



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
DOCTORADO EN CIENCIAS BIOMÉDICAS
INSTITUTO DE INVESTIGACIONES BIOMÉDICAS

**ESTUDIO FUNCIONAL DE LAS SEÑALES DE
PROCESAMIENTO Y POLIADENILACIÓN DE LOS RNA
MENSAJEROS EN *Trichomonas vaginalis***

T E S I S
QUE PARA OPTAR POR EL GRADO DE
DOCTORA EN CIENCIAS

PRESENTA:
QFB. VANESSA FUENTES MORALES

MIEMBROS DEL COMITÉ TUTOR
DRA. MARÍA IMELDA LÓPEZ VILLASEÑOR (TUTOR PRINCIPAL)
INSTITUTO DE INVESTIGACIONES BIOMÉDICAS

DRA. ROSSANA ARROYO VERASTEGUI
CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL IPN

DR. ROBERTO CORIA ORTEGA
INSTITUTO DE FISIOLÓGÍA CELULAR

CIUDAD UNIVERSITARIA, CD. MX. MARZO 2018



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.

Esta tesis se desarrolló bajo la asesoría de la Dra. María Imelda López Villaseñor en el Departamento de Biología Molecular y Biotecnología del Instituto de Investigaciones Biomédicas en la Universidad Nacional Autónoma de México.

El comité tutorial estuvo integrado por la Dra. María Imelda López Villaseñor, la Dra. Rossana Arroyo Verástegui y el Dr. Roberto Coria Ortega.

Esta tesis doctoral fue apoyada por una beca de tiempo completo del CONACyT.

AGRADECIMIENTOS

A la Universidad Nacional Autónoma de México, mi Alma mater que primero me abrigó en la Facultad de Química, y después me dio la oportunidad de estudiar un posgrado de excelencia.

Al Instituto de Investigaciones Biomédicas y al Programa de Doctorado en Ciencias Biomédicas (PDCB) por brindarme una excelente formación académica.

A la Dra. María Imelda López Villaseñor por su excelente guía, consejos y apoyo durante esta tesis. Por ser una tutora invaluable y por su comprensión en los momentos difíciles

A los miembros de mi comité tutorial, integrado por la Dra. Rossana Arroyo Verastegui y el Dr. Roberto Coria Ortega, por sus comentarios y guía durante los tutorales que permitieron enriquecer y mejorar este trabajo.

Al Dr. Roberto Hernández y a la Dra. Ana María Cevallos por sus valiosos comentarios y sugerencias que me permitieron mejorar este trabajo doctoral.

A los miembros del jurado les agradezco por la revisión del manuscrito de tesis y por sus valiosas observaciones y comentarios.

A la M. C. Karla C. López Pacheco por las clonas de cDNA, que permitieron ampliar el análisis bioinformático, así como por el apoyo técnico brindado.

Agradezco la beca otorgada por CONACYT (No. Registro 202742) durante la realización de esta tesis.

*"Si el hombre alcanzara la
mitad de sus deseos,
redoblaría sus inquietudes"*
Benjamín Franklin

ÍNDICE

Resumen	7
Abstract	9
Abreviaturas	10
1. Introducción	
1.1 Clasificación taxonómica de <i>Trichomonas vaginalis</i>	11
1.2 Filogenia molecular	11
1.3 Aspectos biológicos	12
1.4 Metabolismo	14
1.5 Ciclo de vida	15
1.6 Aspectos clínicos de la tricomonosis	17
1.7 Diagnóstico	18
1.8 Epidemiología	19
1.9 Regulación de la expresión génica en <i>T. vaginalis</i>	21
2. Antecedentes	
2.1 Generalidades sobre el proceso de transcripción en eucariontes	23
2.2 Elementos que dirigen la poliadenilación en mamíferos	26
2.3 Señales de poliadenilación en levaduras	27
2.4 Señales de poliadenilación en plantas	28
2.5 Factores proteicos involucrados en el proceso de poliadenilación	28
2.6 Señales de poliadenilación en protozoarios	31
2.7 Vectores para la expresión de genes reporteros en <i>T. vaginalis</i>	34
3. Justificación	38
4. Objetivo general	40
4.1 Objetivos particulares.	40
5. Resultados y Discusión	
5.1 Análisis in silico	
5.1.1. Presencia de la señal propuesta de poliadenilación UAAA en mRNAs de <i>T. vaginalis</i>	42
5.1.2 Presencia de las señales de poliadenilación en la región 3'UTR de varios cDNA	44
5.2 Análisis funcional de la señal de poliadenilación UAAA	50
5.2.1 Identificación del sitio de corte y poliadenilación de los mRNA de <i>cat</i> en la secuencia parental y las mutantes.	51
5.2.2. Evaluación de la población de los transcritos de <i>cat</i>	54
5.2.3 La relocalización del motivo UAAA conlleva a la relocalización del sitio de poliadenilación de los transcritos.	57
5.2.4 Interacción de la señal de poliadenilación UAAA con factores en extractos proteicos de <i>T. vaginalis</i>	62
5.2.5 Actividad de CAT medida en los extractos proteicos	65

5.3	Análisis funcional del sitio de procesamiento del pre-mRNA y la región río abajo rica en U's	66
5.3.1.	Mutagénesis dirigida en la región río abajo rica en U's.	68
6.	Conclusiones	71
7.	Perspectivas	71
8.	Artículo publicado	72
9.	Metodología	82
10.	Soluciones y oligonucleótidos utilizados	96
11.	Referencias	102

RESUMEN

La RNA polimerasa II sintetiza los RNA mensajeros (mRNA) eucariontes. Estos mRNAs deben ser modificados dentro del núcleo, de manera cotranscripcional, para ser exportados y traducidos en el citoplasma. Dentro de estas modificaciones se encuentra la adición de una 7-metilguanósina en el extremo 5' (capping), la remoción de intrones por medio del splicing y la adición de una cola de poli(A) en el extremo 3' del pre-mRNA (poliadenilación). Este último evento es de suma importancia ya que le brinda estabilidad al RNA mensajero, le permite ser exportado del núcleo al citoplasma y participa en el inicio de la traducción.

Para que se lleve a cabo la poliadenilación del transcrito primario, es necesario que sean reconocidos al menos cuatro elementos en *cis* en el RNA. En metazoarios se han identificado y estudiado estos elementos: 1) la señal de poliadenilación, cuya secuencia canónica es el hexanucleótido AAUAAA (y secuencias variantes) localizada de 11-30 nt río arriba del sitio de adición de la cola poli(A), y que es el sitio de unión para el CPSF-160 (cleavage and polyadenylation specificity factor 160 subunit); 2) un elemento río abajo del sitio de corte del pre-mRNA, rico en U's o GU, que es el sitio de unión del factor CstF-64 (cleavage stimulation factor 64 subunit); 3) el sitio de procesamiento del transcrito que preferencialmente se lleva a cabo después del dinucleótido CA; 4) regiones ricas en Us río arriba de la señal de poliadenilación. En algunos organismos se ha descrito también la secuencia UGUA localizada 40-100 nt río arriba del sitio de procesamiento.

Trichomonas vaginalis es un parásito protozoario flagelado de importancia médica, ya que es el agente etiológico de la tricomonosis, una de las enfermedades de transmisión sexual más común del tipo no viral. En este organismo, la mayoría de los estudios sobre la transcripción por la RNA polimerasa II se han enfocado en la región promotora, donde se ha descrito una secuencia Inr [TCA₊₁Y(T/A)] que es reconocida por la proteína IBP39, de la cual no se han identificado homólogos en otras especies. En contraste, el procesamiento del extremo 3' de los mRNAs no había sido estudiado previamente en este organismo.

En nuestro grupo de investigación nos hemos enfocado en el estudio de las señales involucradas en el proceso de poliadenilación en *T. vaginalis*. En un análisis inicial *in silico* se compararon las regiones 3'UTR de 13 cDNAs con sus respectivas secuencias genómicas, lo que permitió identificar algunas señales que podrían participar en la poliadenilación: a) el tetranucleótido UAAA que se encuentra en muchos casos empalmado con el codón de término de la traducción, y se encuentra localizado de 11-30 nt río arriba del sitio de adición del poli(A); b) un sitio de procesamiento con la

secuencia Y(A)₂₋₅ UU y c) regiones río abajo ricas en Us (Espinosa et al. 2002). Para determinar la funcionalidad de las señales propuestas, se construyó un vector de expresión transitorio para *T. vaginalis* (partiendo de una clona genómica de actina) donde se reemplazó la región codificadora de actina por la del gen reportero *cat*. En el presente trabajo se demostró que el tetranucleótido UAAA es la señal de poliadenilación en *T. vaginalis*, utilizando varios enfoques: 1) mutagénesis dirigida sobre el tetranucleótido UAAA en donde se cambiaron todos y cada uno de los cuatro nucleótidos, así como los nucleótidos adyacentes. El efecto de cada una de estas mutaciones sobre la poliadenilación del transcrito se determinó mediante diferentes estrategias: se identificó el sitio de procesamiento de los mRNAs mediante RT-PCR (PCR acoplado a transcriptasa reversa) y secuenciación de los productos amplificados; 2) se analizó la población del mRNA derivado de cada mutante, mediante PCR semicuantitativo; 3) mediante análisis tipo EMSA se determinó el sitio de unión para factores protéicos de *T. vaginalis*; 4) como una evaluación indirecta del efecto de modificar la señal de poliadenilación en el mRNA del gen reportero *cat*, se determinaron los niveles de actividad de la proteína CAT en extractos celulares de cultivos de *T. vaginalis* transfectados con cada uno de los plásmidos estudiados. Se encontró que el tetranucleótido UAAA es la señal que dirige la poliadenilación en *T. vaginalis*, y que cualquier cambio en este motivo afecta dramáticamente tanto el sitio de procesamiento como la cantidad del transcrito; se encontró también que el nivel de actividad de CAT disminuye hasta llegar a ser únicamente el 10% de la actividad del control positivo.

Los datos que se presentan en este trabajo demuestran que el elemento esencial para la poliadenilación de los mRNAs en *T. vaginalis* es el tetranucleótido UAAA, una señal atípica, más corta que la reportada para la mayoría de los eucariontes en donde la secuencia canónica es el hexanucleótido AAUAAA. De manera interesante, en la mayoría de los RNA mensajeros de este organismo el codón de término de la traducción UAA se encuentra formando parte de la señal de poliadenilación.

ABSTRACT

Synthesis of functional mRNA in eukaryotes involves processing of precursor transcripts, including the addition of a poly(A) tail at the 3' end. A multiprotein complex recognizes a polyadenylation signal, generally the hexanucleotide AAUAAA in metazoans, to direct processing of the pre-mRNA. Based on sequence analysis of several cDNAs, we have previously suggested that the UAAA tetranucleotide (which may include the UAA translation stop codon) could be the polyadenylation signal in *Trichomonas vaginalis*, a parasitic protozoon that causes human trichomoniasis. This proposal is analyzed here with the aid of a transient expression system of a reporter gene (*cat* flanked by *T. vaginalis* actin noncoding sequences). When cells were transfected with a plasmid bearing the original 3' untranslated region (UTR) sequence containing the UAAA motif, the resulting *cat* mRNA was polyadenylated similarly to the endogenous actin mRNA. Base changes in the UAAA sequence produced alterations to the polyadenylation site of the reporter mRNAs, while nucleotide substitutions at either side of UAAA did not. Furthermore, relocation of the UAAA motif redirected the processing and polyadenylation of the reporter mRNA. Interaction of *T. vaginalis* proteins with the UAAA motif was shown by electrophoretic mobility shift assays. Based on our findings, we provide evidence that in *T. vaginalis* the UAAA tetranucleotide has a role equivalent to that of the metazoan consensus AAUAAA polyadenylation signal.

ABREVIATURAS

RNA	Acido ribonucleico
DNA	Ácido desoxirribonucleico
mRNA	RNA mensajero
cDNA	DNA complementario al mRNA
Inr	Iniciador para la transcripción en <i>T. vaginalis</i>
UTR	Región no traducida (untranslated región)
nt	Nucleótidos
<i>cat</i>	Gen que codifica para la cloranfenicol acetil transferasa
CPSF	Factor específico de corte y poliadenilación (cleavage and polyadenylation specificity factor)
CstF	Factor de estimulación del corte (cleavage stimulation factor)
RT-PCR	Reacción en cadena de la polimerasa acoplada a la reacción de transcriptasa reversa
µm	Micrómetro
°C	Grados centígrados
CAT	Cloranfenicol Acetil Transferasa
DEPC	Dietilpirocarbonato
dNTP	Deoxinucleótido trifosfato
DSE	Elemento río abajo (downstream element).
EDTA	Ácido etilendiaminotetraacético
hnRNPs	Ribonucleoproteínas heterogéneas nucleares (heterogeneous nuclear ribonuclear proteins)
Kb	Kilobases
min	Minuto
mL	Mililitro
nm	Nanómetro
PABP	Proteína de unión a poli A
PAP	Enzima poly A polimerasa
pb	Pares de bases
PBS	Amortiguador salino de fosfatos
rpm	Revoluciones por minuto
TAE	Amortiguador de Tris-acetato-EDTA
TE	Amortiguador de Tris-EDTA
ORF	Marco abierto de lectura (open reading frame)

1. INTRODUCCIÓN

1.1 Clasificación taxonómica de *Trichomonas vaginalis*

T. vaginalis es un organismo eucarionte unicelular flagelado perteneciente al grupo de los parabasalidos.

Reino: Protista
Subreino: Dimastigota
Superphylum: Tetramastigota
Phylum: Axostylata
Clase: Parabasalea
Orden: Trichomonadida
Familia: Trichomonadidae
Género: *Trichomonas*
Especie: *Trichomonas vaginalis*

1.2. Filogenia molecular

T. vaginalis ha sido tradicionalmente considerado como un organismo de divergencia evolutiva temprana por la posición que presenta al realizar análisis filogenéticos con base en la secuencia del RNA ribosomal de la subunidad grande (Baroin et al. 1988), o de la subunidad pequeña (Sogin M., 1991) (Fig. 1).

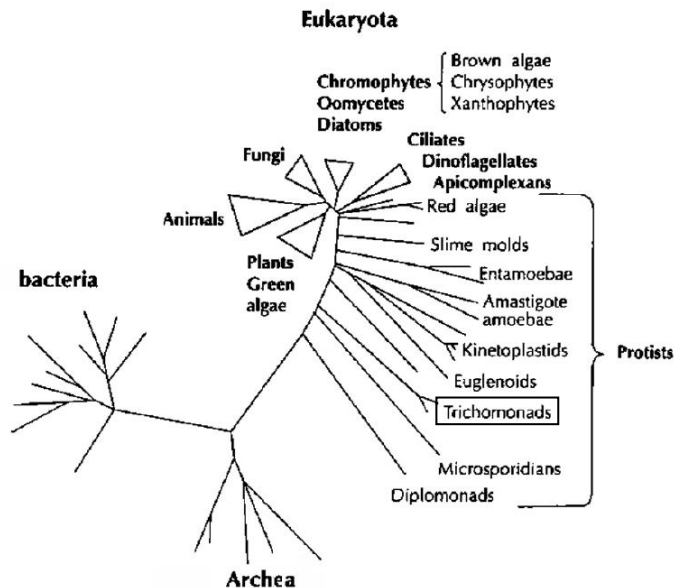


Fig. 1 Árbol filogenético inferido del análisis de la secuencia de la subunidad pequeña del RNA ribosomal de organismos de 3 diferentes reinos. Tomado de Sogin M., 1991.

Para determinar la relación que tienen los parabasálidos con otros organismos protozoarios, se han utilizado diferentes características bioquímicas y moleculares (Yarlett et al. 2001). Con el paso del tiempo se ha demostrado que el modelo mostrado en la figura 1 es el resultado de un artefacto metodológico conocido como “Long Branch Attraction”, el cual no considera si la divergencia es debida a una evolución rápida y reciente, o a una acumulación lenta de cambios a lo largo del tiempo (Dacks et al. 2008). Actualmente es aceptado un modelo filogenético igualitario, producto de análisis filogenéticos más completos, donde los organismos eucariontes se encuentran en un mismo nivel y son agrupados en seis súper-grupos. Los seis súper-grupos eucariontes son: Amoebozoa, Opisthokonta, Excavata, Chromalveolata, Rhizaria y Archaeplastida (Fig. 2). A *T. vaginalis* se le sitúa en el grupo Excavata (Dacks et al. 2008; Simpson et al. 2004).

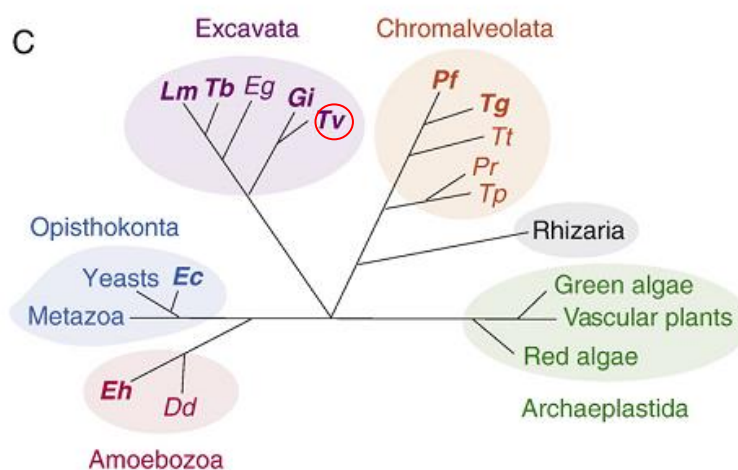


Fig. 2. Arbol filogenético del modelo igualitario de los seis súper-grupos, donde: Dd; *Dictyostelium discoïdium*, Ec; *Encephalitozoon cuniculi*, Eg; *Euglena gracilis*, Eh; *Entamoeba histolytica*, Gi; *Giardia intestinalis*, Lm; *Leishmania major*, Pf; *Plasmodium falciparum*, Pr; *Phytophthora ramorum*, Tb; *Trypanosoma brucei*, Tg; *Toxoplasma gondii*, Tp; *Thalassiosira pseudonana*, Tt; *Tetrahymena thermophila* y Tv; *Trichomonas vaginalis* (Dacks et al. 2008).

1.3 Aspectos biológicos

Trichomonas vaginalis fue descrito por primera vez en 1836 por Alfred Donné a partir de secreciones purulentas de hombres y mujeres, pero no fue sino hasta principios del siglo XX cuando se comenzó a estudiar más detalladamente (Thorburn, 1974). Es un protozooario microaerofílico que posee cinco flagelos, cuatro de los cuales se encuentran en la porción anterior y el quinto flagelo está adosado a la membrana ondulante (Figura 3). La forma de *T. vaginalis* en cultivo es típicamente piriforme, el tamaño de este parásito se sitúa entre los 9 y los 23 μm de largo y 7 μm de ancho. Este organismo infecta el tracto urogenital del humano y es transmitido por contacto sexual causando una enfermedad conocida como tricomonosis.

T. vaginalis pertenece al grupo de los parabasálidos, posee seis cromosomas monocéntricos y es un organismo haploide (Yuh et al. 1997). En la secuencia liberada del genoma de *T. vaginalis* se reporta un tamaño de 160 Mb (Carlton et al. 2007) el cual es muy parecido entre trichomonádidos (86 - 177 Mb). El tamaño del genoma de este parásito es grande en comparación con los genomas secuenciados de otros protozoarios parásitos que se encuentran entre 2.9 y 80 Mb (Zubáková et al. 2008). Sin embargo, la presencia de múltiples secuencias repetidas y transponibles ha dificultado ensamblar el genoma y determinar su tamaño exacto.

En cuanto a su morfología, el núcleo de *T. vaginalis* se localiza en la porción anterior de la célula y como en otros eucariontes está rodeado por poros nucleares. Posee una estructura hialina parecida a una varilla, llamada axostilo (Fig. 3), que bisecta longitudinalmente al protozoario y va del extremo anterior al posterior de la célula terminando en punta esta estructura le brinda soporte (Petrin et al. 1998, Schwebke et al. 2004). El crecimiento de *T. vaginalis* bajo condiciones desfavorables, en presencia de fármacos o a temperaturas bajas produce formas redondeadas, llamadas pseudoquistes, en donde se internalizan los flagelos, no móviles y carecen de una pared quística, se desconoce si estas formas se encuentran en los hospederos (Hussein et al. 2008).

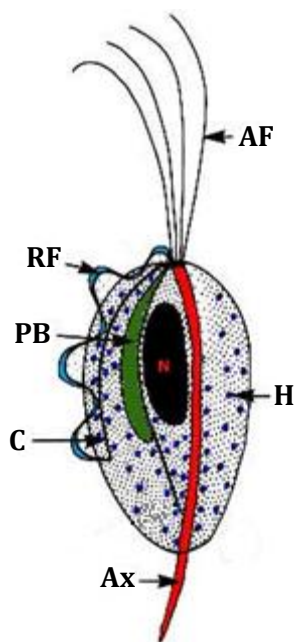


Figura 3. Diagrama de *T. vaginalis*. Este parásito tiene cuatro flagelos anteriores (AF), un flagelo recurrente (RF), núcleo (N), costa (C) que es un apoyo mecánico para la membrana ondulante y es específica de los parabasálidos; cuerpo parabasal (PB), axostilo (Ax) e hidrogenosomas (H).

El citoesqueleto de *T. vaginalis* está compuesto por tubulina y fibras de actina. Estos organismos no poseen mitocondrias ni peroxisomas, y en su lugar es común observar organelos granulosos catalasa negativos, lo que indica que no se trata de peroxisomas. Estos gránulos producen hidrógeno molecular por lo cual son llamados hidrogenosomas, muy importantes en el metabolismo de este protozoario. Derivado de estudios filogenéticos, se ha visto que el hidrogenosoma constituye una forma reducida de la mitocondria (Leitsch, 2016). En *T. vaginalis* hay dos tipos de hidrogenosomas: paracostal y paraxostilar; estos últimos se organizan a lo largo del axostilo en tres hileras paralelas características. También posee gránulos de glucógeno y fagosomas que pueden ser observados por microscopía electrónica de transmisión (Petrin et al. 1998).

1.4 Metabolismo

T. vaginalis posee algunas características metabólicas que son similares a las de bacterias anaeróbicas, pero también algunas que son características de eucariontes superiores.

Los carbohidratos son la fuente de nutrientes preferida por *T. vaginalis*, sin embargo, bajo condiciones donde éstos son limitados, los aminoácidos son empleados para el crecimiento y supervivencia de la célula. El metabolismo de carbohidratos en este parásito es fermentativo bajo condiciones anaeróbicas y aeróbicas debido a que la glucosa no es oxidada completamente. Los productos metabólicos incluyen acetato, lactato, malato, glicerol, CO₂ y, en condiciones anaeróbicas, H₂. El metabolismo de carbohidratos se lleva a cabo en dos compartimientos: el citoplasma y el hidrogenosoma. Este último es análogo a la mitocondria en eucariontes superiores; tiene un diámetro de 0.5 a 1.0 µm y está rodeado por una doble membrana, carece de citocromos y DNA. Los hidrogenosomas son el sitio de oxidación fermentativa de piruvato y en ellos se produce ATP por fosforilación a nivel del sustrato (Petrin et al. 1998). En este organelo se identificó y analizó la β-succinil coenzima A sintetasa (β-SCS), una proteína típicamente localizada en la mitocondria de eucariontes superiores. Otros estudios han mostrado que en el hidrogenosoma se encuentran proteínas relacionadas con la mitocondria, sugiriendo el origen común entre la mitocondria y el hidrogenosoma (Lahti et al. 1992, Roger et al. 1996, Germot et al. 1996, Van Gisven et al. 2007).

T. vaginalis es un parásito obligado ya que carece de la capacidad de llevar a cabo la síntesis de novo de un gran número de macromoléculas, particularmente purinas, pirimidinas y lípidos. Estos nutrientes generalmente son adquiridos de las secreciones vaginales o por fagocitosis de células del huésped o células bacterianas. Además este organismo posee una ruta auxiliar para generar purinas, la cual está mediada por las enzimas nucleósido fosforilasas y cinasas, mientras que las

fosforibosiltransferasas y nucleósido cinasas son capaces de recuperar pirimidinas. El suero es esencial para el crecimiento de los tricomonádidos *in vitro*, ya que provee de lípidos, ácidos grasos, aminoácidos y trazas de metales. *In vitro*, *T. vaginalis* crece de manera óptima a pH de 6 a 6.3, aunque es capaz de crecer en un amplio rango de pHs, especialmente en el cambiante medio de la vagina (Petrin et al. 1998).

