2E and data not shown).

Nuclear run-on analysis of the tRNA-Sec locus. In order to analyze nascent transcripts from the tRNA-Sec gene and several protein-coding genes from the same polycistronic unit on chromosome 6, a nuclear run-on assay was carried out (Fig. 3A). The Pol II genes analyzed were *LmjF.06.0360*, *LmjF.06.0350*, *LmjF.06.0340*, *LmjF.06.0260*, *LmjF.06.0210*, *LmjF.06.0200*, and *LmjF.06.0110*. The *LmjF.06.0370* gene, from the adjacent PGC, was also analyzed. As controls, several Pol III genes (tRNA-Asp, tRNA-Phe, tRNA-Pro, tRNA-Tyr, and 5S rRNA), as well as the 18S rRNA gene, which is transcribed by Pol I, were included.

As shown in Fig. 3A, the intensity of the signal obtained with the tRNA-Sec gene was stronger than that obtained with the protein-coding genes that form part of the same polycistronic unit. Moreover, the tRNA-Sec signal was very similar to that observed with other tRNAs and 5S rRNA. The experiment was repeated five more times, and the results obtained each time were very similar. On average, the signal of the tRNA-Sec gene was ~3.1 times

higher than that of the genes that are located directly upstream and downstream of the tRNA-Sec gene and that are part of the same polycistronic unit (Fig. 3B and C). Although the lengths of the three probes analyzed in the experiment whose results are presented in Fig. 3B and C were very similar (524, 547, and 503 bp for tRNA-Sec, *LmjF.06.0200*, and *LmjF.06.0210*, respectively), the hybridization signals were divided by the probe length. The G+C contents of the three fragments were also very similar: 60.6, 68.1, and 63.2% for tRNA-Sec, *LmjF.06.0200*, and *LmjF.06.0210*, respectively. Thus, these results suggest that tRNA-Sec has another source of transcription, most likely Pol III.

As the promoter sequences for Pol III are located within the coding region, with or without some upstream elements, we wanted to assess if the transcription of tRNA-Sec is initiated at or near its locus. Therefore, we irradiated cells with different UV doses in order to arrest transcription elongation by producing pyrimidine dimers in the DNA (41). Irradiation of cells with incremental doses of UV light resulted in a progressive decrease of nascent transcripts, as assessed by nuclear run-on analysis. As expected, the signals of the protein-coding genes and the 18S rRNA gene were progressively reduced as the UV dose was increased, while the transcription of Pol III-dependent genes was relatively conserved. Interestingly, the pattern of reduction in the signal intensity of the tRNA-Sec gene was very similar to that observed with other tRNA genes and 5S rRNA and different from the pattern seen with the protein-coding genes from the same polycistronic unit (Fig. 4). For example, at a UV dose of 2.5 kJ/m², the hybridization signals for *LmjF.06.0210* and *LmjF.06.0200* were reduced by 54% and 66% of the value for the control, respectively, while that for the tRNA-Sec gene was reduced by only 15%, which is similar to the 10 to 34% reduction observed with the Pol III genes. Therefore, this result indicates that part of the transcription of the tRNA-Sec gene originates from a promoter region located very close to (or within) the tRNA-Sec gene.

It has been reported that Pol II and Pol III show different sensitivities to α -amanitin, since low concentrations of the drug inhibit Pol II transcription but have little impact on Pol III transcription (42, 43). Thus, to further examine the participation of Pol II and Pol III in the transcription of the tRNA-Sec gene, nuclear run-on experiments were carried out with different concentrations of α -amanitin. As in trypanosomatids the noncoding strands are transcribed at different levels, these experiments were performed using single-stranded M13 DNA to analyze the transcription of both strands of the tRNA-Sec gene independently. As shown in Fig. 5A and B, the lowest concentration of α -amanitin (100 μ g/ml) reduced Pol II-mediated transcription of *LmjF.06.0210* and *LmjF.06.0200* by 63 and 65% of the value for the control, respectively. In contrast, the hybridization signal for the tRNA-Sec gene was reduced by only 25%. The transcription of tRNA-Pro, 5S rRNA, and tRNA-Tyr was reduced by 27, 28, and 31%, respectively. Overall, the inhibition curve obtained with the tRNA-Sec gene indicates that the transcription of this gene is a little more sensitive to α -amanitin than the transcription of Pol III genes are but not as sensitive as the transcription of *LmjF.06.0210* and *LmjF.06.0200*. Therefore, these results add further evidence to support the participation of Pol III in the transcription of tRNA-Sec in *L. major*. As expected, the hybridization signal was severely reduced with the highest concentration of α -amanitin (400 μ g/ml) for all genes analyzed, with the exception of the 18S rRNA gene. As reported before in *L. major* (9), the transcription of α -tu-

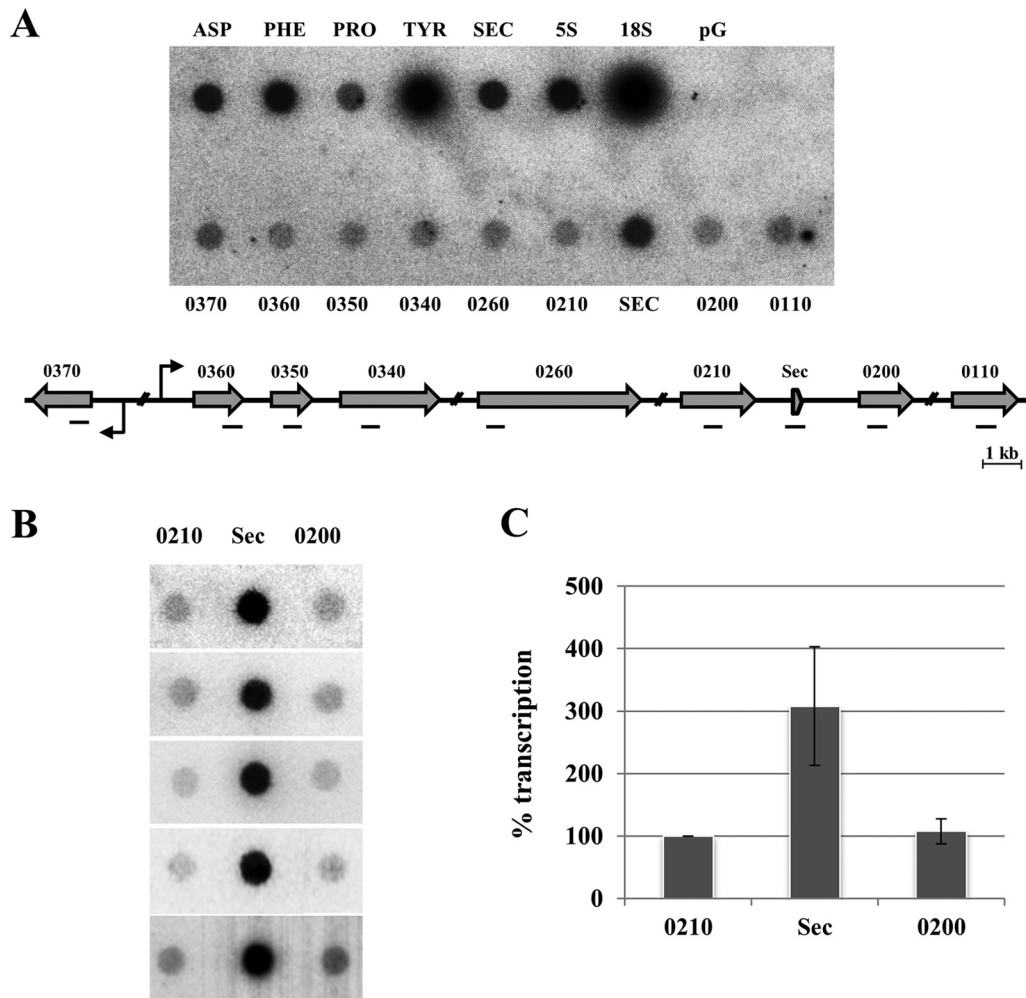


FIG 3 Nuclear run-on analysis of the tRNA-Sec gene. (A) (Top) Labeled nascent RNA from nuclei isolated from *L. major* promastigotes was hybridized to dot blots of double-stranded DNAs (2 μ g) cloned into the pGEM-T Easy vector. (Bottom) A genomic map of the protein-coding genes and the tRNA-Sec gene from *L. major* chromosome 6 is shown. The locations and sizes of the DNA fragments used in the nuclear run-on assays are indicated by the lines below the map. The Pol II genes analyzed were *LmjF.06.0370*, *LmjF.06.0360*, *LmjF.06.0350*, *LmjF.06.0340*, *LmjF.06.0260*, *LmjF.06.0210*, *LmjF.06.0200*, and *LmjF.06.0110*. Also, several genes transcribed by Pol III (tRNA-Asp, tRNA-Phe, tRNA-Pro, tRNA-Tyr, and 5S rRNA), as well as the 18S rRNA gene (transcribed by Pol I), were included. As a control, an empty vector (pG) was also analyzed. (B) Hybridization signals for *LmjF.06.0210*, *LmjF.06.0200*, and tRNA-Sec from five independent nuclear run-on experiments. (C) Relative rates of transcription for *LmjF.06.0210*, *LmjF.06.0200*, and tRNA-Sec. The dots shown in panel B were quantified and plotted, with the signal obtained with *LmjF.06.0210* considered to be 100%. Hybridization signals were divided by the probe length. One-way analysis of variance was performed with the data ($F_2 = 13.7$; $P < 0.001$); multiple comparisons by the Holm-Sidak method indicated that the values are significantly different between tRNA-Sec and *LmjF.06.0200* ($t = 4.46$; $P = 0.002$) and between tRNA-Sec and *LmjF.06.0210* ($t = 4.60$; $P = 0.002$).

bulin, which is encoded by 12 genes in this organism, was not as sensitive to α -amanitin as the transcription of single-copy protein-coding genes. Consistent with previous findings (9), strong antisense transcription was detected for Pol III-transcribed genes, the tRNA-Sec gene, and the 18S rRNA gene (Fig. 5A). It is worth noting that the antisense signals for *LmjF.06.0210* and *LmjF.06.0200* were higher than those observed for most protein-coding genes on chromosomes 1, 3, and 27 (8, 9, 44). However, the signal obtained with the empty vector control was unusually high in these experiments.

To confirm the involvement of Pol III in the transcription of the tRNA-Sec gene, nuclear run-on experiments were performed with tagetitoxin, a Pol III-specific inhibitor (Fig. 6). The hybridization signal for the tRNA-Sec gene was reduced by 70%, while the signals for tRNA-Pro and tRNA-Asp were reduced by 66 and

76%, respectively. A similar reduction was reported for tRNA genes from chromosomes 3 and 27 in *L. major* (9). 5S rRNA transcription was more sensitive to tagetitoxin, as it was reduced by 85%. Under the conditions used, the transcription of *LmjF.06.0210* and *LmjF.06.0200* was slightly reduced by 31 and 37%, respectively, while the transcription of the 18S rRNA was not affected. Therefore, these results confirm that Pol III, in addition to Pol II, transcribes the tRNA-Sec gene in *L. major*.

Are any other tRNA genes transcribed by Pol II in *L. major*? In addition to tRNA-Sec, the *L. major* genome contains two tRNA genes that are independent genes not associated with other Pol III-transcribed genes and that are inserted into clusters of protein-coding genes. The first one is a tRNA-Asp gene (*LmjF.24.TRNAASP.01*), and the second one is a tRNA-Pro gene (*LmjF.24.TRNAPRO.01*); these are located on different PGCs of

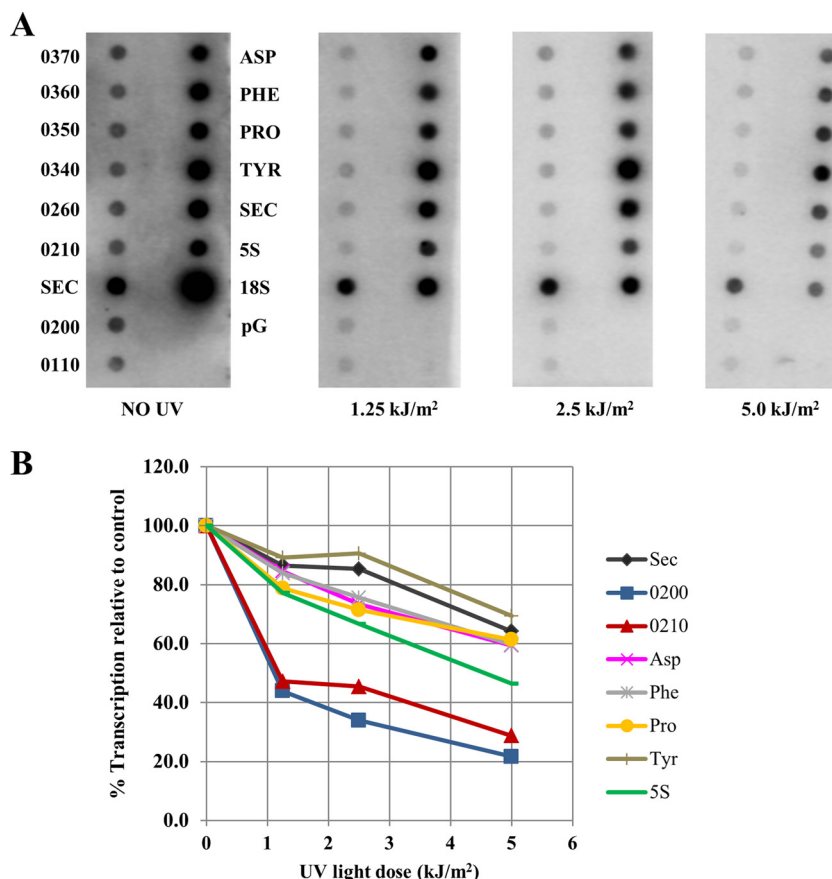


FIG 4 Effect of UV irradiation on transcription of the tRNA-Sec locus. (A) Results of nuclear run-on assays carried out with nuclei isolated from promastigotes that were irradiated with UV light at three different intensities (1.25, 2.5 and 5 kJ/m², as indicated below each panel). After irradiation, cells were incubated for 1.5 h at 28°C to allow the clearing of RNA polymerases engaged prior to irradiation. The genes analyzed were the same ones indicated in the legend to Fig. 3A. (B) The results shown in panel A and from an independent experiment were quantified, and the transcription signal for each gene relative to that for the nonirradiated control was plotted against UV dose. Values represent the means of two experiments.

chromosome 24 (21). To explore the possibility that these two tRNA genes are transcribed by Pol II as part of polycistronic clusters, 5'-RACE assays were performed under conditions similar to those used for the tRNA-Sec genes from *L. major* and *T. cruzi*. Sequencing of dozens of clones showed that none of them contained the miniexon. The majority of the clones started at the 5' end of the mature tRNA, but some of them were shorter. A noteworthy finding was that some of them contained sequences that extended 4 to 5 bases upstream of the tRNA gene, which may represent Pol III transcription start sites. The experiment was repeated several times using different conditions, but the results obtained were the same. Therefore, these results suggest that the single tRNA-Asp and tRNA-Pro genes located on chromosome 24 of *L. major* are not transcribed by Pol II. Consequently, the fact that a single tRNA gene is inserted into a PGC does not seem to be the only requirement for the participation of Pol II in its synthesis. Nevertheless, we cannot rule out the possibility that these two tRNA genes are transcribed by Pol II and that miniexon-containing transcripts are processed very rapidly, so that we were not able to detect them. Alternatively, it is also possible that other unidentified genes transcribed by Pol III are located in the proximities of the tRNA-Asp and tRNA-Pro genes, so that they are not actually single genes but part of clusters of Pol III genes. The intergenic

regions that flank these genes are large enough to contain other Pol III genes (3,212 and 654 bp for the tRNA-Pro gene and 390 and 1,642 bp for the tRNA-Asp gene).

DISCUSSION

It has been shown that the transcription of tRNA-Sec genes is different from the transcription of other tRNA genes in several species analyzed. *L. major* is not the exception, since the data presented here show that the tRNA-Sec gene in this protozoan parasite is transcribed by both Pol II and Pol III. The tRNA-Sec gene is located inside a polycistronic unit on chromosome 6, and Pol II transcribes it as if it were a protein-coding gene. Consequently, some transcripts of the tRNA-Sec gene contain the miniexon at the 5' end and a poly(A) tail at the 3' end, just like the mRNAs synthesized from the protein-coding genes that flank the tRNA-Sec gene. A characteristic AG dinucleotide was recognized as the miniexon addition site in the tRNA-Sec transcript from *L. major* (Fig. 2B), and pyrimidine-rich regions were located between the polyadenylation region of *LmjF.06.0210* and the miniexon addition site of tRNA-Sec and between the polyadenylation region of tRNA-Sec and the miniexon addition site of *LmjF.06.0200*, indicating that typical signals regulate the processing of the polycistronic transcript of the *L. major* tRNA-Sec gene. In *T. cruzi*, the

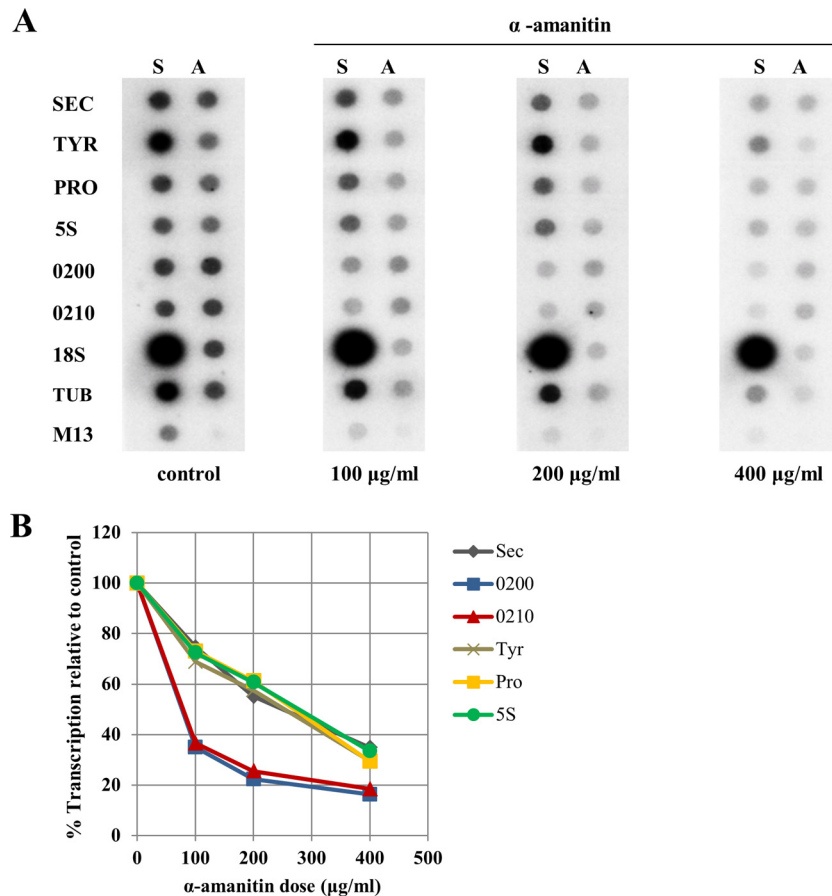


FIG 5 α -Amanitin sensitivity of tRNA-Sec gene transcription. (A) Nuclear run-on RNA was radiolabeled in the presence of different doses of α -amanitin (0, 100, 200, and 400 μ g/ml) and hybridized to filters containing single-stranded DNAs of tRNA-Sec, *LmjF.06.0210*, and *LmjF.06.0200*. The genes used as controls were tRNA-Tyr, tRNA-Pro, 5S rRNA (Pol III), α -tubulin (TUB; Pol II), 18S rRNA (Pol I), and the mp18 vector with no insert (M13). Lanes S, DNA complementary to the sense strand; lanes A, DNA complementary to the antisense strand. (B) Signals obtained with DNA complementary to the sense strand in panel A and from an independent experiment were quantified relative to the signal for the control (α -amanitin dose, 0 mg/ml) and plotted against the α -amanitin dose. Values represent the means from two experiments. All RNA levels were normalized to the level of 18S rRNA.

tRNA-Sec genes are also inserted into a Pol II polycistronic unit, and our data show that their transcripts are processed by *trans* splicing and polyadenylation, as has been reported for the tRNA-Sec genes in *T. brucei* (31). Hence, the involvement of Pol II in the transcription of tRNA-Sec genes seems to be a hallmark of trypanosomatids.

Several lines of evidence indicate that Pol III, in addition to Pol II, participates in transcription of the tRNA-Sec gene in *L. major*: (i) 5'-RACE and RT-PCR analysis showed the presence of transcripts whose sequences extended 5 to 9 bases upstream of the tRNA-Sec gene and transcripts that terminated in the T-residue tract located immediately downstream of the gene; these transcripts correspond to typical tRNA initiation and termination sites, respectively; (ii) in nuclear run-on analysis, the intensity of the signal observed with the tRNA-Sec gene was stronger than that obtained with the protein-coding genes that form part of the same polycistronic unit, and the signal of the tRNA-Sec gene was very similar to that observed with other tRNAs and 5S rRNA; (iii) nuclear run-on experiments with UV-irradiated cells showed that part of the transcription of the tRNA-Sec gene originates from a promoter region located very close to (or within) the tRNA-Sec gene; (iv) the α -amanitin inhibition curve for the tRNA-Sec gene

is more similar to the one observed for Pol III genes than to the one observed for Pol II genes; (v) nuclear run-on experiments with tagetitoxin demonstrated that hybridization signals for the tRNA-Sec gene and other tRNA genes were decreased to similar levels (66 to 76%), while transcription of *LmjF.06.0210* and *LmjF.06.0200* was slightly affected; and (vi) previously reported data obtained by chromatin immunoprecipitation with microarray technology showed the presence of a TATA-binding protein (TBP) enrichment peak in the tRNA-Sec region (45); TBP peaks were present on all Pol III-transcribed genes but not on snoRNA genes, which are polycistronically transcribed by Pol II with the neighboring protein-coding genes (45); although the TBP signal peak found in the region of the tRNA-Sec gene is not as high as the signals found in tRNA gene clusters, it clearly shows the binding of TBP to the tRNA-Sec gene. In *T. brucei* it has been shown that TBP (TRF4) is part of Pol III transcription factor TFIIB (as well as a Pol II transcription factor complex that includes SNAPc and TFIIA) (46).

It has been suggested that the signal observed in a nuclear run-on assay partially represents RNA abundance, as well as the transcription rate, reflecting the rapid rate of processing in *Leishmania*, and that factors such as the size and the G+C content of the fragment and the secondary structure within the RNA may

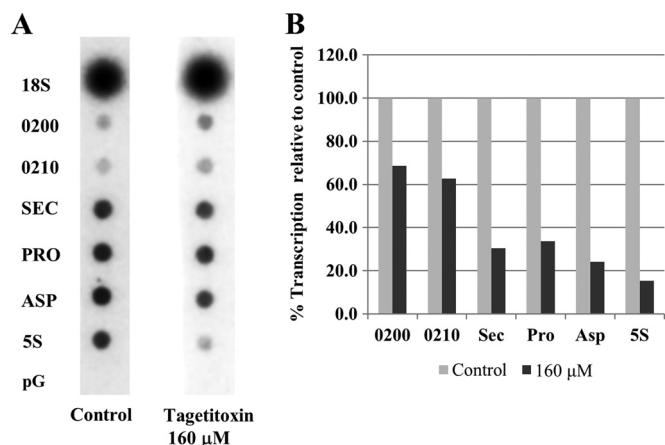


FIG 6 Effect of tagetitoxin on tRNA-Sec gene transcription. (A) Nuclear run-on RNA was radiolabeled in the presence of tagetitoxin at 160 μ M and hybridized to dot blots of double-stranded DNAs (2 μ g) cloned into the pGEM-T Easy vector. The genes analyzed were tRNA-Sec, *LmjF.06.0210*, and *LmjF.06.0200*. Several genes transcribed by Pol III (tRNA-Pro, tRNA-Asp, and 5S rRNA), as well as the 18S rRNA gene (transcribed by Pol I), were also included. As a control, an empty vector (pG) was also analyzed. (B) The results shown in panel A and from an independent experiment were quantified, and the transcription signal for each gene relative to that for the control was plotted. Values represent the means from two experiments. All RNA levels were normalized to the level of 18S rRNA.

also contribute to the observed differences in hybridization signals between fragments (8). However, the fact that the signals obtained with the tRNA-Sec gene were reproducibly higher than those obtained with other genes from the same polycistronic unit and the observation that the signals obtained with Pol III genes and tRNA-Sec gene are comparable (Fig. 3) strongly indicate that Pol III is also involved in transcription of the tRNA-Sec gene in *L. major*.

In *T. brucei*, tRNA-Sec precursors that contain the minixon and a poly(A) tail require further processing to create the mature tRNA-Sec (31). At present, we do not know whether the tRNA-Sec transcripts that contain the minixon and a poly(A) tail in *L. major* are additionally processed to originate mature tRNA-Sec, as in *T. brucei*, or whether they are degraded. The question of why in *T. brucei* the tRNA-Sec gene is transcribed only by Pol II while in *L. major* it is transcribed by both Pol II and Pol III then arises. As *T. brucei* contains two tRNA genes, it is possible that Pol II transcription of these genes generates the levels of tRNA-Sec that the cell requires, while *L. major*, which has a single tRNA-Sec gene, might need the additional participation of Pol III to obtain higher levels of tRNA-Sec. Since the tRNA-Sec genes in both species contain identical internal boxes A and B and they contain a T-residue tract at the 3' end of the genes, it is possible that the upstream box B and box A-like elements, which are not found in *T. brucei*, help recruit Pol III to the tRNA-Sec gene in *L. major*.

To our knowledge, this is the first report of a gene transcribed by more than one RNA polymerase in trypanosomatids. Only a few examples have been reported in other organisms. A recent study of human cells by chromatin immunoprecipitation sequencing revealed that the gene encoding the H1 RNA (a component of RNase P), which is considered a type 3 Pol III gene, can be transcribed by either Pol III or Pol II *in vivo* (47). This gene contains a T-residue tract, which is the typical Pol III termination signal, and a 3' box, a Pol II termination signal present on

snRNAs. Most transcripts that were detected ended at the T-residue tract, suggesting that the RNA transcribed by Pol II is highly unstable. Another example is the human *c-myc* promoter, which is transcribed by Pol II and Pol III *in vitro* and *in vivo* (48). Also, it was recently reported in human that Pol III is able to accurately initiate transcription from Pol II core promoters in *in vitro* transcription assays (49). The ratio of DNA template to nuclear extract determines whether Pol II, Pol III, or both enzymes start transcription from the Pol II promoter, indicating that polymerase specificity is not constant but instead depends on transcription conditions (49). Another study in *S. cerevisiae* yeast showed that Pol II transcribes two sequences found in the promoter and the terminator regions of the rRNA genes that overlap the sequences of the 35S rRNA precursor transcribed by Pol I (50). Inhibition of Pol II transcription decreases Pol III transcription of the 5S rRNA gene. Thus, these findings reveal a complex relationship among all three RNA polymerases in the rRNA loci from yeast (50).

Together, our data support the conclusion that the tRNA-Sec gene in *L. major* is transcribed by Pol II and by Pol III. It would be interesting to measure the ratio of Pol II and Pol III transcripts and to explore if such a ratio changes in different stages of *L. major* growth or under different growth conditions. Our results suggest that the tRNA-Sec gene in *T. cruzi* is transcribed by Pol II. It remains to be investigated whether it is also transcribed by Pol III, as in *L. major*, or if it is transcribed only by Pol II, as in *T. brucei*. Similar to the situation in other organisms, the relationship between Pol II and Pol III in trypanosomatids seems to be more complex than was originally estimated. Further investigations are required to gain insight into this very important area of gene expression.

ACKNOWLEDGMENTS

This work was supported by grant 128461 from CONACyT, grants IN203909 and IN210712 from PAPIIT (UNAM), and project 53 from PAPCA 2013 (FES Iztacala) to S. Martínez-Calvillo and by grants 132312 and 139898 from CONACyT to R. Manning-Cela. N. E. Padilla-Mejía was a Ph.D. student in Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México, and she was the recipient of a doctoral fellowship from CONACyT (fellowship 207150, CVU 216026).

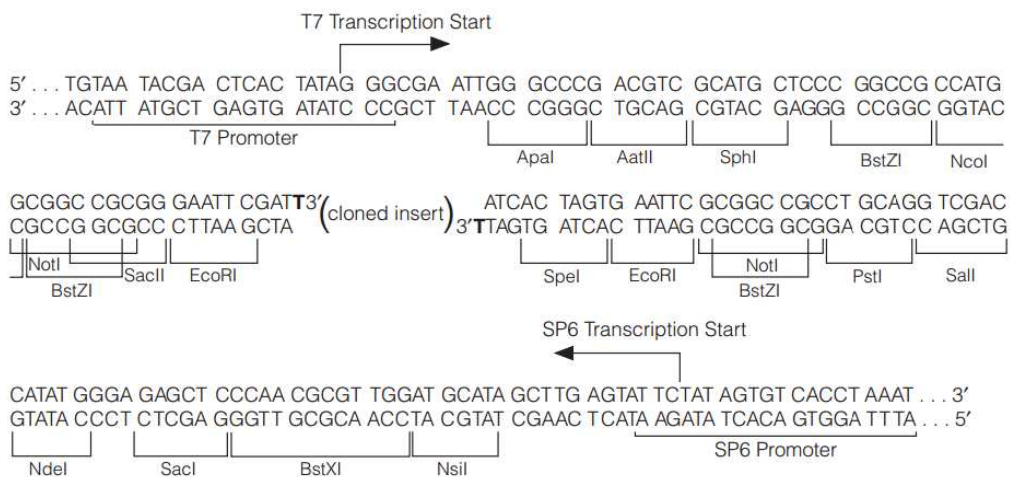
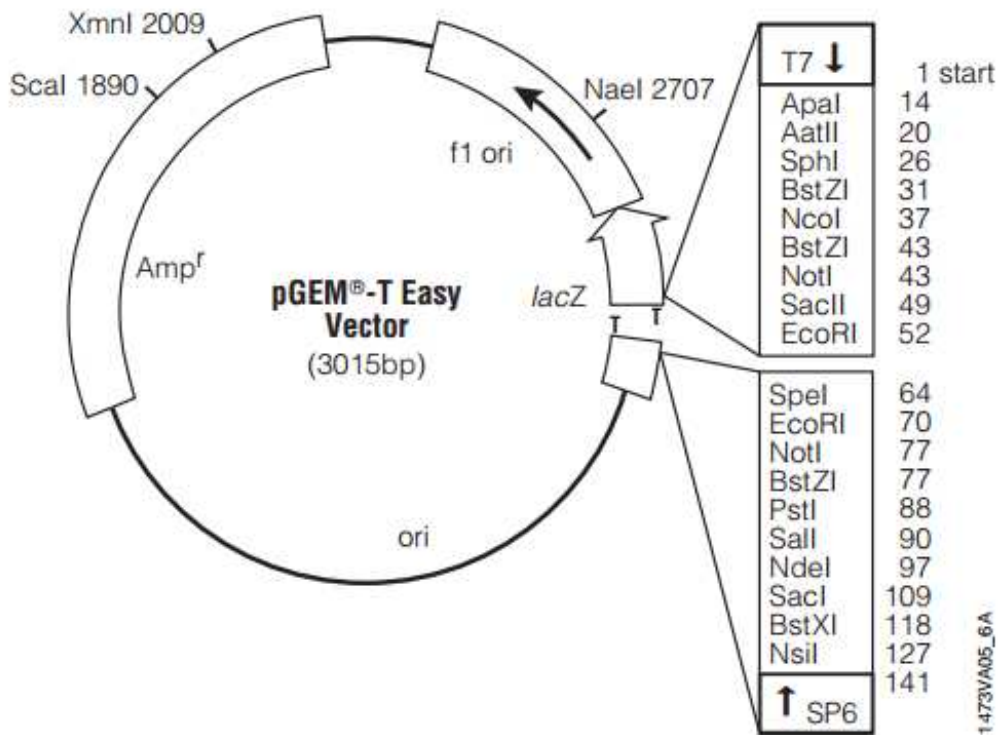
REFERENCES

- Grummt I. 2003. Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus. *Genes Dev* 17:1691–1702. <http://dx.doi.org/10.1101/gad.1098503R>.
- Liu X, Bushnell DA, Kornberg RD. 2013. RNA polymerase II transcription: structure and mechanism. *Biochim Biophys Acta* 1829:2–8. <http://dx.doi.org/10.1016/j.bbagr.2012.09.003>.
- White RJ. 2011. Transcription by RNA polymerase III: more complex than we thought. *Nat Rev Genet* 12:459–463. <http://dx.doi.org/10.1038/nrg3001>.
- Gronal EJ, Evers R, Kosubek K, Cornelissen AW. 1989. Characterization of the RNA polymerases of *Trypanosoma brucei*: trypanosomal mRNAs are composed of transcripts derived from both RNA polymerase II and III. *EMBO J* 8:3383–3389.
- Martínez-Calvillo S, Saxena A, Green A, Leland A, Myler PJ. 2007. Characterization of the RNA polymerase II and III complexes in *Leishmania major*. *Int J Parasitol* 37:491–502. <http://dx.doi.org/10.1016/j.ijpara.2006.11.019>.
- Campbell DA, Thomas S, Sturm NR. 2003. Transcription in kinetoplast protozoa: why be normal? *Microbes Infect* 5:1231–1240. <http://dx.doi.org/10.1016/j.micinf.2003.09.005>.
- Martínez-Calvillo S, Vizuet-de-Rueda JC, Florencio-Martínez LE, Manning-Cela RG, Figueroa-Angulo EE. 2010. Gene expression in trypanosomatid parasites. *J Biomed Biotechnol* 2010:525241. <http://dx.doi.org/10.1155/2010/525241>.

8. Martinez-Calvillo S, Yan S, Nguyen D, Fox M, Stuart K, Myler PJ. 2003. Transcription of *Leishmania major* Friedlin chromosome 1 initiates in both directions within a single region. *Mol Cell* 11:1291–1299. [http://dx.doi.org/10.1016/S1097-2765\(03\)00143-6](http://dx.doi.org/10.1016/S1097-2765(03)00143-6).
9. Martinez-Calvillo S, Nguyen D, Stuart K, Myler PJ. 2004. Transcription initiation and termination on *Leishmania major* chromosome 3. *Eukaryot Cell* 3:506–517. <http://dx.doi.org/10.1128/EC.3.2.506-517.2004>.
10. Siegel TN, Hekstra DR, Kemp LE, Figueiredo LM, Lowell JE, Fenyo D, Wang X, Dewell S, Cross GA. 2009. Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. *Genes Dev* 23:1063–1076. <http://dx.doi.org/10.1101/gad.1790409>.
11. Michaeli S. 2011. Trans-splicing in trypanosomes: machinery and its impact on the parasite transcriptome. *Future Microbiol* 6:459–474. <http://dx.doi.org/10.2217/fmb.11.20>.
12. Murphy WJ, Watkins KP, Agabian N. 1986. Identification of a novel Y branch structure as an intermediate in trypanosome mRNA processing: evidence for trans splicing. *Cell* 47:517–525. [http://dx.doi.org/10.1016/0092-8674\(86\)90616-1](http://dx.doi.org/10.1016/0092-8674(86)90616-1).
13. Sutton RE, Boothroyd JC. 1986. Evidence for trans splicing in trypanosomes. *Cell* 47:527–535. [http://dx.doi.org/10.1016/0092-8674\(86\)90617-3](http://dx.doi.org/10.1016/0092-8674(86)90617-3).
14. Benz C, Nilsson D, Andersson B, Clayton C, Guilbride DL. 2005. Messenger RNA processing sites in *Trypanosoma brucei*. *Mol Biochem Parasitol* 143:125–134. <http://dx.doi.org/10.1016/j.molbiopara.2005.05.008>.
15. Siegel TN, Tan KS, Cross GA. 2005. Systematic study of sequence motifs for RNA trans splicing in *Trypanosoma brucei*. *Mol Cell Biol* 25:9586–9594. <http://dx.doi.org/10.1128/MCB.25.21.9586-9594.2005>.
16. Kolev NG, Franklin JB, Carmi S, Shi H, Michaeli S, Tschudi C. 2010. The transcriptome of the human pathogen *Trypanosoma brucei* at single-nucleotide resolution. *PLoS Pathog* 6:e1001090. <http://dx.doi.org/10.1371/journal.ppat.1001090>.
17. Nilsson D, Gunasekera K, Mani J, Osteras M, Farinelli L, Baerlocher L, Roditi J, Ochsenreiter T. 2010. Spliced leader trapping reveals widespread alternative splicing patterns in the highly dynamic transcriptome of *Trypanosoma brucei*. *PLoS Pathog* 6:e1001037. <http://dx.doi.org/10.1371/journal.ppat.1001037>.
18. LeBowitz JH, Smith HQ, Rusche L, Beverley SM. 1993. Coupling of poly(A) site selection and trans-splicing in *Leishmania*. *Genes Dev* 7:996–1007. <http://dx.doi.org/10.1101/gad.7.6.996>.
19. Gunzl A, Tschudi C, Nakaar V, Ullu E. 1995. Accurate transcription of the *Trypanosoma brucei* U2 small nuclear RNA gene in a homologous extract. *J Biol Chem* 270:17287–17291. <http://dx.doi.org/10.1074/jbc.270.29.17287>.
20. Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, Sisk E, Rajandream MA, Adlem E, Aert R, Anupama A, Apostolou Z, Attipoe P, Bason N, Bausner C, Beck A, Beverley SM, Bianchetti G, Borzym K, Bothe G, Bruschi CV, Collins M, Cadag E, Ciarloni L, Clayton C, Coulson RM, Cronin A, Cruz AK, Davies RM, De GJ, Dobson DE, Duesterhoeft A, Fazelina G, Fosker N, Frasch AC, Fraser A, Fuchs M, Gabel C, Goble A, Goffeau A, Harris D, Hertz-Fowler C, Hilbert H, Horn D, Huang Y, Klages A, Knights A, Kube M, Larke N, Litvin L, et al. 2005. The genome of the kinetoplastid parasite, *Leishmania major*. *Science* 309:436–442. <http://dx.doi.org/10.1126/science.1112680>.
21. Padilla-Mejia NE, Florencio-Martinez LE, Figueroa-Angulo EE, Manning-Cela RG, Hernandez-Rivas R, Myler PJ, Martinez-Calvillo S. 2009. Gene organization and sequence analyses of transfer RNA genes in Trypanosomatid parasites. *BMC Genomics* 10:232. <http://dx.doi.org/10.1186/1471-2164-10-232>.
22. Hatfield DL, Gladyshev VN. 2002. How selenium has altered our understanding of the genetic code. *Mol Cell Biol* 22:3565–3576. <http://dx.doi.org/10.1128/MCB.22.11.3565-3576.2002>.
23. Cassago A, Rodrigues EM, Prieto EL, Gaston KW, Alfonso JD, Iribar MP, Berry MJ, Cruz AK, Thiemann OH. 2006. Identification of *Leishmania* selenoproteins and SECIS element. *Mol Biochem Parasitol* 149:128–134. <http://dx.doi.org/10.1016/j.molbiopara.2006.05.002>.
24. Lobanov AV, Gromer S, Salinas G, Gladyshev VN. 2006. Selenium metabolism in *Trypanosoma*: characterization of selenoproteomes and identification of a Kinetoplastida-specific selenoprotein. *Nucleic Acids Res* 34:4012–4024. <http://dx.doi.org/10.1093/nar/gkl541>.
25. Allmang C, Krol A. 2006. Selenoprotein synthesis: UGA does not end the story. *Biochimie* 88:1561–1571. <http://dx.doi.org/10.1016/j.biochi.2006.04.015>.
26. Geiduschek EP, Tocchini-Valentini GP. 1988. Transcription by RNA polymerase III. *Annu Rev Biochem* 57:873–914. <http://dx.doi.org/10.1146/annurev.bi.57.070188.004301>.
27. Carbon P, Krol A. 1991. Transcription of the *Xenopus laevis* selenocysteine tRNA(Ser)Sec gene: a system that combines an internal B box and upstream elements also found in U6 snRNA genes. *EMBO J* 10:599–606.
28. Park JM, Yang ES, Hatfield DL, Lee BJ. 1996. Analysis of the selenocysteine tRNA[Ser]Sec gene transcription in vitro using *Xenopus* oocyte extracts. *Biochem Biophys Res Commun* 226:231–236. <http://dx.doi.org/10.1006/bbrc.1996.1338>.
29. Marck C, Grosjean H. 2002. tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. *RNA* 8:1189–1232. <http://dx.doi.org/10.1017/S1355838202022021>.
30. Lee BJ, Kang SG, Hatfield D. 1989. Transcription of *Xenopus* selenocysteine tRNA Ser (formerly designated opal suppressor phosphoserine tRNA) gene is directed by multiple 5'-extragenic regulatory elements. *J Biol Chem* 264:9696–9702.
31. Aeby E, Ullu E, Yepiskoposyan H, Schimanski B, Roditi I, Muhlemann O, Schneider A. 2010. tRNasec is transcribed by RNA polymerase II in *Trypanosoma brucei* but not in humans. *Nucleic Acids Res* 38:5833–5843. <http://dx.doi.org/10.1093/nar/gkq345>.
32. Florencio-Martinez L, Marquez-Duenas C, Ballesteros-Rodea G, Martinez-Calvillo S, Manning-Cela R. 2010. Cellular analysis of host cell infection by different developmental stages of *Trypanosoma cruzi*. *Exp Parasitol* 126:332–336. <http://dx.doi.org/10.1016/j.exppara.2010.04.015>.
33. Johnson PJ, Kooter JM, Borst P. 1987. Inactivation of transcription by UV irradiation of *T. brucei* provides evidence for a multicistronic transcription unit including a VSG gene. *Cell* 51:273–281. [http://dx.doi.org/10.1016/0092-8674\(87\)90154-1](http://dx.doi.org/10.1016/0092-8674(87)90154-1).
34. Peacock CS, Seeger K, Harris D, Murphy L, Ruiz JC, Quail MA, Peters N, Adlem E, Tivey A, Aslett M, Kerhornou A, Ivens A, Fraser A, Rajandream MA, Carver T, Norbertczak H, Chillingworth T, Hance Z, Jagels K, Moule S, Ormond D, Rutter S, Squares R, Whitehead S, Rabinowitz E, Arrowsmith C, White B, Thurston S, Bringaud F, Baldauf SL, Faulconbridge A, Jeffares D, Depledge DP, Oyola SO, Hilley JD, Brito LO, Tosi LR, Barrell B, Cruz AK, Mottram JC, Smith DF, Berriman M. 2007. Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. *Nat Genet* 39:839–847. <http://dx.doi.org/10.1038/ng2053>.
35. Rogers MB, Hilley JD, Dickens NJ, Wilkes J, Bates PA, Depledge DP, Harris D, Her Y, Herzyk P, Imamura H, Otto TD, Sanders M, Seeger K, Dujardin JC, Berriman M, Smith DF, Hertz-Fowler C, Mottram JC. 2011. Chromosome and gene copy number variation allow major structural change between species and strains of *Leishmania*. *Genome Res* 21:2129–2142. <http://dx.doi.org/10.1101/gr.122945.111>.
36. Downing T, Imamura H, Decuyper S, Clark TG, Coombs GH, Cotton JA, Hilley JD, de Doncker S, Maes I, Mottram JC, Quail MA, Rijal S, Sanders M, Schonian G, Stark O, Sundar S, Vanaerschot M, Hertz-Fowler C, Dujardin JC, Berriman M. 2011. Whole genome sequencing of multiple *Leishmania donovani* clinical isolates provides insights into population structure and mechanisms of drug resistance. *Genome Res* 21:2143–2156. <http://dx.doi.org/10.1101/gr.123430.111>.
37. Raymond F, Boisvert S, Roy G, Ritt JF, Legare D, Isnard A, Stanke M, Olivier M, Tremblay MJ, Papadopoulou B, Ouellette M, Corbeil J. 2012. Genome sequencing of the lizard parasite *Leishmania tarentolae* reveals loss of genes associated to the intracellular stage of human pathogenic species. *Nucleic Acids Res* 40:1131–1147. <http://dx.doi.org/10.1093/nar/gkr834>.
38. Berriman M, Ghedin E, Hertz-Fowler C, Blandin G, Renaud H, Bartholomeu DC, Lennard NJ, Caler E, Hamlin NE, Haas B, Bohme U, Hannick L, Aslett MA, Shallom J, Marcello L, Hou L, Wickstead B, Alsmark UC, Arrowsmith C, Atkin RJ, Barron AJ, Bringaud F, Brooks K, Carrington M, Cherevach I, Chillingworth TJ, Churcher C, Clark LN, Corton CH, Cronin A, Davies RM, Doggett J, Djikeng A, Feldblyum T, Field MC, Fraser A, Goodhead I, Hance Z, Harper D, Harris BR, Hauser H, Hostetler J, Ivens A, Jagels K, Johnson D, Johnson J, Jones K, Kerhornou AX, Koo H, Larke N, et al. 2005. The genome of the African trypanosome *Trypanosoma brucei*. *Science* 309:416–422. <http://dx.doi.org/10.1126/science.1112642>.
39. El-Sayed NM, Myler PJ, Bartholomeu DC, Nilsson D, Aggarwal G, Tran AN, Ghedin E, Worthey EA, Delcher AL, Blandin G, Westenberger SJ, Caler E, Cerqueira GC, Branche C, Haas B, Anupama A, Arner E, Aslund L, Attipoe P, Bontempi E, Bringaud F, Burton P, Cadag E,

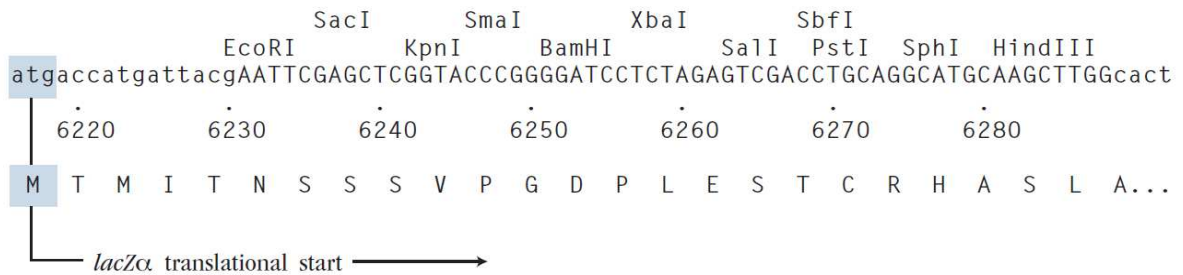
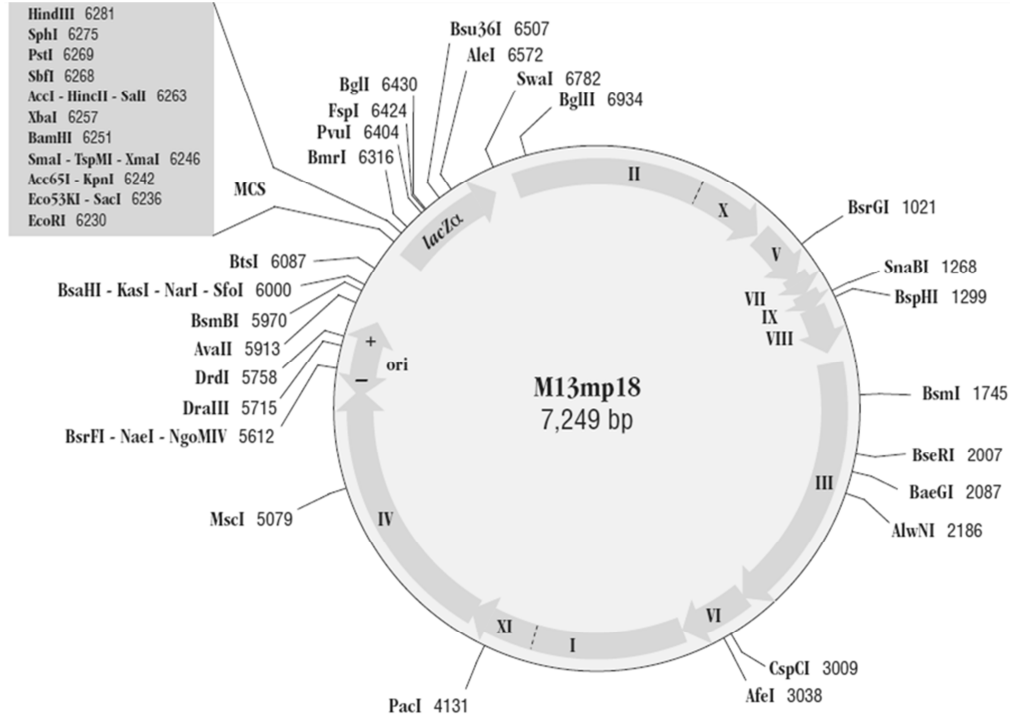
- Campbell DA, Carrington M, Crabtree J, Darban H, da Silveira JF, de Jong P, Edwards K, Englund PT, Fazelina G, Feldblyum T, Ferella M, Frasc AC, Gull K, Horn D, Hou L, Huang Y, Kindlund E, Klingbeil M, Kluge S, Koo H, Lacerda D, Levin MJ, Lorenzi H, Louie T, Machado CR, McCulloch R, McKenna A, et al. 2005. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* 309:409–415. <http://dx.doi.org/10.1126/science.1112631>.
40. Jackson AP, Berry A, Aslett M, Allison HC, Burton P, Vavrova-Anderson J, Brown R, Browne H, Corton N, Hauser H, Gamble J, Gilderthorp R, Marcello L, McQuillan J, Otto TD, Quail MA, Sanders MJ, van Tonder A, Ginger ML, Field MC, Barry JD, Hertz-Fowler C, Berriman M. 2012. Antigenic diversity is generated by distinct evolutionary mechanisms in African trypanosome species. *Proc Natl Acad Sci U S A* 109:3416–3421. <http://dx.doi.org/10.1073/pnas.1117313109>.
 41. Sauerbier W, Hercules K. 1978. Gene and transcription unit mapping by radiation effects. *Annu Rev Genet* 12:329–363. <http://dx.doi.org/10.1146/annurev.ge.12.120178.001553>.
 42. Laird PW, Kooter JM, Loosbroek N, Borst P. 1985. Mature mRNAs of *Trypanosoma brucei* possess a 5' cap acquired by discontinuous RNA synthesis. *Nucleic Acids Res* 13:4253–4266. <http://dx.doi.org/10.1093/nar/13.12.4253>.
 43. Fantoni A, Dare AO, Tschudi C. 1994. 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* 14:2021–2028.
 44. Monnerat S, Martinez-Calvillo S, Worthey E, Myler PJ, Stuart KD, Fasel N. 2004. Genomic organization and gene expression in a chromosomal region of *Leishmania major*. *Mol Biochem Parasitol* 134:233–243. <http://dx.doi.org/10.1016/j.molbiopara.2003.12.004>.
 45. Thomas S, Green A, Sturm NR, Campbell DA, Myler PJ. 2009. Histone acetylations mark origins of polycistronic transcription in *Leishmania major*. *BMC Genomics* 10:152. <http://dx.doi.org/10.1186/1471-2164-10-152>.
 46. Schimanski B, Nguyen TN, Gunzl A. 2005. Highly efficient tandem affinity purification of trypanosome protein complexes based on a novel epitope combination. *Eukaryot Cell* 4:1942–1950. <http://dx.doi.org/10.1128/EC.4.11.1942-1950.2005>.
 47. James Faresse N, Canella D, Praz V, Michaud J, Romascano D, Hernandez N. 2012. Genomic study of RNA polymerase II and III SNAPc-bound promoters reveals a gene transcribed by both enzymes and a broad use of common activators. *PLoS Genet* 8:e1003028. <http://dx.doi.org/10.1371/journal.pgen.1003028>.
 48. Sussman DJ, Chung J, Leder P. 1991. *In vitro* and *in vivo* analysis of the c-myc RNA polymerase III promoter. *Nucleic Acids Res* 19:5045–5052. <http://dx.doi.org/10.1093/nar/19.18.5045>.
 49. Duttke SH. 2014. RNA polymerase III accurately initiates transcription from RNA polymerase II promoters *in vitro*. *J Biol Chem* 289:20396–20404. <http://dx.doi.org/10.1074/jbc.M114.563254>.
 50. Mayan MD. 2013. RNAP-II transcribes two small RNAs at the promoter and terminator regions of the RNAP-I gene in *Saccharomyces cerevisiae*. *Yeast* 30:25–32. <http://dx.doi.org/10.1002/yea.2938>.