1.5 Ciclo de vida

El ciclo de vida de este parásito se da en un único hospedero, el humano, infectando el tracto urogenital. *T. vaginalis* presenta un solo estadio en su ciclo de vida, el trofozoíto, y su reproducción es asexual por fisión binaria (Vanacova et al. 2003) (Fig. 4). Este parásito es transmitido por contacto sexual directo de persona a persona.

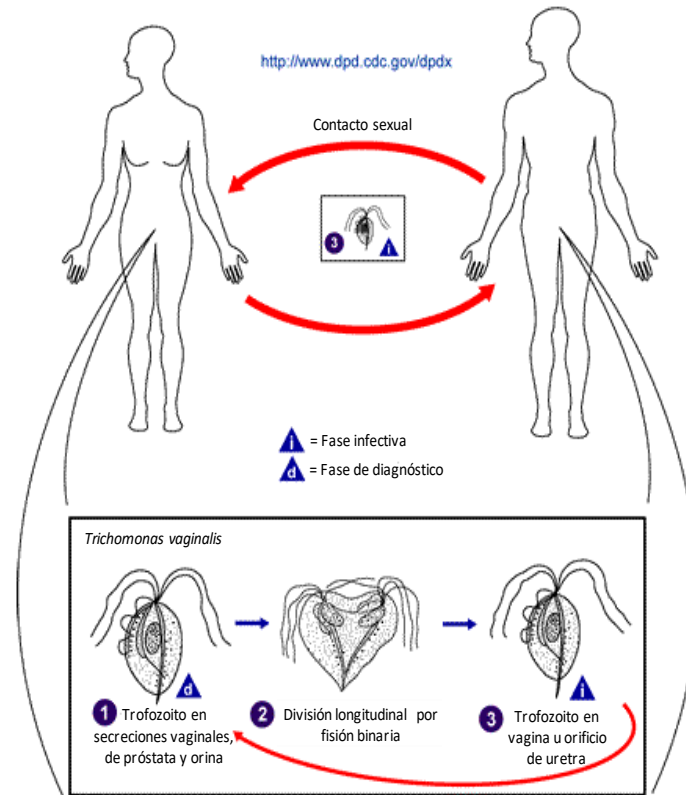


Figura 4. Ciclo de vida de *T. vaginalis*. Tomado y modificado de: <http://www.dpd.cdc.gov/dpdx>

Poco es lo que se sabe acerca del ciclo celular de este parásito, pero el proceso mitótico se ha analizado microscópicamente con más detalle. Se han podido observar cinco fases mitóticas: profase, metafase, anafase (cercana y lejana), telofase (cercana y lejana) y citocinesis. La duración

de la mitosis en este organismo es de 45 min (Gomez-Conde et al. 2000) y se caracteriza por llevarse a cabo de manera cerrada, es decir, la envoltura nuclear se mantiene intacta durante todo el proceso con la formación de un huso mitótico extranuclear. Durante la división nuclear, el núcleo se elonga y adquiere una forma irregular. La cromatina se condensa y se acumula en el núcleo. La elongación del núcleo continúa hasta que sufre un corte por la mitad, produciéndose así dos núcleos hijos (Yusof et al. 2012). Se ha observado la formación de un complejo axostilo/pelta en la fase pre-mitótica que participa activamente en el proceso de división celular, tanto en los cambios de forma y contorsión de la región anterior de la célula como durante la cariocinesis. Por otra parte se ha observado que el flagelo facilita la citocinesis ya que la propulsión generada por el flagelo en cada extremidad del cuerpo celular contribuye a la elongación y consecuente división de la célula (Ribeiro et al. 2000).

1.6 Aspectos clínicos de la tricomonosis

El humano es el hospedero natural de algunos parásitos del género *Trichomonas*, entre los que se encuentra a *T. tenax*, *T. hominis* y *T. vaginalis*. Los dos primeros no causan ninguna enfermedad, sin embargo *T. vaginalis* es el agente causal de una de las enfermedades más comunes de transmisión sexual del tipo no viral: la tricomonosis. La transmisión de este parásito es de persona a persona por contacto sexual, y estudios *in vitro* indican que el periodo de incubación puede ir de 4 a 28 días (Harp et al. 2011). De acuerdo con la severidad de la infección ésta se puede clasificar en: aguda, crónica y asintomática. La infección, una vez establecida, persiste por largos periodos en mujeres pero sólo por corto tiempo en hombres.

En la infección aguda, los principales síntomas en mujeres incluyen flujo vaginal amarillo-verdoso mucopurulento y maloliente. En el 2% de las pacientes infectadas se observa el llamado “cérvix de fresa”, esto es, la vagina y cervix adquieren un aspecto eritromatoso y edematoso causado por una dilatación capilar con la erosión general del epitelio cervical y puntos hemorrágicos; esto puede ser detectado por colposcopia. En la infección crónica los síntomas son más leves, con prurito y disparenia, las secreciones vaginales son escasas y pueden ser mucosas; esta etapa de la infección es importante a nivel epidemiológico ya que durante este estadio, al no presentarse síntomas severos aparentes, las personas infectadas pueden transmitir la enfermedad por la falta de un diagnóstico. Se ha estimado que del 25-50% de las mujeres infectadas son asintomáticas, presentando un pH vaginal y una flora normal, estas se presentan sólo como portadoras del parásito y aproximadamente el 50% de estas mujeres comienza con síntomas después de seis meses de la infección, comenzando con emisiones vaginales (Petrin et al. 2008).

La tricomonosis urogenital en hombres puede ser clasificada dentro de tres grupos: estado de portador asintomático, identificado sólo cuando la pareja ha sido diagnosticada; tricomonosis aguda, caracterizada por uretritis purulenta profusa; y enfermedad sintomática leve, clínicamente indistinguible de otras causas de uretritis no-gonococcal. La duración de la infección es de 10 días o menos en la mayoría de los pacientes hombres; la sintomatología se presenta como una escasa emisión mucopurulenta, disuria y leve prurito o sensación de quemazón inmediatamente después de las relaciones sexuales (Petrin et al. 1998, Leherker et al. 2000).

Aunque *T. vaginalis* fue descubierto en 1836, no fue sino hasta 1916 cuando se identificó a este organismo como el agente causal de la tricomonosis y hasta 1957 se encontró que el 5-nitroimidazol (metronidazol) puede curar la infección causada por este parásito, por lo que se convirtió en el tratamiento primario para tricomonosis (Petrin et al. 1998). La activación del metronidazol ocurre en los hidrogenosomas de *T. vaginalis*, donde la reducción del grupo nitro de la droga por la enzima piruvato:ferredoxin oxidoreductasa resulta en la producción de intermediarios citotóxicos que reaccionan con el DNA del parásito, bloqueando la síntesis de ácidos nucleicos. Debido a la aparición de aislados resistentes a metronidazol, una segunda generación de nitroimidazoles se está usando; un ejemplo es el tinidazol que tiene una vida media en plasma dos veces mayor que el metronidazol, además de que se ha visto que es más efectivo en el tratamiento de los hombres con tricomonosis (Upcroft and Upcroft 2001, Petrin et al. 1998). Recientemente se ha utilizado otra generación de fármacos pertenecientes al grupo de los nitrotiazoles, entre ellos la nitazoxanida. Inicialmente este fármaco se empleaba contra céstodos, y posteriormente se demostró su actividad contra protozoarios, helmintos y bacterias anaerobias (Adagu et al. 2002, White, 2003). Debido a su amplio espectro y poca toxicidad se autorizó su salida al mercado y su empleo en nuestro país tanto en niños como en adultos (Eguiza, 2000).

La patogénesis de *T.vaginalis* está mediada por mecanismos de adhesión a las células del hospedero, mecanismos de evasión de la respuesta inmune y efectos citotóxicos sobre las células del epitelio vaginal. La adhesión de los tricomonátidos a las células epiteliales vaginales es un paso crítico en la patogénesis del parásito, ya que el ambiente vaginal cambia constantemente debido a la influencia hormonal. Al entrar en contacto con estas células el parásito sufre una transformación morfológica de una forma ovoide a una ameboide, formando proyecciones citoplasmáticas que dan lugar a filopodia y pseudopodia que interdigitan con las células blanco. El cambio de la forma de ovoide a la ameboide le confiere al parásito un aumento en la superficie de unión a la célula del

epitelio vaginal, lo cual puede ser necesario para establecerse de manera eficiente (Arroyo et al. 1993), y puede ser considerado un factor de virulencia (Fig. 5A y B).

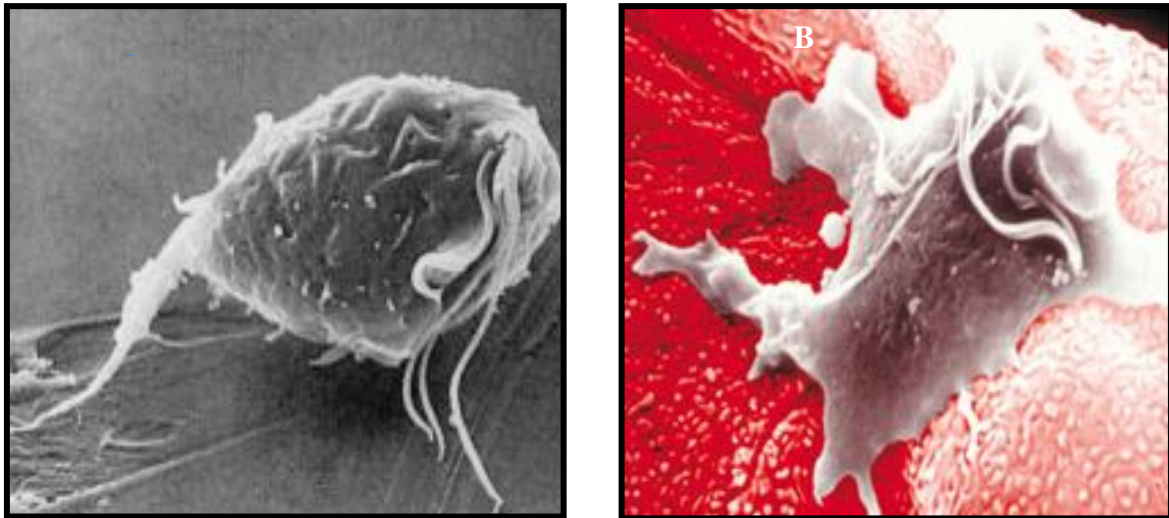


Fig. 5. Panel A: Típica forma ovoide del parásito *T. vaginalis*. Panel B: Transformación morfológica del parásito a una forma ameboide al estar en contacto con una célula de epitelio vaginal. Tomado de: Arroyo R. et al., 1993.

La investigación sobre los mecanismos de patogenicidad de *T. vaginalis* se ha enfocado a los eventos iniciales de la infección, en particular a la adherencia del parásito a las células del epitelio vaginal y a la producción de factores de virulencia extracelulares. La citoadherencia del parásito a las células del epitelio vaginal está mediada por adhesinas, proteínas que se encuentran en la superficie del parásito y aumentan su expresión en presencia de hierro (Alderete et al. 1988, Lehker et al. 2000). Este parásito tiene la capacidad de obtener nutrientes mediante la lisis de eritrocitos. Fiori y col identificaron cinco adhesinas en *T. vaginalis* con capacidad para unirse a la membrana del eritrocito, dos de ellas de manera específica (42 y 140 kD) mientras que las tres restantes participan en la adherencia a las células del hospedero (Harp et al. 2011, Petrin et al. 1998).

Recientemente en nuestro laboratorio se describió la presencia de serin-proteasas en la superficie de *T. vaginalis*, la función que desempeñan dichas proteasas aún no ha sido reportada (Hernández-Romano et al. 2010).

1.7 Diagnóstico

Existe un amplio rango de síntomas clásicos en la tricomonosis, pero estos no son exclusivos y pueden confundirse con otras enfermedades de transmisión sexual (ETS), por lo que utilizarlos para

el diagnóstico no es práctico (Fouts et al. 1980). Existen diferentes pruebas de laboratorio para diagnosticar de una manera confiable y certera la tricomonosis:

El método tradicional de diagnóstico es la observación de la movilidad de los protozoarios en frotis. Esta es una técnica de bajo costo, pero es poco sensible y la pérdida de motilidad de los protozoarios no la hacen muy confiable (Petrin et al. 1998). Los parásitos también pueden ser teñidos con naranja de acridina en preparaciones microscópicas, de esta manera es más sencilla su identificación (Afzan et al. 2010). La prueba del screening ginecológico conocida como “Papanicolau” también permite la identificación de estos parásitos. Por otra parte, el estándar de oro para el diagnóstico de tricomonosis es el cultivo del parásito en medio Diamond, esto debido a que tiene una fácil interpretación y se requieren pocos tricomonádidos para inocular el cultivo. Sin embargo, este ensayo posee varias limitantes, como es el tiempo para dar un diagnóstico (2 a 7 días) y este no está implementado en todos los laboratorios (Petrin et al. 1998).

Otras estrategias utilizadas para el diagnóstico se basan en el uso de anticuerpos, entre ellas están: aglutinación, fijación de complemento, hemaglutinación indirecta, anticuerpos fluorescentes, etc.

Se han desarrollado métodos más confiables y específicos basados en la identificación del DNA del parásito. El uso de técnicas de PCR posee una alta sensibilidad y permiten detectar células no viables y secuencias blanco en muestras clínicas que han sufrido fijación o degradación parcial. (Seña et al. 2007). Comercialmente se encuentra disponible el sistema AffirmVP que utiliza sondas sintéticas para la detección de *Gardnerella vaginalis* y *T. vaginalis* en una misma muestra vaginal (Briselden et al. 1994).

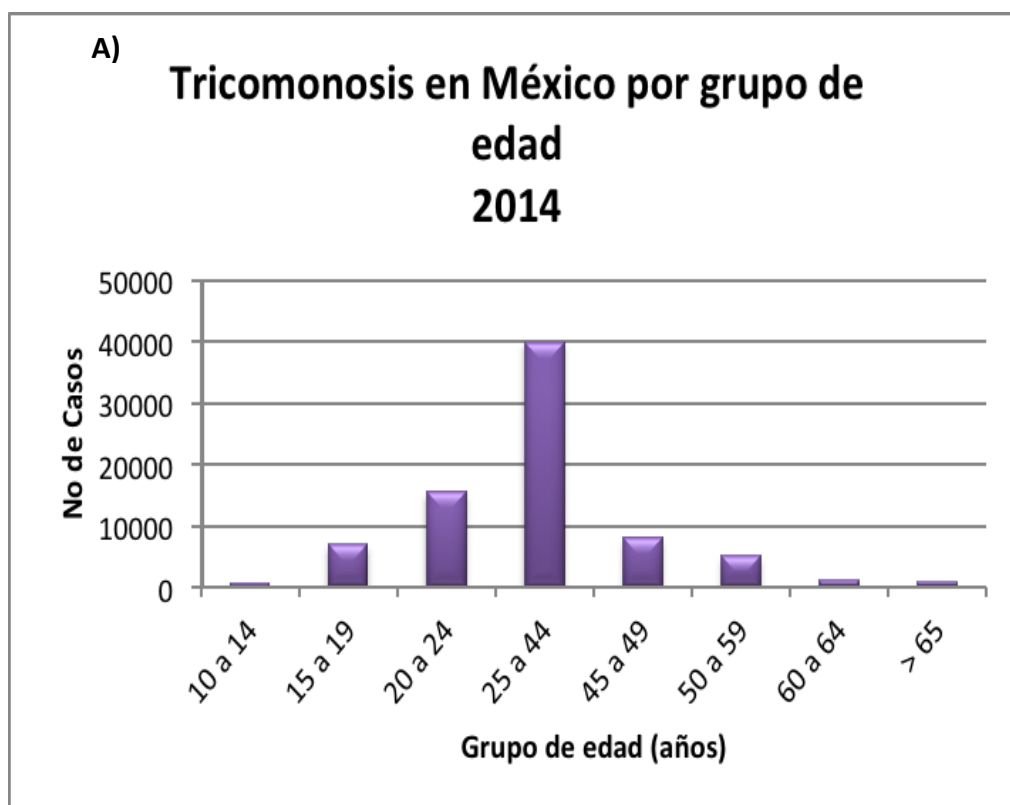
1.8 Epidemiología

La tricomonosis es una de las enfermedades de transmisión sexual del tipo no viral más comunes que infecta alrededor de 143 millones de personas en el mundo cada año, siendo la mayoría mujeres (92%) (WHO, 2016). Se han observado complicaciones en mujeres que sufren la infección por este parásito: la infección durante de la gestación se asocia con parto prematuro, bajo peso del producto al nacer, e incremento de la incidencia de endometritis post-parto. Otras complicaciones incluyen: endometritis, infertilidad, erosión cervical, incremento en la transmisión del VIH, cáncer cervical y enfermedad inflamatoria pélvica (Cotch et al. 1997; Sorvillo et al. 1998; Viikki et al. 2000, Johnston et al. 2008). Las complicaciones asociadas a la tricomonosis en hombres son prostatitis, balanopostitis, epididimitis, enfermedad uretral e infertilidad (Lehker et al. 2000).

En México, la tricomonosis se encuentra dentro de las veinte enfermedades más frecuentes a nivel nacional.

Los estados con mayor incidencia en el año 2015 fueron Tlaxcala, Puebla, Zacatecas, Veracruz, Tamaulipas y Oaxaca

(www.epidemiologia.salud.gob.mx/anuario/2015/incidencia/enfermedad_grupo_edad_entidad_federativa/026.pdf). En la Fig. 6A se muestra que el grupo de edad más afectado son las personas entre 25 y 44 años, probablemente a que este grupo es el que tiene una vida sexual activa, lo cual aumenta el riesgo de contagio. Hasta el año 2006 la cantidad de personas infectadas se mantuvo en un mismo nivel, y aunque esta enfermedad sigue siendo un problema de salud serio, se ha visto que en los últimos años ha disminuido considerablemente, para el 2015 se vio que la frecuencia de esta enfermedad disminuyó hasta aproximadamente un tercio de lo que se reportaba para el año 2002 (Fig. 6B). Estos resultados reflejan que día a día esta enfermedad esta siendo controlada de mejor manera, debido a que actualmente existe más información y educación sexual, acceso a servicios de salud y preocupación de la población por ser atendido.



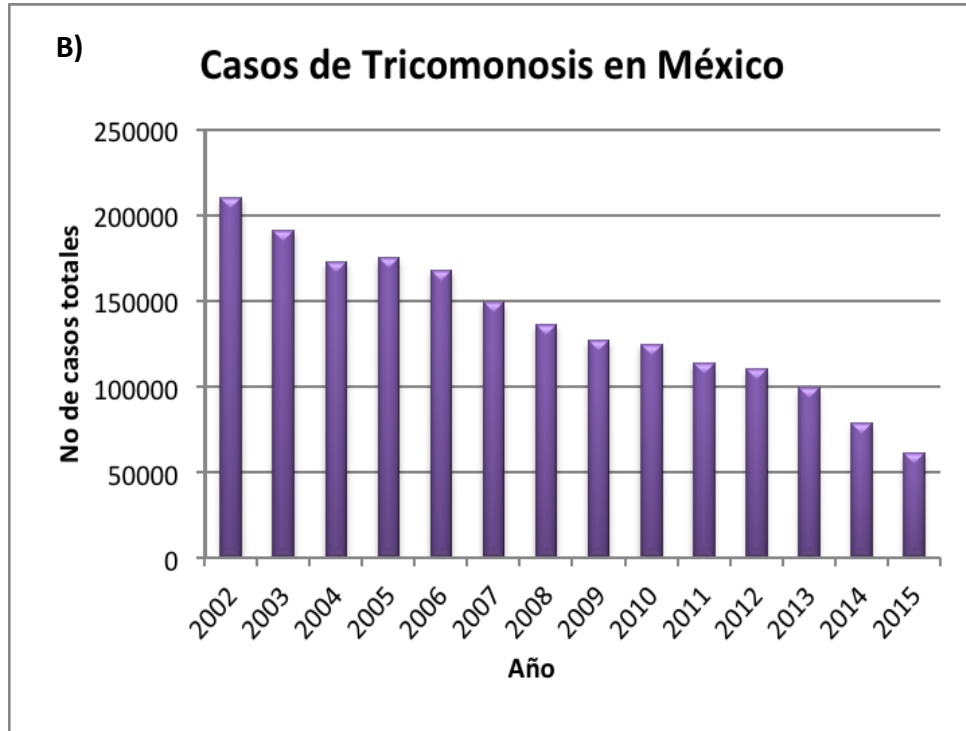


Fig. 6. Casos de tricomonosis reportados en México. A) Número de casos reportados por grupo de edad en el año 2014. B) Número de casos reportados totales en los últimos años. Fuente: Sistema Único de Información para la Vigilancia Epidemiológica/Dirección General de Epidemiología/SSA

1.9. Regulación de la expresión génica en *T. vaginalis*

De manera general, la transcripción en eucariontes es llevada a cabo por tres RNA polimerasas. La RNA polimerasa I (polI) es responsable de las síntesis de RNA ribosomal, polII transcribe los RNAs mensajeros que codifican proteínas y polIII transcribe el RNA ribosomal 5S, los tRNAs y otros RNAs pequeños. La actividad de las tres RNA polimerasas puede ser evaluada mediante su resistencia a α -amanitina. En *T. vaginalis* se ha observado que la polIII es resistente a altas concentraciones de esta toxina, con un 50% de inhibición observado a 250 $\mu\text{g}/\text{mL}$. Esto contrasta con lo reportado para otros eucariontes, donde esta polimerasa exhibe un 50% de inhibición a una concentración de 2-20 $\mu\text{g}/\text{mL}$ de α -amanitina. Se cree que la resistencia de la polIII a la α -amanitina en este protozooario se debe a que los aminoácidos involucrados en la sensibilidad a esta toxina no están conservados (Quon et al. 1994; Liston et al. 1999). En cuanto a la transcripción por polI y polIII, los únicos estudios en este parásito son los que se han llevado a cabo en nuestro laboratorio (Torres-Machorro et al 2010; Franco et al, 2012).

Respecto a la transcripción por polIII, la mayoría de los estudios que se han realizado en este parásito se han enfocado a la región promotora de la transcripción. Los promotores de *T. vaginalis* tienen una estructura distinta a la encontrada en la mayoría de los eucariontes superiores. Éstos tienen una estructura bipartita que consiste en elementos distales que regulan la transcripción y un promotor central (*core*) que contiene el sitio de inicio de la transcripción. En contraste, los promotores de *T. vaginalis* contienen un elemento altamente conservado llamado iniciador (Inr), que tiene la secuencia consenso TCA₊Py(T/A), muy parecida al Inr de metazoarios PyPyA₊N(T/A)PyPy. Este elemento es responsable de la selección del sitio de inicio de la transcripción (Liston y col., 1998) ya que es el sitio de unión para un factor de transcripción específico llamado IBP39 (initiator binding protein, 39 kDa) el cual no tiene homólogos en otras especies. Se ha comprobado que IBP39 interacciona con el dominio carboxilo terminal de la subunidad mayor de la polIII ayudando a que ésta sea reclutada al sitio de inicio de la transcripción (Schumacher et al. 2003, Lau et al. 2006). Recientemente se han descrito dos promotores centrales (*core*): el motivo 3 (M3) y el motivo 5 (M5) los cuales no parecen tener homólogos en otros promotores eucariontes. El motivo M5 tiene la secuencia consenso CC₊TTT, contiene un sitio inusual de inicio de la transcripción y se encuentra principalmente en genes que codifican para proteínas ribosomales. El motivo M3 tiene la secuencia consenso (A/G/T)(A/G)C(G/C)G(T/C)T(T/A/G) que se encuentra a una distancia fija río arriba del M5 o del Inr, y parece participar en la selección del sitio de inicio de la transcripción al actuar de manera sinérgica con los otros motivos, un papel parecido a lo que haría la caja TATA en eucariontes superiores (Smith et al. 2011, Gomez et al 2010).

Los genes de *T. vaginalis* están organizados en unidades de transcripción sencillas. Recientemente se ha comenzado a estudiar el procesamiento co-transcripcional que sufren los pre-mRNA y se ha encontrado la presencia de una metilguanósina (cap) (Simoes-Barbosa et al. 2008b). Anteriormente se pensaba que este parásito carecía de intrones, hasta que con la secuenciación del genoma se demostró la presencia de unos cuantos, así como de la maquinaria de “splicing” en este protozoario (Simoes-Barbosa et al. 2008). La posición de los intrones está conservada en genes ortólogos de diferentes organismos; se han encontrado sólo 65 intrones en los aproximadamente 60,000 genes que se presume posee este organismo (Carlton et al. 2007, Vanacova et al. 2005). Por otro lado, el proceso de poliadenilación de los mRNAs se comenzó a estudiar en el grupo de investigación de la Dra. Imelda López Villaseñor (Espinosa et al, 2002).

En el trabajo que aquí se presenta se logró la caracterización y demostración funcional de las secuencias que dirigen el corte y poliadenilación de los mRNAs en *T. vaginalis*.

2. ANTECEDENTES

2.1 Generalidades sobre el proceso de transcripción en eucariontes.

En eucariontes la transcripción de los genes nucleares es una tarea compartida entre tres RNA polimerasas: RNA Polimerasa I, II y III. La RNA Polimerasa I (polI) es la encargada de transcribir el RNA ribosomal (18S, 5.8S y 28S, nomenclatura canónica); la RNA Polimerasa III (polIII) sintetiza RNAs pequeños tales como tRNAs, 7SL RNA, U6 snRNA y el rRNA 5S; entre ambas polimerasas sintetizan más del 80% del RNA total en la célula (Paule y White 2000). Por otro lado, la RNA Polimerasa II (polII) transcribe una gran porción del genoma y sintetiza los mRNAs que son traducidos en proteínas por el ribosoma en el citoplasma; sintetiza también los RNAs pequeños nucleares o snRNAs (small nuclear RNAs).

La maquinaria de transcripción de polII es muy compleja con cerca de 60 polipéptidos, entre los que se encuentran factores generales de transcripción (GTFs) y un gran número de factores auxiliares involucrados en el reconocimiento de las secuencias promotoras, respuesta a factores regulatorios, y cambios conformacionales que se llevan a cabo durante la transcripción (Woychik et al. 2002, Hahn 2004). Los transcritos sintetizados por la polII sufren varias modificaciones antes de ser exportados del núcleo al citoplasma: capping (adición de un residuo de 7-metil guanosina en el extremo 5'), eliminación de intrones mediante splicing, y poliadenilación (adición de residuos de adeninas en el extremo 3'). Estas modificaciones se llevan a cabo de manera co-transcripcional de tal forma que son reguladas y coordinadas en tiempo y espacio por varios factores proteicos. Además estas modificaciones son un paso esencial para la maduración del mRNA.

La polII es un complejo multimérico que posee un dominio inusual en el extremo carboxilo terminal de la subunidad mayor, conocido como CTD (carboxi terminal domain), muy conservado entre los eucariontes. El CTD comprende secuencias repetidas en tándem del heptapéptido Tyr-Ser-Pro-Thr-Ser-Pro-Ser (YSPTSPS, secuencia consenso). Este dominio adopta una conformación helicoidal y parece tener una estructura flexible; la longitud del CTD es variable entre los diferentes eucariontes, por ejemplo en levadura se encuentran 26 repeticiones del heptapéptido, mientras que los mamíferos tienen 52 repeticiones. El papel que desempeña este dominio es de suma importancia ya que la delección de la mitad de este dominio es letal y se ha visto que participa en múltiples pasos de la transcripción incluyendo la activación, la iniciación, la elongación y la terminación. Cada heptapéptido posee varios sitios potenciales de fosforilación, pudiendo encontrarse 2 formas de

polimerasa polIIA (hipofosforilada) y polIIO (hiperfosforilada), esto debido a su estado de fosforilación (Howe et al 2002).

El cap (7-metil guanosina) se adiciona en el extremo 5' del transcrito una vez que la pol II ha transcrito los primeros 25-30 nucleótidos; esto no requiere de una secuencia específica en el RNA y es dependiente de la fosforilación de la serina 5 en el heptapéptido del CTD (Cho et al., 1997, Gu et al, 2005). La reacción de capping ocurre en tres pasos: 1) remoción del fosfato- γ del primer nucleótido del RNA mediante una 5'-trifosfatasa, 2) transferencia de un grupo guanilil mediante la guanililtransferasa (GTasa), la cual reacciona con un GTP para transferir un grupo GMP al RNA y formar un enlace 5'-5'(GpppN), y 3) La 7-metiltransferasa transfiere un grupo metilo de la S-adenosilmetionina a la posición N7 de la guanina (Howe et al , 2002). El cap es requerido para el splicing, exportación, traducción y la estabilidad del mRNA (Martínez-Rucobo et al, 2015).

Muchos genes de metazoarios contienen intrones que son eliminados mediante el proceso de splicing. El spliceosoma es la maquinaria molecular que lleva a cabo este proceso y está formado por RNA pequeños nucleares (snRNA, small nuclear RNAs) y por más de 100 diferentes proteínas no-spliceosomales (Kaida, 2016). Estos RNAs reconocen e interactúan con secuencias en *cis* en el transcrito de RNA (llamados sitios de splicing) ayudadas por proteínas de la familia SR (ricas en serina/arginina) para eliminar los intrones (Howe 2002, Martínez-Rucobo et al, 2015).

Como parte del procesamiento de los pre-mRNAs, en el extremo 3' de casi todos ellos se añade una cola de poliadenosina monofosfato [poli(A)]; la excepción son los mensajeros de las histonas canónicas, que no se encuentran poliadenilados. La adición de esta cola de poli(A) incrementa la estabilidad del mensajero (Beelman, 1995), regula la remoción del último intrón, participa en la exportación de los mRNAs del núcleo al citoplasma (Eckner, 1991) y también juega un papel importante en el inicio de la traducción (Howe 2002, Cramer et al. 2001, Fong y Bentley 2001). Se ha visto que en humanos la alteración del proceso de transcripción puede estar asociada al desarrollo de algunas enfermedades y al cáncer.

Por experimentos *in vitro* se ha demostrado que durante el proceso de transcripción los transcritos primarios de RNA se extienden más allá del sitio de corte y poliadenilación, lo cual manifiesta que para que se lleve a cabo el procesamiento del extremo 3' de los pre-mRNA eucariontes requiere de 2 pasos fundamentales para su maduración: un corte endonucleolítico sitio-específico, seguido por la adición de 150-200 residuos de adenilato por la poliA polimerasa, en una forma independiente de templado (Bentley, 2014). El proceso de corte y poliadenilación de los transcritos es esencial para la terminación de la transcripción por la pol II y a su vez ejerce influencia sobre el splicing del último

intrón del pre-mRNA (Wahle, 1999, McCracken S, 1997, Bentley, 2014) (Fig 7). En ensayos *in vitro* se han observado estos dos pasos de manera independiente, el corte del pre-mRNA genera dos “moléculas intermedias”: la enzima poliA polimerasa realiza la adición de la cola de poliA sobre la molécula que contiene las señales apropiadas (Sheets et al. 1990, Chen, 1995).

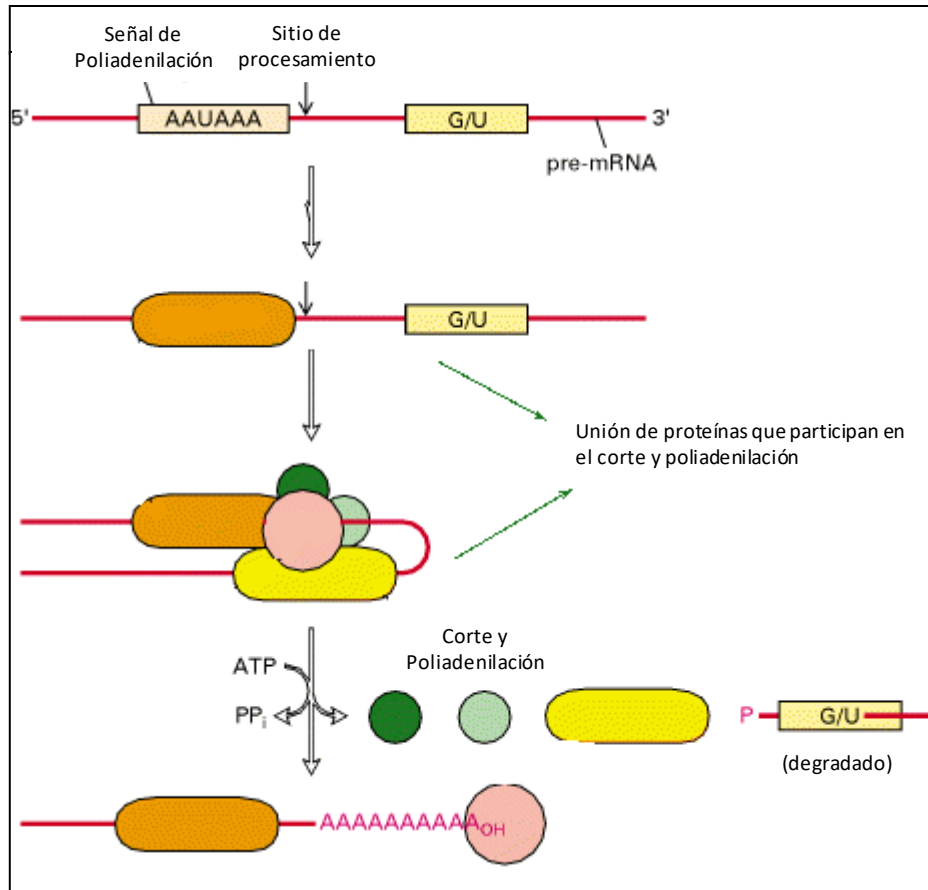


Figura 7: Pasos en el proceso de poliadenilación del mRNA en eucariotes superiores. 1) corte endonucleotídico, 2) poliadenilación. La maquinaria proteica involucrada en el procesamiento se indica de la siguiente manera: en naranja el CSPF-160, en amarillo el CstF-64, en rosa PABP y en verde el resto de la maquinaria de poliadenilación.

El ensamblaje del complejo de procesamiento en el extremo 3' de los pre-mRNA es dirigido por varios elementos en *cis* en el RNA a los que se les conoce como señales de poliadenilación y que son reconocidos por los factores de poliadenilación. Más adelante se describen las señales de poliadenilación de diferentes organismos de manera puntualizada. Además, las señales de poliadenilación son elementos esenciales para la terminación de la transcripción por la RNA pol II (Dye y col., 2001), ya que intervienen en la eficiencia de la terminación (Yonaha y col., 2000).

La terminación de la transcripción es llevada a cabo en un paso posterior a la realización del corte endonucleotídico: el RNA que queda río abajo al sitio de corte, al contener un extremo 5' OH no protegido, es degradado por endonucleasas que, al alcanzar a la RNA pol II, la desestabilizan y de esta manera termina la transcripción (Loya et al, 2016).

Por otro lado, la degradación de los RNA mensajeros es llevada a cabo por endonucleasas, e involucra la eliminación tanto del cap como de la cola de poli(A) de los mensajeros. La velocidad de deadenilación tiene una gran influencia en la velocidad de decaimiento de un RNA mensajero, lo cual se reflejará directamente en la vida media de los mRNA. El tamaño de la cola de poliA varía entre especies, en mamíferos es aproximadamente 250 residuos de adenilato, en levaduras es de 100 residuos, mientras que en insectos es de 150 residuos (Lodis et al, 2005). Esta región es un sitio de unión para proteínas llamadas PABPs (poly A binding proteins) que pueden participar en la estabilidad de los mensajeros (Meyer et al. 2004).

2.2 Elementos que dirigen la poliadenilación en mamíferos

Se han identificado varias secuencias involucradas en dirigir el sitio de poliadenilación. Los siguientes son los elementos canónicos identificados en mamíferos:

1. La señal de poliadenilación (PAS, por sus siglas en inglés) o motivo AAUAAA (secuencia canónica), se encuentra localizada de 10-30 nt río arriba del sitio de corte del pre-mRNA. Esta secuencia fue revelada inicialmente al comparar las secuencias de nucleótidos que preceden al sitio en el que se poliadenilan varios mRNAs. Este motivo es uno de las más conservados y la variante más frecuente es AUUAAA (presente en aproximadamente el 10% de los mRNA), la cual es comparable en función a la de la secuencia canónica. La eliminación de esta secuencia suprime el procesamiento y poliadenilación de los pre-mRNA, y mutaciones en este motivo disminuyen la eficiencia del procesamiento en el extremo 3' (Sheets M, 1990).

2. Elementos río abajo (DSE, downstream elements): esta secuencia se localiza dentro de una región comprendida 30 nucleótidos río abajo del sitio de corte. Es un elemento pobremente conservado y se han descrito dos tipos principales: elementos ricos en Us y elementos ricos en GU. El elemento rico en Us es una corrida corta de Us, mientras que el elemento rico en GU tiene la secuencia consenso YGUGUUY (Y = pirimidina). La señal de poliadenilación se corresponde con un DSE funcionando de forma sinérgica. Sin embargo, la distancia entre el DSE y el sitio de corte puede afectar tanto la posición como la eficiencia del corte (Chou Z, 1994, Gil, 1987, Proudfoot 2000).

3. Sitio de corte: La selección del sitio de corte es determinada principalmente por la distancia entre la señal de poliadenilación (AAUAAA) y el DSE(s). La secuencia de corte no está muy conservada, aunque en el sitio de corte se encuentra un residuo de adenosina en el 70% de mRNAs de vertebrados. En un estudio se demostró que el orden de preferencia para el nucleótido en el sitio de corte es el siguiente: A>U>C>>G (Chen et al, 1995). Por otra parte, es más frecuente que el penúltimo nucleótido con relación al sitio de corte sea un residuo de citosina (en 59% de todos los genes analizados). Por consiguiente, el dinucleótido CA es el motivo consenso que define el sitio de corte y poliadenilación en muchos genes de vertebrados (Chen, 1995, Mandel 2008).

4. Secuencias auxiliares. Se han encontrado algunos otros elementos que pueden modular la eficiencia del procesamiento del extremo 3' en una manera positiva o negativa. Existe una clase de secuencias "enhancer" localizadas río arriba del elemento AAUAAA (llamadas USE), y río abajo del DSE (llamadas auxDSE) que se han encontrado principalmente en virus. Estos USEs y auxDSEs son con frecuencia ricos en U, pero una secuencia consenso aún no se ha propuesto (Zao, 1999, Mandel 2008, Hu, 2005, Bentley 2014).

Ahora bien, cerca del 20% de los sitios de poliadenilación no contienen al hexanucleótido AAUAAA, lo que indica que podría existir otro mecanismo por el que señales no canónicas son reconocidas por la maquinaria de poliadenilación, o incluso pudieran existir otras variantes de factores involucrados en el proceso de poliadenilación (Bentley et al, 2014). Recientemente se ha encontrado que la secuencia UGUAN puede ser un elemento clave en dirigir el procesamiento y poliadenilación de los pre-mRNA de manera independiente del AAUAAA. También se ha encontrado en algunos casos que solo es necesario un elemento rico en A y un DSE "potente" (Mc Cracken, et al 2009).

2.3 Señales de poliadenilación en levaduras.

Las señales de poliadenilación en levadura son mucho menos conservadas que las señales en eucariontes superiores. Al menos tres elementos deben estar presentes para la poliadenilación de los mRNAs de levadura.

1. Elemento de eficiencia: Se localiza a una distancia variable río arriba del sitio de corte y con frecuencia contienen alternados dinucleótidos UA o tramos ricos en U. Se ha encontrado que el hexanucleótido UAYRUA, con la secuencia UAUUAU, es el más eficiente para dirigir el procesamiento del extremo 3' del mRNA (Mandel., 2008).

2. Elemento de posicionamiento: este dirige el corte a un sitio aproximadamente 20 nucleótidos río abajo de la secuencia (UUAAGAAC). Adicionalmente otras secuencias ricas en A han sido caracterizadas como elementos de posicionamiento, siendo los motivos AAUAAA y AAAAAA los más eficientes.

3. El sitio de corte y poliadenilación mapeado en varios genes de levadura muestra que ocurre más frecuentemente en la secuencia PyA_n. Se han encontrado frecuentemente motivos ricos en U antes y después del sitio de poly(A), especialmente en genes que son transcritos con baja eficiencia o bien cuando se requieren elementos de posicionamiento (Mandel, 2008).

2.4 Señales de poliadenilación en plantas

En plantas se requieren al menos tres señales para la poliadenilación de los mRNAs: el elemento cercano-río arriba (NUE), el elemento lejano-río arriba (FUE), y el sitio de corte. El NUE está localizado a una distancia de 10 a 30 nucleótidos río arriba del sitio de corte, siendo un elemento rico en U y A. El FUE es un elemento rico en U y se encuentra aproximadamente 100 nucleótidos río arriba del sitio de corte. En cuanto al corte frecuentemente ocurre en el dinucleótido PyA. Sin embargo, pueden existir múltiples sitios de corte en algunos genes, y el uso de un sitio particular de corte es determinado predominantemente por la posición del NUE (Mandel, 2008).

2.5 Factores proteicos involucrados en el proceso de poliadenilación.

Hace más de cuatro décadas se describió un componente de la maquinaria de corte y poliadenilación presente en extractos nucleares de timo de ternera, con actividad para sintetizar poli(A) a partir de ATP, pero no a partir de ADP (Edmonds, 1992). Actualmente más de 14 proteínas han sido identificadas como parte de la maquinaria de procesamiento del extremo 3' en mamíferos, y más de 20 proteínas en levadura, por lo que se estima que un complejo de cerca de un megadalton es requerido para el procesamiento de los pre-mRNA (Mandel, 2008). Esta maquinaria para el procesamiento del extremo 3' en mamíferos contiene varios sub-complejos, entre los que se encuentran los siguientes: (Figura 8)

1) El factor de corte y poliadenilación o CPSF (cleavage and polyadenylation specificity factor). En mamíferos está conformado por 5 subunidades: CPSF-160, CPSF-30, CPSF-73, CPSF-100 y hFip1.

a. CPSF-160 es la subunidad más grande del CPSF y es una de las más estudiadas, ya que interactúa directamente con el pre-mRNA y dirige tanto el corte como la poliadenilación. Este factor asemeja a la estructura de una propela, se une directamente a la señal de poliadenilación AAUAAA con gran afinidad y tiene un alto porcentaje de similitud con el homólogo de levadura (Cft1p/Yhh1p) (51% de similitud). En levadura se ha visto que interactúa con la señal de

poliadenilación en la región de la segunda propela. El CPSF se asocia con factores de inicio de la traducción como TFIID y durante la etapa de elongación de la transcripción se asocia al CTD de la pol II. Se ha visto también que este factor desempeña un papel en la terminación de la transcripción y participa en la poliadenilación citoplasmática en *Xenopus* (McCracken, 1997, Zhao, 1999, Mandel, 2008).

b. CPSF-73 (su homólogo en levadura es Brr5/Ysh1p). Es un factor esencial para la viabilidad celular. Recientemente se encontró que contiene un dominio metalo- β -lactamasa en su extremo N-terminal siendo un miembro de la familia β -CASP. Se ha demostrado que se une directamente al sitio de corte de una manera dependiente del motivo AAUAAA, sugiriendo que este factor es la endoribonucleasa involucrada en el procesamiento de los pre-mRNA. Se ha visto que tiene la misma función en el procesamiento de los RNA mensajeros de histonas (Mandel, 2008, Callebaut, 2002).

c. CPSF-30. Tiene actividad de unión al RNA en secuencias ricas en poli(U), mediante dominios de dedos de zinc.

d. hFip1. El dominio carboxi-terminal de hFip1, rico en arginina, se une a secuencias ricas en U. Por lo tanto este factor junto con el CPSF-30 están involucrados en reconocer los USEs de tal forma que estabilizan la interacción entre el CPSF-160 y el AAUAAA.

e. CPSF 100. No se conoce claramente el papel que desempeña a pesar de ser muy parecido a CPSF-73 (Colgan, 1997).

2) El factor de estimulación de corte o CstF (cleavage stimulation factor) reconoce al elemento río abajo (DSE), y se requiere específicamente para el corte del transcrito, pero no para la poliadenilación. Este factor comprende tres subunidades de 64 (o su parólogo CstF64 τ), 77 y 50 kDa, y funciona como dímero en el procesamiento del extremo 3'. Contiene un dominio de reconocimiento al RNA (RRM) en su extremo N-terminal, el cual se une con gran afinidad a secuencias ricas en GU. Por otra parte, CstF-77 es el andamio que mantiene la interacción entre CstF50 y CstF64. CstF50 se une al CTD de la polII por lo que se ha propuesto que ayuda al reclutamiento del CstF al complejo de elongación (MacDonald, 1994, Dai, 2011).

3) CFIm: Se requiere para el corte del pre-mRNA, y consiste de CFIm25, CFIm68 y CFIm59; se ha visto que funciona como dímero. CFIm25 tiene un dominio Nudix y se une específicamente al motivo UGUAN, lo cual ayudaría a reclutar la maquinaria de poliadenilación en señales no canónicas. CFIm68 y CFIm59 tienen dominios RRM, pero no se unen al RNA sin la presencia de CFIm25, y el papel que desempeñan es poco claro (Colgan, 1997).

4) CFII_m es un factor que también participa en el corte del pre-mRNA. La PoliA polimerasa (PAP) y la proteína de unión al poliA (PABP) son requeridas para el corte y poliadenilación, esta última controla la longitud de la cola de poliA.

5) Simplekina: es una proteína con función de andamiaje que interacciona con varias proteínas del complejo de poliadenilación tanto en el núcleo como en el citoplasma.

En cuanto a la dinámica de ensamblaje del complejo de poliadenilación, ésta se lleva a cabo de la siguiente manera: *in vivo* los factores de poliadenilación están asociados a la maquinaria de transcripción a través del CTD de pol II. El CPSF es reclutado al complejo de preiniciación a través del TFIID, y junto con CsfF y CFIm se encuentran asociados al CTD de la polII durante la elongación. Cuando la polII transcribe las señales de poliadenilación, ésta pausa su actividad y quizás durante este tiempo se ensamble completamente la maquinaria de procesamiento (Mandel, 2008).

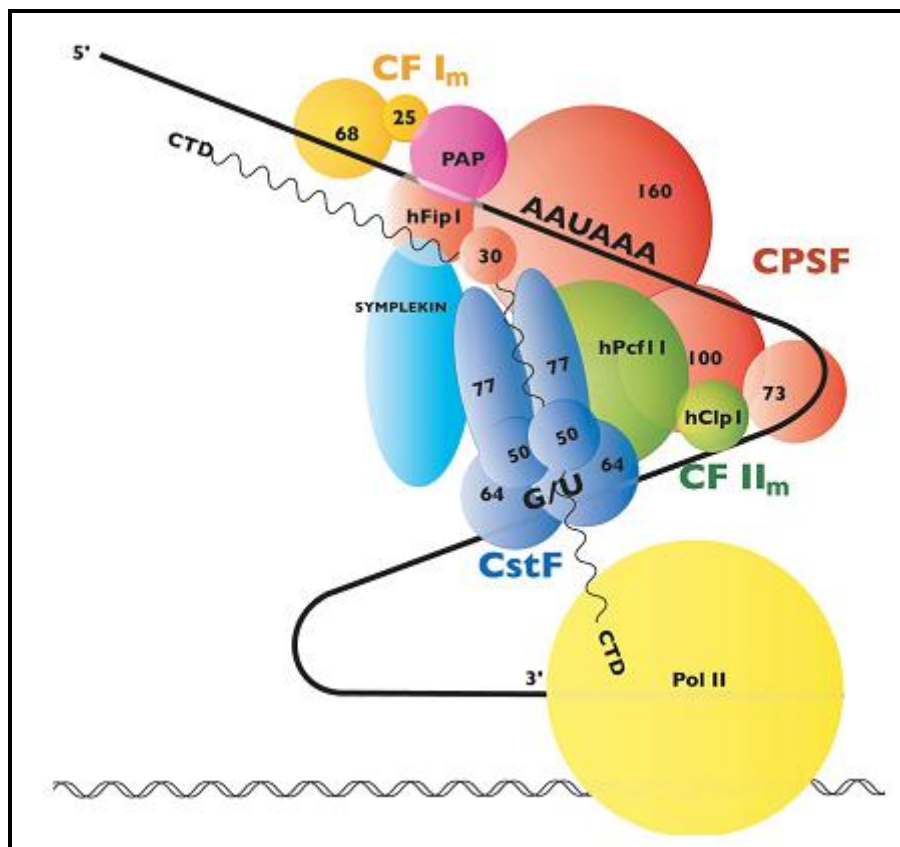


Fig. 8 Modelo de la maquinaria de procesamiento de los pre-mRNA de mamíferos. Tomado de Mandel, 2008.

2.6 Señales de poliadenilación en protozoarios

Las señales de poliadenilación más estudiadas han sido las de mamíferos y levaduras, y poco es lo que se sabe acerca de los elementos en *cis* en el RNA requeridos para que se lleve a cabo el procesamiento y poliadenilación de los pre-mRNA en protozoarios. La mayor parte de la información que se tiene son datos obtenidos por estudios bioinformáticos. A continuación, se mencionan las señales de poliadenilación descritas para algunos protozoarios.