APÉNDICE II: Mapa del vector pGEM-T Easy (Promega)



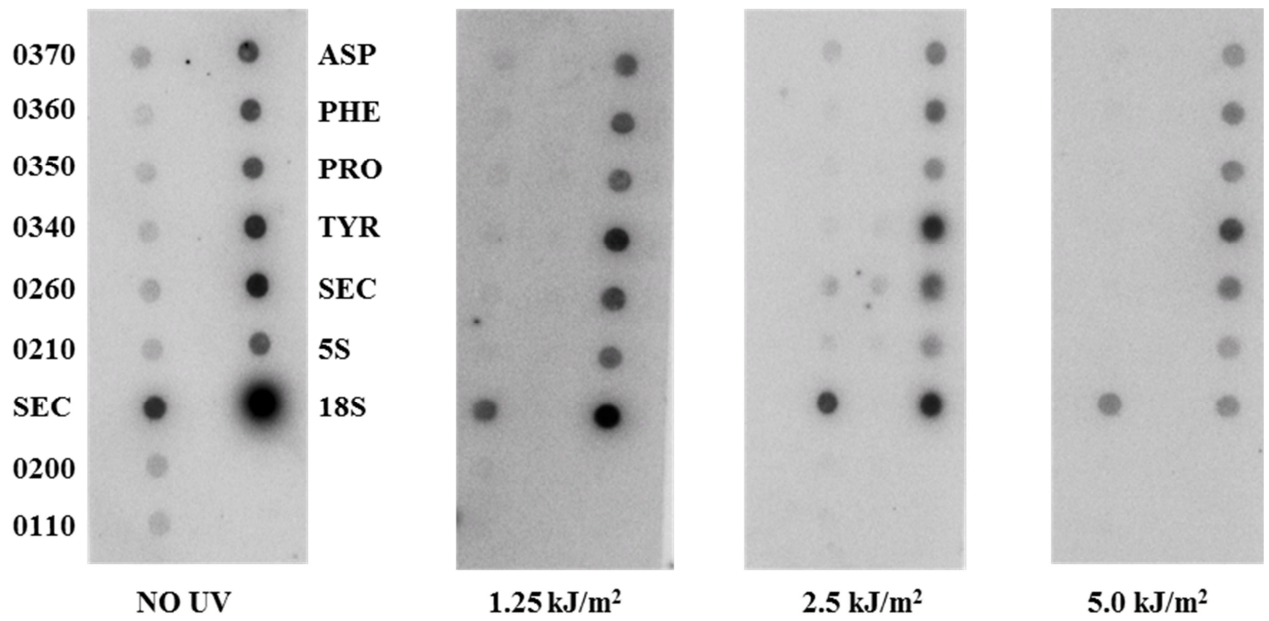
Secuencia del sitio de clonación múltiple del vector

APÉNDICE III: Mapa del vector M13mp18 y M13mp19

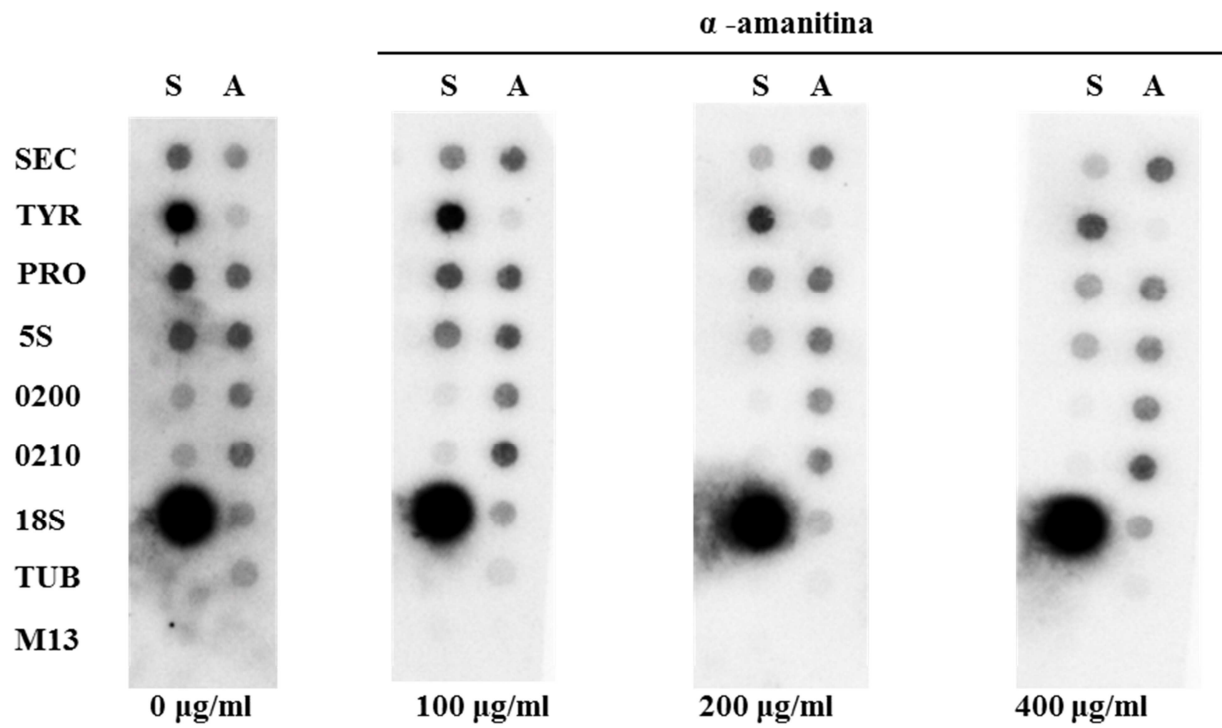


Secuencia del sitio de clonación múltiple de M13mp18. La diferencia en la secuencia entre M13mp18 y M13mp19 radica en el sitio de clonación múltiple. El de M13mp19 corresponde al reverso complementario del mostrado arriba

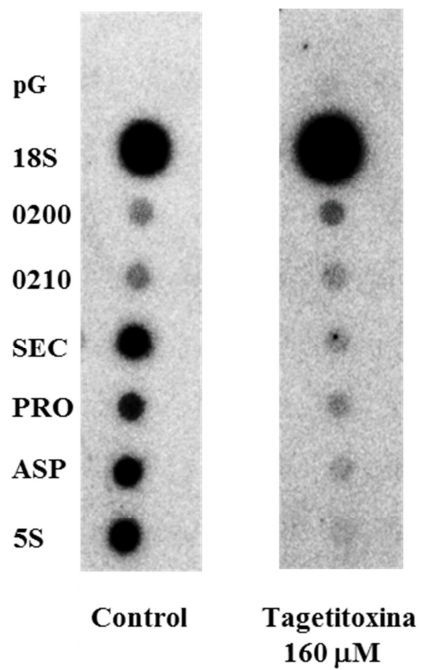
APÉNDICE IV: Repetición del experimento de *run-on* con células irradiadas con luz UV



APÉNDICE V: Repetición del experimento de *run-on* con α -amanitina



APÉNDICE VI: Repetición del experimento de *run-on* con tagetitoxina



APÉNDICE VII: Artículo de la alumna como primer autor relacionado con el presente trabajo

Padilla-Mejía, N.E., Florencio-Martínez, L. E., Figueroa-Angulo, E. E., Manning-Cela, R., Hernández-Rivas, R., Myler, P. J. y Martínez-Calvillo, S. **2009**. Gene organization and sequence analyses of transfer RNA genes in Trypanosomatid parasites. *BMC Genomics*. 10: 232-249.

Research article

Open Access

Gene organization and sequence analyses of transfer RNA genes in Trypanosomatid parasites

Norma E Padilla-Mejía¹, Luis E Florencio-Martínez¹, Elisa E Figueroa-Angulo¹, Rebeca G Manning-Cela², Rosaura Hernández-Rivas², Peter J Myler^{3,4} and Santiago Martínez-Calvillo*¹

Address: 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, México, ²Departamento de Biomedicina Molecular, Centro de Investigación y de Estudios Avanzados del IPN, Apartado Postal 14-740, 07360, México, DF, México, ³Seattle Biomedical Research Institute, 307 Westlake Avenue N, Seattle, WA 98109-5219, USA and ⁴Departments of Global Health, and Medical Education & Biomedical Informatics, University of Washington, Seattle, WA 98195, USA

Email: Norma E Padilla-Mejía - qfedime@yahoo.com.mx; Luis E Florencio-Martínez - luise@campus.iztacala.unam.mx; Elisa E Figueroa-Angulo - efiguer2000@yahoo.com; Rebeca G Manning-Cela - rmanning@cinvestav.mx; Rosaura Hernández-Rivas - roherman@cinvestav.mx; Peter J Myler - peter.myler@sbri.org; Santiago Martínez-Calvillo* - scalv@campus.iztacala.unam.mx

* Corresponding author

Published: 18 May 2009

Received: 17 December 2008

BMC Genomics 2009, 10:232 doi:10.1186/1471-2164-10-232

Accepted: 18 May 2009

This article is available from: <http://www.biomedcentral.com/1471-2164/10/232>

© 2009 Padilla-Mejía et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: The protozoan pathogens *Leishmania major*, *Trypanosoma brucei* and *Trypanosoma cruzi* (the Trityps) are parasites that produce devastating human diseases. These organisms show very unusual mechanisms of gene expression, such as polycistronic transcription. We are interested in the study of tRNA genes, which are transcribed by RNA polymerase III (Pol III). To analyze the sequences and genomic organization of tRNA genes and other Pol III-transcribed genes, we have performed an *in silico* analysis of the Trityps genome sequences.

Results: Our analysis indicated the presence of 83, 66 and 120 genes in *L. major*, *T. brucei* and *T. cruzi*, respectively. These numbers include several previously unannotated selenocysteine (Sec) tRNA genes. Most tRNA genes are organized into clusters of 2 to 10 genes that may contain other Pol III-transcribed genes. The distribution of genes in the *L. major* genome does not seem to be totally random, like in most organisms. While the majority of the tRNA clusters do not show synteny (conservation of gene order) between the Trityps, a cluster of 13 Pol III genes that is highly syntenic was identified. We have determined consensus sequences for the putative promoter regions (Boxes A and B) of the Trityps tRNA genes, and specific changes were found in tRNA-Sec genes. Analysis of transcription termination signals of the tRNAs (clusters of Ts) showed differences between *T. cruzi* and the other two species. We have also identified several tRNA isodecoder genes (having the same anticodon, but different sequences elsewhere in the tRNA body) in the Trityps.

Conclusion: A low number of tRNA genes is present in Trityps. The overall weak synteny that they show indicates a reduced importance of genome location of Pol III genes compared to protein-coding genes. The fact that some of the differences between isodecoder genes occur in the internal promoter elements suggests that differential control of the expression of some isoacceptor tRNA genes in Trityps is possible. The special characteristics found in Boxes A and B from tRNA-Sec genes from Trityps indicate that the mechanisms that regulate their transcription might be different from those of other tRNA genes.

Background

The parasites *Leishmania major*, *Trypanosoma brucei* and *Trypanosoma cruzi*, referred together as Trityps, are trypanosomatid protozoa that cause deadly human diseases known as leishmaniasis, African sleeping sickness and Chagas disease, respectively. Collectively, these pathogens cause millions of deaths in developing countries in tropical and subtropical regions of the world. Analyses of the recently reported genomic sequences of the Trityps revealed a striking feature: their genomes are organized into large directional gene clusters, *i.e.* tens-to-hundreds of protein-coding genes arranged sequentially on the same strand of DNA [1-3]. Transcription of the gene clusters is polycistronic, and mature mRNAs are generated from long precursors by trans-splicing and polyadenylation [4,5]. Most chromosomes contain at least two polycistronic gene clusters (PGCs), which can be either divergently transcribed (towards the telomeres) or convergently transcribed (away from the telomeres). Chromosome 3 from *L. major* contains two convergent PGCs (of 67 and 45 genes) that are separated by a tRNA gene. Interestingly, Pol II-transcription of both PGCs terminates within the tRNA-gene region [6]. The *L. major* nuclear genome is distributed among 36 relatively small chromosomes that range from 0.28 to 2.8 Mb. *T. cruzi* possesses ~28 medium-sized chromosomes, while *T. brucei* has 11 large chromosomes. Regardless of having diverged more than 200 million years ago, the genomes of trypanosomatids show a remarkable conservation of gene order (synteny) [7].

We are interested in the study of transcription by RNA polymerase III (Pol III), which produces small essential RNA molecules, such as tRNA [8]. All tRNAs have sequences of 74–95 bases that fold into a characteristic cloverleaf secondary structure with four constant arms. The acceptor arm binds to a particular amino acid, specified by the anticodon triplet located in the anticodon arm. The anticodon is complementary to an mRNA codon, specific for the amino acid carried by the tRNA. Therefore, tRNAs serve as adaptor molecules that mediate the transfer of information from nucleic acid to protein [9]. Organisms must have at least one tRNA for each of the 20 amino acids. Because different types of relaxed base pairings are allowed at the "wobble" position of the anticodon, certain tRNAs (known as isoacceptors) can read two or more synonymous codons differing by the third base. Consequently, cells do not carry tRNAs with anticodons complementary to all of the 61 possible codons in the genetic code. Interestingly, several organisms contain a large proportion of tRNA genes that have the same anticodon sequence, but differences elsewhere in the tRNA body [10]. The number of these tRNA genes, called isodecoders, varies from very low (10 in yeast) to very high (225–246)

in chimp and human. Thus, the diversity of tRNA genes is higher than originally thought [10].

Most organisms usually contain several hundred tRNA genes distributed randomly over their entire genome. One of the distinctive features of most genes transcribed by Pol III is that their promoter sequences are located within the transcribed region. In the case of tRNA genes, the promoter consists of two conserved elements: Boxes A and B. While Box A is normally positioned close to the transcription start site, the location of Box B is variable, partly because some tRNAs have short introns within the coding region [8,11].

Here we report the *in silico* analysis of tRNA genes in trypanosomatids. We found that, unlike in most other organisms, the distribution of genes in the genomes of *L. major* and *T. brucei* does not seem to be totally random, being confined to a subset of chromosomes. In addition, 14 out of 39 convergent strand-switch regions found in *L. major* contain at least one tRNA gene, which suggests that the use of tRNA genes as signals for termination of transcription of PGCs might be a common process in this parasite. Our analysis also indicated that the majority of the tRNA clusters do not show conservation of gene order among Trityps. Analysis of the putative transcription termination signals in all the tRNA genes showed an average of 5 Ts (+/- 1) in *L. major* and *T. brucei*, and 6 Ts (+/- 2) in *T. cruzi*. Also, special features were found in promoter elements from tRNA-Sec genes from Trityps. Finally, we have identified several tRNA isodecoder genes in the Trityps.

Results and discussion

Number of tRNA genes

Analysis of the GeneDB databases from *L. major*, *T. brucei* and *T. cruzi* (Trityps) revealed the presence of 82, 65 and 115 tRNA genes, respectively (see Table 1 and Additional File 1). By using the tRNAscan-SE program, we confirmed the identity of all the annotated tRNA genes in *L. major* (Fig. 1). However, we found a few discrepancies in the *T. brucei* and *T. cruzi* annotated genomes. In the case of *T. brucei*, it was observed that the tRNA-Sec (Tb04_tRNA-Sec1) gene annotated on chromosome 4 actually corresponds to the sRNA76 (see cluster chr04-V in Fig. 2), which is a tRNA-like molecule that associates to the 7SL RNA in trypanosomatids [12]. By doing a search of the *T. brucei* genome with the tRNA-Sec gene sequence reported previously [13], we located two copies of the genuine gene on chromosome 9 (see clusters chr09-II-III in Fig. 2). One of these tRNA genes is located within an open reading frame (ORF), annotated as a "hypothetical protein, unlikely" (Tb09.160.1080). We also identified one tRNA-Sec gene in *L. major* (see cluster chr06 in Fig. 1). Interestingly, eight tRNA-Sec genes were found in the *T. cruzi*

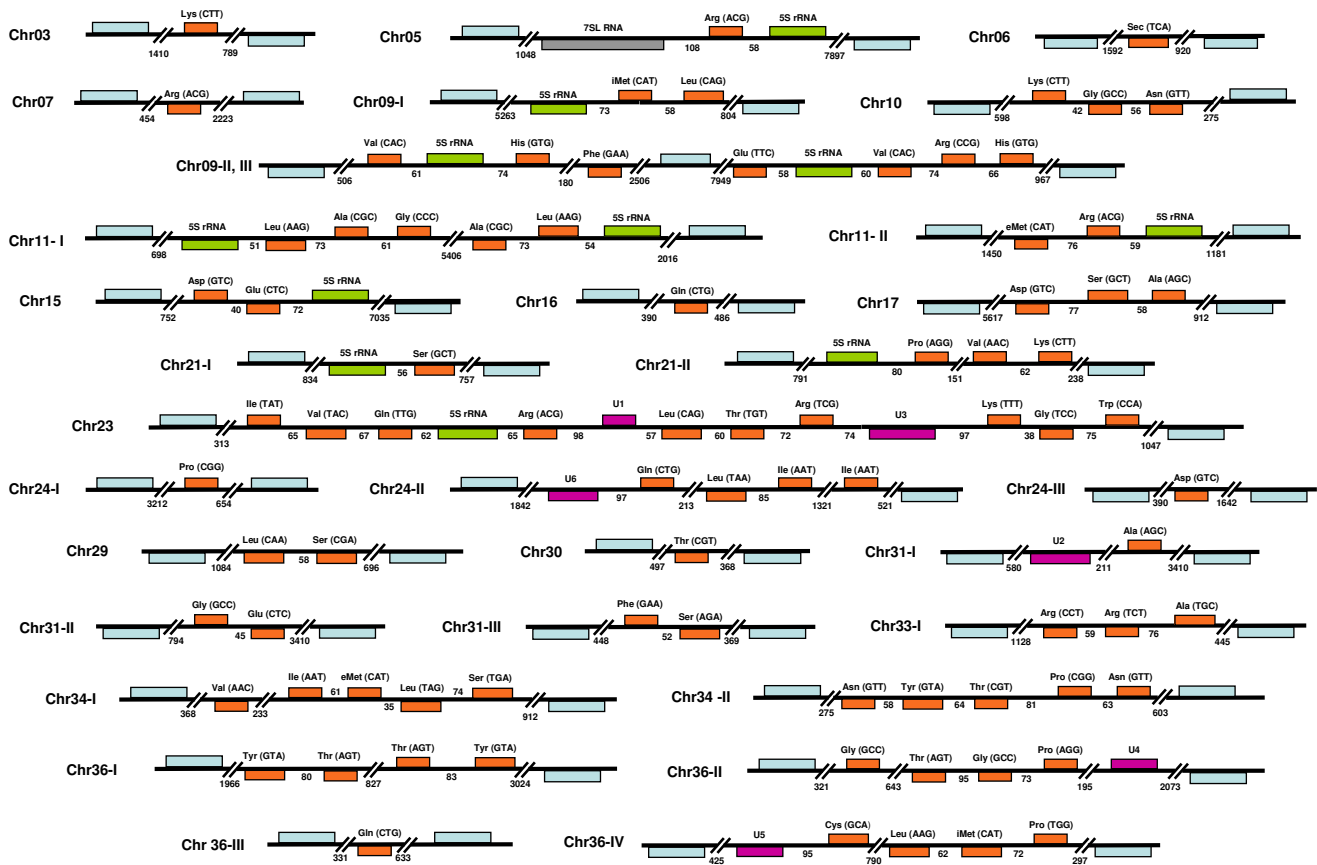


Figure 1

Organization of tRNA genes in *L. major*. The 83 tRNA genes found in the genome of the parasite are shown in orange. The predicted anticodons are indicated in parentheses. 5S rRNA, snRNA and 7SL genes are shown in green, purple and gray, respectively. Genes are drawn to scale, and the sizes of intergenic regions are indicated (in base pairs). Protein-coding genes that flank Pol III-transcribed genes are shown in blue (not to scale). The tRNA-Sec gene on chromosome 6 is located at positions 69,586 to 69,673, in the complement strand. Putative pseudogenes are not shown. For practical purposes, we regarded protein-coding genes as the limits of a particular Pol III locus. For that reason, we considered cluster chr09-II, III as two independent Pol III loci. More of such cases are shown in Fig. 2.

genome (see Additional File 1); they all are organized as independent genes, not clustered with other Pol III genes (data not shown). The tRNA-Sec is a component of a translational mechanism that reads UGA (normally a stop codon) as a selenocysteine codon in selected mRNAs that contain a specific *cis*-acting RNA regulatory sequence in their 3' untranslated regions (3'-UTRs) [14]. The presence of selenoproteins, and all the machinery required for its synthesis, has been demonstrated in *L. major* and *T. cruzi* [13,15].

In *T. cruzi*, it was found that three tRNA genes annotated as Val-CAC (Tc00.1047053457717.10, Tc00.1047053483321.10 and Tc00.1047053506321.220) do not seem to correspond to the assigned amino acid (or to any other).

They showed only 61% identity with Tc00.1047053506459.249, which we consider is the "real" tRNA-Val-CAC gene, since it is 100% identical to the tRNA-Val-CAC gene from *T. brucei* (and shows 98% identity to its orthologue in *L. major*). Interestingly, we observed that they show 75% identity to the sRNA76 from *T. brucei* (data not shown), which suggests that they might actually encode the orthologue of this gene in *T. cruzi*. Alternatively, they may correspond to tRNAs with undetermined or unknown type. tRNA genes with undetermined type have been found in several species, including *Caenorhabditis elegans*, yeast and human, and their function is unknown. Additionally, Tc00.1047053507579.16, annotated as an Ile-TAT gene, seems to be a pseudogene (or an undetermined tRNA); since it shows only 22%

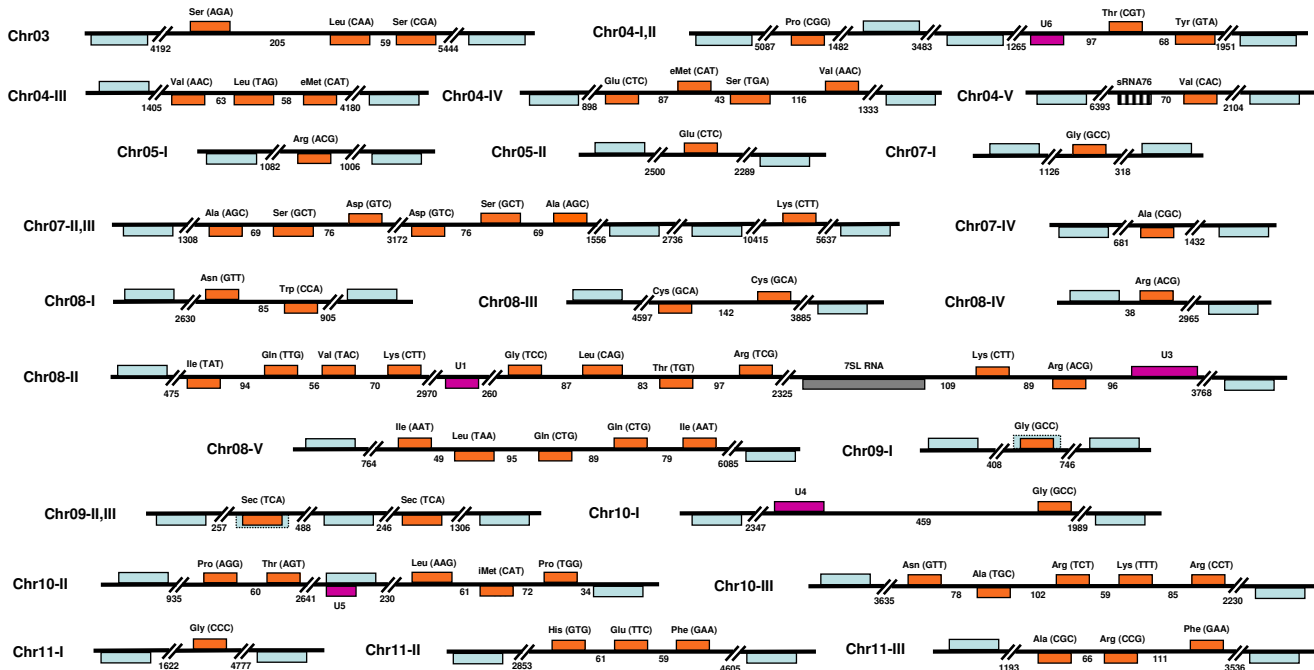


Figure 2
Organization of tRNA genes in *T. brucei*. The 66 tRNA genes, distributed into 23 loci, are indicated in orange. The predicted anticodons are shown in parentheses. snRNA and 7SL genes are shown in purple and gray, respectively. The sRNA76 (misannotated in GeneDB as a tRNA-Sec gene) is shown as a striped box in locus chr04-V. The two tRNA-Sec genes that we found are located in loci chr09-II, III (at positions 505,606 to 505,693 and 509,443 to 509,530 of chromosome 9). The gene from locus chr09-II overlaps a putative Pol II gene (Tb09.160.1080, dotted blue box), annotated as hypothetical protein (unlikely). Also, the tRNA-Gly from locus chr09-I overlaps a putative Pol II gene (Tb09.211.4080, dotted blue box), annotated as hypothetical protein (unlikely). The U5 snRNA gene (locus chr10-II) also overlaps a Pol II gene (Tb10.6k15.2990, sequence orphan), but located on the opposite strand. Genes are drawn to scale, and the sizes of intergenic regions are indicated (in base pairs). Protein-coding genes that flank Pol III-transcribed genes are shown in blue (not to scale). In locus chr08-V, the Leu-TAA gene (Tb_08_TRNA_Leu_2) and the first Gln-CTG gene (Tb_08_TRNA_Gln_2) are annotated in the wrong (opposite) strand in GeneDB. Maps of an incomplete repertoire of tRNA genes from *T. brucei* were previously reported [16].

identity to the other two annotated Ile-TAT genes in *T. cruzi* (Tc00.1047053504427.231 and Tc00.1047053508043.11) (data not shown). Also, we identified an extra copy of tRNA-Ala-TGC on contig 8001 (see Additional File 1).

Thus, our analysis indicates the presence of 120 tRNA genes in *T. cruzi*, excluding four genes that might be undetermined tRNAs or encode orthologues of the sRNA76, and including the eight tRNA-Sec genes and the tRNA-Ala-TGC gene (Table 1 and Additional File 1). In *T. brucei* the number of identified tRNA genes is 66, including the two newly identified tRNA-Sec genes and excluding the gene of the sRNA76 orthologue (Fig. 2 and Table 1). In *L. major* there are 83 tRNA genes (Fig. 1 and Table 1), in addition to a pseudogene that we do not include in our analyses. The number of tRNA genes in trypanosomatids is relatively low, considering that eukaryotic organisms usually contain several hundred tRNA genes. For instance, *C. ele-*

gans has 568 tRNA genes, *Homo sapiens* presents 497 tRNA genes and *Saccharomyces cerevisiae* contain 271 tRNA genes [9,10]. In an extreme case, *Danio rerio* (zebra fish) has ~6000 predicted tRNA genes. On the other hand, the microsporidian parasite *Encephalitozoon cuniculi* has only 44 tRNA genes. Bacterial genomes usually have between 29 and 167 tRNA genes in their genomes [9,10].

tRNA genes from eukaryotes typically contain introns, which are usually located between bases 37 and 38 of the anticodon loop. Archaeal tRNA genes also have introns that can be found at the same location of the anticodon loop or in other regions of the tRNA gene. The size of introns is variable, ranging from 7 to 121 bases [9]. In bacterial genomes, a very small number of tRNA-gene introns have been reported, but they correspond to self-splicing introns (group I autocatalytic introns). Analysis of the tRNA genes in *Trityps* indicated that only the tRNA-Tyr genes contain an intron; which was previously reported in

Table 1: Repertoire of tRNA genes and codon usage in Trityps.

Amino acid	Codon	tDNA anticodon	<i>L. major</i> gene copy number	<i>T. brucei</i> gene copy number	<i>T. cruzi</i> gene copy number	Anticodon ²	<i>L. major</i> Codon usage %	<i>T. brucei</i> Codon usage %	<i>T. cruzi</i> Codon usage %
Ala	GCT	AGC	2	2	2 (98%) ¹	IGC	1.83	2.08	1.61
	GCC	GGC	0	0	0		3.61	1.82	1.62
	GCA	TGC	1	1	2 (97%) ¹	UGC	2.02	2.32	2.16
	GCG	CGC	2	2 (97%) ¹	2	CGC	3.62	2.07	2.36
Arg	CGT	ACG	4	3	4 (98%) ¹	ICG	1.05	1.59	1.60
	CGC	GCG	0	0	0		3.18	1.45	1.75
	CGA	TCG	1	1	2	UCG	0.73	0.9	1.58
	CGG	CCG	1	1	2	CCG	1.38	1.21	1.85
	AGA	TCT	1	1	2	UCU	0.27	0.68	1.56
	AGG	CCT	1	1	2	CCU	0.55	1	1.86
Asn	AAT	ATT	0	0	0	GUU	0.53	1.82	1.37
	AAC	GTT	3	2	4		2.07	1.92	1.45
Asp	GAC	GTC	3	2	2	GUC	3.42	2.27	1.60
	GAT	ATC	0	0	0		1.44	2.81	1.57
Cys	TGT	ACA	0	0	0	GCA	0.42	1.09	1.78
	TGC	GCA	1	2	2		1.46	1.13	2.34
Gln	CAA	TTG	1	1	2	UUG	0.76	1.69	1.66
	CAG	CTG	3	2	4	CUG	3.32	2.1	1.8
Glu	GAA	TTC	1	1	1	UUC	1.14	3.17	2.17
	GAG	CTC	2	2	4 (95%) ¹	CUC	3.32	3.81	2.21
Gly	GGT	ACC	0	0	0	GCC	1.25	2.27	1.51
	GGC	GCC	4	3	4		3.36	1.49	1.95
	GGA	TCC	1	1	2	UCC	0.64	1.56	2.16
	GGG	CCC	1	1	2	CCC	1.19	1.39	1.64
His	CAT	ATG	0	0	0	GUG	0.65	1.13	1.37
	CAC	GTG	2	1	4 (98%) ¹		2.04	1.3	1.64
Ile	ATT	AAT	3	2	4	IAU	0.82	1.9	1.42
	ATC	GAT	0	0	0		1.88	1.16	1.10
	ATA	TAT	1	1	2	UAU	0.27	1	0.75
Leu	TTA	TAA	1	1	2	UAA	0.16	0.98	0.89
	TTG	CAA	1	1	2	CAA	1.1	1.96	2.33
	CTT	AAG	3	1	4	IAG	1.11	2.23	1.59
	CTC	GAG	0	0	0		2.5	1.56	1.27
	CTA	TAG	1	1	2	UAG	0.47	0.82	0.54
	CTG	CAG	2	1	2	CAG	3.83	1.85	1.95
Lys	AAA	TTT	1	1	2	UUU	0.54	2.06	1.94
	AAG	CTT	3	3	4	CUU	2.78	2.66	1.88
Met	ATG	CAT	4	3	6	CAU	2.25	2.34	2.12
Phe	TTT	AAA	0	0	0	GAA	1.04	2.05	2.17
	TTC	GAA	2	2	4		1.91	1.59	1.49
Pro	CCT	AGG	2	1	2	IGG	0.86	1.11	1.05
	CCC	GGG	0	0	0		1.24	1.11	1.03

Table 1: Repertoire of tRNA genes and codon usage in Trityps. (Continued)

	CCA	TGG	1	1	2	UGG	1.05	1.39	1.44
	CCG	CGG	2	1	2	CGG	2.61	1.18	1.57
Ser	TCT	AGA	1	1	2	<u>IGA</u>	1.02	1.26	1.37
	TCC	GGA	0	0	0		1.69	1.25	1.24
	TCA	TGA	1	1	2 (98%) ¹	UGA	0.73	1.35	1.41
	TCG	CGA	1	1	2	CGA	2.17	1.16	1.2
	AGT	ACT	0	0	0	<u>GCU</u>	0.73	1.51	1.21
	AGC	GCT	2 (98%) ¹	2	2		2.6	1.35	1.60
Thr	ACT	AGT	3	1	2	<u>IGU</u>	0.68	1.3	1.13
	ACC	GGT	0	0	0		1.77	1.21	1.18
	ACA	TGT	1	1	2	UGU	1.04	1.74	1.74
	ACG	CGT	2	1	2	CGU	2.48	1.48	1.98
Trp	TGG	CCA	1	1	2	CCA	1.07	1.09	2.33
Tyr	TAT	ATA	0	0	0	<u>GUA</u>	0.4	1.13	0.84
	TAC	GTA	3	1	2		1.99	1.41	1.04
Val	GTT	AAC	2	2	2	<u>IAC</u>	0.86	2.29	1.49
	GTC	GAC	0	0	0		1.92	1.14	1.26
	GTA	TAC	1	1	2	UAC	0.54	1.26	0.83
	GTG	CAC	2	1	1	CAC	3.82	2.87	2.65
SeC	TGA	TCA	1	2	8				
TOTAL			83	66	120				

¹ One isodecoder tRNA was identified in *L. major*, one in *T. brucei* and six in *T. cruzi*. The percentage of identity between isodecoders is shown between parentheses. The rest of the genes with the same anticodon are 100% identical.

² Underlined anticodons are those that recognize more than one codon; for example, the Tyr anticodon GUA recognizes two codons, TAT and TAC, since anticodon Tyr-AUA is missing in trypanosomatids. A1 of the anticodon is modified to inosine (I).

T. brucei [16]. The intron is 11 bases long in *L. major* and *T. brucei*, and 13 bases long in *T. cruzi*; and as in other organisms, it is located between bases 37 and 38 (data not shown). Thus, introns are very rare in Trityps, not only in protein-coding genes, but also in tRNA genes.

Isoacceptor tRNA species

Analyses of the anticodon sequences of the tRNA genes in Trityps showed the presence of 46 isoacceptor types in each of the three species (Table 1) [17]. These 46 isoacceptor types are able to read the 61 codons that specify the canonical amino acids, in addition to Sec, the 21st amino acid. The number of isoacceptor species found in Trityps is similar to that found in other organisms (from 41 to 55 isoacceptors) [10]. It is important to mention that the two methionine isoacceptors, the initiator and elongator, have been identified (see below), but for practical purposes these two isoacceptors will be considered as one.

Sixteen anticodons were not found in the tRNA genes of trypanosomatids, even though their corresponding codons are present in the protein-coding genes of these organisms [17]. For example, the tRNA with anticodon Ile-GAU is not present in the genome of trypanosomatids,

but the codon AUC is present in their protein sequences (Table 1). As mention above, this is possible because some tRNAs are able to recognize more than one codon by allowing flexible base-pairing between the first nucleotide of the anticodon and the third position of the codon (tRNA wobble recognition). Analysis of the data shown in Table 1 indicates that C3 or U3 in the codon are recognized by G1 or A1 of the anticodon (A is converted to the nucleotide inosine in the mature tRNA, which can pair with U3 or C3 in the codon). Thus, trypanosomatids use the A1- or G1-sparing strategy as a decoding mode [9]. This anticodon-choice pattern is similar to that of other eukaryotes such as *C. elegans*, *H. sapiens* and *A. thaliana*. Other eukaryotic organisms, like yeast and *D. melanogaster*, use the A1- or G1 and C1-sparing strategy [9]. As observed previously [17], the spared anticodons are used equally, since 50% (8/16) of the U3 and C3 codons are read by A1 (or I1) and the remaining 50% are read by G1. In most four-fold degenerate codon families (*i.e.* Leu, Val, Ser, Pro, Thr, Ala and Arg, but not Gly) A1 reads the codons containing U3 and C3, since the corresponding tRNAs with G1 are not present in the Trityps. We found the same for the Ile family, which contains three codons. On the other hand, all the two-fold degenerate families

use G1 to read U3 or C3, given that the tRNAs with A1 are missing. The families that use this strategy are: Asn, Asp, Cys, Gly (although it is fourfold-degenerated), His, Phe, Ser (AGU and AGC codons) and Tyr.

The genomes of *L. major*, *T. brucei* and *T. cruzi* contain four, three and six tRNA-Met genes, respectively [16,17]. Further analysis of these genes indicated that in *L. major* two of them (LmjF09.TRNAMET.01 and LmjF36.TRNAMET.01) correspond to initiator tRNAs (iMet) and two (LmjF11.TRNAMET.01 and LmjF34.TRNA.01) correspond to the elongator form (eMet) (Fig. 3). In *T. brucei*, one iMet (Tb10_tRNA_Met_1) and two eMet genes (Tb04_tRNA_Met_1, Tb04_tRNA_Met_2) were found, although only one of each type was previously reported [16]. Finally, two iMet (Tc00.1047053508231.92, Tc.00.1047053506251.88) and four eMet genes

(Tc00.1047053504055.87, Tc00.1047053504055.91, Tc00.1047053506435.327 and Tc00.1047053506435.345) were located in *T. cruzi*. As shown in Fig. 3, the two types of tRNA-Met possess specific features, and most of them were found in the genes from the Trityps. One of the main characteristics is the highly conserved A:T base pair that is present in all tRNA-iMet in eukaryotes at position A1:U72 (A1:U71 in Trityps), whereas a G:C pair is found in tRNA-eMet [9]. In yeast, it has been reported that the A1:U72 base pair is the most important determinant for a tRNA-Met to play the role of iMet, since it is necessary for binding to Initiation Factor 2 (eIF2) [18]. When this sequence is mutated, the tRNA-iMet is able to bind to Elongation Factor Tu (EF-Tu) and participates in translation elongation. It is likely that this base pair has a similar function in trypanosomatids.

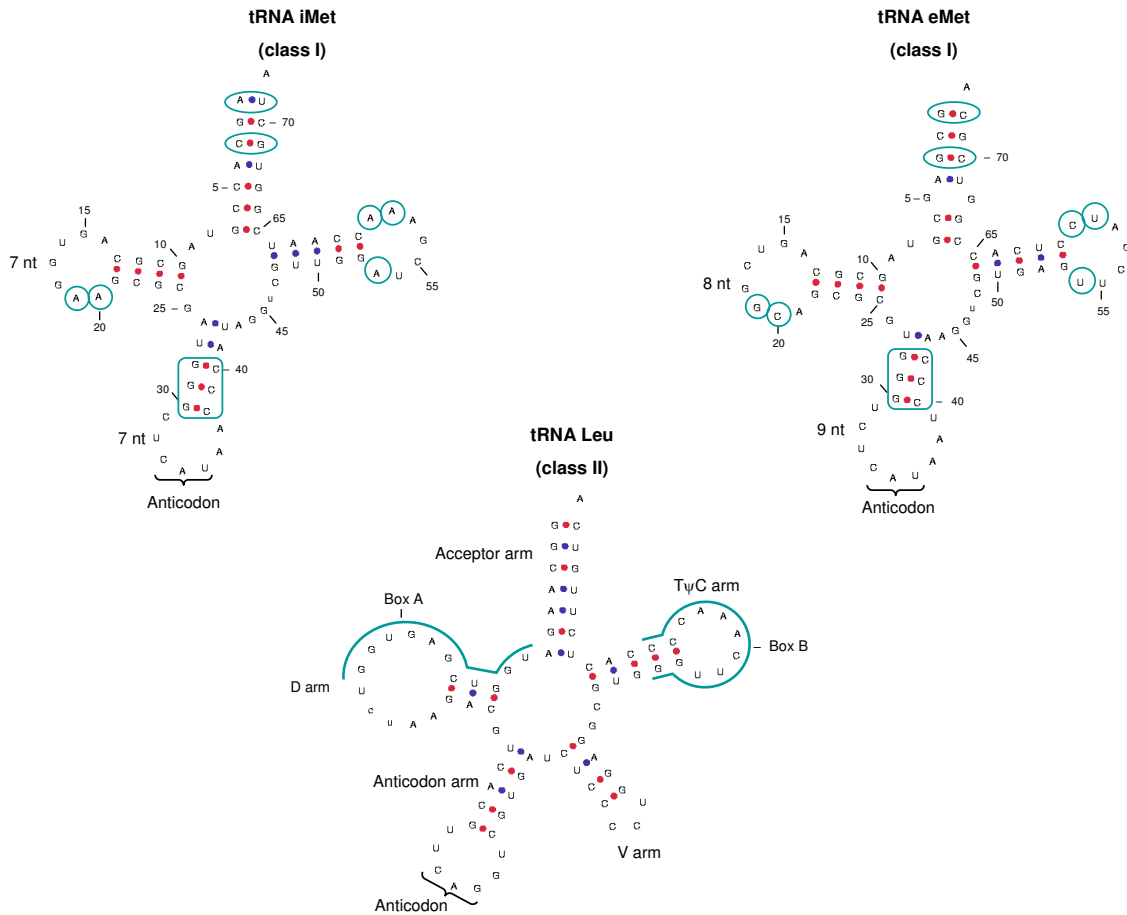


Figure 3
Secondary structures of initiator and elongator tRNA-Met in *L. major*. Important features for iMet and eMet tRNAs function are indicated. These two molecules correspond to class I tRNAs. The secondary structure of a class II tRNA (Leu) is also presented. The names of the different arms are shown. The position of the internal control elements (Boxes A and B) is also indicated.

Other determinants for iMet function are: A53, A58 and A59 in the T ψ C loop (U54, U59 and C60 in eMet); the base pair C3:G69 (G3:C70 in eMet); bases A19 and A20 (G19 and C20 in eMet); a D-loop composed of 7 nucleotides (8 nt in eMet); and an anticodon loop of 7 nt (9 nt in eMet) [9,19]. All these features are conserved in the Tritryps (Fig. 3). Another distinctive characteristic of the iMet tRNA is the presence of three consecutive G:C pairs at the bottom of the anticodon stem, which is conserved not only in eukaryotes, but also in eubacteria and archaeobacteria. In *E. coli*, it has been shown that mutations in these three consecutive G bases reduce the efficiency of initiation of protein synthesis, by affecting the interactions between the tRNA and the ribosomal P site; thus, these bases are essential to discriminate between initiator and elongator tRNA-Met [9,20]. Tritryps iMet tRNAs have these conserved G:C pairs but, surprisingly, we found them in the eMet tRNAs as well (Fig. 3). Thus, in the Tritryps these base pairs are not a discriminator between iMet and eMet tRNAs, and these organisms must use other features of the iMet to direct it to the P site of the ribosome.

In several organisms it has been observed that there is a correlation between tRNA gene copy number and codon usage [21,22]. Apparently, selection on synonymous codon positions causes co-adaptation of codon usage and tRNA content, in order to optimize the effectiveness of protein synthesis [23]. In Tritryps, it has been reported that bias in codon usage correlates with tRNA gene copy number and with protein expression level [17]. This conclusion was made after analyzing around 60,000 codons from highly expressed (tandem duplicated) protein-coding genes from the three parasites. We conducted a similar analysis, but including all the 8272 protein coding genes from *L. major* (5,249,748 codons), and 5119 randomly selected genes from *T. brucei* (2,620,035 codons), as well as 1779 genes from *T. cruzi* (986,435 codons). We plotted codon usage (see Table 1) against the number of tRNA genes for each isoacceptor, for the three species, and a possible correlation was evaluated by the Spearman test (Figure 4, panels A-C). The data indicated a positive correlation between these variables for *L. major* ($r_s = 0.80$) and *T. brucei* ($r_s = 0.64$), which indicates that, similar to other organisms, codon usage patterns seem to be co-adapted with the relative abundance of the corresponding tRNAs in these parasites. However, in the case of *T. cruzi*, the obtained Spearman value ($r_s = 0.35$) indicated a low degree of correlation between the two variables. This may reflect the fact that the *T. cruzi* strain used for the sequencing project is a hybrid of two strains, and some of their genes might be duplicated, while others might not be; as shown in Fig. 4C and Table 1, in *T. cruzi* the vast majority of the isoacceptor species are encoded by either two or four genes (only two isoacceptors have one gene, and none of them has three genes), whereas in *L. major* and *T. brucei*

a high number of the isoacceptors are encoded by a single gene. The correlation analysis was repeated, but now plotting the percentage of codon usage *versus* the number of tRNA genes per amino acid (Fig. 4, panels D-F). This time, the Spearman value was high in *T. cruzi* ($r_s = 0.78$), indicating a strong correlation between both parameters. As before, strong correlations were found in *L. major* ($r_s = 0.84$) and *T. brucei* ($r_s = 0.85$).

Organization of tRNA genes

In *L. major*, the 83 tRNA genes are distributed among 31 loci, on 19 different chromosomes (Fig. 1 and Additional File 1). Most tRNA genes are organized into clusters of 2 to 10 genes, on either top or bottom strand, which may contain other Pol III-transcribed genes. For example, in the locus located on chromosome 23 (chr23 in Fig. 1) there are 10 tRNA genes, a 5S rRNA gene and the U1 and U3 snRNA genes. Locus IV on chromosome 36 (chr36-IV) has four tRNA genes and the U5 snRNA gene. The eleven 5S rRNA genes found in the *L. major* genome are distributed in six chromosomes, and are always associated to tRNA genes (Fig. 1 and Additional File 1). Only eight loci contain single tRNA genes (chr03, chr06, chr07, chr16, chr24-I, chr24-III, chr30 and chr36-III). In most cases, intergenic regions that separate Pol III-transcribed genes are short, with an average size of 202 bases (Fig. 1). However, they can be as small as 35 bases (intergenic region between tRNA-eMet and tRNA-Leu on cluster chr34-I) and as long as 5406 bases (intergenic region between tRNA-Gly and tRNA-Ala on cluster chr11-I). Intergenic regions between the protein-coding genes that flank the clusters and the first or last Pol III gene of the cluster are normally longer than the ones that separate Pol III genes. The average length of such regions is 1490 bp, with a minimum of 238 bp (intergenic region between tRNA-Lys and the "right" Pol II gene on cluster chr21-II) and a maximum of 7949 pb (intergenic region between tRNA-Glu and the "left" Pol II gene on cluster chr09-III) (Fig. 1). The mean length of intergenic regions between protein-coding genes in the *L. major* genome is 2045 bp [1].

In *T. brucei*, the 66 tRNA genes are located on 26 loci, on eight different chromosomes (Fig. 2 and Additional File 1). As in *L. major*, in *T. brucei* the number of tRNA genes per cluster ranges from 2 to 10. Eleven of the 66 tRNA genes are single genes in *T. brucei* (loci chr04-I, chr05-I, chr05-II, chr07-I, chr07-III, chr07-IV, chr08-IV, chr09-I, chr09-II, chr09-III and chr11-I). Similarly to *L. major*, intergenic regions that separate Pol III genes in *T. brucei* are short in most cases (average length is 327 bp, ranging from 43 to 3172 bp) (Fig. 2). Regarding intergenic regions between the protein-coding genes that flank the clusters and the first or last Pol III gene of the cluster, the mean size is 2473 bp (showing a range from 34 to 10415 bp). In the *T. brucei* genome the average length of intergenic

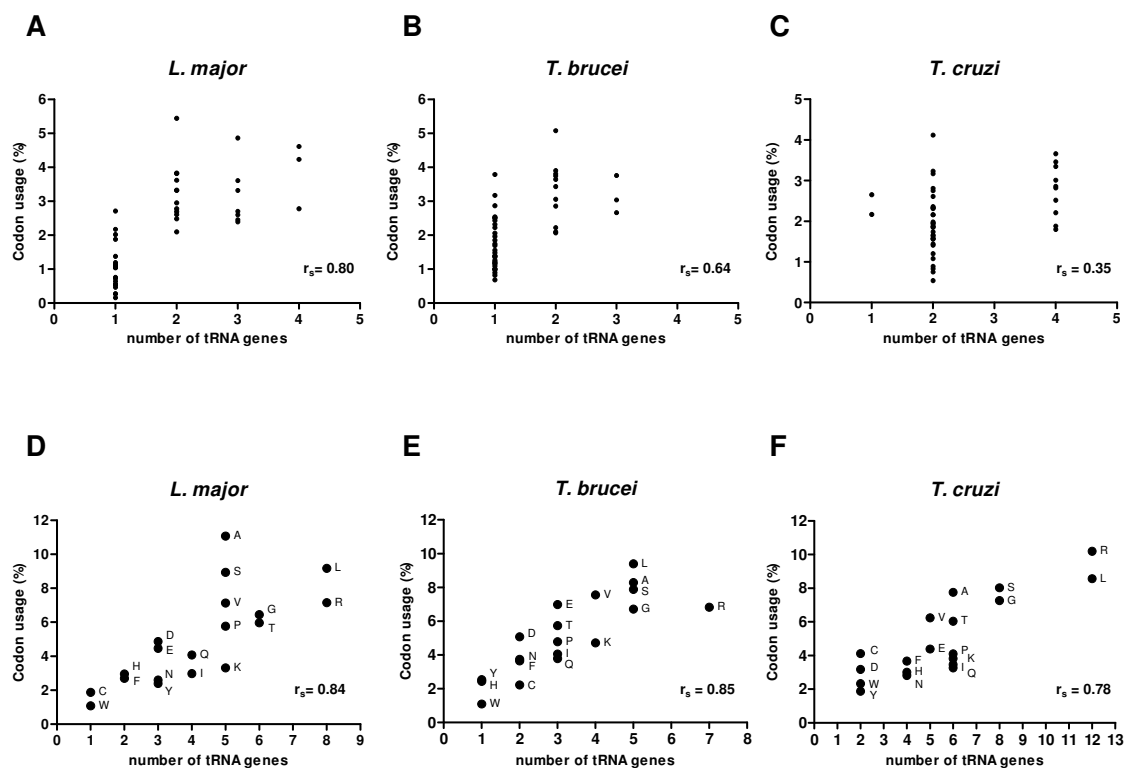


Figure 4

Correlation between number of tRNA genes and codon usage in *Trityps*. Panels A-C show the correspondence between codon usage and the number of tRNA genes for each isoacceptor (45 types), for the three species (see Table 1). Panels D-F show the correlation between codon usage and the number of tRNA genes per amino acid. Correlation coefficients were evaluated by performing a Spearman test. r_s values are indicated in each panel; $p < 0.0001$ for panels A, B, D and E; $p < 0.02$ for panel C; $p < 0.0001$ for panel F.

regions between protein-coding genes is 1279 bp [2]. Since it has not been possible to assemble fully adjacent sequences for the chromosomes of *T. cruzi*, at the present we are unable to determine the genomic organization of the tRNA genes in this parasite. In contrast to *L. major*, the 5S rRNA genes in *T. brucei* and *T. cruzi* are organized into tandem arrays, which are not associated to tRNA genes [24,25].

In most eukaryotic organisms, tRNA genes seem to be dispersed randomly throughout the genome. However, in human cells the distribution is non-random, since more than 25% of the tRNA genes are located in a region of only about 4 Mb on chromosome 6. This region represents only 0.1% of the human genome, but contains an almost complete set of tRNA genes. Moreover, 280 out of 497 tRNA genes (more than half) are found on either chromo-

some 1 or chromosome 6 [26]. The distribution of genes in the *L. major* genome does not seem to be totally random, since half of the chromosomes do not contain even a single tRNA gene. Additionally, 60 tRNA genes (72%) are located in only 7 chromosomes (9, 11, 23, 24, 31, 34 and 36), which represent only 26% of the genome (Fig. 1 and Additional File 1). In *T. brucei*, 40 (61.5%) of the tRNA genes are found in just 3 chromosomes (4, 7 and 8), which is only about 24% of the genome (Fig. 2 and Additional File 1). tRNA genes in *S. cerevisiae*, though dispersed in the linear genome, co-localize with 5S rDNA at the nucleolus. Nucleolar localization requires tRNA gene transcription, because inactivation of the internal promoter eliminates its nucleolar location [27]. It remains to be tested whether tRNA genes in trypanosomatids show such a specific cellular localization.

In *Schizosaccharomyces pombe* and *C. elegans*, tRNA genes are often clustered in centromeres [28]. These tRNA genes contribute to centromere function by defining domain boundaries important for centromere activity [29]. Putative centromeric regions have been reported in a few chromosomes in *T. cruzi* and *T. brucei* [30], where they localize to strand-switch regions that separate divergent PGCs. While these regions do not seem to contain tRNA genes, two clusters of tRNA genes in *L. major* (chr09-II and chr10), and one cluster in *T. brucei* (chr04-I) are located in divergent strand-switch regions (Figs. 1 and 2), and thus might be candidates to contain centromeric regions. Therefore, it is possible that in trypanosomatids, like *S. pombe* and *C. elegans*, some tRNA genes might be important for centromeric activity.

Spatial relation between Pol III and Pol II genes

We have previously shown that transcription of two convergent PGCs on *L. major* chromosome 3 terminates on the convergent strand-switch area, within the tRNA-gene region [6]. Interestingly, 14 of the 39 convergent strand-switch regions (35.9%) in the *L. major* genome contain at least one tRNA gene (Fig. 1), representing 45.2% of the 31 tRNA loci. A similar situation was found in *T. brucei*, where 34.6% of the tRNA loci are located within convergent strand-switch regions (Fig. 2). This suggests that the use of tRNA genes as signals for termination of transcription of convergent clusters of protein-coding genes might be a common process in trypanosomatids. Indeed, recent evidence suggest that this is the case for tRNA clusters located within PGCs, since peaks of acetylated histone H3 are found immediately downstream of the tRNA cluster in all cases [31]. Acetylated histones are markers for open chromatin in all eukaryotes and have been found at the 5' end of all polycistronic gene clusters in *L. major*.

Synteny of Pol III genes

It has been found that the genomes of the Trityps are highly syntenic, that is to say, they show conservation of gene order, with the *T. brucei* and *L. major* genomes containing 110 blocks of syntenicity spanning 19.9 and 30.7 Mb, respectively [7]. Many of these syntenicity blocks correspond to intact PGCs, which are transcribed by Pol II. In contrast, the majority of the tRNA clusters do not show syntenicity, but a few of them do show conservation (Fig. 5). Among the latter, the most remarkable example is a cluster of 13 Pol III genes that is highly syntenic; corresponding to chr23 in *L. major*, chr08-II from *T. brucei*, and the cluster located on contig Tc6288 from *T. cruzi* (Fig. 5A). Surprisingly, the order of the genes in this cluster is identical between *T. brucei* and *T. cruzi*, although the U1 snRNA, the 7SL RNA and the tRNA-Leu genes are located on different strands. Most of the 13 Pol III-transcribed genes are present in the *L. major* cluster, but their order is not identical to either of the other two clusters (Fig. 5A). Additionally, a 5S rRNA

gene replaced a 7SL RNA gene and a tRNA-Trp gene replaced one of the tRNA-Lys genes. Another Pol III-gene cluster that exhibits syntenicity is chr24-II in *L. major*, chr08-V in *T. brucei* and contig Tc6223 from *T. cruzi* (Fig. 5B). Here, we found tRNA genes for Ile, Leu and Gln that are syntenic among the three species. A second copy of a tRNA-Ile is conserved between *L. major* and *T. brucei*. One difference is that the *L. major* cluster contains an U6 snRNA gene that is replaced by a tRNA-Gln gene in *T. brucei*. Other syntenic tRNA clusters are chr33-I from *L. major*, chr10-III from *T. brucei* and contig Tc8001 from *T. cruzi* (Fig. 5C), and chr34-I from *L. major*, chr04-III from *T. brucei* and contig Tc4886 from *T. cruzi* (Fig. 5D). Several of the protein-coding genes that flank these four syntenic tRNA-gene clusters are also syntenic among Trityps (Fig. 5). As in Trityps, an overall weak syntenicity of Pol III-transcribed genes has been observed between two species of the oomycete *Phytophthora* [32], indicating a reduced importance of genome location of Pol III genes compared to protein-coding genes.

Consensus sequences of promoter elements

One of the distinctive features of most genes transcribed by Pol III is that their promoter sequences are internal, located within the transcribed region. In the case of tRNA genes, the promoter consists of two conserved elements: Boxes A and B [33]. While Box A is normally positioned close to the transcription start site, the location of Box B is variable, partly because some tRNAs have short introns within the coding region. In contrast, tRNA genes in prokaryotic cells contain promoter elements similar to those found in protein-coding genes: the start-point (usually a purine), the -10 sequence (the TATA Box) and the -35 sequence (the hexamer) [34]. Consensus sequences of trypanosomatid tRNA promoter elements (Fig. 6A) were determined by analyzing the sequences of all tRNA genes in *L. major*, *T. brucei* and *T. cruzi* and comparing them to the sequences of Boxes A and B from *S. cerevisiae* [9]. The tRNAs were divided into two classes, depending on the size of the variable loop (Fig. 3). Class I tRNAs have a short variable loop of 4 or 5 nucleotides, whereas class II tRNAs possess a long variable arm, with a double helical stem of 3 to 7 base pairs and a loop of 3 to 5 nucleotides (Fig. 3) [9]. In the Trityps, 43 genes belong to class II (all Leu and Ser tRNA genes, but excluding tRNA-Sec genes), and 215 genes are class I. Since we observed sequence differences between class I and class II tRNA genes, we analyzed them separately. Half of the bases from the consensus sequence of Box A (positions 1, 2, 4, 7, 10 and 11) are identical between class I and class II genes (and identical to the *S. cerevisiae* consensus sequence) (Fig. 6A). However, position 5 is different between both classes, since class I genes may have any nucleotide, whereas class II genes always have a C. Also, class I tRNA genes present C or T at position 6, while class II genes always have A or

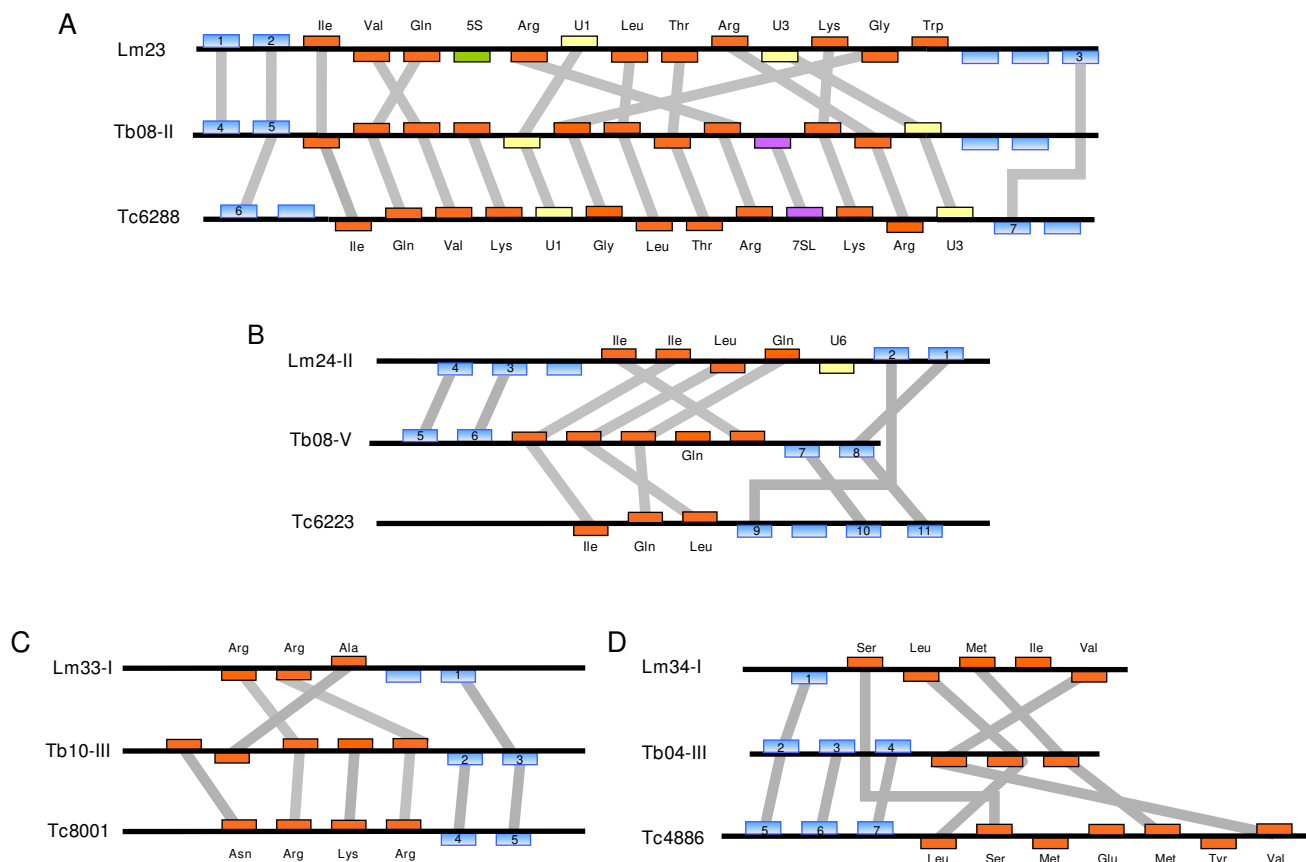


Figure 5
Comparative order of Pol III-transcribed genes in Trityps. Clusters that present some degree of synteny among Trityps are shown. In *L. major*, they correspond to loci chr23 (panel A), chr24-II (panel B), chr33-I (panel C) and chr34-I (panel D). The order of genes in clusters chr24-II and chr34-I was inverted compared to the maps shown in figure 1. The corresponding orthologous regions in *T. brucei* and *T. cruzi* are indicated. Orthologous genes are joined by grey lines. Figure is not to scale.