En *Giardia spp.* la secuencia AGUAAA se ha propuesto como señal de poliadenilación y está localizada de 20-35 nt río arriba del sitio de adición del poliA y entre 6 y 19 nucleótidos río abajo del codón de término de la traducción, dando 3'UTR muy cortos. Se ha observado que en los transcritos de ENC6, el cual codifica para una proteína involucrada en la enquistación, se emplean diferencialmente las señales de poliadenilación dando transcritos inusualmente largos. El motivo AGUAAA se localiza a 87 y a 107 nucleótidos río abajo del codón de término de la traducción, utilizándose con mayor frecuencia la segunda señal de poliadenilación, lo cual sugiere que el 3'UTR puede tener un papel en la regulación génica, esto debido al proceso de diferenciación que sufre este organismo (Que X, 1996).

En el diplomonádido *Spironucleus salmonicida*, se analizaron 134 ESTs y se observó que los 3'UTR de este organismo son muy cortos entre 9 y 14 nt de longitud. El único elemento común que se encontró fue el codón de término de la traducción con la secuencia TRA (UAA, UGA), por lo que se ha propuesto que ésta podría ser la señal de poliadenilación en este organismo (Andersson JO, 2007).

En *Entamoeba histolytica* se ha propuesto a la señal consenso UA(A/U)UU como la señal de poliadenilación, localizada en la posición -25 a -15 relativa al sitio de corte y poliadenilación; esta secuencia puede incluir el codón de término de la traducción. Además, mediante un análisis *in silico* y de la expresión de mRNA se identificaron 16 probables factores de corte y poliadenilación, entre ellos 4 factores que no presentan homología con los factores descritos en mamíferos y levaduras y que pudieran contribuir a una eficiente reacción de corte y poliadenilación (López-Camarillo y col., 2005, Morrison, 2007).

En *Dictyostelium discoideum*, mediante análisis bioinformáticos se propuso al motivo AAUAAA (y variantes) como la señal de poliadenilación localizada 25 nucleótidos río arriba del sitio de corte, algo inusual es que algunos de estos motivos pueden encontrarse dentro de la región codificadora. El sitio de corte es preferentemente una U, y las secuencias que rodean al sitio de corte son ricas en A (Rivero F, 2002).

En *Plasmodium* las señales y mecanismos de procesamiento del extremo 3' están poco definidas. Se sabe que un número importante de mensajeros de *Plasmodium* no están poliadenilados, mientras que para otros transcritos se han reportado múltiples sitios de poliadenilación. Algunos transcritos de *Plasmodium* se poliadenilan río abajo del un pentámero AAUAA y río arriba de secuencias ricas en G/U (Cann, H 2004).

En los tripanosomátidos (*Trypanosoma cruzi*, *Trypanosoma brucei* y *Leishmania*) el proceso de poliadenilación no se lleva a cabo por el reconocimiento de secuencias consenso como en otros eucariontes, sino por un proceso acoplado al trans-splicing, donde el sitio aceptor de splicing (AG) se encuentra río abajo del sitio de poliadenilación (Texeira, S.M.R., 1998., Ullu E., y col., 1996). La presencia de una secuencia rica en polipirimidinas es crucial, ya que el dinucleótido AG es usado como sitio aceptor de splicing sólo si se sitúa río abajo de la secuencia rica en polipirimidinas (Schurch, N. y col., 1994). Por lo tanto, la secuencia rica en polipirimidinas dirige la poliadenilación del transcrito río arriba y el trans-splicing del transcrito río abajo mediante el cual se adiciona el SL (splice leader).

A continuación se muestra la comparación de las señales de poliadenilación descritas en protozoarios (Tabla 1).

Tabla 1. Señales de poliadenilación en algunos protistas.

Especie	Motivo	Otros motivos
<i>Giardia lamblia</i>	AGUAAA	---
<i>S. salmonicida</i>	URA	----
<i>Entamoeba histolytica</i>	UA(A/U)UU	----
<i>Dictyostelium discoideum</i>	AAUAAA	Sitio de corte U, regiones ricas en As
<i>Trypanosoma cruzi</i>	Ausente, proceso acoplado a trans-splicing	----
<i>Plasmodium falciparum</i>	No todos los mRNA están poliadenilados, algunos usan AAUAA	Regiones ricas en G/U

Procesamiento y poliadenilación del mRNA en *Trichomonas vaginalis*

Uno de los proyectos que se realizan en el laboratorio de la Dra. Imelda López Villaseñor es el estudio de las señales involucradas en el procesamiento de los mRNA de *T. vaginalis*. Este trabajo se inició con la tesis doctoral de la Dra. Norma Espinosa, en donde se compararon las regiones 3' no traducidas de 13 clonas de cDNA de distintos genes con su respectiva secuencia genómica. Se encontró que las regiones 3' UTR de los mRNAs son muy cortas, y no se encontró la señal de poliadenilación consenso AAUAAA. Al alinear estas secuencias, se encontró que el único motivo

consenso es el tetranucleótido UAAA (que generalmente contiene el codón de término de la traducción UAA); esta señal se localiza entre 11 y 30 nucleótidos río arriba del sitio de corte, por lo que se propuso como la posible señal de poliadenilación en *T. vaginalis*. También se encontró que el sitio de corte podría estar definido por el motivo $Y\downarrow(A)_{2-5}UU$; adicionalmente, se encontraron secuencias ricas en U río abajo del sitio de corte y poliadenilación. En la Fig. 9 se hace una comparación de las señales propuestas para la poliadenilación en *T. vaginalis* con las señales descritas para este proceso en animales y levaduras (Espinosa y col., 2002).

Es importante resaltar que la señal de poliadenilación propuesta para *T. vaginalis* difiere de las señales canónicas, principalmente en dos aspectos: el primero es su longitud, ya que se proponen cuatro nucleótidos en lugar de los seis nucleótidos de la secuencia canónica. En segundo lugar, la señal propuesta en *T. vaginalis* contiene al codón de término de la transcripción UAA.

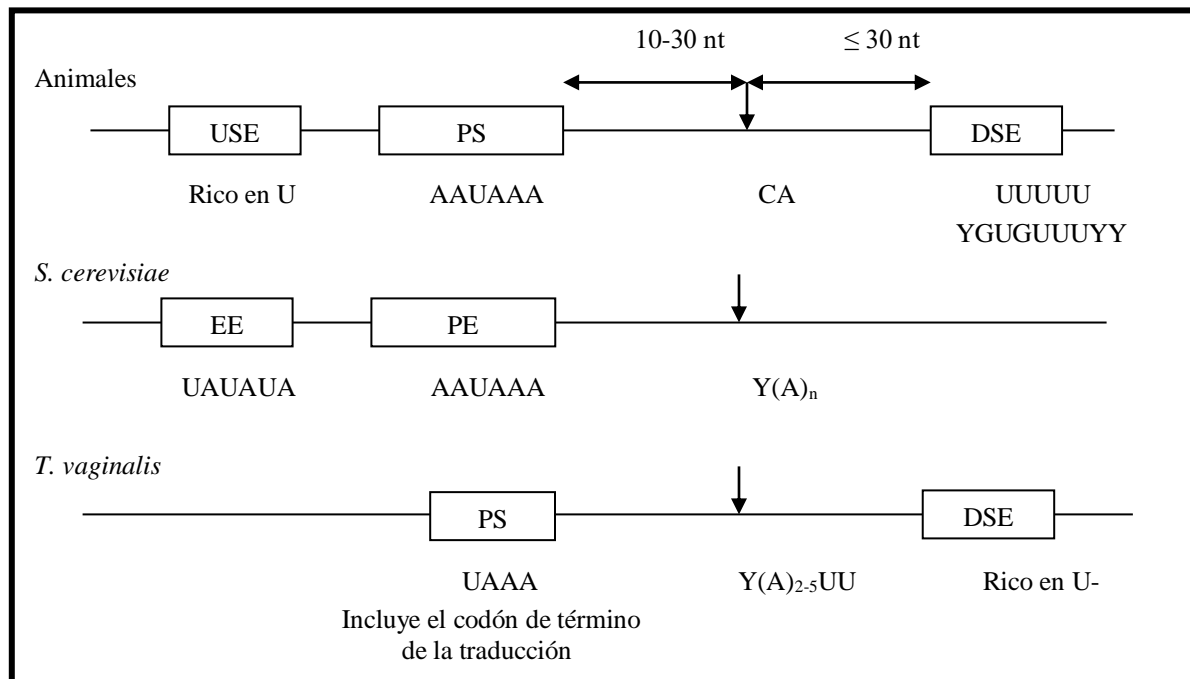


Fig. 9 Comparación de las señales de poliadenilación de animales, *S. cerevisiae* y las propuestas para *T. vaginalis*. USE, elemento río arriba; DSE, elemento río abajo; EE, elemento de eficiencia; PE, elemento de posicionamiento; PS, señal de poliadenilación y nt, nucleótido.

Como ya se señaló, las señales descritas para la poliadenilación de mensajeros en *T. vaginalis* fueron propuestas inicialmente de manera bioinformática con base en el alineamiento de varios cDNAs con sus secuencias genómicas. Sin embargo, hacía falta la demostración funcional de cada uno de estos elementos. Durante el trabajo de tesis de la M.C. María Guadalupe Barrera se comenzó este abordaje experimental mediante la construcción de vectores de expresión transitorios para *T.*

vaginalis, utilizando genes reporteros en donde se pudieran modificar las señales propuestas para la poliadenilación y así evaluar su efecto en los transcritos y en los productos finales.

Un gen reportero es una herramienta que permite cuantificar de manera rápida y sensible los cambios transcripcionales. Este debe poseer varias características: el producto del gen reportero no debe encontrarse en la célula huésped, el análisis del producto debe hacerse de manera rápida y sencilla, los resultados deben ser reproducibles y el producto no debe afectar a la célula.

Uno de los reporteros más comúnmente utilizados es el gen *cat* que codifica para la enzima cloranfenicol acetiltransferasa (CAT), una enzima procariótica muy estable. La actividad de la enzima CAT puede evaluarse *in vitro* ya que modifica covalentemente al cloranfenicol, transfiriéndole un grupo acetilo del acetilCoA en la posición C3 donde se encuentra el grupo hidroxilo primario del cloranfenicol. El producto de esta reacción, 3-acetil cloranfenicol, sufre un rearrreglo en el cual el grupo acetilo se transfiere al C1 formando el 1-acetil cloranfenicol; este rearrreglo permite que se lleve a cabo un segundo ciclo de acetilación al C3, dando como resultado el 1,3-diacetil cloranfenicol. La actividad de CAT puede calcularse con base en el porcentaje de cloranfenicol que es convertido a productos mono y diacetilados, medido a través de la incorporación de una marca radiactiva, y cuantificado mediante un contador de líquido de centelleo.

Existen otros genes reporteros que se utilizan frecuentemente, entre los cuales están: la luciferasa, en la cual se mide su capacidad de actuar sobre la luciferina y producir un destello de luz; el gen *lacZ*, cuya enzima reacciona con el sustrato X-gal produciendo un color azul oscuro; y la proteína verde fluorescente (GFP), la cual produce un color verde en aquellas células que la expresan.

2.7 Vectores para la expresión de genes reporteros en *T. vaginalis*

En el laboratorio de la Dra. López Villaseñor se han construido vectores para la expresión de genes reporteros en *T. vaginalis* (trabajo desarrollado por L.I.B.B. Jessica Guzmán). Uno de estos vectores es el pAC22CAT-1. Este plásmido contiene al gen *cat* (que codifica para la enzima cloranfenicol acetil transferasa, CAT) flanqueado por las regiones 5' y 3' intergénicas de un gen de actina propio del parásito. Por mutagénesis dirigida se cambió el codón de término de la traducción del gen *cat* (TAA → TAG) y la región 3' UTR del gen de actina parental se conservó de manera íntegra, incluyendo el codón de término de la traducción UAA que forma parte de la señal propuesta de poliadenilación (Guadalupe Barrera, tesis de Maestría, 2007). Este vector (llamado ámbar) y sus derivados fueron utilizados a lo largo del presente trabajo de tesis (Fig 10).

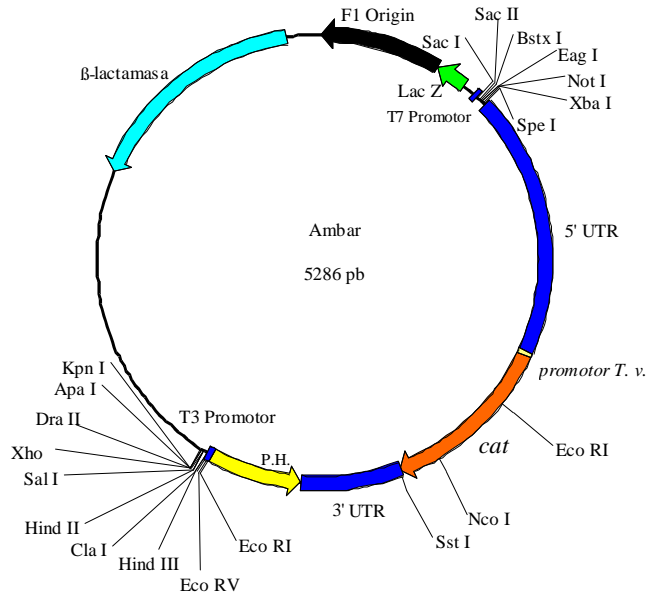


Fig. 10 Vector de expresión Ámbar que contiene al gen *cat* flanqueado por las regiones 5' y 3' de actina de *T. vaginalis*.

En la Fig. 11 se muestra parte de la secuencia nucleotídica del vector Ámbar donde se observa al gen reportero *cat* (en azul), flanqueado por las regiones intergénicas de actina. Se resaltan las señales propuestas de poliadenilación, que fueron modificadas en este trabajo para probar su funcionalidad.

```

5' -AC TAGTGT TTTT ATATATAATT ATTCGACGG AGCTTAGCAA
ATGAAGACCG TAGCGAAATC AATCTATCAG AATATGTCAG GCACTGGCTA TTCTCATTAT
TTTTGTGATG TGTCGGCATA TGCTAAAAAA GTTAAATAAT TGTGATTGGA CACTTTTAGA
TGTTCAATAT TGATCTAAAC ATAAAAATCAA TATATAACAG TAACAATCAA ATAGTGT TTA
TGAAAAGATA GACGAACTAT CTTTTGCGTT ACTATAATTC GAATGATATC ATATTTTTTTA
ACTTTTTTTAG CGAAGACGTA GTTTTTTCGAC ATAATATTTT TGAATAACGA TATAGTTTCT
CCAAAAATTT ATGTTTTTGGG GAACTTAAAC GAATGTCAGA ATATTGGAAA AAAAATTATG
AGGACATTTT CGTAATCAAA TAACAGATTT GTAACCTATA TGTACGTTTA ATTCTGAAAT
CAAGCGTAAT TTCGAAATGG GTAAGATAAT ATCGTGATGT CTGATGTGAA ATTAGTAAAA
AGAATAAAAA TTTTATGAAT GGTTCCTTGA TATGAATACC ATTCATGATA ATCGTTATAA
TTTATAAAAAC ATTTTGTAGT TCAGTGCAAT CAATATTTGT GAATTGTCAA ATTTATTCTC
AAAATTTTGA GAGCAATTGA TTGGAATAAAA GATATAGAAA ATATATTTTAT TGATAGAAAAT
AGCAAAATTG GCCAGATATA ATATCGGAAA TAAAATGACA TCACTTCACA TTGTCCCTTT
GAAAATTTGA TCTTTGAGGA AAATATTCCG GAAACCACAT ATATATGTAT AATATCGAAA
GATTAAAAAA ATTAAGAAGG GACTTAGATG TCCTGGTTAT AAATTTTGAG GACTCTAACC
GTTGTCTAAA GAAGTGAAAT GATTGGAATG TTAACCTCAT TTCATTTTCT GCCATGAACA

```

```

AAAAAATCAC TGGATATAACC ACCGTTGATA TATCCCAATC GCATCGTAAA GAACATTTTG
AGGCATTTCA GTCAGTTGCT CAATGTACCT ATAACCAGAC CGTTCAGCTG GATATTACGG
CCTTTTTAAA GACCGTAAAG AAAAATAAGC ACAAGTTTTA TCCGGCCTTT ATTCACATTC
TTGCCCGCCT GATGAATGCT CATCCGGAAT TCCGTATGGC AATGAAAGAC GGTGAGCTGG
TGATATGGGA TAGTGTTTAC CCTTGTTACA CCGTTTTCCA TGAGCAAAC TAAACGTTTT
CATCGCTCTG GAGTGAATAC CACGACGATT TCCGGCAGTT TCTACACATA TATTTCGCAAG
ATGTGGCGTG TTACGGTGAA AACCTGGCCT ATTTCCCTAA AGGGTTTTATT GAGAATATGT
TTTTTCGTCTC AGCCAATCCC TGGGTGAGTT TCACCAGTTT TGATTTAAAC GTGGCCAATA
TGGACAACCTT CTTGCCCCC GTTTTACCA TGGGCAAATA TTATACGCAA GGCGACAAGG
TGCTGATGCC GCTGGCGATT CAGGTTTCATC ATGCCGCTCTG TGATGGCTTC CATGTCGGCA
GAATGCTTAA TGAATTACAA CAGTACTGCG ATGAGTGGCA GGGCGGGGCG TAGGAGCTCA
CTTAAACAAT TTCGGTATAA TATATTTCAA AATTAAGTAA CCTTTTTTCT TTAGTGAACC
TATTTTGGAA ATCACTTCAT TAATATTGTG TTTTAATGGT TTATATGTGA CCAATACGAT
GTACCCTGAA CATTCTTTT CCAATATGAT TTTTACCTTT CACCAAATAC AAGTATCATG
CTAATTAAG TGCCTGATTT GCATTGTATT CTTTAACTAT ACATAATTGA ATCAATCAAT
AACACTTGAT AATTGTGTAT TATTCCTTGT TGATTTAAAA GAATATCTTT TGATATTTTG
AAGTAAATTA AATTATTCTT CATCAAATCA TAGAGATAAA CTAAAAACAAG ACATTTATAG
TAAGTTAATC AATTTATTAA TAAAAATACA AATTTAAATA ATACAAAATA CTAATTGAAT
TACTCGTTGA CCATTCCAAT CTTCTTTAGA ACATTGAGAA GTTGAGGATC AATTCCTTCA
TCCAAGTCTT CGTAGTGAGC AATCTGTGAG TATGAGCTAG AGTCGTCAGA TCTATTGTCA
TCTCTATCCA AAGGACTTGA AATTGGTCTA GGAATTGCAG AAACATCAAT AGGAACGATG
TTTGCATTTT CGCATTCTT CATCATATAA TCATTTTCGT GGCTCAATCT ATCATCGGTT
ATCTCAACGA ATGACAGCAA TCTCTTGAC TCCTCGTACG ATCTCTTGTA GCTCAACCTG
TTGTGCTGGT CCTTGACTTT GTCGCTGTTT TCGTTTGAGC TGAATTC-3'

```

Figura 11. Secuencia parcial del vector Ámbar. En rojo se muestran las regiones 5' y 3' intergénicas de un gen de actina de *T. vaginalis*, en azul se muestra la región codificadora del gen *cat*, en negro la secuencia linker entre *cat* y la región 3' intergénica de actina, en amarillo se señala el codón de término de la traducción (UAG) de *cat*, en un cuadro negro se resalta la señal propuesta de poliadenilación en *T. vaginalis* (UAAA), con doble subrayado la señal propuesta para el procesamiento de los transcritos (Y↓AAAAUU), y en cursivas la región rica en U's río abajo del sitio de procesamiento.

“La formulación de un problema, es más importante que su solución”
Albert Einstein

3. JUSTIFICACIÓN

El proceso de poliadenilación ha sido estudiado exhaustivamente en mamíferos y levaduras, sin embargo, en protozoarios es poco lo que se conoce con detalle acerca de este proceso. Uno de estos organismos en donde se ha estudiado un poco más este proceso son los tripanosomátidos, ya que la poliadenilación se encuentra asociada al trans-splicing del mini-exon (splice leader). En la mayoría de los protozoarios apenas se ha comenzado el estudio de las señales de poliadenilación mediante análisis bioinformáticos. Previamente, mediante análisis informáticos, nuestro grupo de trabajo propuso las posibles señales de poliadenilación para el protozoario *T.vaginalis* (Espinosa *et al*, 2002, Barrera-Andrade, 2007), pero aún quedaba pendiente demostrar experimentalmente la funcionalidad de estas señales. Este objetivo fue abordado en el presente trabajo.

*“Cada día sabemos más y
entendemos menos”*
Albert Einstein

4. OBJETIVO GENERAL

Demostrar experimentalmente la funcionalidad de las señales propuestas para la poliadenilación de los RNA mensajeros en *T. vaginalis*:

- señal de poliadenilación: UAAA
- señal para el corte del y procesamiento del mRNA: Y↓AAAAUU
- elemento río abajo al sitio de corte, rico en U

4.1 Objetivos particulares

- 1) Actualización de los análisis *in silico* para la detección del motivo UAAA propuesto como señal de poliadenilación en genes de *T. vaginalis* que codifican para mRNA.
- 2) Analizar funcionalmente la señal de poliadenilación UAAA, mediante mutagénesis dirigida de esta señal utilizando ensayos transcripcionales con el gen reportero *cat*.
- 3) Determinar el sitio de procesamiento de los transcritos de *cat*.
- 4) Medir los niveles de actividad de la proteína reportera CAT, como un efecto indirecto de las mutaciones en la señal de poliadenilación.
- 5) Analizar funcionalmente la secuencia propuesta como sitio de corte del mRNA, y el elemento río abajo rico en U (DSE).
- 6) Definir si una secuencia de RNA que contiene a la señal de poliadenilación es capaz de ser reconocida por factores presentes en extractos proteicos de *T. vaginalis*.

*“Nuestra recompensa se
encuentra en el esfuerzo
y no en el resultado. Un
esfuerzo total
es una victoria completa”*
Mahatma Gandhi

5. RESULTADOS Y DISCUSIÓN

5.1 Análisis *in silico*

En el 2002, nuestro grupo de investigación propuso las señales para la poliadenilación de mRNAs en *T. vaginalis* a partir del análisis de 13 secuencias de cDNA obtenidas de secuencias reportadas y de genes que se estaban trabajando en el laboratorio (Espinosa et al. 2002). A la fecha existe un mayor número de secuencias de cDNAs reportadas ya que en el 2007 se publicó el genoma de este protozooario (Carlton et al. 2007). Gracias a ésto, el análisis *in silico* que se realizó inicialmente en el 2002 pudo ser ampliado. Para llevarlo a cabo, se consideró si el motivo UAAA es la señal de poliadenilación en *T. vaginalis*, entonces estará presente en todos los 3'UTR de los mRNA analizados.

Para este análisis se utilizaron dos enfoques:

- a) Si el motivo UAAA es la señal de poliadenilación de los genes que codifican para proteínas, entonces en de la base de datos de trichdb (www.trichdb.org/trichdb/) se podrán identificar señales potenciales de poliadenilación en la región 3' intergénica de estos genes.
- b) Se buscará el motivo UAAA en todas las regiones 3' UTR de cDNAs accesads en la base de datos del NCBI de *T. vaginalis*. De esta manera se podrán identificar señales potenciales de poliadenilación en la región 3' UTR, así como el sitio preciso de adición de la cola de poliA en los mRNAs.

5.1.1 Presencia de la señal propuesta de poliadenilación UAAA en mRNAs de *T. vaginalis*

Asumiendo que UAAA es la señal de poliadenilación en los mRNA de *T. vaginalis*, esta secuencia debería estar presente en la región 3' de todos los genes que codifican para proteínas en este organismo. Valiéndonos del genoma y de las bases de datos de ESTs (Expressed Sequence Tags, por sus siglas en inglés) disponibles actualmente, se buscó la presencia del motivo UAAA en la región 3' no-codificante de 780 genes. Para este análisis bioinformático, se utilizó la secuencia de aquellos genes de los que se tiene evidencia de ser expresados, lo cual se determinó mediante el de alineamiento de los ESTs con su respectivo gen, verificando que tuvieran un 100% de identidad. Es importante resaltar que en estos casos, sólo podemos identificar si está presente o no el motivo UAAA en estos genes, ya que se cuenta con evidencia de su expresión pero no ha sido determinado el sitio de adición del poliA de manera experimental.

De los 780 genes empleados en el análisis bioinformático, se examinó una región de 403 nt, que incluye el codón de paro de la traducción y 400 nt río abajo de éste. Se definió este tamaño de la región analizada debido a que los 3'UTR de los genes de *T. vaginalis* son generalmente muy cortos. Los resultados mostraron que de los genes analizados el 99.5% contienen al menos un UAAA dentro de la región analizada, esto quiere decir que solo 4 genes no tienen al tetranucleótido dentro de los primeros 400 nt río abajo del codón de paro. La Fig. 12 muestra la posición del primer tetranucleótido que fue encontrado en la región 3' intergénica.

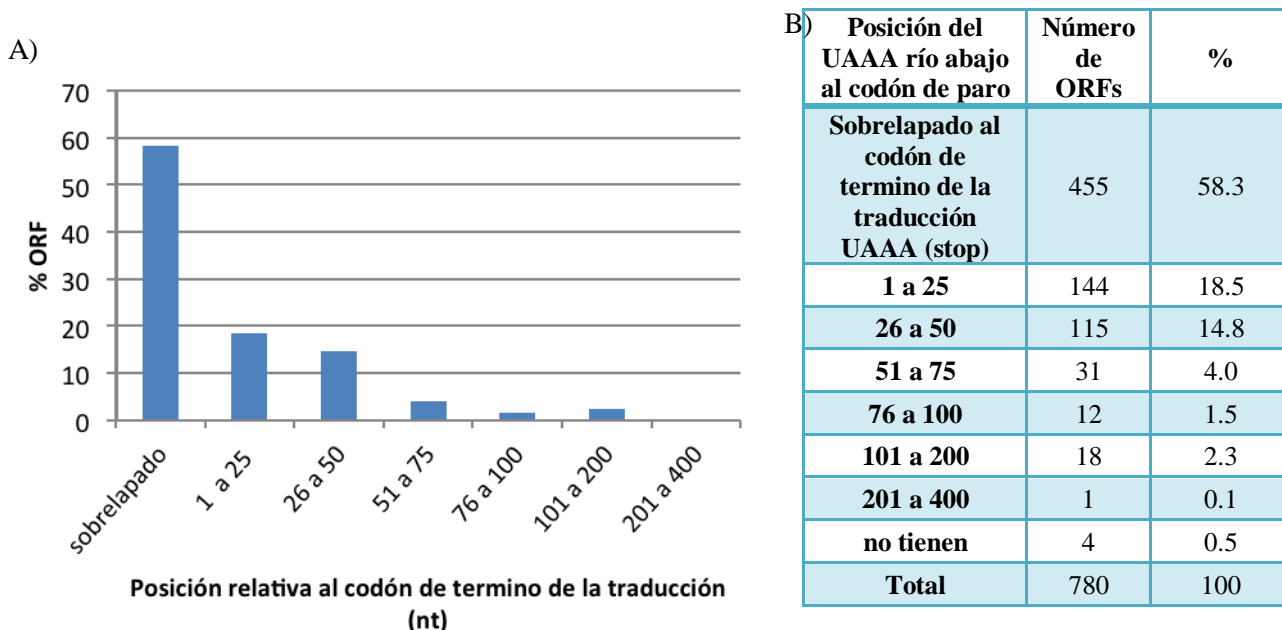


Fig. 12. Frecuencia de localización del primer UAAA en la región 3' intergénica de 780 genes analizados de *T. vaginalis*. A) Gráfica de la frecuencia del UAAA en ventanas de 25 nt; B) Tabla del porcentaje de genes que contienen al UAAA en las regiones analizadas. ORF: Marco abierto de lectura (Open Reading Frame).

Dentro del análisis realizado se incluyó la frecuencia del uso de los tres codones de término de la traducción: UAA, UAG, UGA. Se encontró que el 89.1% de los genes emplean al codón de paro UAA, 4.5% el codón UAG y 6.4% el UGA. En cuanto a la distribución del tetranucleótido UAAA, se puede observar que en más de la mitad de los genes analizados este motivo se encuentra sobrelapado con el codón de paro UAA; lo anterior indica que en los genes analizados existe una señal potencial de poliadenilación de los mRNAs.

Se ha reportado que *T. vaginalis* posee regiones 3'UTR cortas, por lo tanto, se espera encontrar las señales de poliadenilación cercanas al codón de término de la traducción. En este análisis

encontramos que el 91.6% de los genes, tienen al menos un motivo UAAA dentro de los primeros 50 nt río abajo del codón de paro. El análisis fue extendido con la finalidad de identificar si se encontraba la señal canónica de poliadenilación AAUAAA y/o algunas variantes como el pentanucleótido AAUAA que también había sido propuesto como posible señal de poliadenilación en este organismo (Bentley, 2014). Sin embargo, encontramos que únicamente el tetranucleótido UAAA es el motivo común.

En este análisis encontramos que el tetranucleótido UAAA está presente en prácticamente todos los genes. Sin embargo, y de acuerdo con los resultados que se presentarán mas adelante, sabemos que la presencia de este motivo no es suficiente para dirigir el proceso de poliadenilación, sino que es necesario que se encuentre en un contexto determinado por secuencias adicionales como son: el sitio de procesamiento del RNA mensajero y la región río abajo rica en uracilo.

5.1.2 Presencia de las señales de poliadenilación en la región 3'UTR de varios cDNA

Como ya se mencionó, el análisis bioinformático del cual se desprende la propuesta de las señales de poliadenilación en *T. vaginalis* se inició con tan solo 13 secuencias de cDNA. Con el propósito de ampliar este análisis teórico, se estudiaron 67 secuencias de cDNA que fueron alineadas con su secuencia genómica, lo que nos permite identificar el sitio preciso de adición del poliA en el mRNA (fig. 13). Se utilizaron solamente aquellos cDNAs que cumplieron los siguientes criterios:

- 1) contienen parte de la región codificadora y la región 3'UTR
- 2) contienen la secuencia de poliA.

Cysteine CP2 **GCTCTCTAAA**TTGCGA**TAAA**AATGT**A**CGATGAAGTACTTTGATA**TAAA**ATAGTTTTGGGAAACGCGATT**T**

Alanyl-tRNA **TCTCTCTAAGTTCGGTTTCACACAGTAA**GTCTTTTTTTTTAGATAGAATTTTT**TAAA**ATTTAATTTTTTT**T**

Adhesin AP51-3 **TCAAGGCTGTCAAGGCTGCTAACGGTGAAAAGCTC****TAAA**GAATTTTGCTCATT**T**

Cysteine CP4 **GTGAAGCTACAATGGCCTGCATCCACAGGACAAG****TAAA**CATTTTTAGTGATTACATACG**CT**

Glucose-6-phosphate **AAAGTGCAAGCACTACTATTTGACTTCTC****TAAATAAA**CAGCTTTCATTATGATCGATCTTTTC**T**

RasC **GCCTAATTGTT****CATGTTTGCTTTTA****TAAA**ACGATATTGCCAAATT-----
TAGATTATGATTTCTATGATGTTTGTACATTCTATTAGATTGTTTCTATACATA**TAAA**GCAATAATATACTCTATTTCTGCG**G**

Glucosamine **GGCTTTCAGAAGATACCCATTC****TAAA**TTTGACTTAATTTTTTTTTAAGTT**TAAA**AAAACCTCTAATTT**T**

Sir2_like **TTGACGACCTTAATTAG**CTGTATATTAGATATTTTCTACTTTTTGTTTT**TAAATAAA**TATAATCTGTCATATT**TG**

Alcoh_dehydrog2 **AAATGGAATGATGACCTCGAGTAA**GCATGTATTACT**TAAA**TCTTAAGTTTCAACTCCCTC**AC**

Centrin1 **CTTCCCTCTCTAA**GATTCTTTGTTAATGCACAACCAAAATGCGTTCTCACAACCTGTGAAAC---
TCTAATTCTACCAATGTATTCAAAGCACAATATATTCTGTTATACTCAAATTTTATTCTTGAAA**TAAA**TGTAAGAAGTTATGAT**CT**

Flavoprot **TTCTAAA**GAAGATTTCCAAAAATAATTTTAATAATTTTCGATGTTTCGAAACGGCTCGCTTTGG**TAAA**AAAAGTTT--
TAATAATCTTCTCAATTTTTTTCCAAATATTTTTTCCATTGAAAAGAATTT**TAAA**ATTGAATTA**TAAA**ACAGTTTTCTTGATA**T**

TBP-associat prot **TATATAGACGAAAGTGTGAATTGTCGATTAA**ATTTTGTTTGTA**TAAATAAA**ATATGTTTATTCA**T**

Secreted ade prot **CGAACCATCAGCTGA**CTACGAACCACTTACTAAGCGGATTTT**TAAA**CACTTCGCTTTCTTAATT**T**

Rsp GTPase **AATAATAGCTGTTGTGTCATTGCCTAA**TTTACCCTTTTGATTA**TAAA**AATGCATGATTTTTTTGGAGTTTTAG**C**

AP65-3 **CGATCCACAGGCTGAATACGAGCAACTCTACTAA**GCAGTTTT**TAAA**ACTCTTTCAATGTCTTT**G**

Adhesin AP33-3 **AAGAGAATGGGCAAGATC****TAAA**TTTCTCA**TAAA**AGTTCACAATTCAATAACCTAATT**C**

Tran EF1-b **CATCGCTTTCAGAAGATC****TAAATAAA****TAAA**C**TAAA**GGATTATTGATAGCCC**TAAA**CATATTAATTTTAATAC**T**

Adhesin AP33-1 **AGGAGATGAAGAGAATGGGCAAGATC****TAAA**GTTTCCATATCTTAT**TAAA**TCGATTTGAATGTT**C**

Adhesin AP33-2 **ATGAAGAGATTAGGAAAGATC****TAAA**TCTACTGCTTACTT**T**

Saposin-prt3 **CTGCAAGGACTAA**TGATTCCATTCCTA**A**GATTATCAACAA**TAAA**AAAATCCATCGAA**TAAA**TATACGGTC**T**

Glutamate dehyd **GCTATGCTTGCCACGGCTGCTGCTAA**GCTAGCGAA**TAAA**ATTTTCTAGTTTTTAGT**T**

AAATCC**TTGATTTTTTTT**AA**CTTTTT**TCATA**TTTTAT**

AAATAC**TTTTTT**ATTT**CACTCTAC****TTCTTA**TT**AGATAGA**

AAATAA**TTT**CACATA**AAAT**CTAAAGTCATGAC**TTCTTTTT**AAATCTACAATCA**ACT**

AAATATCTAA**TTTTTTT**AT**TATATAAAATTTTT**CCAA**TTT**CATTACATGCCAAG**TT**

AAATACA**TTGA****TTTTTT**ATATG**TTTCTTTT**AA**TCTA**TTATATCTA**TTCTTT**ATATGT

AAATATGAG**TTTTTT**GAAAGATCA**ATTTTT**CCAA**TT**ACA**TTA**TT**CAACA**TT**GA**

AAATATAAC**TTCCA****TTTTT**CCT**CTGT**TT**CTAACACATACTCA**TT**AAAAAA**TT**CC**

AAATCTCTCA**TTTTTT**CTGAGGA**TTT**TACAA**ACTGGGCGAGA**ATA**CAATGAAA**

AAATACTA**TTTTT**ACA**TTT**ACAATCCGTTCAAACACAAAATACGAA**T**

AAATTGAAGCAAGCTAGCGCAACACAAACATA**TTTTT**ACAT**CTTTT**ATCA**AT**

AAATTTGACA**TTCA****TTTTT**CAAAAATGGAAGAG**TTTATT**GAGCG**TTT**AAAGTCA**ATT**

AAATAT**TTTTT**AT**TTT**CAAAAAATCCTCCGACAAAATGATCAAGAACACAGA**CAAT**

AAATAAAAGC**TTTTTCTTTT**AT**CTTTT**CACCTGACAT**CTTGACA**AGTCAG**ATTG**

AAATTA**ACTACA****TTTTTT**CATATATA**TTTTT**AA**TTT**CTAAATAAAGATATAT**GA**

AAATCCTGA**TTTTTTTT**AAAGTA**TTCTCAT**TTATCGTGTGGAAAG**TT**ACAAAATCA**T**

AAAAATAAATCATCA**TTTTTCTCA****TTTTTTTT**CAA**TT**AAAAAGTCACA**TTCTCA**

AACTCACAAT**TTACT****TTTTT**AAAAGA**TTTTT**GC**TTATT**CTCGTGGCAAGAA**TTTA**

AAAAACT**TTTTTTCT**TACAAGTCAA**TA**CTCTAA**CA****TTT**AAACTCTAAAT**CTT**

AAAAACTCATAC**TTCT**AAATAGGAAAG**CTTCTCT**TAAGAAAATGC**ATTCTTTTTT**

AAAAAACTTG**TTTTTT**AT**TT**CATACAA**TT**AAGATAGC**TTGAGATGC**TT**TTTTGATTT**

AATCATCGAA**CAATTTTT**TAA**TTTTT**CTTAA**ACTATTA**ACCAATGGC**TTTATTTT**

Adhesin AP51-2 **TGCTAACGGTGAGAAGCTCATCTAAAGAA**GTTTGCAAAT**TAAA**ATTC**TAAA**GTTATCAACATCTGATT**C** **AAAAAAAAAATC**TAA**TTTTTT**AAAACATAATCA**TT**AAC**TG**TT**AGTGCTCA****TTTT**

Alcohol dehydrogenase **TCTCATCAAGCCAGTCGTCCTCATCGAATACAACGACAAGCTT****TAAA**AT**TAAA**AAAGTATAATTTCT**T** **AATA**GAATAGA**TT**CAC**TTCTTTTT**AACAAAGAC**TT**ATAATGGAAGTT**CCTGA**

Cysteine CP1 **GAAGCCTCTGCTGCTTGCATCCCAACAGTTTCAGCC****TAAA**CC**TAAA**AATATTACTGTT**C** **AAAAAACAATT****TTTTTT**AATA**TT**CATCTAAATACCAATCTCTAATGATGATA**TTTAGAA****TTT**GA

Eukaryotic release factor 1 **CTCTAAA**TTGGA**TAAA**GTATTTTGCAGAAATTAAGCT**TAAA**CTCATTATATTTAATTT**CATT****T** **AAATA**GAT**TTCC****TTTTT**ATGA**TT**CTA**TTTTAAGA****TTT**CAATAAA**TTTT**ATGAA**TTT**G**TTT**

Legumain cysteine protease 1 **TGCAGACATCAAGGCTGCTATTGATGCCATCTGCTAA**GTGCGGGAA**TAAA**AGTATCGTGATTCTG**G** **ATAA****TTTTTTTTTTTTT**A**TTTT**CTTATCA**TTTT**CATATAAAAGTTGCAATTA**TTT**C

Legumain cysteine protease 2 **ACTTTACTTCTAA**TCATAATTTTTTATTTAATTTCTTTTTTATT**TAAA**GAGT**AT**ATGCAGTT**T** **AAAATG**ATAG**TTTT**ACTCTG**TTT**AAC**TTTTTAAATC****TTTTTT**AAAA**TTTTTTTT**CGCCAA

Sir2 **TATGATGCCTAAATAGATATTTATTTTAGCCAGAAAGA****TAAA**CAAAA**TAAA**ATGTTAATTTTATGGATTTTTT**CAC** **AAATA**GAGCTAAACAATGCC**TT**TATAG**TTTTAA****TTT**AAATCG**TTTT**G**TTTTT**ATCAT

14-3-3 protein A isoform **TCGTGATAACGTTCAACTCTGGACACAGGACAACGAGT****TAAA**TTATGATTTAATGCTCAATG**T** **AATA**TTAAGGA**TTTTTTTT**TATATTCGACCAA**TTTT**C**TTT**ACATAAAT**TGTA**

Rab1a GTPase **CTGGATGCTGCTAAA**TACAATATCTTTACTAATAATTT**TAAA**AACA**TAAA**GCGGCGATATGTCGCACTATAC**C** **AATA**TCAT**TT**GATAATA**TTT**GATGA**TT**ACT**TT**AAAATGACAAAAATATAA**TTT**AAT**TT**CT

Rab6a GTPase **TAATGCCAACTCAAACAACAATGCCAAAAGTGGTGGCTGCTGTTAA****TAAA**AATA**TTT**AATTTCTTTCTTTTTTAT**T** **AAATA**ACAAACATAAATGATGAAGATCCGGGTGTAATGAAAGTGGAGACTCTACG**T**

Rab6b GTP **TAAA**AGTATTACCAACATTTCTTTAACTCCTATTATACTTTTTGATT**TAAA**TC**TAAA**ATAT**TAAA**AATATCAGT**C** **AGAT**TATAACAA**TTT**A**TTT**CA**TT**GAT**TT**AATGCTAAGTACATATAAGAGGA**TT**AACTA

Rab11 GTPase **AAAGAGCCCAGAAGAGAAGAAGGGATGCTGCTGAA**ACTTATTA**TAAA**TATGCTAACTATTACACTTT**CATT** **AAAAAAAAAAAA**CAAAACAAAATA**TTTT**A**TTTT**ATAT**TTT**ATC**TTTT**AA**TTT**ACAAAAA**CTTAA**

RNA-binding protein **CCAGTTAGCGAGTAAA**TCTCACA---
GTTAACTCTTCGAAGTACCCTTCGATTCCCAAACCGATGACGTTAATTAGATTTATTTTAGAA**TAAA**ATATAATCT**T**GTATGTCAC **AAAACA**ATC**TT**ACGCA**TTTT**AAAGA**TTTTTTTT**GAAGA**TTTT**CATATC**TT**GACAAA**TT**C

Saposin protein 1 **CACCAGCCAAGTCTGCCAGTTCTTAAGATCTGCCAGTAAA**ACCTTCTGACTAATCTCTAAGTT**T** **AAAGT**TTGA**TTTT**A**TTT**ATC**TT**CTCTAACT**TT**GAAA**TT**C**TTTT**ATC**TTTT**ATCCAA

Surface integral membrane protein P270 **GAGGAGTCAGGATCAGGAGTTGAAGTCTAA**TTTATTA**TAAA**TATATTCTATTTAATTAGAATTATGA**T** **AATA**TA**TTTT**TAGA**TTTTT**ATATA**TT**CAACATCTCTAATAA**TT**CTCATG**TTTT**ATTTA

T1 MYB24 mRNA **GATTATAATCAGTTTCGATCTTTGGAACGATTTTTCATTCTTC****TAAA**CGTTTTAAGTTAAGTGTGCTT**C** **AATG**CACCAATC**TTTTT**AATCCCAAGCGTAAACGAAGGATGA**TTTTT**C**TTTTT**C

Centrin 2 **TACGAAGAATTTGTTCATATCATGAAGAAGACATCACTCTTC****TAAA**ATATTTTTATTT**TTTT****T** **ATACT**AAATACTCATAATCGTA**TTT**AGATC**TT**CACCGC**TTTT**CAAATG**TT**GCA**TT**

Prolyl-tRNA synthetase **CGCTCGTTCATAC****TAAA**TTTTTTTT**TAAA**TATAACTCGCATTGAAA**TAAA**GATATCTTACAATTTACT**T** **ACTC**TCCAAA**TT**AATG**TTT**TACA**TT**ACT**TTTT**CTA**TT**AATATAAAATTA**TTT**ACTGTA

Rab small **CATTGGTAAGAAAGGCTGCTGTTAA**TTTCTATTATATATTCTTTTGGAGTGATTA**TAAA**GATCTTTCGAATTTAA**T** **TTAA**GGGTGCACCCGTATGTACG**TTT**A**TTT**GAGAGGAGGAGGA**TT**ACCCTCTC

64kDa histone H1 **GGTAA**GTTTGTCAATTATAGAAAGAGAAGTTAAGGGAATTGTTA**TAAA**AATACC**TAAA**GAGTTATTCACTTTTAT**T** **ATAT**TGC**TTAA****TTTT**CTCA**TTTTTTTT**AGATATACAACGAGAAAA**TTT**

Alpha-actinin **TAACTCACTTTACAACCA****TAAA**TTTAGTTATTGAATTTGAT**TAAA**AAA**TT**TACCCAGATTT**TT**C **TTAC**TTTTACCCTCCCATGTAA**TTTT**CTGGGGAAAACGATAA**TT**CA**TTTTTTTT**AT

Adhesin AP51-1 **GTGAGAAGGCTGTTAAGGCTGCCAAGGGCGAAAAGTTC****TAAA**ATCATCTCA**TAAA**TTTTTCTCTAG**C** **AAATA**AACGA**TTTTT**CGTCTCTCTC**TT**AACAA**TT**AA**TT**C**TT**ACT**TTTT**AA

Longevity protein **GAAAAGGCCGAATAG**ATTGTTTTATTTTAC**TAAA**CGTTTCTTTTATCTA**TT** **TTGT**TTAATAACTGCCATG**TTTTT**GAT**TT**AAC**TT**ACGAAA**ACTA**

RRas GTPase GGGCTGTTGCAATATCGAGTAAAATTGTAAATTTCTGAAC TATGAATTTTGCGGCCTTTTATTTATAAAATTT
Saposin prot2 GCTTTGCTAAA AAA TTAAT TAAA AGTGACAAACACACTTTATTGCGATTT AAAA AAAATCATCAAGTTT CTTA TTTTTA TTTTTT AATGATTT AGATTAGT
Syntaxin 16/TLG2 GGTACTGTTGTCATTATTAAGAAGCAAAGAAATCTTC TAAA TAA TTTTTGCCTATTATCTT -ATT TTCA TTTTTTT AAA TTTTTT GATTAAGTAGAAA

mgll1 gene TCATCGCTGATTTCAAACAGGGCCTTGACGCTCTTTTATAA CTCTACTTAGTTTCTTGACTT TAATT AATT TTAGAAAA TTTTTAAA TTTT CTC TTTTTT CACCCTTTA TTTT CCCTATAA
Ferredoxin TGCTTGGCTATCACACTCAGTGGTGAAAACGATGGTGTGTTTTTCGAGCTC TAA TAATTGAAAGTTTAT T AAATT GTTTTT GA TTTTTT CCAATAC TTAAG TTACAT TCAAAAATGAATCGC TTTA
Succinyl-CoA AAAGTTC TAAA ATCATCTCA TAAATT TTTCTCTAGCCAAA TAAA CGATTTTTCTGCTCTCTCTTAACAAAT T AAATT CTTAC TTTT AAATAAAAACAACAGGAAAATAA TTA AAAATAGAA TTACAA
Superoxide ACCACGTCAACTGGAAC TTTGTTGAGCAGAAC TAAAGGCTGCTGGTTTG TAA AACGTTCC TTTCTCATT G AAAATT ATTAAC TTTTTTTT C AAGTAATGGTA TT CAGTATTTT GAGC TTTG TTT
Calmodulin ATTCGTCCACCTCATGCTTACATCA TAA ATATTTTGGAA TTTTTATT TTTCT TAAA TTTACAATTATATTT T AATC GAAATAAAC TTTT CTGCA TTTT CTGGAATG TT CGATGTAGAAAAA TTA
Sarco-endo ATPase TGAA TAAA TTTCCGATAATTTTTTTTGGTTTTT TAAA AAATTTAATTGATTGAAAAAATTATAATAC AAAAACT TAC TTCT TTTT GTTT CCTT GATCA TT AATGGATGAAGA TTTAGATT CGA

Act1 TTC TAAA AAGTTTTTTTTTCGTTGTCTC AATT TAATCAAAAAATA TTTTTT CAT TATTAC TT GATAT
Act4 TTC TAAA ACAATTTGAAATGTTTCAGATTCC AAAATT TCTAATTTTTTTT AAC TTTTTTTT ACTC TTTTTT CCAA
Act5 TTC TAAA CAATTTCCGGTATAATATATTT C AAAATT AACTAACC TTTTTT CTTT AGTGAACCTA TTTT GGA
Act8 TTC TAAA CTATATTTTATTAGCTTCCAGTTAAC AAATT TAATCAG TTTTTT AAAATA TTTTTT CAATA TTTT GA
Act10 TTC TAAA TTTTCGGTTTTATTTGAT T AATT TTCTCTC TTTTTT AATAA TTTT CTCATC TT ATTATA
Act11 TTC TAAA TTTTC TAAA AATATCGGTATATCCTTTAG AAATT TTATC TTTTTTTT AC TTTT NNAAAAATCTAAATG
URF AAC TAAA AATTG TAAA TGTTCAGGAGTCATTGCTCAGTAAC AAAAAATT ACTAATC TTTTTT ACATCTACAA TTAAGATAATAAT

TVAG_121680 TAATATATTTTGTGGATAAAAATTTTCATGTTTGATATA TAAA TGTATCTATTTCGT ATAC AAC TTT CAATCTA TTA AATAAGTA TTTT ATGCTAGTAATAATTGAAT
GCATAAAAAAACACAAAA TTTT CATTTTTAT
TVAG_048120 TAATTCATTTTTATAATTGTATGTA TAAA TTTATTTGTTTATTCCTT ATTT GAACCCAATGCA TTT CAT TCTCCATCCTCAATATCTAA TTCTCATATCAGAAGA
TTTTCTCAA TTAGCAG TTTTTT GTA TTTT CTAGATA TTTT ATATGTGGA
TVAG_266050 TAATTTGTTTATTATATATCTTTTTTGCTTCTAT TAAA GATTTATTAATATTATT T AATC TGCAGACTTA TTTTTT AACATGCTCCTCATCAAAGATAAG TTT CAAAAA
TVAG_395750 TGA-14nt-TTTCTTTCTTTGTTTGAAATAAATATTGTATTC TAAA ACTCTATTCTTTCCCTTTTTAGTG AGTT GG TTTAATATAGATAGA TTTTTT CAGC TTTTTT CCGTTGAGTCA TT

TVAG_468150 TAA TACTACTA TAAA TGTGTTATTATAATG T ATTATAC TTTTCA TTTCTACCTA TTGATTGCCTA TTG TTTTTT ATGATA TTTTT ATT

TVAG_488140 TAAACTATTTTTTTCGAA TAAA GTCATGTCAAT T AAAAAACA TACA TTTTT CCTTCTTTCTAC TTTGATTGATT CATAAC TTTATATCTCTGT T GAT

TVAG_494460 TGAA TAAA TATGTATTTAATTGTTA G ATCTAATC TTTATA TTTTT CCAAACCACATG TTTT AGTAAGCCCTCAATATTATAT

Fig 13. Alineamiento de mRNA de *T. vaginalis* con sus respectivas secuencias genómicas. Código de colores: Región codificadora (letra azul); región 3' intergénica (letra negra); señal propuesta de poliadenilación (sombreado verde); sitio de procesamiento de los transcritos definido como el último nucleótido antes de la cola de poliA (sombreado fucsia); región 3' adyacente al sitio de corte (sombreado amarillo); región río abajo rica en U's (sombreado azul), nucleótidos distintos entre cDNA y secuencia genómica (sombreado rojo).