G. Regarding Box B, position 8 is different between both classes: an A is always present in class II genes, while class I genes may have any nucleotide. Around 20% of the class I tRNA genes in Trityps have an additional nucleotide in Box A, between positions 9 and 10 (marked with an asterisk in Fig. 6A). These tRNA genes are: eMet-CAT, Asn-GTT, Ile-AAT, Ile-TAT, Lys-CTT, Phe-GAG and Tyr-GTA. Regarding class II tRNAs, the Leu-TAA genes have a T between positions 9 and 10 of Box A (Fig. 6A).

A few exceptions to the consensus sequence were found among Trityps. For Box A from class I genes, these include the following: four tRNA-Val genes (LmjF09.VAL.01, LmjF09.VAL.02, Tb08_tRNA_Val_1 and Tc00.1047053506459.249) have a C at position 2 (instead of A or G); all four tRNA-Val-TAC genes (LmjF23.VAL.01, Tb08_tRNA_Val_1, Tc00.1047053504427.233 and Tc00.1047053508043.13) present an A at position 3 (instead of G); all six tRNA-Ala-AGC genes (LmjF17.TRNA

ALA.01, LmjF31.TRNAALA.01, Tb07_tRNA_Ala_1, Tb07_tRNA_Ala_2, Tc00.1047053510057.40 and Tc00.1047053508909.130) have an extra base (an A) between positions 8 and 9; and the six tRNA-Val-AAC genes from Trityps (LmjF21.TRNAVAL.01, LmjF34.TRNAVAL.01, Tb_04_tRNA_Val_1, Tb_04_tRNA_Val_2, Tc00.1047053506435.363 and Tc00.1047053504055.95) also have an additional base (a G) between positions 8 and 9.

Concerning class II genes, the exceptions to Box A consensus sequence are the four tRNA-Leu-TAA genes present in the Trityps genomes (LmjF24.TRNALEU.01, Tb08_tRNA_Leu_2, Tc00.1047053510721.13 and Tc00.1047053511909.9), which have a T at position 2 (instead of G or A). In regard to Box B, the genes that do not have the consensus sequence are: LmjF09.TRNAHIS.01 and LmjF09.TRNAHIS.02 present a C at position 4; Tb07_tRNA_Ala_3 has a T at position 5;

A

	Box A											
	1	2	3	4	5	6	7	8	9	*	10	11
Trityps class I	T	G/A	G	C/T	N	C/T	A	G/A	C/T	C/T	G	G
Trityps class II	T	G/A	G/C	C/T	C	G/A	A	G	T	T	G	G
<i>S. cerevisiae</i>	T	G/A	G	C/T	N	N	A	G/A	H	H	G	G

	Box B										
	1	2	3	4	5	6	7	8	9	10	11
Trityps class I	G/A	G	T	T	C	G/A	A	N	H	C	C/T
Trityps class II	G	G	T	T	C	G/A	A	A	C/T	C	C
<i>S. cerevisiae</i>	N	G	T	T	C	G/A	A	N	C/T	C	N

B

	tRNA-Sec box A													
<i>H. sapiens</i>	T	G	<u>A</u>	T	C	C	T	C	A	G	T	G	G	13
<i>X. laevis</i>	T	G	<u>A</u>	C	C	C	T	C	A	G	T	G	G	13
<i>D. melanogaster</i>	T	G	<u>A</u>	A	C	T	T	C	G	G	T	G	G	13
<i>C. elegans</i>	T	G	<u>A</u>	A	C	C	A	T	G	G	C	G	G	13
<i>P. falciparum</i>	T	G	<u>A</u>	G	T	T	A	G	C	A	T	G	G	13
Trityps	T	G	<u>A</u>	G	C	T	C	A	G	C	T	G	G	13
	*	*	*								*	*		

Figure 6

Consensus sequences for Box A and Box B from tRNA genes in Trityps. Sequences from class I and class II tRNA genes from Trityps are compared to the *S. cerevisiae* consensus sequences (panel A). In Box A, some tRNA genes contain an extra base between positions 9 and 10 (marked with an asterisk) (see text). Conserved positions are shown in bold type. Panel B shows a comparison of Box A from selenocysteine tRNA genes from the indicated species. Conserved nucleotides are indicated with an asterisk. An A in the third position (in bold and underlined) seems to be specific to Box A from Sec genes. H represents C, T or A.

Tc00.1047053511241.10 presents an A at position 10; and tRNA-Ala-AGC and tRNA-iMet genes have an A at position 3 (data not shown).

Analysis of the promoter sequences from tRNA-Sec genes in Trityps indicated that Box A contains an additional A between bases 2 and 3, compared with the consensus sequences (see Fig. 6B). This insertion was previously reported in tRNA-Sec genes from other organisms [35] (Fig. 6B). Regarding Box B, tRNA-Sec genes from Trityps present two changes compared to the highly conserved consensus sequence: a C at position 1 (instead of a G) and a G at position 11 (in lieu of C) (data not shown). In other species, the sequence of Box B from tRNA-Sec is identical

to the corresponding consensus sequence. In *Xenopus laevis*, transcription of tRNA-Sec genes is directed by three extragenic domains (a TATA Box, a proximal sequence element and an activator element) and Box B. Apparently, Box A is not part of the promoter [36,37]. Since both internal control elements from Sec genes in Trityps differ from the corresponding consensus sequences, it is possible that synthesis of tRNA-Sec is regulated only by external elements in these parasites. We are currently exploring this possibility.

tRNA isodecoder genes

Sequence analysis of isoacceptor tRNAs in several organisms indicated the presence of tRNA isodecoder genes

(tRNA genes having the same anticodon but different sequences elsewhere in the tRNA body) [10]. In eukaryotes, the number of isodecoder genes ranges from 10 (yeast) to 246 (chimp), while in bacterial genomes the number of isodecoders varies from 0 to 26 [10]. By comparing the sequences of isoacceptor tRNAs in the Trityps, one isodecoder gene was found in *L. major* (tRNA-Ser-GCT), and one was found in *T. brucei* (tRNA-Ala-CGC) (Fig. 7 and Table 1). Since there are only two copies of the corresponding isoacceptor class in each case, we have arbitrarily designated LmtRNA-Ser.01 (LmjF17.TRNASER.01) and TbtRNA-Ala.01 (Tb07_tRNA_Ala_3) as the isodecoders. Sequence identity between the isodecoder and the "majority member" is 98% in *L. major* and 97% in *T. brucei*. While in *L. major* the sequence difference locates near Box A, in *T. brucei* one of the two observed differences lies in one of the conserved bases of Box B (Fig. 7). In *T. cruzi*, six isodecoder genes were identified: TctRNA-Glu.01 (Tc00.1047053506435.336), TctRNA-Ala.01 (Tc00.1047053510057.40), TctRNA-Ala.03 (Tc00.1047053475029.40), TctRNA-Arg.01 (Tc00.1047053504427.243), TctRNA-His.01 (Tc00.1047053511241.10) and TctRNA-Ser.01 (Tc00.1047053510057.50) (Table 1 and Fig. 7). In four cases (TctRNA-Glu.01, TctRNA-Ala.03, TctRNA-Arg.01 and TctRNA-His.01), sequence differences were located to variable nucleotides from Box B (Fig. 7). As in Trityps, sequence variations between human tRNA isodecoders have been located within internal control elements [10]. In such cases, changes were found in variable nucleotides from Boxes A and B. Thus, the occurrence of changes within internal control elements in tRNA isodecoder genes suggests that differential regulation of Pol III transcription is possible in Trityps; the fact that the highly conserved C at position 10 of Box B from TbtRNA-Ala.02 is changed to T in the corresponding tRNA isodecoder (TbtRNA-Ala.01) (Fig. 7) strongly supports this possibility. Sequence changes in isodecoders are not only restricted to internal control elements, but they might be present all along the tRNA body (Fig. 7) [10]. Therefore, the diversity of tRNA genes is much higher than originally thought. The functional meaning of such diversity has yet to be investigated.

Signals for transcription termination

A cluster of several T residues in the coding DNA strand acts as a signal to terminate Pol III transcription [33]. The cluster of Ts is usually located within the first 30 bases following the gene. In human and mice, Pol III needs four Ts to end transcription, and tRNA genes that have five or more Ts are very rare in these species. On the other hand, in the genomes of *S. pombe* and *S. cerevisiae* the majority of the tRNA genes have six and seven Ts, respectively [38,39]. Interestingly, they do not have any single gene whose termination signal is shorter than five Ts. For a par-

ticular species, termination efficiency tends to increase with the length of the T run. In *L. major*, it has been shown that transcription of the tRNA located on chromosome 3 terminates within a tract of four Ts [6]. To gain insight into Pol III termination signals in trypanosomatids, we decided to analyze the sequences downstream of all the tRNA genes. A cluster of Ts of variable length was found on every single tRNA gene in the Trityps (see Additional File 2); the distance between the end of the gene and the run of Ts varies from zero to seven bases. In *L. major*, the mean length of the run of Ts is 4.87 bases, with a minimum of four and a maximum of nine Ts (Fig. 8, panels A and D). Similar results were obtained in *T. brucei*, where the average T-run length is 4.89 bases (ranging from four to ten Ts) (Fig. 8, panels B and D). In the tRNA genes from *T. cruzi*, however, the stretches of Ts are longer, showing a mean length of 6.56, with two genes presenting a run of 16 consecutive T residues (Fig. 8C and 8D).

The presence of a second stretch of Ts that acts as a potential "back up" termination signal is a common feature in tRNA genes from eukaryotes [38]. The second run of Ts is normally located within the first 30 bp downstream of the first one. In the case of *S. cerevisiae*, *S. pombe*, *H. sapiens* and *Mus musculus*, the percentages of tRNA genes that have a back up T-run are 44, 53, 31 and 33%, respectively [38]. Analysis of the sequences downstream of the T-runs in *T. cruzi* showed that 58 tRNA genes (48.3%) have a back up T-run, whose length is between 4 to 15 bases (Additional File 2). Therefore, the percentage of tRNA genes with a second run of Ts in *T. cruzi* is very similar to that found in *S. cerevisiae* and *S. pombe*. Interestingly, only 13 tRNA genes in *L. major* (15.6%) and 18 genes in *T. brucei* (27%) present a back up T-track (which is from 4 to 10 bases long) (see Additional File 2); these percentages are even smaller than those found in mammals. Thus, in *L. major* and *T. brucei* a single and short run of Ts seems to be sufficient to achieve proper transcription termination in the majority of the tRNA genes. On the other hand, *T. cruzi* seems to require longer T-runs, in addition to a second T stretch, to direct transcription termination. This indicates that the mechanism of Pol III transcription termination in *L. major* and *T. brucei* might be different from that one in *T. cruzi* and other eukaryotes. Sequences downstream and upstream of the run of Ts might contribute to the strength of the terminator, as observed in some tRNAs in *S. cerevisiae* [38].

Although most tRNA genes are clustered in trypanosomatids, the presence of runs of Ts located downstream of all the tRNA genes suggests that they are transcribed as monocistrons, which is common among eukaryotes. In prokaryotic cells genes encoding tRNAs are transcribed in either a monocistronic or polycistronic manner. In the latter case, an RNA containing several tRNA precursors in

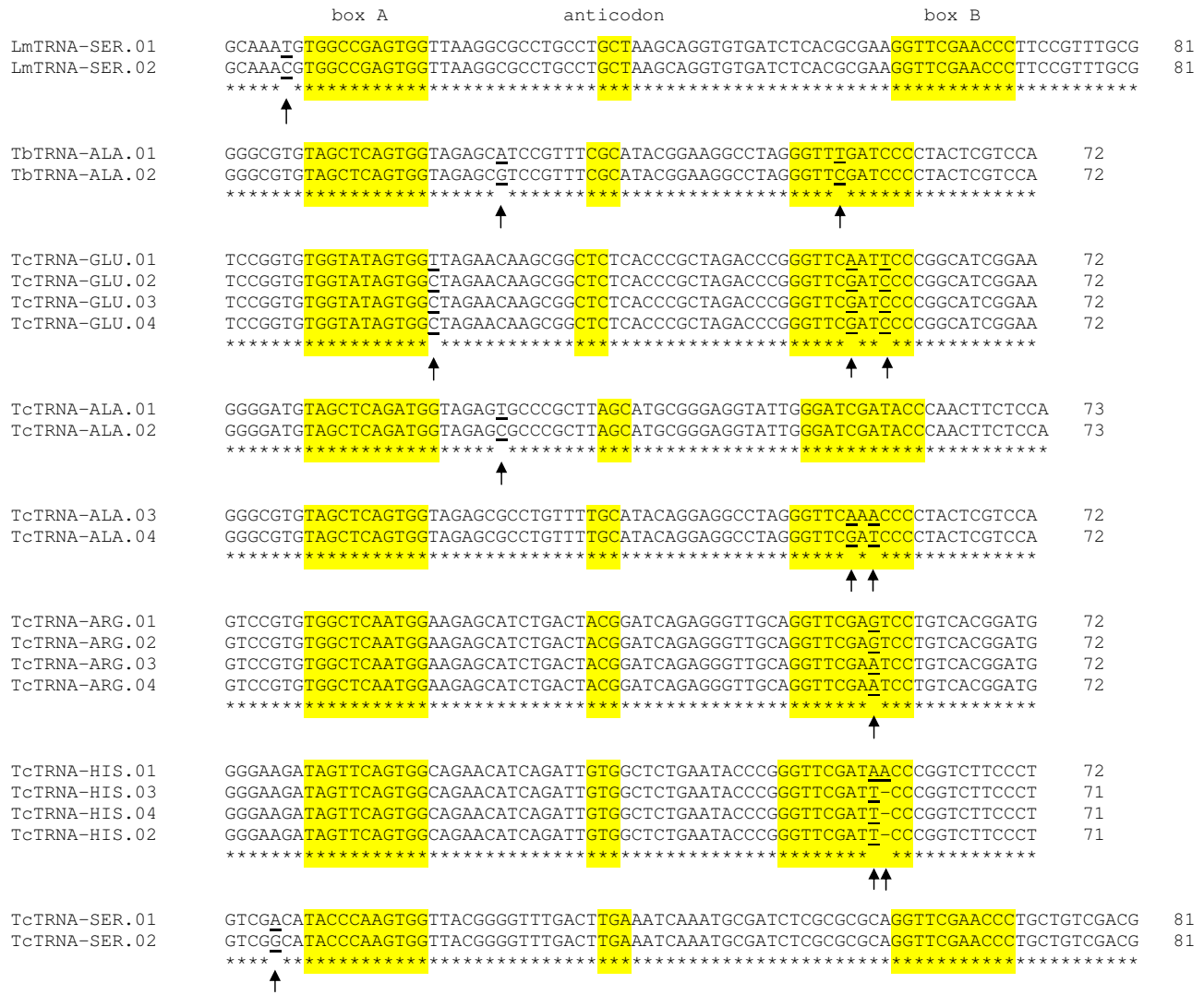


Figure 7
Sequence comparison of tRNA isodecoder genes in the Trityps. One isodecoder tRNA gene was found in *L. major* (LmTRNA-Ser) and *T. brucei* (TbTRNA-Ala), and six were found in *T. cruzi* (TcTRNA-Glu, -Ala.01, -Ala.03, -Arg, -His and -Ser). Bases that show variation are indicated in bold and underlined, and marked with an arrow. The position of internal control elements (Boxes A and B) and the anticodon are indicated. The genes included in this figure are the following (GeneDB names in parentheses): LmTRNA-SER.01 (LmjF17.TRNASER.01), LmTRNA-SER.02 (LmjF21.TRNASER.01), TbTRNA-ALA.01 (Tb07_tRNA_Ala_3), TbTRNA-ALA.02 (Tb11_tRNA_Ala_1), TcTRNA-GLU.01 (Tc00.1047053506435.336), TcTRNA-GLU.02 (Tc00.1047053504055.89), TcTRNA-GLU.03 (Tc00.1047053508999.180), TcTRNA-GLU.04 (Tc00.1047053510959.8), TcTRNA-ALA.01 (Tc00.1047053510057.40), TcTRNA-ALA.02 (Tc00.1047053508909.130), TcTRNA-ALA.03 (Tc00.1047053475029.40), TcTRNA-ALA.04 (gene located in contig 8001, not annotated in geneDB), TcTRNA-ARG.01 (Tc00.1047053504427.243), TcTRNA-ARG.02 (Tc00.1047053506619.59), TcTRNA-ARG.03 (Tc00.1047053508043.23), TcTRNA-ARG.04 (Tc00.1047053511191.29), TcTRNA-HIS.01 (Tc00.1047053511241.10), TcTRNA-HIS.03 (Tc00.1047053508087.5), TcTRNA-HIS.04 (Tc00.1047053508861.10), TcTRNA-HIS.02 (Tc00.1047053506663.10), TcTRNA-SER.01 (Tc00.1047053510057.50), TcTRNA-SER.02 (Tc00.1047053508909.120).

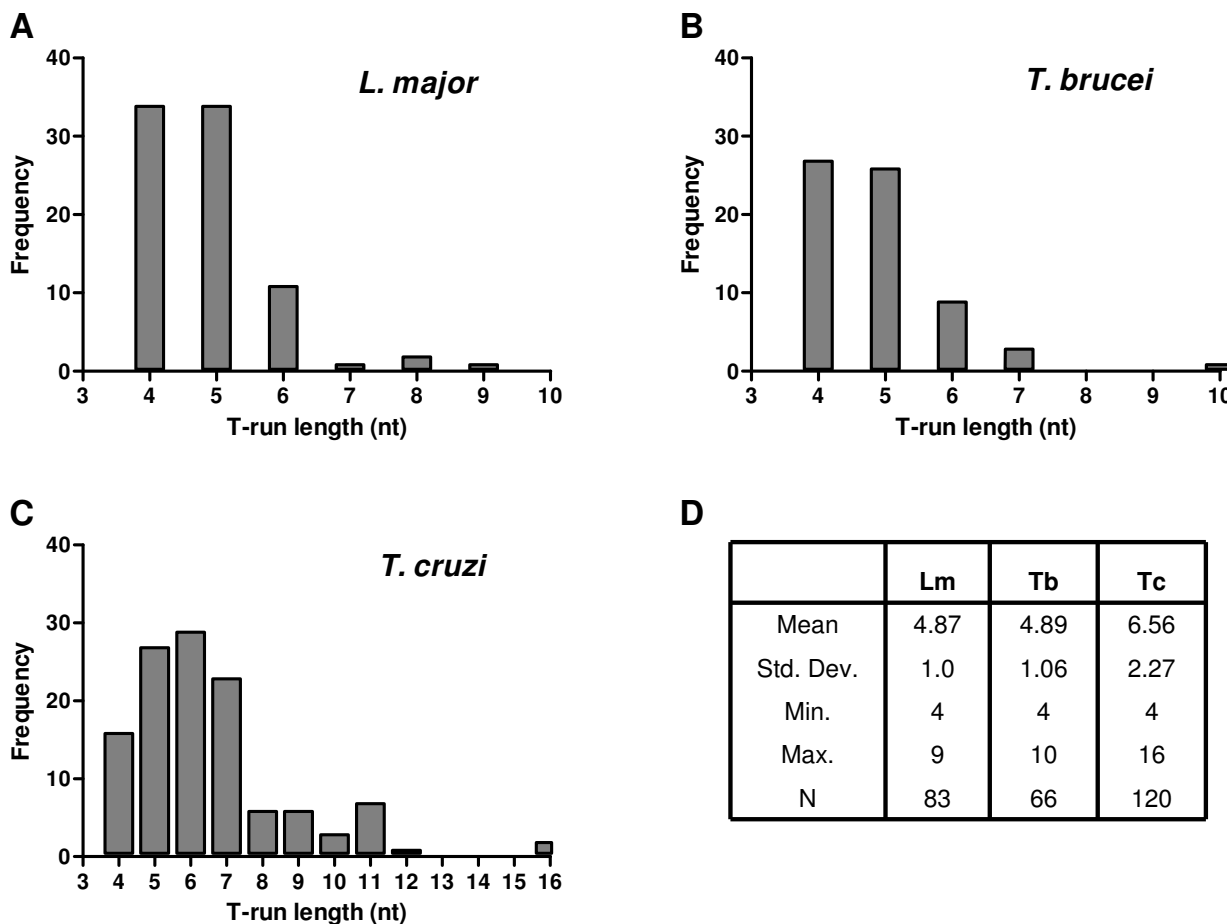


Figure 8
Length distribution of termination signals from tRNA genes in Trityps. The size of the run of Ts in every tRNA gene from *L. major* (panel A), *T. brucei* (panel B) and *T. cruzi* (panel C) was plotted against frequency. Statistical data is shown in panel D.

tandem is processed to yield functional tRNAs [40]. In plants, dicistronic transcripts containing a tRNA and a snoRNA have been found [41]. Moreover, the presence of precursor RNA molecules that contain both a tRNA and a mRNA has been reported in *E. coli* [42].

Conclusion

In comparison to most eukaryotic organisms, Trityps present a low number of tRNA genes. A total of 46 isoacceptor types were identified, which are able to read the 61 codons that specify the canonical amino acids, in addition to Sec. Trypanosomatids use the A1- or G1-sparing strategy as a decoding mode, by allowing flexible base pairing between G1 or A1 of the anticodon and C3 or U3 in the codon. Most tRNA genes in Trityps are organized into clusters (from 2–10 genes) that may contain other Pol III genes. Some of the clusters show a remarkable conservation of gene order among Trityps. The distribution of

tRNA genes in the genomes of *L. major* and *T. brucei* does not seem to be totally random. We also found that 14 of the 39 convergent strand-switch regions present in the *L. major* genome are separated by at least one tRNA gene, which raise the possibility that other tRNA genes (in addition to the one present on chromosome 3) are involved in termination of Pol II transcription of convergent PGCs in this parasite. A run of Ts of variable length was found downstream of all the 269 tRNA genes present in the Trityps. In *T. cruzi* the clusters of Ts are larger than in *L. major* and *T. brucei* (an average of 6 Ts versus 5 Ts, respectively); moreover, the presence of a back up T run is more common in *T. cruzi* than in the other two Trityps. Analysis of the internal promoter elements allowed us to establish consensus sequences for Boxes A and B of class I and class II tRNA genes. Interestingly, special characteristics were found in Boxes A and B from tRNA-Sec genes in Trityps, which suggests that the mechanisms that regulate their

transcription might be different from those of other tRNA genes. Lastly, we have identified several tRNA isodecoder genes in the *Trityps*, especially in *T. cruzi*. The fact that in some cases the sequence differences occur within the internal promoter elements suggests the possibility of differential expression of tRNA genes in these organisms.

Methods

All tRNA genes annotated in the *L. major*, *T. brucei* and *T. cruzi* genome databases <http://www.genedb.org> (versions 2.1) were analyzed with the tRNAscan-SE program <http://lowelab.ucsc.edu/tRNAscan-SE/> [43] to verify the presence and identity of the tRNAs. In addition to the examination of all the features associated to the typical tRNA cloverleaf structure, we also analyzed the presence of internal promoter elements (Boxes A and B) and T-tracts at the 3' end of the tRNA gene (which should be present in tRNA genes, but not necessarily in pseudogenes). Sequence comparisons among putative tRNA isoacceptors were performed using the ClustalW2 program <http://www.ebi.ac.uk/Tools/clustalw2/index.html>. Information for the genomic and synteny maps was obtained from the GeneDB databases. BLAST searches were performed in these databases to locate the tRNA-Sec genes [13], as well as the sRNA76 [12]. Codon usage data was obtained from the Sanger site <http://www.sanger.ac.uk/> for *L. major*, and from the Kazusa web page <http://www.kazusa.or.jp/codon/> for *T. brucei*. Codon usage data for *T. cruzi* was calculated by analyzing coding sequences (obtained from the Sanger site) in the codon usage page from SMS http://www.bioinformatics.org/sms2/codon_usage.html. The Spearman correlation analysis and the descriptive statistical analysis of T-run data were performed with the GraphPad Prism5 program <http://www.graphpad.com>.

Authors' contributions

NPM performed data collection, synteny analysis, sequence comparisons, location of tRNA-Sec genes and the statistical analysis of the data. She also carried out the analysis of the transcription terminations signals and helped draft the manuscript. LEFM analyzed the tRNA genes with the tRNAscan-SE program and helped draw the genome maps. EEFA helped collect data from the *T. brucei* databases, and analyzed the sequences to obtain the consensus sequences for boxes A and B. RGMC was responsible for data collection from the *T. cruzi* database and participated in the codon usage analysis. RHR helped draft the manuscript and participated in the statistical analysis of the data. PJM 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.

Additional material

Additional file 1

Table S1. List of tRNA genes in the *Trityps*. In *L. major* and *T. brucei*, genes are shown by chromosomal location. Other Pol III-transcribed genes that are associated to tRNA genes are also listed. In *T. brucei*, the names of the genes correspond to the temporary systematic name. For those genes that already have a permanent systematic name in the GeneDB database, it is indicated between parentheses. tRNA genes in *T. cruzi* are organized by amino acid.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-232-S1.xls>]

Additional file 2

Table S2. tRNA-gene transcription termination signals in *Trityps*. The sequences shown start immediately downstream of the end of the tRNA genes. Clusters of Ts are shown in bold type. Those genes that present a second (back up) T-run of 4 or more bases are highlighted in yellow.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-232-S2.xls>]

Acknowledgements

We thank Ana M Cevallos-Gaos for fruitful discussions. This work was supported by grants 1 R01 TW007255-01 from The Fogarty International Center of NIH, 47543 from CONACyT and IN203606-3 from PAPIIT to S. Martínez-Calvillo. N.E. Padilla-Mejía (Posgrado en Ciencias Biológicas, UNAM) was recipient of a fellowship from CONACyT.