En la figura 13 se presenta la secuencia nucleotídica que abarca parte de la región codificadora del gen y la región 3' intergénica, de acuerdo con el código de colores descrito en el pie de figura. En 66 de los cDNAs analizados se encontró al tetranucleótido UAAA en la región 3' UTR. En este análisis se encontró un caso peculiar: el cDNA de la sintaxina 16/TGL2-like (número de acceso AY344235), ya que este no contiene el motivo UAAA en la región 3'UTR que es de solamente 21 nucleótidos. Sin embargo, al buscar río arriba se encontró la secuencia UAAA dentro de la región codificadora, formando parte de los últimos 2 codones previos al codón de término de la traducción TAA (...TCT AAA TAA...). Este caso es similar a lo que se ha reportado en *Dictyostelium discoideum* donde a partir de un análisis bioinformático se propusieron las señales de poliadenilación, y varias de éstas se localizan dentro de la región codificadora. Con los resultados anteriores se construyó un diagrama tipo Logo de secuencias (<http://weblogo.berkeley.edu/logo.cgi>) en donde se muestran los motivos más frecuentes en la región 3' UTR y río abajo al sitio de corte (fig. 14).

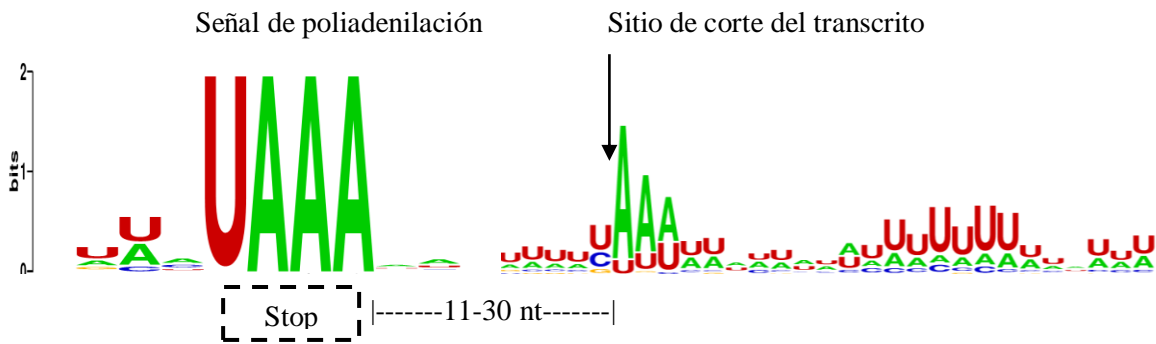


Fig. 14. Logo obtenido con base en el alineamiento de secuencias relativas al motivo UAAA y al sitio de corte y poliadenilación de 67 secuencias de cDNA de *T. vaginalis*.

Como se puede ver en la figura 14, el motivo UAAA se encuentra en el 100% de los cDNAs analizados; se puede ver también que esta secuencia se limita al tetranucleótido ya que no hay otros nucleótidos alrededor de éste que se encuentren presentes de manera predominante. Es importante mencionar que la distancia entre la señal de poliadenilación UAAA y el sitio de corte va de 11 a 30 nucleótidos. A diferencia de lo encontrado para el motivo UAAA, en el caso del sitio de procesamiento no se observa una secuencia que prevalezca en un 100%; sin embargo, existe preferencia por ciertos nucleótidos que cumplen con la propuesta anterior de nuestro laboratorio: $Y\downarrow A_{2-5}UU$. Por último, para la región río abajo al sitio de corte, se observa que existen regiones en

donde predominan las corridas de Us. Estos datos refuerzan las propuestas anteriores de nuestro laboratorio (Espinosa *et al*, 2002).

5.2 Análisis funcional de la señal de poliadenilación UAAA

En nuestro laboratorio se ha utilizado una estrategia para el estudio funcional de secuencias implicadas en la regulación de la expresión genética y el procesamiento de transcritos de *T. vaginalis*, la cual consiste en la construcción de vectores de expresión (plásmidos) que contienen un gen reportero flanqueado por las secuencias reguladoras a estudiar, y el análisis de los transcritos correspondientes al reportero. Las secuencias por estudiar pueden ser modificadas por mutagénesis dirigida y los transcritos derivados de estas mutaciones pueden ser analizados y comparados entre sí. Esta estrategia fue utilizada para el estudio tanto de regiones promotoras como de regiones involucradas en el procesamiento de la región 3' de los transcritos (Franco et al. 2012).

Como se mencionó en la sección de antecedentes, los primeros ensayos dirigidos a la disección funcional de la señal de poliadenilación en *T. vaginalis* fueron realizados como parte de la tesis de Maestría de la M. en C. María Guadalupe Barrera; estos datos preliminares sugirieron que el tetranucleótido UAAA desempeña un papel importante en dirigir la poliadenilación de los mRNA. Uno de los objetivos de este trabajo de tesis, por tanto, consistió en corroborar y completar el análisis funcional de la señal propuesta de poliadenilación de *T. vaginalis*.

Para confirmar que la señal de poliadenilación se restringe únicamente al motivo UAAA, se construyeron mutaciones adicionales a las realizadas previamente (Fig. 15). El vector empleado contiene al gen reportero *cat* con las regiones regulatorias 5' y 3' de un gen de actina de *T. vaginalis* (Fig. 10).

Las construcciones utilizadas en esta primera parte del estudio fueron las siguientes:

- a) El plásmido parental (Fig. 10 de la sección antecedentes)
- b) Cambiar los cuatro nucleótidos del motivo UAAA
- c) Cambiar de manera individual cada uno de los nucleótidos del motivo UAAA
- d) Cambiar el nucleótido río arriba y adyacente al tetranucleótido UAAA, para delimitar la señal (UUAAA → GUAAA).
- e) Insertar la señal de poliadenilación canónica de eucariontes superiores: AAUAAA.
- f) Construir una revertante de la mutante puntual UCAA, revirtiendo a la secuencia original UAAA. Con esta mutante se podrá demostrar que los cambios observados en la mutante se deben exclusivamente al cambio en ese nucleótido.

En la fig. 15 se muestran las distintas mutaciones utilizadas. Todos los plásmidos fueron verificados por secuenciación.

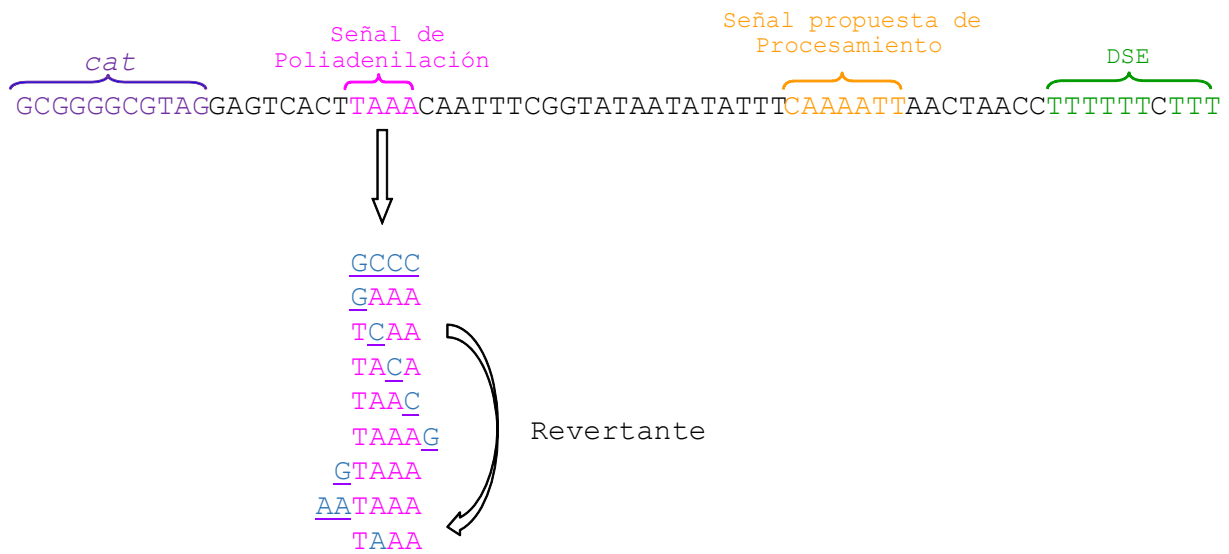


Fig. 15. Cambios realizados por mutagénesis dirigida en el tetranucleótido TAAA. Los nucleótidos que fueron cambiados se encuentran en color morado y están subrayados. Las primeras cinco mutaciones (GCCC, GAAA, TCAA, TACA, y TAAC fueron realizadas por la M. en C. María Guadalupe Barrera. Se muestra la secuencia de DNA.

5.2.1 Identificación del sitio de corte y poliadenilación de los mRNA de *cat* en la secuencia parental y las mutantes.

Para determinar el efecto que producen sobre los transcritos los cambios en el motivo UAAA y en las secuencias aledañas, se identificó el sitio de procesamiento de los transcritos del gen reportero *cat* derivados de las diferentes construcciones. Como se describe en Materiales y Métodos, se utilizaron cultivos de *T. vaginalis* crecidos hasta fase logarítmica, se transformaron con los plásmidos parental y mutantes, se extrajo mRNA y se sintetizó el cDNA. El cDNA se clonó en Zero Blunt TOPO y se secuenciaron varias clonas derivadas de cada ensayo.

Inicialmente se determinó el sitio de procesamiento del plásmido que contiene el motivo UAAA parental (pcat-UAAA). De este análisis se obtuvieron 22 clonas de cDNA y en todos los casos analizados la adición de la cola de poliA ocurre en un mismo sitio, esto es, 22 nt río abajo del motivo UAAA (Fig. 16 panel A). Este sitio de poliadenilación coincide con la posición en donde se poliadenilan los mRNAs de actina, gen del cual se derivó el vector reportero. El procesamiento del transcrito ocurre en la secuencia C↓AAAAUU, que coincide con la propuesta para el sitio de corte derivada de nuestro análisis teórico: [Y↓(A_{2.5})UU].

En contraste con los transcritos derivados del plásmido parental (pcat-UAAA), las secuencias de las clonas de cDNA derivadas de los plásmidos mutantes mostraron que cualquier cambio en el motivo UAAA modifica el sitio de poliadenilación de los mRNA de *cat* (Fig. 16 panel B). Es importante señalar que la obtención de clonas de cDNA de cada una de las pcat-mutantes fue muy difícil, esto posiblemente debido a una baja abundancia y/o inestabilidad de estos transcritos.

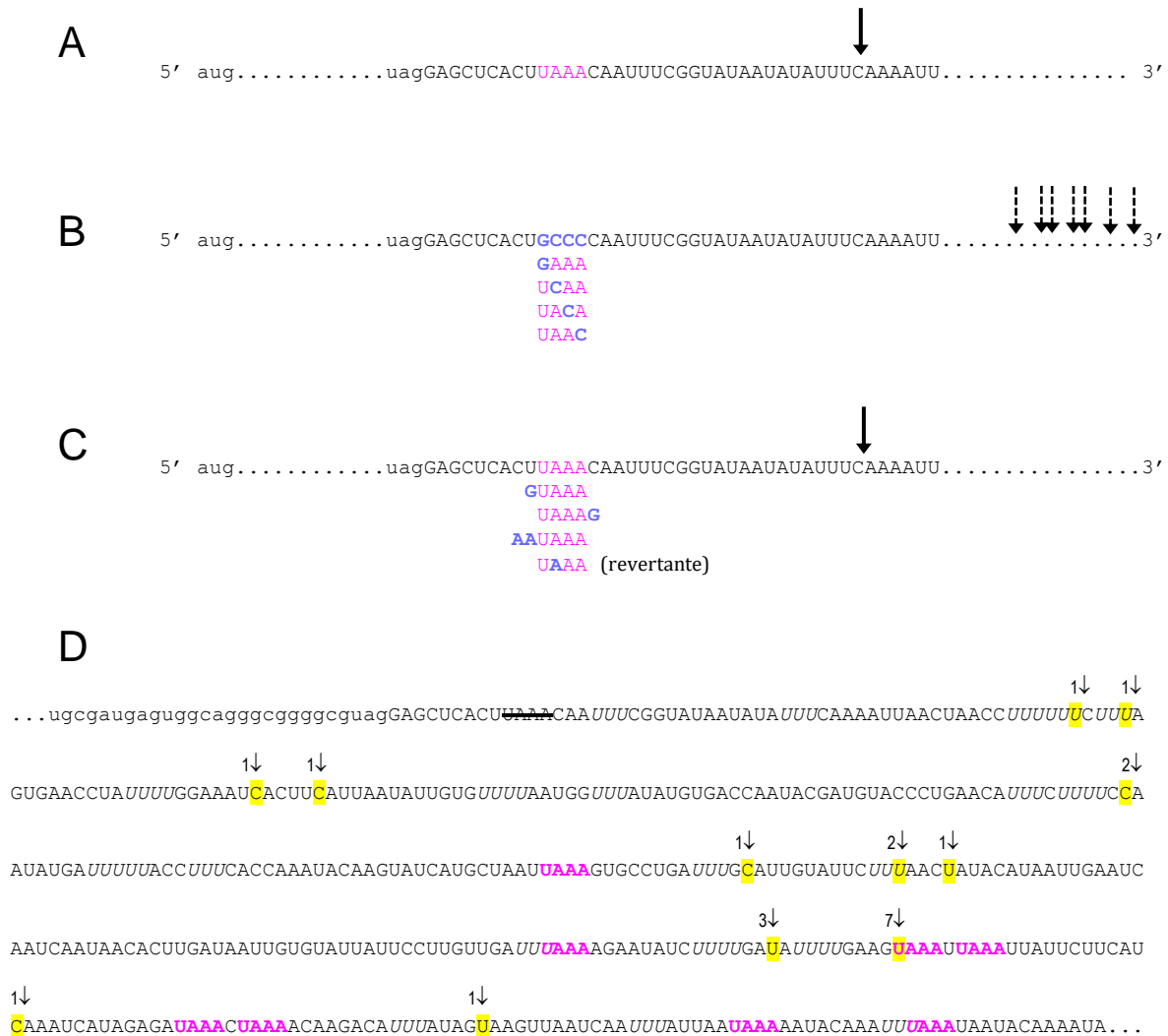


Fig. 16 Representación esquemática del sitio de poliadenilación de los mRNA de *cat*. A) Plásmido parental (pcat-UAAA); B) mutantes que modifican el motivo UAAA (pcat-mutante); C) Mutantes que delimitan al motivo UAAA, y revertante (UCAA → UAAA), las flechas indican el sitio de procesamiento; D) Representación de los sitios de poliadenilación de las mutantes en la secuencia de la región 3' transcrita del gen reportero *cat*. Las letras minúsculas corresponden a la secuencia codificadora de *cat*, las letras tachadas representan el sitio mutado de acuerdo con el panel B. Las letras sombreadas en amarillo corresponden al último nucleótido identificado antes de la cola de poliA; la flecha indica el sitio de procesamiento, y el número junto a flecha indica el número de clonas identificadas que fueron procesadas en ese sitio. En rosa se resaltan las secuencias UAAA, y en cursivas las regiones ricas en U's que se localizan en la región 3'.

Al analizar el sitio de procesamiento de los cDNAs obtenidos de las pcat-mutantes, a primera vista pareciera que la poliadenilación se lleva a cabo en sitios al azar (Fig. 16 panel D). Sin embargo, al examinar estos resultados se observó lo siguiente:

- a) Los mRNA derivados de las pCAT-mutantes poseen regiones 3' UTRs más largos y heterogéneos debido a que son poliadenilados en distintas posiciones, todas ellas río abajo al sitio en donde se poliadenilan los transcritos de la clona parental. Es importante mencionar que en este ensayo únicamente se pueden identificar a los transcritos poliadenilados (aislados mediante oligo dT), no pudiendo detectarse a la población de transcritos no poliadenilados.
- b) No se encontraron entre los sitios de procesamiento de los transcritos de *cat* derivados de los plásmidos pcat-mutantes predilección por una secuencia específica. Esto sugiere que las cuatro bases del tetranucleótido UAAA son igual de relevantes para definir el sitio de poliadenilación en un contexto apropiado (apartado 5.2.3).
- c) Se identificaron varios motivos UAAA río abajo del motivo UAAA parental (mostrados en rosa en la Fig. 16 panel D). Sin embargo, ninguno de estos motivos es capaz de dirigir la poliadenilación de manera exitosa a un sitio único de procesamiento.

Los datos anteriores indican que la secuencia UAAA es la señal de poliadenilación en *T. vaginalis*. Sin embargo, por sí mismo el motivo UAAA no es suficiente para dirigir el procesamiento y la poliadenilación de los transcritos sino que parece requerir de otras secuencias que le brinden un contexto funcional. Estas señales podrían ser las que definen el sitio de procesamiento y la región rica en U's, como se había propuesto de manera teórica. Llama la atención que de las 22 clonas de cDNA derivadas de los plásmidos pcat-mutantes, 16 clonas muestran poliadenilación río abajo y cerca de un tetranucleótido UAAA. En particular, siete de éstas clonas corresponden a la poliadenilación en un mismo sitio: U↓AAAUU, que corresponde al sitio teórico propuesto para el procesamiento de los mRNA en *T. vaginalis*. Sin embargo, consideramos que al no existir río abajo regiones ricas en U, no se logran procesar allí de manera eficiente todos los transcritos. Así mismo, debe mencionarse también que algunos mRNA fueron procesados en sitios próximos a un tetranucleótido UAAA; sin embargo, estos mRNA están flanqueados por regiones ricas en U's, las cuales se han identificado como sitios de interacción para factores de poliadenilación en otros eucariontes (Mandel, 2008). Estas secuencias pudieron favorecer la poliadenilación en estos sitios. Es importante mencionar que ninguno de los sitios en donde se poliadenilaron los transcritos de las pcat-mutantes posee las tres señales propuestas para el procesamiento y poliadenilación de los mRNA de *T. vaginalis*.

Los datos de la figura 16 panel C nos permiten concluir que la funcionalidad del motivo UAAA se limita a este tetranucleótido, ya que el cambio de las bases adyacentes río arriba o río abajo al mismo (pcat-gUAAA y pcat-UAAAg) no afectan el sitio de poliadenilación. Tampoco se observan cambios cuando se utiliza un vector al que se le introdujo la señal de poliadenilación empleada por metazoarios (pcat-aaUAAA). Estos resultados no son inesperados, ya que todas estas construcciones contienen al motivo UAAA. Adicionalmente, el uso del plásmido revertante (UCAAA→UAAA) nos permitió corroborar que los cambios en la poliadenilación observados en los transcritos de los plásmidos pcat-mutantes se deben únicamente a la mutación en el motivo UAAA, ya que con esta construcción se recupera el fenotipo observado con el plásmido parental.

5.2.2. Evaluación de la población de los transcritos de *cat*

Con el fin de determinar cómo se modifica la población de transcritos poliadenilados de *cat*, derivados del plásmido parental y de las mutantes en la señal de poliadenilación (Fig. 16A, 16B y 16C), se llevó a cabo la siguiente estrategia:

Se utilizó un sistema de dos plásmidos que funcionaran como reporteros transcripcionales simultáneos e independientes. Uno de los plásmidos es el ya descrito pcatUAAA, sobre el cual se hicieron todas las mutaciones al motivo UAAA y que contiene al reportero *cat*. El segundo plásmido (pgfpUAAA, que contiene al reportero *gfp*) se utilizó como control interno de transcripción; fue construido de manera similar a pcatUAAA, de manera que la región codificadora de *gfp* se encuentra flanqueada por las regiones intergénicas 5' y 3' del gen de actina de *T. vaginalis*. La construcción del pgfpUAAA fue realizada por la LIBB. Jessica Guzmán. (Fig. 17).



Fig. 17. Plásmido reportero *gfp*-UAAA. En letras minúsculas se muestran el primero y el último codón de la región codificadora del reportero *gfp*. En mayúsculas se muestra la región 3' intergénica de actina. La flecha indica el sitio de corte y poliadenilación de los mRNA. En negritas se indica la señal de poliadenilación TAAA, que coincide con el codón de término de la traducción TAA.

Para validar el sistema, los dos plásmidos reporteros pcat-UAAA y pgfp-UAAA se cotransfectaron en cultivos de *T. vaginalis* en fase logarítmica de crecimiento. Debido a que las regiones 3' y 5' de ambos plásmidos son iguales, se esperaba encontrar una abundancia de transcritos semejante para *cat* y para *gfp*. Se extrajo RNA total y se realizó el ensayo de RT-PCR multiplex (ver metodología). El tamaño esperado para cada amplificación es de ~235 pb para *cat* y ~200 pb para *gfp* (Fig. 18, panel A).

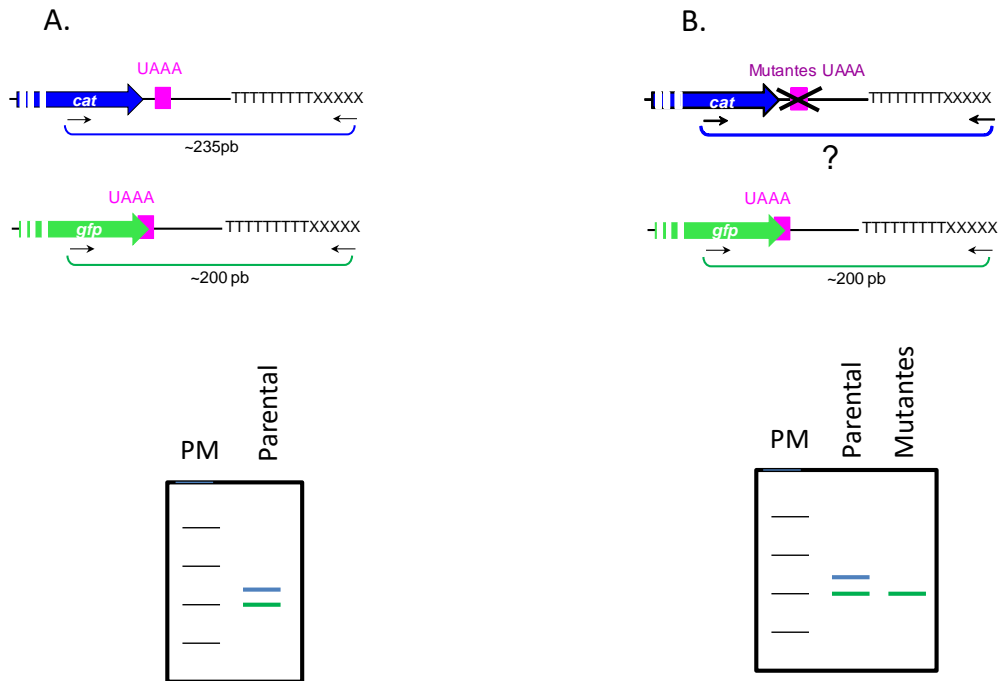


Fig. 18 Estrategia para evaluar la población de transcritos mediante PCR acoplado a transcriptasa reversa. Para ambos sistemas se utilizó un oligonucleótido reversa (dT) que se ancla en la cola de poliA, y un oligonucleótido específico para *cat* ó para *gfp*.