References

- Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, Sisk E, Rajandream MA, Adlem E, Aert R, Anupama A, Apostolou Z, Attipoe P, Bason N, Bauser C, Beck A, Beverley SM, Bianchetti G, Borzym K, Bothe G, Bruschi CV, Collins M, Cadag E, Ciarlioni L, Clayton C, Coulson RM, Cronin A, Cruz AK, Davies RM, De GJ, Dobson DE, Duesterhoeft A, Fazelina G, Fosker N, Frasch AC, Fraser A, Fuchs M, Gabel C, Goble A, Goffeau A, Harris D, Hertz-Fowler C, Hilbert H, Horn D, Huang Y, Klages S, Knights A, Kube M, Larke N, Litvin L, Lord A, Louie T, Marra M, Masuy D, Matthews K, Michaeli S, Mottram JC, Muller-Auer S, Munden H, Nelson S, Norbertczak H, Oliver K, O'neil S, Pentony M, Pohl TM, Price C, Purnelle B, Quail MA, Rabinowitsch E, Reinhardt R, Rieger M, Rinta J, Robben J, Robertson L, Ruiz JC, Rutter S, Saunders D, Schafer M, Schein J, Schwartz DC, Seeger K, Seyler A, Sharp S, Shin H, Sivam D, Squares R, Squares S, Tosato V, Vogt C, Volckaert G, Wambutt R, Warren T, Wedler H, Woodward J, Zhou S, Zimmermann W, Smith DF, Blackwell JM, Stuart KD, Barrell B, Myler PJ: **The genome of the kinetoplastid parasite, *Leishmania major*.** *Science* 2005, **309**:436-442.
- Berriman M, Ghedin E, Hertz-Fowler C, Blandin G, Renauld H, Bartholomeu DC, Lennard NJ, Caler E, Hamlin NE, Haas B, Bohme U, Hannick L, Aslett MA, Shallom J, Marcello L, Hou L, Wickstead B, Alsmark UC, Arrowsmith C, Atkin RJ, Barron AJ, Bringaud F, Brooks K, Carrington M, Cherevach I, Chillingworth TJ, Churcher C, Clark LN, Corton CH, Cronin A, Davies RM, Doggett J, Djikeng A, Feldblyum T, Field MC, Fraser A, Goodhead I, Hance Z, Harper D, Harris BR, Hauser H, Hostetler J, Ivens A, Jagels K, Johnson D, Johnson J, Jones K, Kerhornou AX, Koo H, Larke N, Landfear S, Larkin C, Leech V, Line A, Lord A, Macleod A, Mooney PJ, Moule S, Martin DM, Morgan GW, Mungall K, Norbertczak H, Ormond D, Pai G, Peacock CS, Peterson J, Quail MA, Rabinowitsch E, Rajandream MA, Reitter C, Salzberg SL, Sanders M, Schobel S, Sharp S, Simmonds M, Simpson AJ, Tallon L, Turner CM, Tait A, Tivey AR, Van AS, Walker D, Wanless D, Wang S, White B, White O, Whitehead S, Woodward J, Wortman

- J, Adams MD, Embley TM, Gull K, Ullu E, Barry JD, Fairlamb AH, Opperdoes F, Barrrell BG, Donelson JE, Hall N, Fraser CM, Melville SE, El-Sayed NM: **The genome of the African trypanosome *Trypanosoma brucei***. *Science* 2005, **309**:416-422.
3. El-Sayed NM, Myler PJ, Bartholomeu DC, Nilsson D, Aggarwal G, Tran AN, Ghedin E, Worthey EA, Delcher AL, Blandin G, Westenberg S, Caler E, Cerqueira GC, Branche C, Haas B, Anupama A, Arner E, Aslund L, Attipoe P, Bontempi E, Bringaud F, Burton P, Cadag E, Campbell DA, Carrington M, Crabtree J, Darban H, da Silveira JF, de JP, Edwards K, Englund PT, Fazelina G, Feldblyum T, Ferella M, Frasch AC, Gull K, Horn D, Hou L, Huang Y, Kindlund E, Klingbeil M, Kluge S, Koo H, Lacerda D, Levin MJ, Lorenzi H, Louie T, Machado CR, McCulloch R, McKenna A, Mizuno Y, Mottram JC, Nelson S, Ochaya S, Osoegawa K, Pai G, Parsons M, Pentony M, Pettersson U, Pop M, Ramirez JL, Rinta J, Robertson L, Salzberg SL, Sanchez DO, Seyler A, Sharma R, Shetty J, Simpson AJ, Sisk E, Tammi MT, Tarleton R, Teixeira S, Van AS, Vogt C, Ward PN, Wickstead B, Wortman J, White O, Fraser CM, Stuart KD, Andersson B: **The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease**. *Science* 2005, **309**:409-415.
 4. Campbell DA, Thomas S, Sturm NR: **Transcription in kinetoplastid protozoa: why be normal?** *Microbes Infect* 2003, **5**:1231-1240.
 5. Martinez-Calvillo S, Yan S, Nguyen D, Fox M, Stuart K, Myler PJ: **Transcription of *Leishmania major* Friedlin chromosome 1 initiates in both directions within a single region**. *Mol Cell* 2003, **11**:1291-1299.
 6. Martinez-Calvillo S, Nguyen D, Stuart K, Myler PJ: **Transcription initiation and termination on *Leishmania major* chromosome 3**. *Eukaryot Cell* 2004, **3**:506-517.
 7. El-Sayed NM, Myler PJ, Blandin G, Berriman M, Crabtree J, Aggarwal G, Caler E, Renauld H, Worthey EA, Hertz-Fowler C, Ghedin E, Peacock C, Bartholomeu DC, Haas BJ, Tran AN, Wortman JR, Alsmark UC, Angiuoli S, Anupama A, Badger J, Bringaud F, Cadag E, Carlton JM, Cerqueira GC, Creasy T, Delcher AL, Djikeng A, Embley TM, Hauser C, Ivans AC, Kummerfeld SK, Pereira-Leal JB, Nilsson D, Peterson J, Salzberg SL, Shallow J, Silva JC, Sundaram J, Westenberg S, White O, Melville SE, Donelson JE, Andersson B, Stuart KD, Hall N: **Comparative genomics of trypanosomatid parasitic protozoa**. *Science* 2005, **309**:404-409.
 8. Geiduschek EP, Kassavetis GA: **The RNA polymerase III transcription apparatus**. *J Mol Biol* 2001, **310**:1-26.
 9. Marck C, Grosjean H: **tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features**. *RNA* 2002, **8**:1189-1232.
 10. Goodenbour JM, Pan T: **Diversity of tRNA genes in eukaryotes**. *Nucleic Acids Res* 2006, **34**:6137-6146.
 11. Paule MR, White RJ: **Survey and summary: transcription by RNA polymerases I and III**. *Nucleic Acids Res* 2000, **28**:1283-1298.
 12. Beja O, Ullu E, Michaeli S: **Identification of a tRNA-like molecule that copurifies with the 7SL RNA of *Trypanosoma brucei***. *Mol Biochem Parasitol* 1993, **57**:223-229.
 13. Cassago A, Rodrigues EM, Prieto EL, Gaston KW, Alfonso JD, Iribar MP, Berry MJ, Cruz AK, Thiemann OH: **Identification of *Leishmania* selenoproteins and SECIS element**. *Mol Biochem Parasitol* 2006, **149**:128-134.
 14. Allmang C, Krol A: **Selenoprotein synthesis: UGA does not end the story**. *Biochimie* 2006, **88**:1561-1571.
 15. Lobanov AV, Gromer S, Salinas G, Gladyshev VN: **Selenium metabolism in *Trypanosoma*: characterization of selenoproteomes and identification of a Kinetoplastida-specific selenoprotein**. *Nucleic Acids Res* 2006, **34**:4012-4024.
 16. Tan TH, Pach R, Crausaz A, Ivans A, Schneider A: **tRNAs in *Trypanosoma brucei*: genomic organization, expression, and mitochondrial import**. *Mol Cell Biol* 2002, **22**:3707-3717.
 17. Horn D: **Codon usage suggests that translational selection has a major impact on protein expression in trypanosomatids**. *BMC Genomics* 2008, **9**:2.
 18. Drabkin HJ, Estrella M, Rajbhandary UL: **Initiator-elongator discrimination in vertebrate tRNAs for protein synthesis**. *Mol Cell Biol* 1998, **18**:1459-1466.
 19. Astrom SU, von Pawel-Rammingen U, Bystrom AS: **The yeast initiator tRNA^{Met} can act as an elongator tRNA(Met) in vivo**. *J Mol Biol* 1993, **233**:43-58.
 20. Mandal N, Mangroo D, Dalluge JJ, McCloskey JA, Rajbhandary UL: **Role of the three consecutive G:C base pairs conserved in the anticodon stem of initiator tRNAs in initiation of protein synthesis in *Escherichia coli***. *RNA* 1996, **2**:473-482.
 21. Percudani R, Pavesi A, Ottonello S: **Transfer RNA gene redundancy and translational selection in *Saccharomyces cerevisiae***. *J Mol Biol* 1997, **268**:322-330.
 22. Rocha EP: **Codon usage bias from tRNA's point of view: redundancy, specialization, and efficient decoding for translation optimization**. *Genome Res* 2004, **14**:2279-2286.
 23. Duret L: **tRNA gene number and codon usage in the *C. elegans* genome are co-adapted for optimal translation of highly expressed genes**. *Trends Genet* 2000, **16**:287-289.
 24. Hasan G, Turner MJ, Cordingley JS: **Ribosomal RNA genes of *Trypanosoma brucei*: mapping the regions specifying the six small ribosomal RNAs**. *Gene* 1984, **27**:75-86.
 25. Hernandez-Rivas R, Martinez-Calvillo S, Romero M, Hernandez R: ***Trypanosoma cruzi* 5S rRNA genes: molecular cloning, structure and chromosomal organization**. *FEMS Microbiol Lett* 1992, **71**:63-67.
 26. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Showlken R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendt MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubinfeld M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la BM, Dedhia N, Blocker H, Hornischer K, Nordseik G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kasprzyk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de JP, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ: **Initial sequencing and analysis of the human genome**. *Nature* 2001, **409**:860-921.
 27. Thompson M, Haeusler RA, Good PD, Engelke DR: **Nucleolar clustering of dispersed tRNA genes**. *Science* 2003, **302**:1399-1401.
 28. Takahashi K, Murakami S, Chikashige Y, Niwa O, Yanagida M: **A large number of tRNA genes are symmetrically located in fission yeast centromeres**. *J Mol Biol* 1991, **218**:13-17.
 29. Scott KC, White CV, Willard HF: **An RNA polymerase III-dependent heterochromatin barrier at fission yeast centromere I**. *PLoS ONE* 2007, **2**:e1099.
 30. Obado SO, Bot C, Nilsson D, Andersson B, Kelly JM: **Repetitive DNA is associated with centromeric domains in *Trypanosoma brucei* but not *Trypanosoma cruzi***. *Genome Biol* 2007, **8**:R37.

31. Thomas S, Green A, Sturm NR, Campbell DA, Myler PJ: **Histone acetylations mark origins of polycistronic transcription in *Leishmania major*.** *BMC Genomics* 2009, **10**:152.
32. Tripathy S, Tyler BM: **The repertoire of transfer RNA genes is tuned to codon usage bias in the genomes of *Phytophthora sojae* and *Phytophthora ramorum*.** *Mol Plant Microbe Interact* 2006, **19**:1322-1328.
33. Geiduschek EP, Tocchini-Valentini GP: **Transcription by RNA polymerase III.** *Annu Rev Biochem* 1988, **57**:873-914.
34. Brun YV, Sanfacon H, Breton R, Lapointe J: **Closely spaced and divergent promoters for an aminoacyl-tRNA synthetase gene and a tRNA operon in *Escherichia coli*. Transcriptional and post-transcriptional regulation of *glxX*, *valU* and *alaW*.** *J Mol Biol* 1990, **214**:845-864.
35. Lee BJ, Kang SG, Hatfield D: **Transcription of *Xenopus* selenocysteine tRNA^{Ser} (formerly designated opal suppressor phosphoserine tRNA) gene is directed by multiple 5'-extragenic regulatory elements.** *J Biol Chem* 1989, **264**:9696-9702.
36. Carbon P, Krol A: **Transcription of the *Xenopus laevis* selenocysteine tRNA(Ser)^{Sec} gene: a system that combines an internal B box and upstream elements also found in U6 snRNA genes.** *EMBO J* 1991, **10**:599-606.
37. Park JM, Yang ES, Hatfield DL, Lee BJ: **Analysis of the selenocysteine tRNA(Ser)^{Sec} gene transcription in vitro using *Xenopus* oocyte extracts.** *Biochem Biophys Res Commun* 1996, **226**:231-236.
38. Braglia P, Percudani R, Dieci G: **Sequence context effects on oligo(dT) termination signal recognition by *Saccharomyces cerevisiae* RNA polymerase III.** *J Biol Chem* 2005, **280**:19551-19562.
39. Matsuzaki H, Kassavetis GA, Geiduschek EP: **Analysis of RNA chain elongation and termination by *Saccharomyces cerevisiae* RNA polymerase III.** *J Mol Biol* 1994, **235**:1173-1192.
40. Vold BS, Green CJ, Narasimhan N, Strem M, Hansen JN: **Transcriptional analysis of *Bacillus subtilis* rRNA-tRNA operons. II. Unique properties of an operon containing a minor 5 S rRNA gene.** *J Biol Chem* 1988, **263**:14485-14490.
41. Kruszka K, Barneche F, Guyot R, Ailhaas J, Meneau I, Schiffer S, Marchfelder A, Echeverria M: **Plant dicistronic tRNA-snoRNA genes: a new mode of expression of the small nucleolar RNAs processed by RNase Z.** *EMBO J* 2003, **22**:621-632.
42. van Delft JH, Schmidt DS, Bosch L: **The tRNA-tufB operon transcription termination and processing upstream from tufB.** *J Mol Biol* 1987, **197**:647-657.
43. Lowe TM, Eddy SR: **tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence.** *Nucleic Acids Res* 1997, **25**:955-964.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

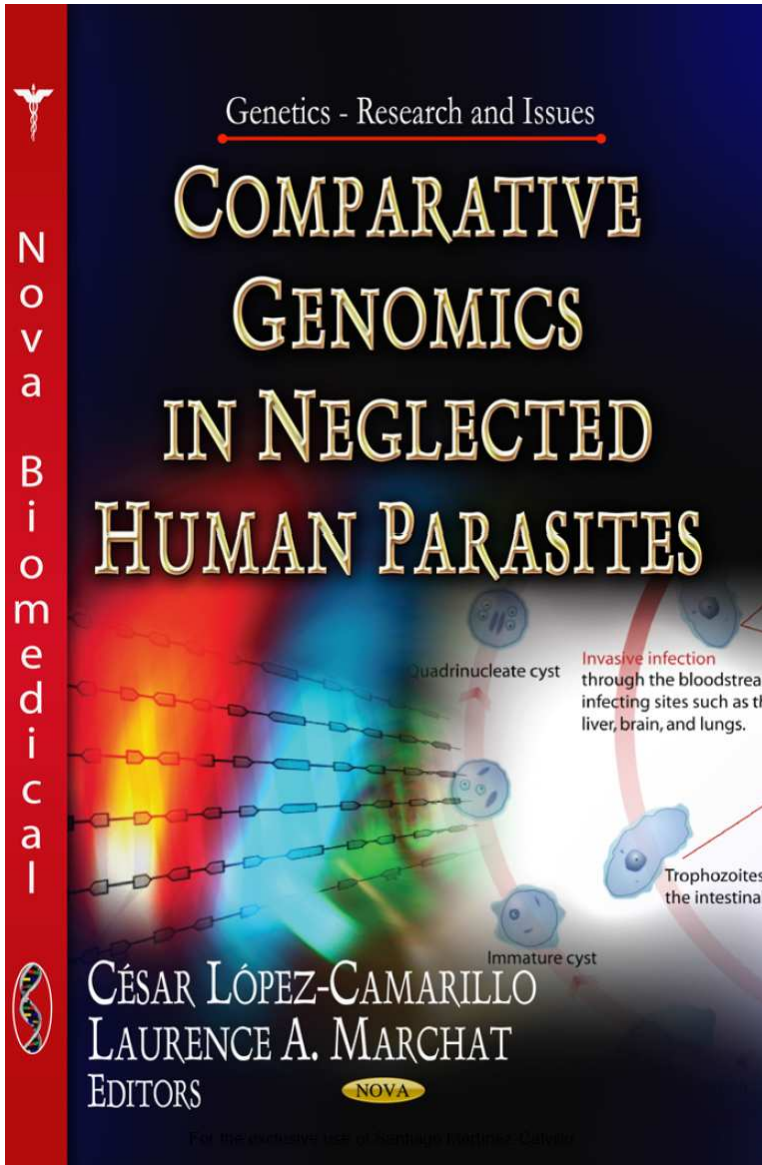
- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp



APÉNDICE VIII: Capítulo de libro de la alumna como primer autor relacionado con el presente trabajo

Padilla-Mejía, N.E., Gómez-Hurtado, C. M., Sánchez-Santamaría, I. I., Florencio-Martínez, L. E., Manning-Cela, R. G. y Martínez-Calvillo, S. **2013**. Comparative genomics of *Leishmania* parasites. En Comparative Genomics in Neglected Human Parasites. Editores: Lopez-Camarillo, C. y Marchat, L.A. Nova Science Publishers, Inc., New York, New York, USA. pp 59-80.



In: Comparative Genomics in Neglected Human Parasites ISBN: 978-1-62417-874-0
Editors: César López-Camarillo and L. A. Marchat © 2013 Nova Science Publishers, Inc.

Chapter 4

COMPARATIVE GENOMICS OF *LEISHMANIA*
PARASITES

Norma E. Padilla-Mejía^{a,b}, Claudia M. Gómez-Hurtado^a,
Inti I. Sánchez-Santamaría^a, Luis E. Florencio-Martínez^a,
Rebeca G. Manning-Cela^c and Santiago Martínez-Calvillo^{a,b}

^aUnidad de Biomedicina, Facultad de Estudios Superiores Iztacala, Universidad Nacional Autónoma de México, Col. Los Reyes Iztacala, Tlalnepantla, Edo. de México, Mexico

^bPosgrado en Ciencias Biológicas, Facultad de Estudios Superiores Iztacala, Universidad Nacional Autónoma de México, Col. Los Reyes Iztacala, Tlalnepantla, Edo. de México, Mexico

ABSTRACT

The parasites of the genus *Leishmania* are trypanosomatid protozoa that produce a spectrum of diseases ranging from mild skin or mucosal lesions to fatal visceral leishmaniasis. The genome sequences of five different species of *Leishmania* have been reported to date: *L. major*, *L. infantum*, *L. braziliensis*, *L. mexicana* and *L. donovani*. Analyses of the sequences revealed that the genomes of these parasites are organized into large directional gene clusters, *i.e.* tens-to-hundreds of protein-coding genes arranged sequentially on the same strand of DNA. A remarkable conservation of gene order (synteny) was observed in the genomes of the different *Leishmania* species. Interestingly, an unexpectedly small number of species-specific genes was identified. However, the analyses showed gene and chromosome copy number differences between species, indicating that increased gene copy number may cause changes in gene expression that might influence disease tropism. Contrary to what occurs in other eukaryotes, transcription in *Leishmania* and other trypanosomatids is polycistronic, and mature messenger RNAs (mRNAs) are generated from primary transcripts by *trans*-splicing and

* Corresponding author: Email: scalv@campus.iztacala.unam.mx, Phone: (+52) (55) 5623-1333 ext. 39782, Fax: 00 (+52) (55) 5623-1138.

polyadenylation. Little is known about either the DNA sequences or the proteins that are involved in transcription initiation in *Leishmania*. Bioinformatic analyses of the genome databases of these parasites led to the identification of a small number of proteins involved in gene expression. However, functional studies have revealed that trypanosomatids have more general transcription factors than originally estimated. Analysis of the *Leishmania* databases showed the presence of a relatively low number of transfer RNA (tRNA) genes, which are organized into clusters of two to 10 genes that may contain other Pol III-transcribed genes. Also, with the exception of *L. braziliensis*, the *Leishmania* genome does not contain active retrotransposons. Interestingly, *L. braziliensis* possesses components of an RNA-mediated interference pathway, which is not present in other *Leishmania* species.

1. INTRODUCTION

The members of the Trypanosomatidae family (order Kinetoplastida) are flagellated protozoa of both biological and medical importance. This family includes parasites of the genus *Leishmania* and *Trypanosoma*, which exhibit complex life cycles with several developmental stages that alternate between vertebrate and invertebrate hosts. *Leishmania* is the causative agent of leishmaniasis, a disease with a wide range of clinical manifestations that arise from infections with different species of the parasite [1]. There are three main clinical forms of leishmaniasis: cutaneous, mucocutaneous and visceral [2; 3]. Cutaneous leishmaniasis is the most common form and it typically produces ulcers, which heal spontaneously, on the face and arms of the infected person. However, ulcers cause serious disability and leave severe and permanently disfiguring scars. Mucocutaneous leishmaniasis is the most disfiguring clinical form of the disease, since it causes extensive destruction of the oral-nasal and pharyngeal cavities with dreadful disfiguring lesions and mutilation of the face. Visceral leishmaniasis is characterized by fever, substantial weight loss and swelling of the spleen and liver, and it is fatal in the absence of treatment [4]. *Leishmania* parasites are transmitted to humans and other vertebrates through the bites of infected sandflies of the genus *Lutzomyia* (in the New World) and *Phlebotomus* (in the Old World). The metacyclic promastigote, the infective form of the parasite, invades the macrophages and differentiates into amastigotes, which are the proliferative forms within the vertebrate host. The parasite replicates within the insect vector as a non-infective procyclic promastigote [5]. According to the World Health Organization (WHO), leishmaniasis is endemic in 98 countries or territories in tropical and sub-tropical areas of the world [6]. Approximately 350 million people are considered at risk of contracting leishmaniasis, and some two million new cases occur yearly. Visceral leishmaniasis causes an estimated over 50,000 deaths annually, a rate surpassed among parasitic diseases only by malaria. Thus, the *Leishmania* parasites represent an important global health problem for which there is no vaccine and few effective medicines. The recent publication of the genome sequences of several species of *Leishmania* will not only help to improve vaccine and drug design, but also to understand diverse molecular aspects and the evolution of the parasite.

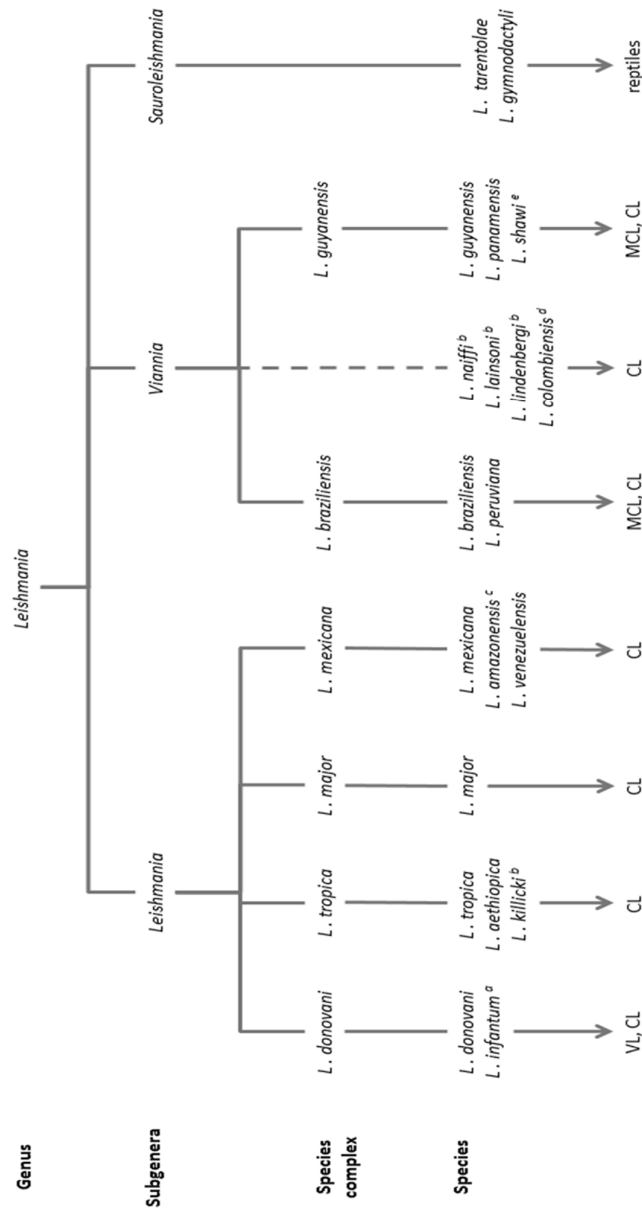
2. DIVERSITY OF *LEISHMANIA* SPECIES

In 1903, William Leishman and Charles Donovan independently described *Leishmania* parasites, but they were previously observed by David D. Cunningham in 1885 and Peter Borovsky in 1898 [7]. The parasite was first cultivated in 1904 by Leonard Rogers [8]. To date, more than 30 different species of *Leishmania* have been described in five different continents, and at least 20 of them cause disease in human [6]. Molecular methodologies allow accurate species identification. Isoenzyme electrophoresis is currently the reference identification technique, and correlations have been established between clinical forms and zymodemes (parasite populations with common isoenzyme patterns) for some species. Kinetoplast DNA (kDNA) digested with restriction enzymes has also helped to identify species and parasite populations (schizodemes). Other DNA techniques, more sensitive and easier to use, will probably prevail in the future.

Depending on which parts of the sandfly gut are colonized by the parasite, the genus *Leishmania* is divided into three subgenera: *Leishmania*, *Viannia* and *Sauroleishmania* [9-11]. Species classified within the *Leishmania* subgenus produce cutaneous and visceral leishmaniasis in both the New World and the Old World, while species included in the *Viannia* subgenus cause cutaneous and mucocutaneous leishmaniasis only in the New World. Species that belong to the *Sauroleishmania* subgenus infect reptiles and are not pathogenic to humans. The three *Leishmania* subgenera are divided into several complexes that include different species (figure 1).

Leishmania and the rest of the trypanosomatids diverged early from the main eukaryotic lineage. As a consequence, this group of organisms shows numerous unusual features at the genetic, biochemical and cytological level [5; 12; 13]. For example, *Leishmania* presents a single branched mitochondrion that extends the whole length of the cell and can fill up to 12% of the cell volume. The mitochondria contain the kDNA, which is the most unusual structure in the organelle (see below). *Leishmania* extensively decodes its mitochondrial transcripts through RNA editing, a process that involves specific uridine insertion and deletion that can double the size of the primary transcript, and that is essential for creation of translatable open reading frames [14; 15].

Another organelle that is exclusive to *Leishmania* and other trypanosomatids is the glycosome, a peroxisome-related structure that contains enzymes involved in several metabolic pathways, including glycolysis. Glycosomes are bounded by a single membrane and have a protein-dense matrix [16]. Another unique characteristic of trypanosomatids is the presence of subpellicular microtubules, which is a layer of microtubules that run longitudinally underneath the plasma membrane [17]. Remodelling of this cytoskeleton facilitates the distinct morphologies of the various developmental forms of *Leishmania*. In trypanosomatids, the flagellum consists of a 9+2 microtubule configuration axoneme and a paraflagellar rod, a unique lattice-like structure that lies alongside the axoneme. The flagellum emerges from a flagellar pocket, a small invagination of the plasma membrane that constitutes the only site of exocytosis and endocytosis in the cell [18].



3. GENOME ORGANIZATION

The *Leishmania* genome is organized into large directional gene clusters (DGCs), *i.e.* tens-to-hundreds of protein-coding genes arranged sequentially on the same strand of DNA. This striking gene organization was initially observed on *L. major* Friedlin chromosome 1, the first entirely sequenced chromosome in trypanosomatids, which contains 85 genes organized into two divergent DGCs, with the first 32 genes clustered on the bottom strand and the remaining 53 genes grouped on the top strand [19]. The publication of the complete genome of *L. major* Friedlin revealed that the genes in all the chromosomes are organized into large DGCs [20]. A similar gene organization was reported for *T. brucei* [21] and *T. cruzi* [22], indicating that the presence of DGCs is another unusual feature that is present in all trypanosomatids. Recently, the genome sequences of other four species of *Leishmania* were reported: *L. infantum* (strain JPCM5), *L. braziliensis* (M2904), *L. mexicana* (U1103) and *L. donovani* (BPK282/0c14) [23-25]. These reports provided resources for comparative genomic studies, and information about the genes that might have a role in the tissue-specific development of the different species of *Leishmania*.