Los productos de amplificación se analizaron en un gel de acrilamida al 5%. Para corroborar la identidad de las bandas de amplificación, se hicieron digestiones con enzimas de restricción específicas para cada fragmento como se muestra en la figura 19. El tamaño de los productos de digestión obtenidos corresponde con el tamaño esperado.

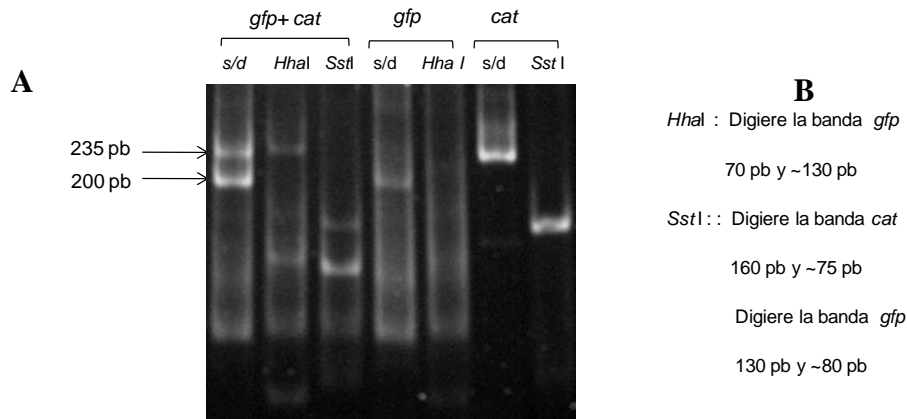


Fig. 19. Validación del sistema para estimar los transcritos reporteros de *cat* y *gfp*. A) Productos de amplificación por RT-PCR multiplex derivados de los plásmidos pcatUAAA y pgfpUAAA, analizados en gel de acrilamida al 5%. ***gfp+ cat***: productos de amplificación utilizando oligonucleótidos para *cat* y para *gfp*; ***gfp***: amplificación usando solamente oligos para *gfp*; ***cat***: amplificación utilizando solamente oligonucleótidos para *cat*. Las enzimas de restricción usadas fueron *Hha I* y *Sst I*; s/d indica muestras sin digerir. B) Se indican los tamaños esperados para cada digestión.

Una vez que se determinó que el sistema de evaluación co-transcripcional funciona adecuadamente, se realizó un RT-PCR multiplex semicuantitativo. Para cada ensayo se incluyó el plásmido pgfpUAAA control y cada uno de los plásmidos pcat mutantes, como se muestra en la figura 20. Se realizaron amplificaciones a 16, 20, 24 y 30 ciclos; dado que a 16 y 20 ciclos no se observaron bandas de amplificación, solamente se muestran los resultados obtenidos a 24 y 30 ciclos (Fig. 20).

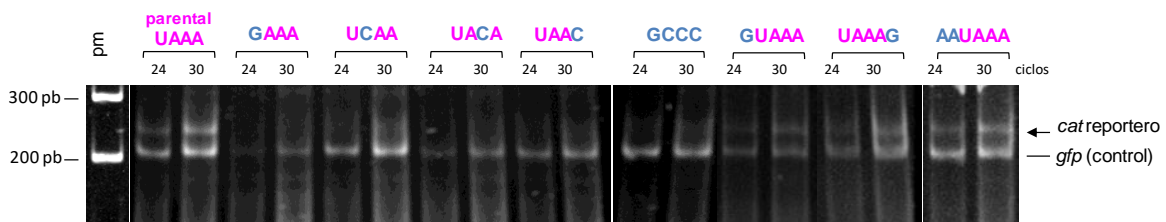


Fig. 20. Análisis de la población de mRNAs de *cat*. Los productos de amplificación del ensayo RT-PCR multiplex semicuantitativo fueron verificados por electroforesis en un gel de acrilamida al 5% teñido con bromuro de etidio. La banda inferior (~200 pb) corresponde al control interno *gfp*, mientras que la banda superior (~235 pb) corresponde a *cat*. Se encuentran de color azul los nucleótidos que fueron cambiados en el motivo UAAA, del lado izquierdo se encuentran todas las mutantes del tetranucleótido, mientras que del lado derecho están aquellas que delimitan al motivo UAAA. pm (marcador de peso molecular de 100 pb).

En todos los casos la banda de amplificación de *gfp* se observa bien definida. En el caso de la expresión de mRNA de *cat*, éste se expresa como una población aparentemente homogénea de transcritos cuando proviene del vector parental pcat-UAAA, o bien cuando su expresión proviene de plásmidos donde se encuentra íntegro el motivo UAAA (pcat-gUAAA, pcat-UAAAg, pcat-

aaUAAA). Esta banda bien definida refleja una población homogénea de mRNA, sugiriendo un solo sitio de poliadenilación en estos transcritos. En contraste con lo anterior, los productos de RT-PCR derivados de los plásmidos que carecen del motivo UAAA (GAAA, UCAA, UACA, UAAC y GCCC), se observa como una zona tenue de amplificación difusa de *cat*, sugiriendo que esta población de mRNA es heterogénea en la región 3'UTR, de acuerdo a los datos obtenidos anteriormente (Fig. 16D). Más aún, la poca abundancia de estas bandas de amplificación podría ser un reflejo de la producción de transcritos no poliadenilados que no serían amplificados y visualizados mediante esta estrategia.

Estos resultados demuestran que cualquier cambio en el motivo UAAA causa una disminución en la abundancia relativa de los mRNA poliadenilados, como consecuencia de un sitio pobremente definido para el procesamiento y poliadenilación de los transcritos.

5.2.3 La relocalización del motivo UAAA conlleva a la relocalización del sitio de poliadenilación de los transcritos.

Los resultados anteriores indican que el motivo UAAA es fundamental para dirigir la poliadenilación de los mRNA de *T. vaginalis*, pero no es suficiente para definir el sitio de poliadenilación; los elementos adicionales que hemos propuesto son la señal de procesamiento $Y\downarrow A_{2-5}UU$ y el elemento río abajo rico en U's (DSE), que podrían tener un papel importante en definir el sitio de poliadenilación de los mRNA de este organismo. Para corroborar el papel que juega el motivo UAAA en el procesamiento de los mRNA, se planteó mover esta señal y relocalizarla en dos sitios río abajo al sitio original, que estuvieran localizados en el contexto adecuado en relación a una posible señal de procesamiento y un elemento rico en U's. Se partió de la mutante pcat-UcAA, y mediante mutagénesis dirigida se construyeron dos nuevos plásmidos: pcat-UAAA74 (relocalización cercana a 74 nt río abajo del codón de *cat*) y pcat-UAAA118 (relocalización lejana a 118 nt río abajo del codón de paro de *cat*). De esta manera se hicieron dos reconstrucciones de la señal de poliadenilación en un contexto similar al parental, como se muestra en la figura 21.

A. Relocalización cercana pcat-UAAA74



B. Relocalización lejana: pcat-UAAA118

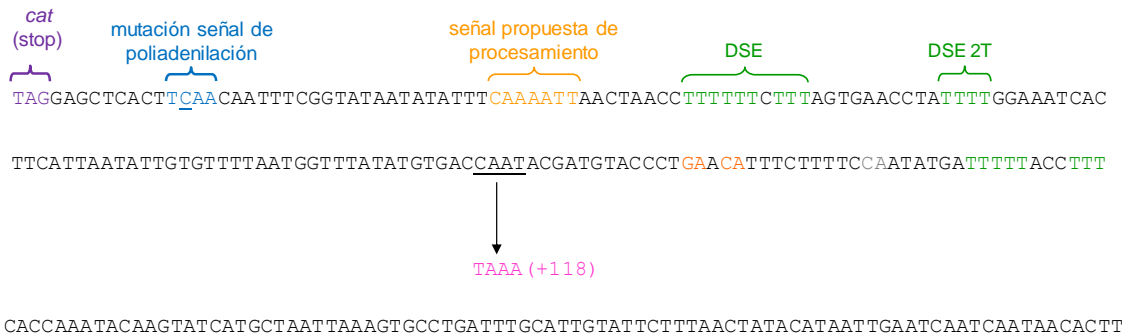


Fig. 21. Relocalización de la señal de poliadenilación UAAA. A) Relocalización cercana (74 nt río abajo al codón de paro de *cat*). La secuencia GAAA se cambió a TAAA mediante mutagénesis dirigida. B) Relocalización lejana (118 nts río abajo al codón de paro de *cat*), la secuencia CAAT se cambió a TAAA mediante mutagénesis dirigida. En ambos casos en verde se muestran las regiones ricas en U's y en naranja los potenciales sitios de procesamiento de los mRNA.

Se espera que en estos dos casos, el motivo UAAA dirija la poliadenilación al sitio de corte más cercano, ya que se encuentra en un contexto favorable en cuanto a secuencia. Como se describe en el apartado de Metodología, los plásmidos se transflectaron a células de *T. vaginalis* en cultivo y a partir del RNA total se realizó un ensayo de RT-PCR. Los tamaños esperados de los productos de amplificación, en el caso de que existiera una relocalización del sitio de procesamiento, serían ~280 pb para la relocalización cercana (pcat-UAAA74) y ~ 345 pb para la relocalización lejana (pcat-UAAA118) (Fig. 21).

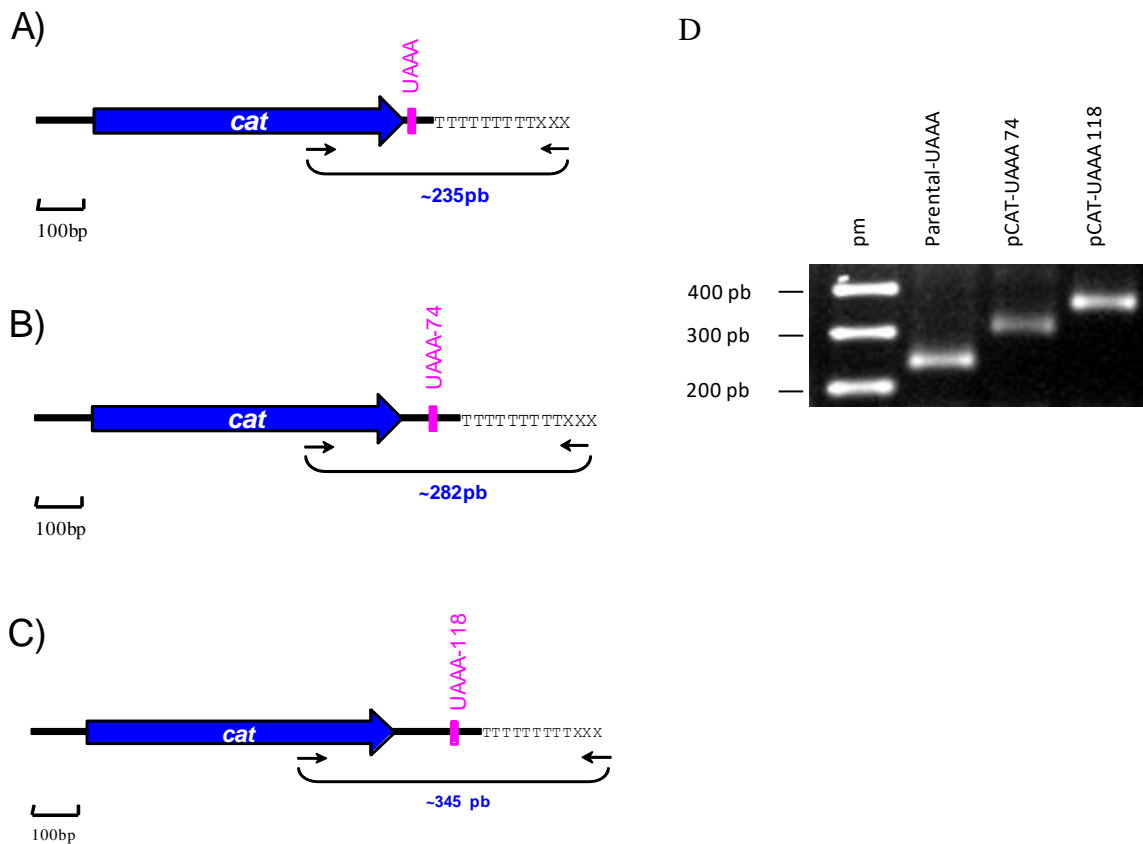


Fig. 22. A) Esquema del ensayo de RT-PCR usado para estimar el sitio de poliadenilación de los mRNA de *cat* derivados de los plásmidos pcat-UAAA74 (B) y pcat-UAAA118 (C), comparados con el plásmido parental-UAAA (A). D) Productos del ensayo de RT-PCR derivado de los plásmidos pcat-UAAA74 y pcat-UAAA18, analizados mediante un gel de agarosa 2.2 % teñido con bromuro de etidio, pm (marcador de peso molecular de 100 pb).

En la figura 22 observamos que los productos de amplificación por RT-PCR derivados de pcat-UAAA74 y pcat-UAAA118, aparecen como una banda discreta del tamaño esperado; esto sugiere que el procesamiento del mRNA de *cat* se dirigió a la región esperada, con el predominio de una población homogénea.

Para conocer el sitio exacto de procesamiento en ambos casos, los productos de amplificación se clonaron en el vector PGEM-T Easy (Promega) y se secuenciaron varias clonas de cDNA seleccionadas al azar. Los resultados del análisis de las secuencias se muestran en la figura 23. Encontramos que a pesar de haber observado una banda definida en la electroforesis de la Figura 22, los productos del RT-PCR derivados de pcat-UAAA74 y pcat-UAAA118 se encuentran procesados en varios sitios muy cercanos entre sí. Esto se puede deber a que existe más de un sitio posible de procesamiento en la región en la cual fue relocalizada la señal de poliadenilación UAAA.

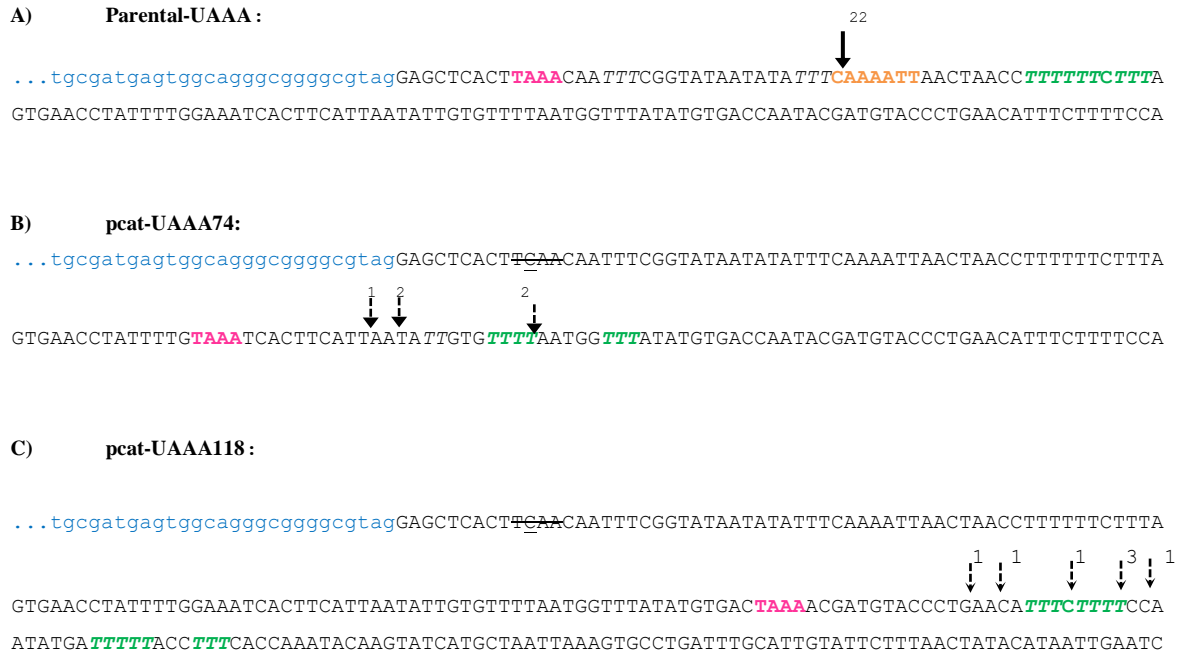


Fig. 23. Identificación del sitio de procesamiento de los mRNA de *cat* derivados de los plásmidos indicados. A) Secuencia del vector parental en donde se observa un solo sitio de procesamiento. B) Sitios de procesamiento y poliadenilación de los cDNAs derivados del plásmido pcat-UAAA74 (relocalización cercana). C) Sitios de procesamiento y poliadenilación de los cDNAs derivados del plásmido pcat-UAAA118 (relocalización lejana). El sitio de procesamiento y poliadenilación de los transcritos se indica con una flecha. El número de clonas identificadas en cada sitio de procesamiento se indica en la parte superior de las flechas. En rosa se muestra al motivo TAAA. En letras minúsculas y azules se indica parte de la región codificadora del gen *cat*. En letra verde y cursiva se indican las regiones ricas en U's. En naranja se indica el sitio de procesamiento de la secuencia parental.

Los resultados anteriores muestran que el hecho de colocar el tetranucleótido UAAA a 74 y 118 nt río abajo a su posición original permitió dirigir el corte y la poliadenilación del mRNA siempre y cuando exista un contexto adecuado, es decir, una señal de procesamiento y regiones ricas en U's. Hay que resaltar que en estos casos la poliadenilación de los transcritos se llevó a cabo en varios sitios de corte posibles que cumplen de manera laxa con la secuencia teórica propuesta por nosotros: $Y\downarrow(A_{2-5})UU$. Por lo tanto, decidimos construir un nuevo plásmido mediante mutagénesis dirigida en el cual se introdujo la señal procesamiento CAATT; a este plásmido, derivado del pcat-UAAA118, se le llamó pcat-UAAA118CS (Fig. 23).

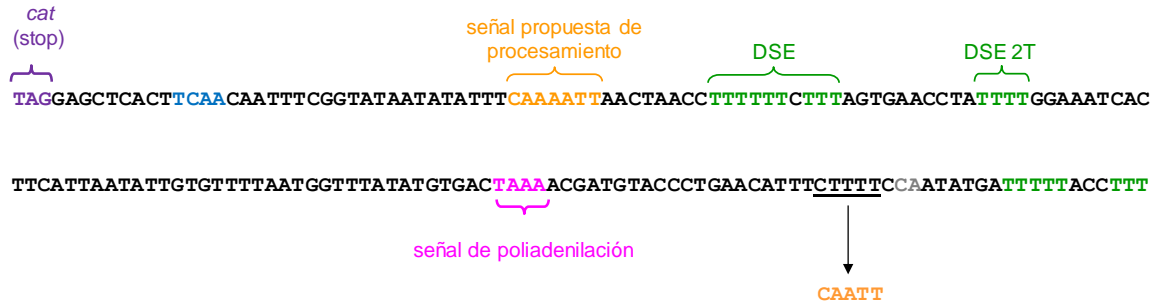


Fig. 24. Secuencia parcial del plásmido pcat-UAAA118CS. En este caso se introdujo la señal de procesamiento CAATT en el plásmido pcat-UAAA118. Los nucleótidos subrayados (CTTTT) se cambiaron por el motivo CAATT, señalado en amarillo. Este motivo se localiza 20 nt río abajo de la señal de poliadenilación UAAA.

En el caso del plásmido pcat-UAAA118CS se reconstruye el contexto y los motivos propuestos para dirigir el proceso de poliadenilación en *T. vaginalis*. La construcción de este plásmido cat-UAAA118CS se verificó mediante secuenciación, y con ella se transfectaron células de *T. vaginalis* en cultivo; los transcritos del reportero *cat* se analizaron por RT-PCR, de la manera ya descrita. En el análisis de RT-PCR se espera un producto de amplificación de ~345 pb (Fig. 25).

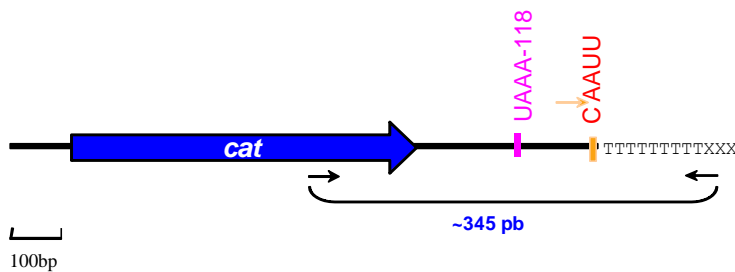


Fig. 25. Esquema del ensayo de RT-PCR usado para estimar el sitio de poliadenilación de los mRNA de *cat* derivados de los plásmidos pcat-UAAA118CS.

Este producto de amplificación se clonó y se secuenciaron varias clonas de cDNA seleccionadas al azar (Fig. 26). Se encontró que en todos los casos, el sitio de corte y poliadenilación de los transcritos se llevó a cabo en el sitio propuesto C↓AATT.

pcat-UAAA118CS:

```
...tgcgatgagtgccagggcgggcgtagGAGCTCACTTCAACAATTTCCGTATAATATATTTCAAATAACTAACCTTTTTTCTTTA
                                                                                                     6 ↓
GTGAACCTATTTGGAAATCACTTCATTAATATTGTGTTTTAATGGTTTATATGTGACTAAACGATGTACCCTGAACATTTCAATTCCA
ATATGA TTTTACC TTTACCAAATACAAGTATCATGCTAATTAAAGTGCCTGATTTGCATTGTATTCTTTAACTATACATAATTGAATC
```

Fig. 26. Sitio de procesamiento de los transcritos derivados del plásmido pcat-UAAA118CS. La flecha indica el sitio de procesamiento, y el número indica la cantidad clonas de cDNA procesadas en ese sitio. En rosa se muestra al tetranucleótido TAAA y con letras minúsculas azules se muestra parte de la región codificadora del gen *cat*. En verde y cursivas se muestran las regiones ricas en U's. En naranja se muestra el sitio de procesamiento C↓AATT.

Los resultados anteriores confirman que el motivo UAAA es la señal necesaria para dirigir la poliadenilación de los transcritos en *T. vaginalis*. Sin embargo, esta señal no es suficiente por ella misma, sino que requiere de elementos específicos que definen el sitio de corte, como son el sitio de procesamiento Y↓(A₂₋₅)UU y una región rica en Us

5.2.4 Interacción de la señal de poliadenilación UAAA con factores en extractos proteicos de *T. vaginalis*

Se ha reportado que la señal de poliadenilación en eucariontes superiores es el hexanucleótido AAUAAA y es sitio de unión para el factor CPSF-160, el cual es parte de la enorme maquinaria que participa en la poliadenilación (Wahle, 1999; Mandel, 2008). En el caso de *T. vaginalis*, hasta la fecha no se han reportado los factores que podrían estar participando en este proceso. Por lo anterior, como primer paso para vislumbrar la maquinaria de poliadenilación, es necesario determinar si la señal de poliadenilación UAAA es reconocida por proteínas propias del parásito que pudieran ser los factores de poliadenilación. La técnica que se empleó para este propósito fueron ensayos de retardo en la movilidad electroforética de una sonda (EMSA). Es importante recalcar que en este caso de estudio, las sondas utilizadas deberán ser de RNA ya que ésta molécula es el sustrato del reconocimiento de los factores para el corte del transcrito y su poliadenilación. La sonda que se usó es un transcrito de RNA sintetizado por medio de transcripción *in vitro*, como se describe en el apartado de Metodología (Fig. 27).