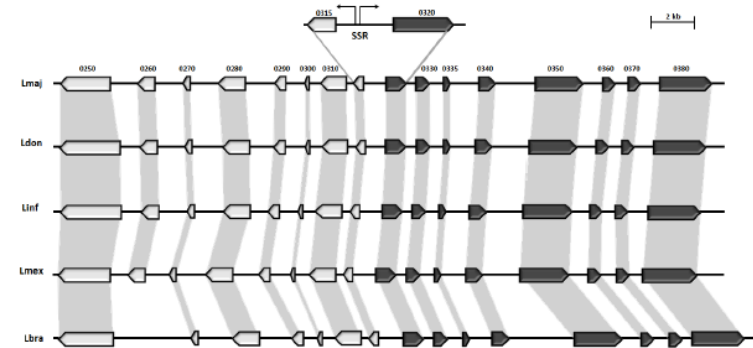


Figure 2. Synteny of protein-coding genes on chromosome 1 from *Leishmania*. The first eight genes of each of the two directional gene clusters present on this chromosome are shown for *L. major* (Lmaj), *L. donovani* (Ldon), *L. infantum* (Linf), *L. mexicana* (Lmex), and *L. braziliensis* (Lbra). The names of the genes, indicated only for *L. major*, are abbreviated (for instance, 0250 corresponds to LmjF01.0250). Orthologous genes are joined by grey lines. In the enlargement of the strand switch region (SSR) from *L. major*, the transcription start sites for both DGCs are denoted with arrows.

The genomes of the different species of *Leishmania* show a noteworthy conservation of gene order (synteny), despite an estimated divergence of 46 million years [26]. Gene order is conserved for more than 99% of genes between *L. major*, *L. infantum* and *L. braziliensis* [23]. To exemplify synteny, figure 2 shows a schematic representation of the strand switch region (SSR) and the first eight genes of the two DGCs present on chromosome 1 from several species of *Leishmania*. As observed, the order of the genes is exactly the same in all the species. The only exception is the orthologue of gene 0260 (LmjF01.0260), which was lost in

the *L. braziliensis* genome (figure 2). Sequence conservation of coding regions is high among different species of *Leishmania*; for example, the average amino acid identity between *L. major* and *L. infantum* is 92%, and the average nucleotide identity is 94% [23].

Table 1. Summary of *Leishmania*, *T. brucei* and *T. cruzi* genomes.

	<i>L. major</i>	<i>L. infantum</i>	<i>L. braziliensis</i>	<i>L. mexicana</i>	<i>L. donovani</i>	<i>T. brucei</i>	<i>T. cruzi</i>
Chromosome number	36 [20]	36 [24]	35 [24]	34	36	11 ^a	41 [27]
Genome size (bp)	32,855,089 [24]	32,101,728 [24]	31,997,773 [24]	32,108,741	32,444,998	26,075,396	~60,000,000
Overall G+C content (%)	59.7 [24]	59.3 [24]	57.8 [24]	59.7	NR	46.4	51
Coding G+C content (%)	62.5 [24]	62.45 [23]	60.38 [22]	NR ^c	NR	50.9	53.4
Predicted protein-coding genes	8,412 [24]	8,214 [24]	8,357 [24]	8,250	NR	9,068	~12,000
Pseudogenes	97 [23]	41 [23]	161 [23]	99	13	904	3,590
Intergenic regions mean length (bp)	2,045 [20]	2,049 [26]	1,976 [26]	NR	NR	1,279	1,024
Gene density (genes/Mb)	252 [20]	235 [26]	258 [26]	NR	NR	317	385
tRNA genes	83 [14]	77 ^b	75 ^b	84 ^b	76 ^b	66 [14]	120 [14]

^a Only mega chromosomes

^b These numbers correspond to the genes annotated in GeneDB (*L. infantum*, 67 genes; *L. braziliensis* 66; *L. mexicana*, 83, *L. donovani*, 64), plus the genes that we found using tRNAScan SE

^c NR: not reported

*Reference numbers are shown in brackets

The size of the haploid genomes in the diverse species of *Leishmania* is very similar, ranging from 31.9 to 32.8 megabases (Mb) in *L. braziliensis* and *L. major*, respectively (table 1). These genomes are slightly larger than the one from *T. brucei* (26 Mb), but smaller than the one from *T. cruzi* (~60 Mb). The number of predicted protein-coding genes is 8412 in *L. major*, 8241 in *L. infantum*, 8357 in *L. braziliensis* and 8250 in *L. mexicana* [24]. The estimated number of protein-coding genes is 9068 in *T. brucei* [22] and ~12000 in *T. cruzi* [27]. The majority of the predicted genes (~70%) encode hypothetical proteins of unknown function. These may represent genes that have parasite-specific functions, or which are sufficiently diverged as to have no significant sequence similarity to their functional homologs in other species [28].

In the *Leishmania* genomes, most protein-coding genes are separated by relatively short intergenic regions (mean length of 2045 bp in *L. major*). The density of protein-coding genes in the *L. major* genome is 252 genes per Mb, which is similar to what has been found in other trypanosomatids (317 genes per Mb in *T. brucei* and 385 genes per Mb in *T. cruzi*), but considerably higher than what has been reported in human (~8 genes per Mb) [29]. The overall G+C content of the genomes of the various species of *Leishmania* (from 57.8 to 59.7%) is higher than G+C content in *T. brucei* (46.4%) and *T. cruzi* (51%), and the G+C content of the coding regions is even higher (62.5% in *L. major*) (table 1). Similarly to other eukaryotes, the ends of the chromosomes in *Leishmania* contain the telomeric repeat

GGGTTA. Chromosomes contain short subtelomeric regions composed of variable repetitive elements that are responsible for a major part of the polymorphisms observed between homologous chromosomes. In contrast, subtelomeric regions in *T. brucei* are long and encode species-specific genes. DNA repeats represent 9-10% of *Leishmania* genome. Also, conserved amino acid repeats, which have been hypothesized to have a role in pathogenicity, are present in 3-4% of the predicted proteins in *Leishmania* [24]. The vast majority of protein-coding genes in trypanosomatids lack introns; in fact, *cis*-splicing has only been demonstrated for the gene encoding the poly-A polymerase [30].

Trypanosomatids do not seem to contain DNA transposons. However, analysis of the *T. cruzi* and *T. brucei* genomes confirmed the presence of abundant long terminal repeat (LTR) and non-LTR retrotransposons [21; 31]. They account for ~5% and 2% of the *T. cruzi* and *T. brucei* genomes, respectively. In contrast, *L. major* and *L. infantum* do not contain active retrotransposons. Nevertheless, they have remnants of extinct *ingi/L1Tc*-like retrotransposons called DIREs [32]. Recently, two new families of degenerated retrotransposons were identified in *Leishmania*: *SIDER1* and *SIDER2* [33; 34]. These sequences are predominantly located within the 3'-UTR of *Leishmania* mRNAs, and it was shown that *SIDER2* acts as an instability element, since *SIDER2*-containing mRNAs are generally expressed at lower levels compared to the non-*SIDER2* mRNAs [33].

Surprisingly, recent findings showed the presence of potentially active SLACS/CZAR retrotransposons in the *L. braziliensis* genome [23; 35]. These elements are associated with tandemly repeated spliced leader genes, showing an arrangement similar to that of the SLACS or CZAR elements in *T. brucei* or *T. cruzi*, respectively. Moreover, the telomeric regions of *L. braziliensis* contain a family of 20-30 previously unknown transposable elements called "telomere-associated transposable elements" (TATEs). Each TATE includes putative reverse transcriptase, phage integrase, and DNA and/or RNA polymerase sequences. TATEs are present on at least 12 chromosomes, inserted in the middle of the GGGTTA telomeric repeats [36].

4. SPECIES-SPECIFIC GENES AND RNA INTERFERENCE

Sequence comparisons among the different species of *Leishmania* showed a surprisingly small number of species-specific genes, in spite of the broad differences in pathogenicity and immune response [23; 24]. In *L. mexicana* it was reported the presence of only two unique genes, which encode predicted proteins of unknown function. Interestingly, one of these genes is predicted to be a pseudogene in other *Leishmania* species, but it is the orthologue of an intact gene in *T. brucei*. *L. major* possesses 14 specific genes: 13 encode proteins of unknown function and one encodes a PFPI-like peptidase. In *L. braziliensis* there are 67 unique genes, of which 54 are predicted to encode proteins of unknown function. Species from the *L. donovani* complex (*L. donovani* and *L. infantum*) contain 19 specific genes, of which 15 encode proteins of unknown function; these genes might play an important role in the potential for visceralization associated with the *L. donovani* complex [25]. The location of the species-specific genes in *Leishmania* is also unexpected, since they are dispersed all over the genome rather than clustered in subtelomeric regions or breakpoints of DGCs, as reported in other trypanosomatids [22].

To determine if species-specific genes are present in different strains of the same species, second strains of *L. major* and *L. mexicana* were sequenced [24]. It was found that the second strain of *L. mexicana* (M379) contains intact copies of both unique genes identified in *L. mexicana* U1103. Similarly, the strain LV39 from *L. major* contains 13 of the 14 specific genes found in *L. major* Friedlin. Moreover, 17 different strains of *L. donovani* isolated from patients from India, Nepal and Ethiopia contained the 19 unique genes found in the *L. donovani* complex [25]. Thus, specific genes in *Leishmania* are conserved between strains of the same species, even when they were isolated from different geographical regions. However, extensive variation in gene and chromosome copy number was reported among different species and strains of *Leishmania* (see below).

In *Leishmania*, establishment of pseudogenes and gene loss seem to be the main factors that influence the evolutionary formation of new species. The species specificity of most genes can be attributed to the deterioration of a gene in the other species. In other words, in the species that lack the functional gene there is a degenerate sequence (pseudogene) in the corresponding region of synteny [23; 24]. For example, one of the unique genes in *L. infantum*, the one that encodes a putative phosphatidylinositol 3-kinase (PI3K), is predicted to encode a pseudogene in the other species of *Leishmania*. Also, A2 is one of the *L. donovani*-specific genes that has been proposed to be required for parasite survival in visceral organs. *L. major* and *L. tropica* contain only A2 pseudogenes [37].

RNA interference (RNAi) is a mechanism of gene silencing present in many eukaryotes. In this process, short double-stranded RNA molecules induce the sequence-specific degradation of homologous mRNA. It is thought that RNAi is not only involved in the control of endogenous mRNAs, but also has a defensive role against transposon activity and viruses [38]. Numerous proteins participate in RNAi, including Dicer and Argonaute. Several years ago a functional RNAi pathway was identified in *T. brucei* [39]. In contrast, *T. cruzi*, *L. major* and *L. donovani* lack Dicer and Argonaute activities and associated genes [40]. These findings led to the assumption that *T. brucei* was the only trypanosomatid with active RNAi mechanisms. Consequently, it was surprising when genes involved in the RNAi pathway were identified in *L. braziliensis* [23; 35] and other species of the *Viannia* subgenus, such as *L. guyanensis* and *L. panamensis* [41]. It was later demonstrated that in fact *L. braziliensis* shows strong RNAi activity with reporter and endogenous genes [41]. Functional and evolutionary studies of RNAi genes established that RNAi must have been lost twice independently in trypanosomatids: once in the lineage leading to *T. cruzi*, and a second time after the separation of the *Viannia* subgenus from the remaining *Leishmania* species [41]. In *L. braziliensis*, the Argonaute gene is located on chromosome 11. Sequence analysis of the syntenic regions in *L. major* and *L. infantum* revealed highly degenerate Argonaute pseudogenes, which supports the hypothesis that RNAi-related genes were lost in members of the *Leishmania* subgenus [36].

5. GENE AND CHROMOSOME COPY NUMBER VARIATION

Importantly, differences in gene content and chromosome copy number between species and strains of *Leishmania* have been recognized as a major source of genomic variation. Regarding chromosome architecture, *L. major*, *L. donovani* and *L. infantum* have 36

chromosomes, whereas *L. braziliensis* has only 35 chromosomes due to the fusion of chromosomes 20 and 34. The genome of *L. mexicana* is organized into 34 chromosomes, as the result of two fusion events involving chromosomes 8 and 29, and chromosomes 20 and 36 [42]. *Leishmania* is generally considered to be diploid, although several chromosomes have been described as aneuploid [43; 44]. Homologous chromosomes can vary in size, as it has been described for the homologous copies of chromosome 1 in *L. major* Friedlin, which differ in size by approximately ~30 kb [45].

Chromosome copy number has been analyzed for different species of *Leishmania* by chromosome Read Depth and other methodologies such as allele frequency distribution plots, FISH and FACS [24; 25; 46]. Analysis of two *L. major* strains, Friedlin and LV39, showed that most chromosomes in this species are disomic, except for chromosome 31 (reported as tetrasomic), chromosome 1 (trisomic) and chromosome 34 (reported as supernumerary). Several trisomic chromosomes have been reported in *L. infantum* JPCM5 (nine chromosomes), *L. donovani* LV9 (three chromosomes) and *L. mexicana* U1103 (three chromosomes). In *L. donovani* BPK206/0, five chromosomes are trisomic and two chromosomes are tetrasomic. In different species, several chromosomes appear to be of intermediate status, being neither disomic nor trisomic. These patterns may result from a mixture of individual cells within a population with monosomic, disomic and trisomic chromosomes. Interestingly, in the strain M2904 of *L. braziliensis* 30 of 35 chromosomes are clearly trisomic, three are tetrasomic and one is hexasomic (chromosome 31). Thus, in contrast to other species of *Leishmania*, *L. braziliensis* M2904 is primarily triploid [24]. It is worth mentioning that chromosome 31 is the only one that is supernumerary in all the species and isolates of *Leishmania* analyzed, including the homologous chromosome 30 in *L. mexicana*. The functional meaning of this finding has yet to be determined. In conclusion, *Leishmania* parasites can contain disomic, trisomic, tetrasomic or supernumerary chromosomes, depending on the species, and the pattern of aneuploidy seems to be characteristic for a specific strain.

It is also important to mention that multicopy gene arrays are present more frequently in disomic chromosomes than in supernumerary chromosomes in four species of *Leishmania* analyzed (*L. braziliensis*, *L. major*, *L. mexicana* and *L. infantum*). It has been suggested that these disomic chromosomes may have persisted as non-supernumerary because of fitness constraints that are gene-dosage related. Since gene-dosage sensitivity is known to be important in selection of copy number changes, in the *Leishmania* species there may be selection against whole-chromosome duplications for those chromosomes that have a higher proportion of dose-sensitive genes [24].

Analysis of allelic variation among different *Leishmania* species by heterozygous single-nucleotide polymorphisms (SNPs) showed a remarkably low level of heterozygosity in *L. major* (297 SNPs) and *L. infantum* (629 SNPs) compared with *L. mexicana* (12,531 SNPs) and *L. braziliensis* (44,588 SNPs). This finding could result from a higher degree of inbreeding within *L. major* and *L. infantum* relative to *L. braziliensis* and *L. mexicana*. The high level of heterozygosity observed in *L. braziliensis* could in some measure be attributable to its triploidy and the supernumeracy nature of many of its chromosomes, since redundancy of essential genes may allow a higher rate of neutral mutation than that of the diploid genomes of *L. major* and *L. infantum* [24].

Among different strains of *L. donovani*, 3549 SNPs were identified [25]. About 82% of them were located in non-coding regions, 7% were synonymous mutations in coding regions

and 11% caused changes at the protein sequence level. Remarkably, some SNPs seem to be related to resistance to the antimonial drug sodium stibogluconate (SSG). Also, it is interesting that SSG-resistant parasites contain on average 44% fewer rDNA transcription units and 77% more mini-exon units than sensitive parasites. Moreover, *L. donovani* SSG-resistant lines show more copies of an episome that contains the mitogen-activated protein kinase (MAPK) locus.

6. MAJOR GENE FAMILIES

Another characteristic feature of the *Leishmania* genome is the presence of tandem arrays of duplicated genes, which facilitate increased protein expression in the absence of the regulated transcriptional control that is found in other organisms. The number of tandem arrays differs among species: 132 in *L. mexicana* U1103, 200 in *L. major* Friedlin, 207 in *L. infantum* JPCM5 and 214 in *L. braziliensis* [24]. The number of genes per tandem array also varies among *Leishmania* species.

From the total number of protein-coding genes, only 56 were present as multicopy genes in all species of *Leishmania* [24]. These include the amastin genes, which encode small surface glycoproteins of unknown function that are also present in *T. cruzi*. The majority of amastin genes is expressed in intracellular amastigotes, and may contribute to their survival in the human host. In *L. major*, amastin genes are distributed on seven different chromosomes. The largest tandem array, located on chromosome 34, contains 23 amastin genes that alternate with the highly conserved tuzin genes (that encode proteins of unknown function) [47]. *L. infantum* and *L. braziliensis* have amastin gene arrays at the same locations; however, the large amastin-tuzin gene array seems to be shorter in *L. braziliensis* [23]. The *Leishmania* glycoprotein GP63, also known as leishmanolysin or major surface protease (MSP), is a surface metalloprotease involved in parasite survival, infectivity and virulence [48]. In *L. major*, it is encoded by a tandem array of seven genes located on chromosome 10. Interestingly, GP63 tandem array is composed of ~28 genes in *L. braziliensis* [23].

PSA, also called PSA-2 or GP46, is a *Leishmania* family of membrane-bound or secreted proteins, whose main signature consists in a specific array of leucine-rich repeats flanked by conserved cysteine-rich domains. It is overexpressed in metacyclic promastigotes, where it is involved in resistance to complement-mediated lysis. In *L. major*, 32 PSA genes have been identified on chromosomes 5, 9, 12, 21 and 31. Similarly, 14 and 8 PSA genes were found on the same chromosomes in *L. infantum* and *L. braziliensis*, respectively. The PSA subfamily present on chromosome 12 is the most complex in *L. major* and *L. infantum*, since it is composed of tandem arrays of 24 genes in the former and seven genes in the latter. In contrast, in *L. braziliensis* there is only one PSA gene on this locus [49]. Other multicopy genes present in the *Leishmania* genome include alpha- and beta-tubulin, HSP70, HSP83, kinesins, several translation initiation and elongation factors, protein kinases, RNA helicases, protein phosphatases, amino acid permeases and several hypothetical proteins [20].

7. GENE EXPRESSION IN *LEISHMANIA*

Eukaryotic cells usually have three distinct classes of nuclear RNA polymerases (Pol): Pol I, II, and III. Each class of polymerase transcribes a different kind of RNA. Pol I synthesizes 18S, 5.8S and 28S ribosomal RNAs (rRNAs), and Pol II transcribes mRNAs and most of the small nuclear RNAs (snRNAs). Pol III is involved in the production of small essential RNAs, such as tRNAs, 5S rRNA and some snRNAs.

Contrary to what occurs in other eukaryotes, Pol II transcription in *Leishmania* and other trypanosomatids is polycistronic [50-52]. Most chromosomes contain at least two PGCs, which can be either divergently transcribed (towards the telomeres) or convergently transcribed (away from the telomeres). Mature nuclear mRNAs are generated from primary transcripts by *trans*-splicing and polyadenylation (figure 3) [53; 54]. *Trans*-splicing is a process that adds a capped 39-nucleotide miniexon or spliced leader (SL) to the 5' termini of the mRNAs [55; 56]. All the genes that are part of a PGC are transcribed at the same level, as a consequence of polycistronic transcription. However, the mature mRNAs of adjacent genes might show very different concentrations and/or stage-specific expression. This is because gene expression in trypanosomatids is mainly regulated posttranscriptionally at the level of mRNA processing and stability [13; 57]. Sequences in the 3' untranslated region (3'-UTR) of an mRNA play a key role in gene expression. For example, the 3'-UTR from the amastin mRNA in *L. infantum* has a 450-bp region that confers amastigote-specific gene expression by a mechanism that increases the mRNA translation [58].

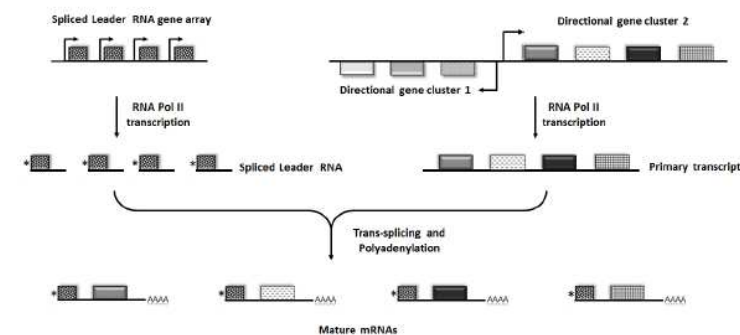


Figure 3. Transcription and processing of mRNAs in *Leishmania*. Pol II transcription initiates upstream of the first gene of the DGCs (arrows), generating primary transcripts (shown only for DGC2) that are processed by *trans*-splicing and polyadenylation to generate the mature mRNAs. By *trans*-splicing, a capped spliced leader RNA is added to the 5' end of every mRNA. In the spliced leader locus (located on a different chromosome) each gene possesses a Pol II promoter region (arrows). The cap in the spliced leader RNA is indicated with an asterisk at the 5' end of the RNA. The four As located at the 3' end of the mature mRNAs represent the poly-A tail.

In trypanosomatids, identification of Pol II promoters for protein-coding genes has proven to be a difficult goal, complicated by a relatively low Pol II transcriptional activity and rapid processing of the primary transcripts. Nevertheless, transcriptional analysis of chromosome 1 from *L. major* showed that Pol II transcription of the entire chromosome initiates in the strand-switch region (between the two divergent PGCs) and proceeds bidirectionally towards the telomeres (figure 2) [52]. Several transcription start sites were mapped for both PGCs within a <100-bp region that contains long G-tracts (or C-tracts), but do not contain a TATA box or any other typical Pol II core promoter elements. Thus, while in most eukaryotes each gene possesses its own promoter, a single region seems to drive the expression of the entire chromosome 1 in *L. major* [52]. Similar studies performed on chromosome 3 from *L. major* confirmed that Pol II transcription initiates only upstream of the first gene of a PGC [59]. Since most genes are organized into large PGCs in trypanosomatids, the number of regions where transcription of Pol II initiates in these organisms is very low (only a few per chromosome) compared to other eukaryotes. There is no substantial sequence homology among the transcription initiation regions on *L. major* chromosomes 1 and 3. Hence, rigorously conserved sequence recognition sites do not appear to be required for Pol II transcription initiation in *Leishmania*. Interestingly, a recent ChIP-chip study in *L. major* showed that H3 histones acetylated at K9/K14, a marker for sites of active transcription initiation in other eukaryotes, are found at all divergent strand-switch regions in the parasite [60]. Moreover, peaks for two transcription factors, TBP and SNAP50, were also associated with divergent strand-switch regions [60].

In trypanosomatids, the only Pol II promoter that has been extensively characterized is the one driving the expression of the SL RNA [61-63]. In *L. tarentolae* it consists of two domains: the -60 element (from -67 to -58, relative to the TSS) and the -30 element (from -41 to -31). In trypanosomatids, transcription factors could not be identified by standard *in silico* analysis because they are extremely divergent [64; 65]. Consequently, their identification relied on biochemical, structural and functional analyses. Regarding Pol II transcription, several general transcription factors that participate in SL RNA synthesis have been identified in *T. brucei*. These include TBP, TFIIB, SNAPc, TFIIA and TFIIF [66-70]. Thus, these recent findings indicate that trypanosomatids possess more general transcription factors than initially estimated from *in silico* studies.

8. GENOMIC ORGANIZATION OF TRANSFER RNA AND RIBOSOMAL RNA GENES

Analysis of the *L. major* genome databases showed the presence of 83 tRNA genes distributed among 31 loci, on 19 different chromosomes [20; 71]. Most tRNA genes are organized into clusters of two to 10 genes, on either top or bottom strand, which may contain other Pol III-transcribed genes. For example, in the locus located on chromosome 23 there are 10 tRNA genes, a 5S rRNA gene and the U1 and U3 snRNA genes (figure 4). In *T. brucei*, a total of 66 tRNA genes were located on 26 loci, on eight different chromosomes. As in *L. major*, in *T. brucei* the number of tRNA genes per cluster ranges from two to 10. In *T. cruzi*, 120 tRNA genes were identified [71]. Thus, the number of tRNA genes in trypanosomatids is relatively low, considering that eukaryotic organisms usually contain several hundred tRNA

genes. For instance, *C. elegans* has 568 tRNA genes and *Homo sapiens* presents 497 tRNA genes [72; 73].

In contrast to protein-coding genes, the majority of the tRNA clusters do not show synteny in trypanosomatids, but a few of them do show conservation [71]. Among the latter, the most outstanding example is the cluster of 13 Pol III genes, located on chromosome 23 in *L. major*, that is highly syntenic. Interestingly, the order of the genes in this cluster is identical between *T. brucei* and *T. cruzi*, although three genes are located on different strands [71]. The majority of the 13 genes is present in the *L. major* cluster, but their order is not identical to either of the other two clusters (figure 4, compare *L. major* with *T. brucei*). Additionally, a 5S rRNA gene replaced a 7SL RNA gene and a tRNA-Trp gene replaced one of the tRNA-Lys genes.

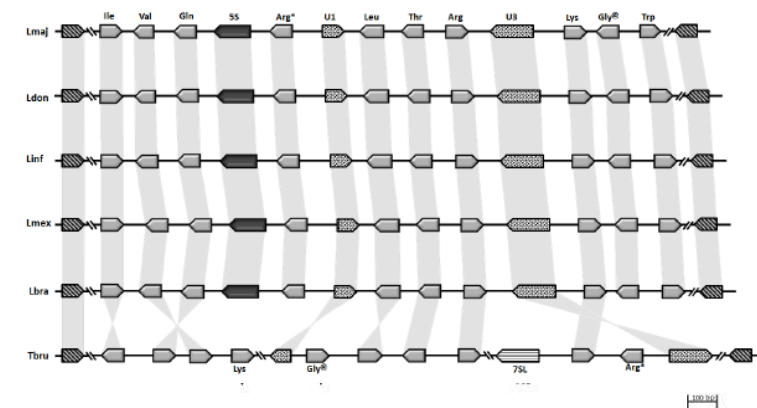


Figure 4. Comparative order of Pol III-transcribed genes in *Leishmania* and *T. brucei*. The order of the 13 Pol III-transcribed genes present on chromosome 23 in *Leishmania* is shown for *L. major* (Lmaj), *L. donovani* (Ldon), *L. infantum* (Linf), *L. mexicana* (Lmex), *L. braziliensis* (Lbra) and *T. brucei* (Tbru). The loci include several tRNA genes (Ile, Val, Gln, Arg, Leu, Thr, Lys, Gly and Trp), a 5S rRNA gene (5S), two snRNA genes (U1 and U3) and a 7SL RNA gene (7SL). All the species of *Leishmania* show exactly the same order. *T. brucei* shows some degree of synteny. Orthologous genes are joined by grey lines (with the exception of Gly* and Arg* from *T. brucei*, which were not joined to simplify the figure). Flanking protein-coding genes are shown as diagonally striped boxes (not to scale).

Many tRNA genes have not been annotated in the genome databases of *L. infantum*, *L. braziliensis*, *L. mexicana* and *L. donovani*. Thus, in order to compare to *L. major*, we performed *in silico* searches of tRNA genes and other Pol III-transcribed genes. Our results showed that the vast majority of tRNA clusters are highly syntenic among *Leishmania* species. For example, the order of the Pol III-transcribed genes located on chromosome 23 is identical in all the species of *Leishmania* (figure 4). However, several differences were observed. For instance, a cluster of two tRNA genes and a 5S rRNA gene located on chromosome 15 seems to be absent in the corresponding syntenic region of *L. infantum* and *L. donovani*. Also, while cluster I on chromosome 11 [71] contains five tRNA genes and two 5S rRNA genes in *L. major* and *L. mexicana*, it contains only three tRNA genes in *L. infantum*,

L. donovani and *L. braziliensis*. Moreover, several other clusters lack at least one tRNA gene in *L. braziliensis*.

In most eukaryotic organisms, tRNA genes seem to be dispersed randomly throughout the genome. However, in human cells the distribution is non-random, since more than 25% of the tRNA genes are located in a region of only about 4 Mb on chromosome 6 that represents only 0.1% of the human genome [29]. The distribution of genes in the *L. major* genome does not seem to be totally random, since half of the chromosomes do not contain even a single tRNA gene. Additionally, 60 tRNA genes (72%) are located on only seven chromosomes (9, 11, 23, 24, 31, 34 and 36), which represent only 26% of the genome [71].

It was previously shown that transcription of two convergent PGCs on *L. major* chromosome 3 terminates on the convergent strand-switch region, within a tRNA-gene region [59]. Interestingly, 14 of the 39 convergent strand-switch regions (35.9%) in the *L. major* genome contain at least one tRNA gene, representing 45.2% of the 31 tRNA loci. A similar situation was found in *T. brucei*, where 34.6% of the tRNA loci are located within convergent strand-switch regions. This suggests that the use of tRNA genes as signals for termination of transcription of convergent clusters of protein-coding genes might be a common process in trypanosomatids.

The eleven 5S rRNA genes found in the *L. major* genome are distributed on six chromosomes, and are always associated to tRNA genes [20]. A similar organization is present in other species of *Leishmania*. In contrast, the 5S rRNA genes in *T. brucei* and *T. cruzi* are organized into tandem arrays that are not associated to tRNA genes [74; 75]. In trypanosomatids, as in other organisms, the coding regions of the 18S, 5.8S and 28S rRNAs occur as tandem repeats that are clustered at one or several loci. However, a distinctive property of rRNA genes in trypanosomatids is the fragmentation of the 28S-like rRNA into multiple independent molecules: 24S α , 24S β , S1, S2, S4 and S6 [76; 77]. Interestingly, in *L. major* there are two copies of the S4 gene (also known as LSU ϵ) in the majority of the rRNA repeats [78]. In the different species of *Leishmania*, rRNA genes are located on chromosome 27. The number of rRNA repeats varies among species and strains, but it seems to be low (~10), comparing to other eukaryotes (~150) [25; 78; 79].

9. MITOCHONDRIAL GENOME

As mentioned above, *Leishmania* and other trypanosomatids have only one mitochondrion, and it contains a single kDNA network that is condensed into a disk-shaped structure, positioned in a specialized region of the mitochondrial matrix near the flagellar basal body. The kDNA network is composed of two classes of catenated circular DNA molecules, maxicircles and minicircles [80]. Each mitochondrion has approximately 50 copies of maxicircles, with a size between 20-40 kb, depending on the species. The number of minicircles (~2 kb) ranges from 5,000 to 10,000 per organelle [81]. Maxicircles encode 18 protein-coding genes (components of the respiratory chain), two ribosomal RNAs, and some guide RNAs (gRNAs), which are small RNA molecules that participate in RNA editing. These genes are located on both strands of a ~17 kb-long conservative region [82]. The rest of the maxicircle, known as the divergent region, is composed of repeated sequences.

The minicircle molecules encode from one to five gRNAs. They contain one to four conserved regions and an equal number of variable regions, depending on the species. In *L. tarentolae*, minicircles are organized into a ~170 bp conserved region that contains the origins of replication for both strands, and a variable region that defines the specific minicircle sequence class [83]. The CSB-3 sequence (GGGGTTGGTGTGA) (conserved in minicircles from all trypanosomatids) provides a relative position and polarity marker for gRNA genes. A region of bent DNA is situated immediately adjacent to the conserved region. All gRNA genes so far identified are localized within the variable region.

Maxicircle sequences have been reported for *L. tarentolae* and *L. donovani* 1S LdBob. Sequence comparisons showed that the gene order (with the 12S rRNA gene on one side and the ND5 gene on the other) is identical between both species of *Leishmania*, and identical to that of other trypanosomatids [82]. As expected, gene sequences are very similar between *L. tarentolae* and *L. donovani*, with nucleotide identity levels close to 90% [84]. Sequence conservation also included small pre-edited regions of 5'-edited and internally-edited mitochondrial genes, as well as some regions of extensively-edited genes. However, one important difference between both maxicircles is a full-length minicircle insertion found in the 3' region of the ND1 gene of *L. donovani* 1S LdBob [84]. The insertion is 99.1% identical to one of the minicircles from *L. infantum* [85]. The insert has not been observed in other strains and species of *Leishmania*, which suggests that it represents a distinctive characteristic of the strain 1S LdBob of *L. donovani*.

The mitochondrial genome of *Leishmania* is transcribed by a nuclear-encoded mitochondrial RNA polymerase belonging to a family of single-subunit RNA polymerases, as it occurs in other eukaryotes [86]. Neither mitochondrial promoters nor transcription factors have been identified in trypanosomatids [87; 88]. In maxicircles, transcription of the top strand starts ~1.2 kb upstream of the 12S rRNA gene [89], while initiation sites have not been located on the bottom strand. Numerous polycistronic transcripts have been detected, indicating that maxicircles are transcribed polycistronically. Similarly, polycistronic transcripts from minicircles have also been observed [90]. Like maxicircles, no promoters have been identified for minicircles. However, each gRNA transcription unit is flanked by 18-bp imperfect inverted repeats that have been proposed to function in gRNA expression [88].

CONCLUSION

New technologies have allowed the sequencing of the whole genomes of five different species of *Leishmania*. Analyses of the sequences showed that the genomes of these parasites are organized into large DGCs. Comparisons of the genomic sequences revealed a high degree of synteny, not only for protein-coding genes, but also for tRNA genes. Polycistronic transcription of DGCs generates long primary transcripts that are processed by *trans*-splicing and polyadenylation to produce mature mRNAs. Although an unexpectedly small number of species-specific genes was found, gene and chromosome copy number differences were observed between species. Thus, in the absence of regulated transcriptional activity, *Leishmania* seems to raise mRNA abundance by increasing gene-copy number, which can be achieved via gene duplications on disomic chromosomes, or through the formation of supernumerary chromosomes. Changes in gene expression might influence disease tropism.

The *Leishmania* genome contains a small number of tRNA and rRNA genes, and do not contain active retrotransposons (with the exception of *L. braziliensis*). Interestingly, a functional RNAi pathway was identified only in *L. braziliensis* and other members of the *Viannia* subgenus. The potential to manipulate gene expression by RNAi might be very useful for gene function studies in *Leishmania*. Thus, whole-genome sequencing has offered new data that is reshaping research initiatives for leishmaniasis. Further analyses will provide the basis for more detailed molecular studies that could help to improve drug and vaccine design.

ACKNOWLEDGMENTS

This work was supported by grants 128461 from CONACyT and IN210712 from PAPIIT (UNAM) to S. Martínez-Calvillo, and by grants 132312 and 139898 from CONACyT, and "proyectos multidisciplinares" from CINVESTAV-IPN to R. Manning-Cela. N.E. Padilla-Mejía was recipient of a doctoral fellowship from CONACyT.

REFERENCES

- [1] Herwaldt BL. Leishmaniasis. *Lancet*, 1999 354,1191-1199.
- [2] Davies, CR; Kaye, P; Croft, SL; Sundar, S. Leishmaniasis: New Approaches To Disease Control. *British Medical Journal*, 2003 326, 377-382.
- [3] Antinori, S; Schifanella, L; Corbellino, M. Leishmaniasis: New Insights From An Old And Neglected Disease. *European Journal Of Clinical Microbiology And Infection Diseases*, 2011 Doi 10.1007/S10096-011-1276-0.
- [4] Hailu, A; Musa, AM; Royce, C; Wasunna, M. Visceral Leishmaniasis: New Health Tools Are Needed. *Public Library Of Science Medicine*, 2005 2, E211.
- [5] Handman E. Cell Biology Of *Leishmania*. *Advances In Parasitology*, 1999 44, 1-39.
- [6] WHO. Control Of The Leishmaniases, Technical Report Series 949. Geneva: World Health Organization; 2010. P. 1-186.
- [7] Row R. *Leishmania donovani* And *Leishmania tropica*. *British Medical Journal*, 1912 1, 717-718.
- [8] Thomson, JG; Sinton, JA. *Leishmania donovani*, In Cultures, Recovered From Spleen During Life And From The Bone Marrow Of The Cadaver. *Proceedings Of The Royal Society Of Medicine*, 1921 14, 21-23.
- [9] Banuls, AL; Hide, M; Tibayrenc, M. Evolutionary Genetics And Molecular Diagnosis Of *Leishmania* Species. *Transactions Of The Royal Society Of Tropical Medicine And Hygiene*, 2002 96 Suppl 1, S9-13.
- [10] Bates PA. Transmission Of *Leishmania* Metacyclic Promastigotes By Phlebotomine Sand Flies. *International Journal For Parasitology*, 2007 37, 1097-1106.
- [11] Schonian, G; Mauricio, I; Cupolillo, E. Is It Time To Revise The Nomenclature Of *Leishmania*? *Trends In Parasitology*, 2010 26, 466-469.
- [12] Matthews, KR. The Developmental Cell Biology Of *Trypanosoma brucei*. *Journal Of Cell Science*, 2005 118, 283-290.
- [13] Haile, S; Papadopoulou, B. Developmental Regulation Of Gene Expression In Trypanosomatid Parasitic Protozoa. *Current Opinion In Microbiology*, 2007 10; 569-577.
- [14] Simpson, L; Thiemann, OH; Savill, NJ; Alfonzo, JD; Maslov, DA. Evolution Of RNA Editing In Trypanosome Mitochondria. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 2000 97; 6986-6993.
- [15] Schnauffer, A; Ernst, NL; Palazzo, SS; O'Rear, J; Salavati, R; Stuart, K. Separate Insertion And Deletion Subcomplexes Of The *Trypanosoma brucei* RNA Editing Complex. *Molecular Cell*, 2003 12, 307-319.
- [16] Parsons M. Glycosomes: Parasites And The Divergence Of Peroxisomal Purpose. *Molecular Microbiology*, 2004 53, 717-724.
- [17] Gull K. The Cytoskeleton Of Trypanosomatid Parasites. *Annual Review Of Microbiology*, 1999 53, 629-655.
- [18] Field, MC; Carrington M. The Trypanosome Flagellar Pocket. *Nature Reviews. Microbiology*, 2009 7, 775-786.
- [19] Myler, PJ; Audleman, L; Devos, T; Hixson, G; Kiser, P; Lemley, C; et al. *Leishmania major* Friedlin Chromosome 1 Has An Unusual Distribution Of Protein-Coding Genes. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 1999 96, 2902-2906.
- [20] 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, 436-442.
- [21] El-Sayed, NM; Myler, PJ; Bartholomeu, DC; Nilsson, D; Aggarwal, G; Tran, AN; et al. The Genome Sequence Of *Trypanosoma cruzi*, Etiologic Agent Of Chagas Disease. *Science*, 2005 309, 409-415.
- [22] Berriman, M; Ghedin, E; Hertz-Fowler, C; Blandin, G; Renaud, H; Bartholomeu, DC; et al. The Genome Of The African Trypanosome *Trypanosoma brucei*. *Science*, 2005 309, 416-422.
- [23] 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. *Nature Genetics*, 2007 39, 839-847.
- [24] Rogers, MB; Hillel, 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 Research*, 2011 21, 2129-2142.
- [25] Downing, T; Inamura, 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 Research*, 2011 21, 2143-2156.
- [26] Lukes, J; Mauricio, IL; Schonian, G; Dujardin, JC; Soteriadou, K; Dedet, JP; et al. Evolutionary And Geographical History Of The *Leishmania donovani* Complex With A Revision Of Current Taxonomy. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 2007 104, 9375-9380.
- [27] Weatherly, DB; Boehlke, C; Tarleton, RL. Chromosome Level Assembly Of The Hybrid *Trypanosoma cruzi* Genome. *Bmc Genomics*, 2009 10, 255.

- [28] Myler, PJ; Beverley, SM; Cruz, AK; Dobson, DE; Ivens, AC; McDonagh, PD; et al. The *Leishmania* Genome Project: New Insights Into Gene Organization And Function. *Medical Microbiology And Immunology*, 2001 190, 9-12.
- [29] Lander, ES; Linton, LM; Birren, B; Nusbaum, C; Zody, MC; Baldwin, J; et al. Initial Sequencing And Analysis Of The Human Genome. *Nature*, 2001 409, 860-921.
- [30] Mair, G; Shi, H; Li, H; Djikeng, A; Aviles, HO; Bishop, JR; et al. A New Twist In Trypanosome RNA Metabolism: Cis-Splicing Of Pre-mRNA. *RNA*, 2000 6, 163-169.
- [31] Bringaud, F; Ghedin, E; El-Sayed, NM; Papadopoulou, B. Role Of Transposable Elements In Trypanosomatids. *Microbes And Infection*, 2008 10, 575-581.
- [32] Ghedin, E; Bringaud, F; Peterson, J; Myler, P; Berriman, M; Ivens, A; et al. Gene Synteny And Evolution Of Genome Architecture In Trypanosomatids. *Molecular And Biochemical Parasitology*, 2004 134, 183-191.
- [33] Bringaud, F; Muller, M; Cerqueira, GC; Smith, M; Rochette, A; El-Sayed, NM; et al. Members Of A Large Retroposon Family Are Determinants Of Post-Transcriptional Gene Expression In *Leishmania*. *Public Library Of Science Pathogens*, 2007 3, 1291-1307.
- [34] Requena, JM; Folgueira, C; Lopez, MC; Thomas, MC. The Sider2 Elements, Interspersed Repeated Sequences That Populate The *Leishmania* Genomes, Constitute Subfamilies Showing Chromosomal Proximity Relationship. *Bmc Genomics*, 2008 9, 263.
- [35] Teixeira, SM; De Paiva, RM; Kangussu-Marcolino, MM; Darocha, WD. Trypanosomatid Comparative Genomics: Contributions To The Study Of Parasite Biology And Different Parasitic Diseases. *Genetics And Molecular Biology*, 2012 35, 1-17.
- [36] Smith, DF; Peacock, CS; Cruz, AK. Comparative Genomics: From Genotype To Disease Phenotype In The Leishmaniases. *International Journal For Parasitology*, 2007 37, 1173-1186.
- [37] Zhang, WW; Mendez, S; Ghosh, A; Myler, P; Ivens, A; Clos, J; et al. Comparison Of The A2 Gene Locus In *Leishmania donovani* And *Leishmania major* And Its Control Over Cutaneous Infection. *Journal Of Biological Chemistry*, 2003 278, 35508-35515.
- [38] Fire, A; Xu, S; Montgomery, MK; Kostas, SA; Driver, SE; Mello, CC. Potent And Specific Genetic Interference By Double-Stranded RNA In *Caenorhabditis elegans*. *Nature*, 1998 391, 806-811.
- [39] Ngo, H; Tschudi, C; Gull, K; Ullu, E. Double-Stranded RNA Induces mRNA Degradation In *Trypanosoma brucei*. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 1998 95, 14687-14692.
- [40] Robinson, KA; Beverley, SM. Improvements In Transfection Efficiency And Tests Of RNA Interference (RNAi) Approaches In The Protozoan Parasite *Leishmania*. *Molecular And Biochemical Parasitology*, 2003 128, 217-228.
- [41] Lye, LF; Owens, K; Shi, H; Murta, SM; Vieira, AC; Turco, SJ; et al. Retention And Loss Of RNA Interference Pathways In Trypanosomatid Protozoans. *Public Library Of Science Pathogens*, 2010 6, E1001161.
- [42] Britto, C; Ravel, C; Bastien, P; Blaineau, C; Pages, M; Dedet, JP; et al. Conserved Linkage Groups Associated With Large-Scale Chromosomal Rearrangements Between Old World And New World *Leishmania* Genomes. *Gene*, 1998 222, 107-117.

- [43] Bastien, P; Blaineau, C; Pages, M. *Leishmania*: Sex, Lies And Karyotype. *Parasitology Today*, 1992, 8: 174-177.
- [44] Martinez-Calvillo, S; Stuart, K; Myler, PJ. Ploidy Changes Associated With Disruption Of Two Adjacent Genes On *Leishmania major* Chromosome 1. *International Journal For Parasitology*, 2005 35, 419-429.
- [45] Sunkin, SM; Kiser, P; Myler, PJ; Stuart, K. The Size Difference Between *Leishmania major* Friedlin Chromosome One Homologues Is Localized To Sub-Telomeric Repeats At One Chromosomal End. *Molecular And Biochemical Parasitology*, 2000 109, 1-15.
- [46] Sterkers, Y; Lachaud, L; Crobu, L; Bastien, P; Pages, M. Fish Analysis Reveals Aneuploidy And Continual Generation Of Chromosomal Mosaicism In *Leishmania major*. *Cellular Microbiology*, 2011 13, 274-283.
- [47] Rochette, A; McNicoll, F; Girard, J; Breton, M; Leblanc, E; Bergeron, MG; et al. Characterization And Developmental Gene Regulation Of A Large Gene Family Encoding Amastin Surface Proteins In *Leishmania* Spp. *Molecular And Biochemical Parasitology*, 2005 140, 205-220.
- [48] Mauricio, IL; Gaunt, MW; Stothard, JR; Miles, MA. Glycoprotein 63 (Gp63) Genes Show Gene Conversion And Reveal The Evolution Of Old World *Leishmania*. *International Journal For Parasitology*, 2007 37, 565-576.
- [49] Devault, A; Banuls, AL. The Promastigote Surface Antigen Gene Family Of The *Leishmania* Parasite: Differential Evolution By Positive Selection And Recombination. *Bmc Evolutionary Biology*, 2008 8, 292.
- [50] Johnson, PJ; Kooter, JM; Borst, P. Inactivation Of Transcription By UV Irradiation Of *T. brucei* Provides Evidence For A Multicistronic Transcription Unit Including A VSG Gene. *Cell*, 1987 51, 273-281.
- [51] Mottram, JC; Murphy, WJ; Agabian, N. A Transcriptional Analysis Of The *Trypanosoma brucei* Hsp83 Gene Cluster. *Molecular And Biochemical Parasitology*, 1989 37, 115-127.
- [52] Martinez-Calvillo, S; Yan, S; Nguyen, D; Fox, M; Stuart, K; Myler, PJ. Transcription Of *Leishmania major* Friedlin Chromosome 1 Initiates In Both Directions Within A Single Region. *Molecular Cell*, 2003 11, 1291-1299.
- [53] Campbell, DA; Thomas, S; Sturm, NR. Transcription In Kinetoplastid Protozoa: Why Be Normal? *Microbes And Infection*, 2003 5, 1231-1240.
- [54] Martinez-Calvillo, S; Vizuet-de-Rueda, JC; Florencio-Martinez, LE; Manning-Cela, RG; Figueroa-Angulo, EE. Gene Expression In Trypanosomatid Parasites. *Journal Of Biomedicine And Biotechnology*, 2010, 525241.
- [55] Parsons, M; Nelson, RG; Watkins, KP; Agabian, N. Trypanosome mRNAs Share A Common 5' Spliced Leader Sequence. *Cell*, 1984 38, 309-316.
- [56] Kooter, JM; Borst, P. Alpha-Amanitin-Insensitive Transcription Of Variant Surface Glycoprotein Genes Provides Further Evidence For Discontinuous Transcription In Trypanosomes. *Nucleic Acids Research*, 1984 12, 9457-9472.
- [57] Clayton, CE. Life Without Transcriptional Control? From Fly To Man And Back Again. *The Embo Journal*, 2002 21, 1881-1888.
- [58] Boucher, N; Wu, Y; Dumas, C; Dube, M; Sereno, D; Breton, M; et al. A Common Mechanism Of Stage-Regulated Gene Expression In *Leishmania* Mediated By A Conserved 3'-Untranslated Region Element. *Journal Of Biological Chemistry*, 2002 277, 19511-19520.

- [59] Martínez-Calvillo, S; Nguyen, D; Stuart, K; Myler, PJ. Transcription Initiation And Termination On *Leishmania major* Chromosome 3. *Eukaryotic Cell*, 2004 3, 506-517.
- [60] Thomas, S; Green, A; Sturm, NR; Campbell, DA; Myler, PJ. Histone Acetylations Mark Origins Of Polycistronic Transcription In *Leishmania major*. *Bmc Genomics* 2009 10, 152.
- [61] Saito, RM; Elgort, MG; Campbell, DA. A Conserved Upstream Element Is Essential For Transcription Of The *Leishmania tarentolae* Mini-Exon Gene. *The Embo Journal*, 1994 13, 5460-5469.
- [62] Gunzl, A; Ullu, E; Dorner, M; Frago, SP; Hoffmann, KF; Milner, JD; et al. Transcription Of The *Trypanosoma brucei* Spliced Leader RNA Gene Is Dependent Only On The Presence Of Upstream Regulatory Elements. *Molecular And Biochemical Parasitology*, 1997 85, 67-76.
- [63] Luo, H; Bellofatto, V. Characterization Of Two Protein Activities That Interact At The Promoter Of The Trypanosomatid Spliced Leader RNA. *Journal Of Biological Chemistry*, 1997 272, 33344-33352.
- [64] Gunzl, A; Vanhamme, L; Myler, PJ. Transcription In Trypanosomes: A Different Means To The End. Barry, JD; McCulloch, R; Mottram, JC; Acosta-Serrano, A. (Ed.) In: Trypanosomes: After The Genome. Wyonham: Horizon Bioscience; 2007; P. 177-208.
- [65] Hernandez-Rivas, R; Florencio-Martinez, LE; Martinez-Salazar, M; Martinez-Calvillo, S. Gene Expression And Transcriptional Machinery In Trypanosomatid And Apicomplexa Parasites. Terrazas, LI. (Ed.) In: Advances In The Immunobiology Of Parasitic Diseases. Kerala, India: Research Signpost; 2007; P. 313-37.
- [66] Ruan, JP; Arhin, GK; Ullu, E; Tschudi, C. Functional Characterization Of A *Trypanosoma brucei* TATA-Binding Protein-Related Factor Points To A Universal Regulator Of Transcription In Trypanosomes. *Molecular And Cellular Biology*, 2004 24, 9610-9618.
- [67] Palenchar, JB; Liu, W; Palenchar, PM; Bellofatto, V. A Divergent Transcription Factor TFIIIB In Trypanosomes Is Required For RNA Polymerase II-Dependent Spliced Leader RNA Transcription And Cell Viability. *Eukaryotic Cell*, 2006 5, 293-300.
- [68] Schimanski, B; Brandenburg, J; Nguyen, TN; Caimano, MJ; Gunzl, A. A TFIIIB-Like Protein Is Indispensable For Spliced Leader RNA Gene Transcription In *Trypanosoma brucei*. *Nucleic Acids Research*, 2006 34, 1676-1684.
- [69] Lecordier, L; Devaux, S; Uzureau, P; Dierick, JF; Walgraffe, D; Poelvoorde, P; et al. Characterization Of A TFIIH Homologue From *Trypanosoma brucei*. *Molecular Microbiology*, 2007 64, 1164-1181.
- [70] Lee, JH; Nguyen, TN; Schimanski, B; Gunzl, A. Spliced Leader RNA Gene Transcription In *Trypanosoma brucei* Requires Transcription Factor TFIIH. *Eukaryotic Cell*, 2007 6, 641-649.
- [71] Padilla-Mejía, 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.
- [72] Marck, C; Grosjean, H. tRNomics: Analysis Of tRNA Genes From 50 Genomes Of Eukarya, Archaea, And Bacteria Reveals Anticodon-Sparing Strategies And Domain-Specific Features. *Rna*, 2002 8, 1189-1232.

- [73] Goodenbour, JM; Pan, T. Diversity Of tRNA Genes In Eukaryotes. *Nucleic Acids Research*, 2006 34, 6137-6146.
- [74] Hasan, G; Turner, MJ; Cordingley, JS. Ribosomal RNA Genes Of *Trypanosoma brucei*: Mapping The Regions Specifying The Six Small Ribosomal RNAs. *Gene* 1984 27, 75-86.
- [75] Hernandez-Rivas, R; Martinez-Calvillo, S; Romero, M; Hernandez R. *Trypanosoma cruzi* 5S rRNA Genes: Molecular Cloning, Structure And Chromosomal Organization. *The Fems Microbiology Letters*, 1992 71, 63-67.
- [76] Spencer, DF; Collings, JC; Schnare, MN; Gray, MW. Multiple Spacer Sequences In The Nuclear Large Subunit Ribosomal RNA Gene Of *Crithidia fasciculata*. *The Embo Journal*, 1987 6, 1063-1071.
- [77] Hernandez, R; Diaz De Leon, F; Castaneda, M. Molecular Cloning And Partial Characterization Of Ribosomal RNA Genes From *Trypanosoma cruzi*. *Molecular And Biochemical Parasitology*, 1988 27, 275-279.
- [78] Martínez-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. *Molecular And Biochemical Parasitology*, 2001 116, 147-157.
- [79] Inga, R; De Doncker, S; Gomez, J; Lopez, M; Garcia, R; Le Ray, D; et al. Relation Between Variation In Copy Number Of Ribosomal RNA Encoding Genes And Size Of Harbouring Chromosomes In *Leishmania* Of Subgenus *Viannia*. *Molecular And Biochemical Parasitology*, 1998 92, 219-228.
- [80] Liu, B; Liu, Y; Motyka, SA; Agbo, EE; Englund, PT. Fellowship Of The Rings: The Replication Of Kinetoplast DNA. *Trends In Parasitology*, 2005 21, 363-369.
- [81] Lukes, J; Guilbride, DL; Votycka, J; Zikova, A; Benne, R; Englund, PT. Kinetoplast DNA Network: Evolution Of An Improbable Structure. *Eukaryotic Cell*, 2002 1, 495-502.
- [82] Simpson, L; Wang, SH; Thiemann, OH; Alfonzo, JD; Maslov, DA; Avila, HA. U-Insertion/Deletion Edited Sequence Database. *Nucleic Acids Research*, 1998 26, 170-176.
- [83] Simpson, L. The Genomic Organization Of Guide RNA Genes In Kinetoplastid Protozoa: Several Conundrums And Their Solutions. *Molecular And Biochemical Parasitology*, 1997 86, 133-141.
- [84] Nebohacova, M; Kim, CE; Simpson, L; Maslov, DA. RNA Editing And Mitochondrial Activity In Promastigotes And Amastigotes Of *Leishmania donovani*. *International Journal For Parasitology*, 2009 39, 635-644.
- [85] Tripp, CA; Myler, PJ; Stuart, KD. Nucleotide Sequence Of A Minicircle From *Leishmania infantum*. *Molecular And Biochemical Parasitology*, 1993 62, 319-320.
- [86] Clement, SL; Koslowsky, DJ. Unusual Organization Of A Developmentally Regulated Mitochondrial RNA Polymerase (Tbntn1p) Gene In *Trypanosoma brucei*. *Gene*, 2001 272, 209-218.
- [87] Lukes, J; Hashimi, H; Zikova, A. Unexplained Complexity Of The Mitochondrial Genome And Transcriptome In Kinetoplastid Flagellates. *Current Genetics*, 2005 48, 277-299.
- [88] Aphasizhev, R; Aphasizheva, I. Mitochondrial RNA Processing In Trypanosomes. *Research In Microbiology*, 2011 162, 655-663.

- [89] Michelotti, EF; Harris, ME; Adler, B; Torri, AF; Hajduk, SL. *Trypanosoma brucei* Mitochondrial Ribosomal RNA Synthesis, Processing And Developmentally Regulated Expression. *Molecular And Biochemical Parasitology*, 1992 54, 31-41.
- [90] Grams, J; McManus, MT; Hajduk, SL. Processing Of Polycistronic Guide RNAs Is Associated With RNA Editing Complexes In *Trypanosoma brucei*. *The Embo Journal*, 2000 19, 5525-5532.