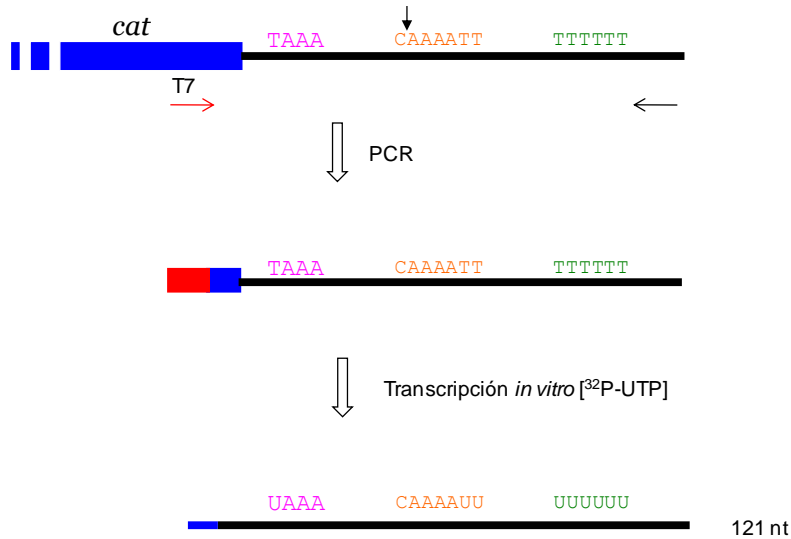


Fig. 27. Esquema del procedimiento para sintetizar la sonda de RNA usada en los ensayos EMSA. Para generar el DNA sustrato de la reacción de transcripción, se diseñó un oligonucleótido que tuviera la secuencia del promotor de T7 y que tuviera también una región complementaria al gen *cat*. Mediante PCR se generó el DNA templado usado en la transcripción *in vitro*. En esta reacción se utilizó (α - 32 P)UTP para obtener una sonda radiactiva de RNA para ser utilizada en los ensayos de retardo.

Se utilizaron dos sondas diferentes de RNA para los ensayos tipo EMSA: una sonda “parental” que lleva la señal de poliadenilación UAAA, y una sonda a la que llamamos “mutante” que en su lugar tiene la secuencia GCCC (Fig. 28 panel A); se utilizaron también extractos de proteínas totales de *T. vaginalis*, como se describe en el apartado de Metodología. El reconocimiento de una sonda por factores presentes en el extracto proteico se visualiza por una diferencia en su movilidad electroforética, ya que la unión de una sonda a proteínas retardará su movilidad al ser comparada a con el desplazamiento de la sonda sola.

A)



B)

C)

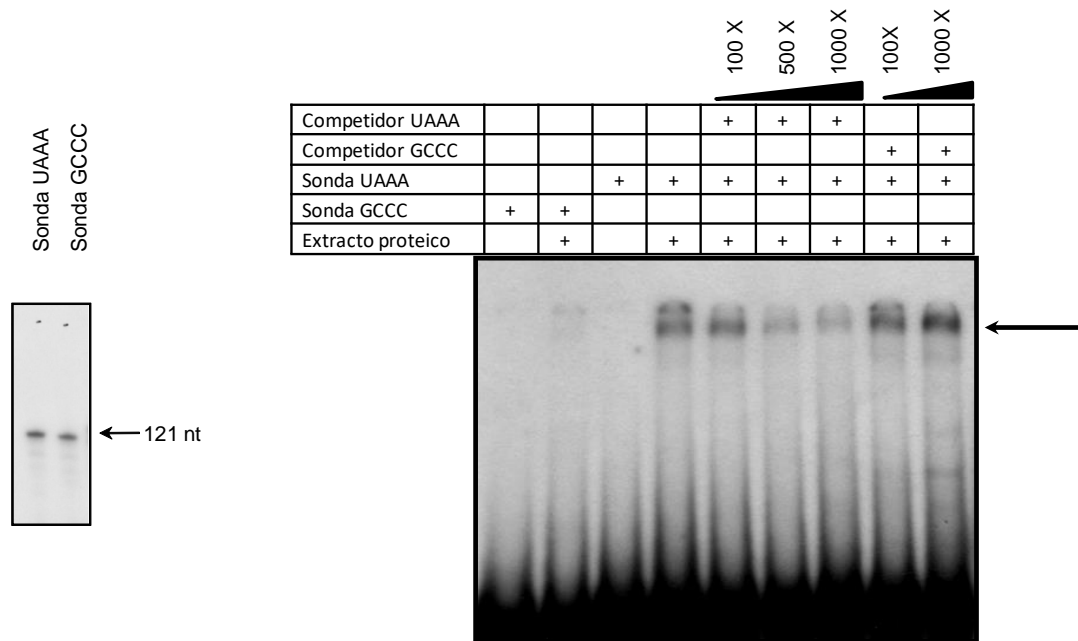


Fig. 28. Interacción de la señal de poliadenilación UAAA con extractos proteicos de *T. vaginalis*. A) Secuencia de las sondas de RNA usadas para los ensayos EMSA. Contiene 39 nt de la región codificadora de *cat* y 79 nt río abajo del codón de paro que corresponden a la región intergenica 3' de actina de *T. vaginalis*. B) Integridad de las sondas de RNA, evaluadas mediante electroforesis en un gel de acrilamida (Urea 8M) y autorradiografía. C) Ensayos de retardo de la movilidad electroforética (EMSA). Las competencias se hicieron con 100, 500 y 1000X de transcritos no radiactivos UAAA y GCCC.

En la Fig. 29 panel C se puede observar la formación de un complejo de alto peso molecular con la sonda UAAA y factores proteicos de *T. vaginalis*. Esta unión es específica ya que es posible competirla usando un transcrito UAAA no marcado radiactivamente. En contraste, cuando en el ensayo de competencia se usa el transcrito no radiactivo GCCC, no se observa una disminución en la intensidad de la banda. Estos resultados indican la presencia de al menos una proteína que es capaz de reconocer la señal de poliadenilación UAAA, y que pudiera ser equivalente al factor CPSF-160 descrito en otros eucariontes. Es importante señalar que en los ensayos EMSA no fue posible competir por completo la banda de retardo, lo que pudiera deberse a que la sonda empleada lleva, además de la señal UAAA, otras señales que participan en la poliadenilación como son el sitio de procesamiento y la región rica en U's, las cuales también son sitios reconocidos por factores proteicos en otros organismos.

Nuestro grupo de investigación se encuentra interesado en estudiar la maquinaria que participa en la poliadenilación de los mRNA de este parásito. En un intento por identificar proteínas ortólogas que formen parte de esta maquinaria, se han realizado análisis bioinformáticos por medio de los cuales se han identificado algunas de estas proteínas como son: una proteína con homología a CPSF-160,

la cual sería la encargada de reconocer a la señal de poliadenilación; una proteína con homología a CstF que podría reconocer a la región rica en Us, Simplekina y poliA-polimerasa, entre otras.

5.2.5 Actividad de CAT medida en los extractos proteicos

Una estrategia indirecta para analizar el efecto que tienen los cambios en la poliadenilación de transcritos, fue evaluar la actividad bioquímica de la enzima cloranfenicol acetil transferasa (CAT), codificada por el gen reportero *cat*, en extractos proteicos de cultivos de *T. vaginalis* transfectados con los distintos plásmidos pcat. En todos los ensayos de transfección se incluyó un control positivo (plásmido pcat-UAAA) y un control negativo (células transfectadas con solución amortiguadora), como se describe en la sección de Metodología. Los datos de este apartado son resultado del presente trabajo y del trabajo de tesis de M. C. Ma. Guadalupe Barrera.

En la tabla 5 se muestran los resultados obtenidos de la medición de la actividad bioquímica de CAT en los extractos provenientes de células transfectadas con las distintas construcciones plasmídicas: parental, mutantes puntuales del motivo UAAA, mutantes que delimitan el motivo (GUAAA, UAAAG), introducción de la señal canónica (AAUAAA) y el plásmido revertante UAAA.

Tabla 5. Actividad de CAT en cultivos de *T. vaginalis* transfectados con los plásmidos pcat

Construcciones	Actividad de CAT % (SD)
Parental UAAA	100
GCCC	11.7 (4.05)
GAAA	10.2 (8.5)
UCAA	9.9 (3.1)
UACA	10.3 (5.8)
UAAC	14.5 (9.3)
GUAAA	146.9 (7.9)
UAAAG	101.6 (7.8)
AAUAAA	127.4 (22.4)
Revertante UAAA	122.3 (31.5)

Como se observa en la tabla 5, las células transfectadas con plásmidos que llevan cambios en el tetranucleótido UAAA tienen una actividad de CAT que va del 10 al 15 % de la actividad mostrada por el plásmido parental UAAA. En contraste, las construcciones que delimitan al tetranucleótido (GUAAA, UAAAG), la revertante y la que lleva el hexanucleótido AAUAAA tienen actividades semejantes al parental. Dado que todos los plásmidos poseen el mismo promotor, asumimos que los niveles de transcripción son similares para todos los casos. La única diferencia entre los transcritos es la señal de poliadenilación que, de acuerdo con nuestros resultados anteriores, al mutarse produce transcritos aberrantes. Proponemos que estas mutantes producirían un menor número de mRNAs traducibles y, por lo tanto, una menor abundancia de la enzima CAT reflejada en un menor nivel de actividad.

Estos resultados muestran la importancia que tiene el tetranucleótido UAAA en el proceso de poliadenilación ya que la presencia de éste permite que se lleve a cabo un procesamiento adecuado de los mRNA lo cual influye directamente en una población estable de transcritos, sin embargo, también se observó que es necesario que esta señal este dentro de un contexto apropiado (sitio de procesamiento y el elemento río abajo rico en U's).

5.3. Análisis funcional del sitio de procesamiento del pre-mRNA y la región río abajo rica en U's

El análisis funcional de la señal propuesta de procesamiento $Y\downarrow A_{2-5}UU$ lo inicié como parte de mi trabajo de tesis de licenciatura. Se utilizó el mismo vector de expresión usado en el análisis funcional del tetranucleótido UAAA, que contiene al gen reportero *cat* (parental-UAAA) y que además contiene río abajo la secuencia CAAAATT. Se ha visto ya que esta secuencia es utilizada como sitio de procesamiento de los transcritos de *cat*, los cuales se poliadenilan después del residuo $C\downarrow$.

Para completar el análisis funcional de la señal propuesta de procesamiento, se construyeron las mutaciones descritas en la Fig. 29, con el propósito de determinar si hay alguna base preferencial para el corte y la poliadenilación de los transcritos. Debe tenerse en consideración que esta señal propuesta de procesamiento es más grande que la reportada para otros organismos (Wahle, 1999). Por lo anterior, resulta interesante determinar si realmente es requerida una señal de procesamiento tan grande en *T. vaginalis*. Al igual que las mutaciones sobre la señal de poliadenilación, las

mutaciones sobre la señal propuesta de procesamiento se realizaron mediante mutagénesis dirigida. Todos los plásmidos fueron verificados mediante secuenciación.

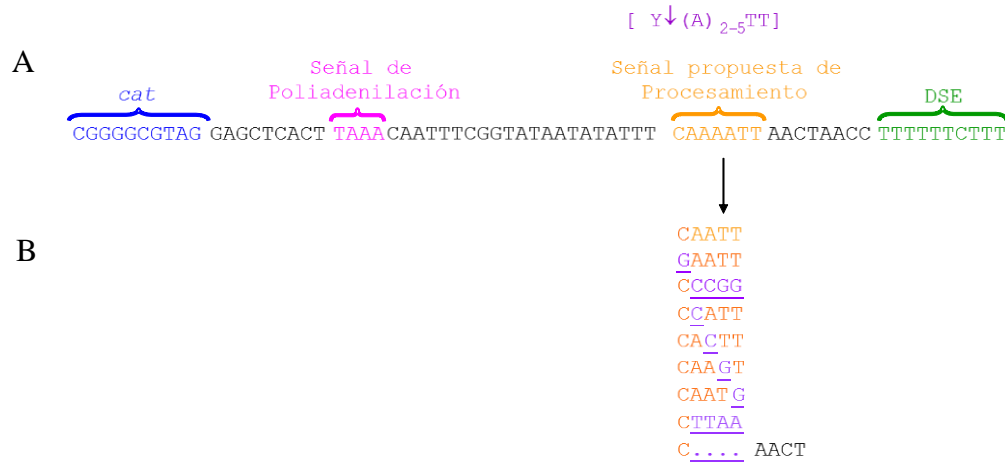


Fig. 29. Mutaciones realizadas sobre la señal propuesta de procesamiento para los mRNA de *T. vaginalis*. A) Secuencia parcial del plásmido parental. B) Mutaciones realizadas sobre la señal de procesamiento, se encuentran subrayados los nucleótidos que fueron cambiados, la última mutación realizada es una delección del tetranucleótido AATT (marcada con puntos suspensivos).

Los plásmidos para analizar la señal de procesamiento fueron transfectadas en células de *T. vaginalis* en cultivo, y el análisis de los transcritos se realizó de la manera descrita para el análisis de la señal de poliadenilación. Los productos obtenidos a partir del ensayo de 3' RACE fueron clonados y secuenciados, y el análisis de las estas secuencias se muestra en la figura 30.

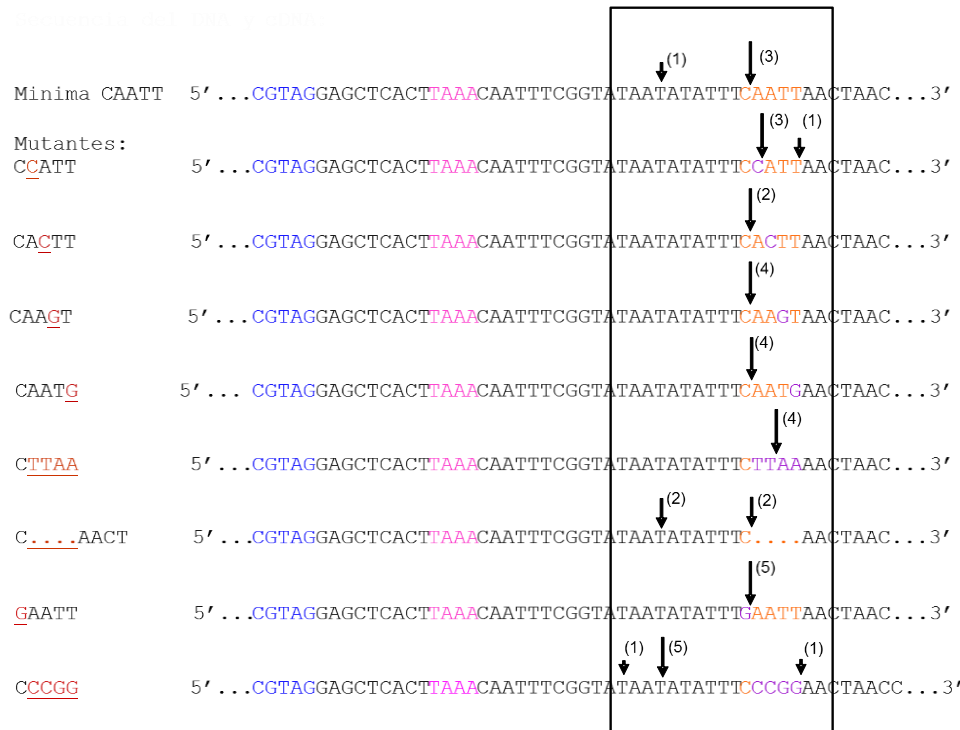


Fig. 30. Secuencia de los cDNAs derivados de los plásmidos mutantes en el sitio de procesamiento. La flecha indica el sitio de corte y poliadenilación, y el número indica la cantidad de secuencias identificadas en ese sitio.

De los resultados mostrados en la figura 30, se puede concluir lo siguiente:

1. En todos los casos, el procesamiento y poliadenilación de los transcritos se lleva a cabo en un nucleótido adyacente y río arriba a una Adenina.
2. El nucleótido preferente sobre el cual se lleva a cabo el corte es una pirimidina (C o T), aunque también puede procesarse en una G, aunque con mucha menor frecuencia.
3. La distancia de la señal de poliadenilación al sitio de procesamiento se acota a una región entre 11 y 30 nucleótidos, como se había propuesto previamente. Esto se determina también en la mutante donde se realizó la delección del tetranucleótido AATT, ya que a pesar de generar una nueva señal de corte y procesamiento similar a la propuesta (CAACT), el procesamiento de los mRNA no siempre se lleva a cabo en el mismo sitio debido a que dicha delección modifica la distancia entre la señal de poliadenilación y la región rica en U's.

De manera adicional al análisis anterior, se evaluó la actividad bioquímica de la proteína reportera CAT como un efecto indirecto a las mutaciones realizadas. Contrario a lo observado con las mutantes de la señal de poliadenilación, ninguna de estas mutaciones tiene un efecto negativo sobre la actividad de CAT, ya que la actividad obtenida es similar al parental (datos no mostrados). Esto pudiera deberse a que las mutaciones realizadas al sitio de procesamiento solo afectan el sitio de corte mientras que la eficiencia de poliadenilación de los transcritos se mantiene, ya que se mantiene un sitio funcional para su procesamiento.

5.3.1. Mutagénesis dirigida en la región río abajo rica en U's.

Se ha reportado la presencia de varios motivos que participan en la poliadenilación de los mRNA. Además de los ya mencionados, otro motivo importante para el proceso de poliadenilación es el elemento río abajo rico en U's (en inglés downstream element- DSE), el cual es reconocido por el factor CstF-64 en otros eucariontes.

El plásmido parental (UAAA) con el que se hicieron las mutantes en la señal de poliadenilación y en la señal de procesamiento, contiene dos regiones ricas en U's: la primera se encuentra a 14 bases (DSE) y la segunda a 34 bases (DSE 2T) río abajo al sitio de procesamiento del mensajero. Estas

distancias se encuentran dentro del rango de distancias que ha sido reportado previamente para otros organismos, que puede de ir de 30 a 100 bases.

Para analizar si estos elementos participan en el proceso de poliadenilación de los mRNA de *T. vaginalis*, se llevó a cabo una mutagénesis dirigida en ambas señales. En el caso de este tipo de elementos, no es útil realizar mutaciones puntuales ya que las secuencias de estos motivos no están conservadas, en estos casos es necesario truncar las corridas de U's. Por lo tanto, en una primera construcción se sustituyó el elemento DSE (TTTTTCTTT → TAACATCTAT), mientras que en una segunda mutación se cambió también el DSE 2T (TTTT → TGGT) (Fig. 32).

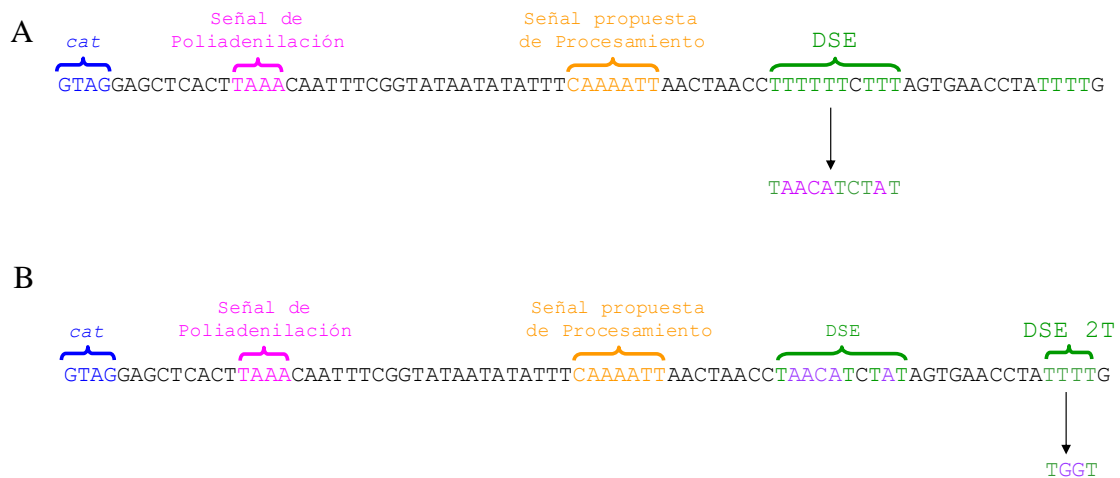


Fig. 31. Mutagénesis dirigida a las dos regiones ricas en U propuestas: DSE y DSE 2T. A) Secuencia parental del elemento DSE: la flecha indica la mutación realizada. B) Secuencia de la mutante DSE sobre la que se hizo el segundo cambio (DSE2T), señalado con una flecha.

La mutación de ambas construcciones fue verificada por secuenciación. Siguiendo la misma estrategia antes descrita, se analizó mediante 3' RACE el efecto de estas mutaciones sobre el proceso de poliadenilación de los RNA mensajeros de *cat*. El análisis de las secuencias se muestra en la figura 32:

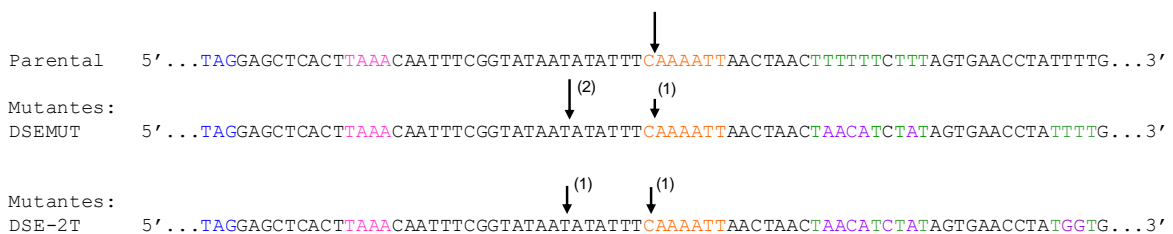


Fig. 32. Sitio de procesamiento de los transcritos derivados de los plásmidos mutantes en la región rica en U's. La flecha indica el sitio de corte y poliadenilación, y el número indica el número de clonas de cDNA procesadas en ese sitio.

Al comparar los sitios de poliadenilación de las mutantes DSEMUT y DSE-2T con respecto al parental, observamos que algunos mRNA están siendo procesados en el mismo lugar y otros están siendo procesados río arriba en la secuencia T↓ATATT, la cual curiosamente fue uno de los sitios más usados en las mutantes de la señal de procesamiento. De la misma manera, en el caso de las mutantes de la señal de procesamiento, al medir la actividad de la proteína reportera CAT obtenida de parásitos transfectados con estas dos mutantes, se observan niveles parecidos al parental (datos no mostrados).

Estos resultados no nos permiten obtener una conclusión clara sobre cuál es el papel de este elemento río abajo en el proceso de poliadenilación. Debe tenerse presente que este tipo de elementos son difíciles de analizar ya que se encuentran en una región extensa río abajo del sitio de corte del mRNA y además tienen poca conservación en su secuencia.

6. CONCLUSIONES

- El tetranucleótido UAAA es la señal de poliadenilación en el protozooario *T. vaginalis*.
- El tetranucleótido UAAA es reconocido por factores proteicos de *T. vaginalis* hasta ahora no caracterizados.
- Por sí mismo, el tetranucleótido UAAA no es suficiente para dirigir el procesamiento de los pre-mRNA; es necesario que se encuentre en un contexto apropiado que incluye un sitio de corte dado por la secuencia $Y\downarrow A_{2-5}UU$, localizada a una distancia de 11-30 pares de bases, y una secuencia río abajo rica en U's.

7. PERSPECTIVAS

- Identificar y caracterizar los factores proteicos que interaccionan con el motivo UAAA para identificar a los miembros de la maquinaria proteica que intervienen en el procesamiento de los pre-mRNA.
- Realizar cambios en la distancia en la región rica en U's río abajo de la señal de procesamiento para completar el análisis de esta región.

8. ARTÍCULO PUBLICADO



Functional Analysis of Sequence Motifs Involved in the Polyadenylation of *Trichomonas vaginalis* mRNAs

Vanessa Fuentes,^a Guadalupe Barrera,^a Joaquín Sánchez,^b Roberto Hernández,^a and Imelda López-Villaseñor^a

Departamento de Biología Molecular y Biotecnología, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, México D. F., México,^a and Facultad de Medicina, Universidad Autónoma del Estado de Morelos, Cuernavaca, Morelos, México^b

Synthesis of functional mRNA in eukaryotes involves processing of precursor transcripts, including the addition of a poly(A) tail at the 3' end. A multiprotein complex recognizes a polyadenylation signal, generally the hexanucleotide AAUAAA in metazoans, to direct processing of the pre-mRNA. Based on sequence analysis of several cDNAs, we have previously suggested that the UAAA tetranucleotide (which may include the UAA translation stop codon) could be the polyadenylation signal in *Trichomonas vaginalis*, a parasitic protozoan that causes human trichomoniasis. This proposal is analyzed here with the aid of a transient-expression system of a reporter gene (*cat* flanked by *T. vaginalis* actin noncoding sequences). When cells were transfected with a plasmid bearing the original 3' untranslated region (UTR) sequence containing the UAAA motif, the resulting *cat* mRNA was polyadenylated similarly to the endogenous actin mRNA. Base changes in the UAAA sequence produced alterations to the polyadenylation site of the reporter mRNAs, while nucleotide substitutions at either side of UAAA did not. Furthermore, relocation of the UAAA motif redirected the processing and polyadenylation of the reporter mRNA. In addition, a pre-mRNA cleavage site for polyadenylation was defined. Interaction of *T. vaginalis* proteins with the UAAA motif was shown by electrophoretic mobility shift assays. Based on our findings, we provide evidence that in *T. vaginalis* the UAAA tetranucleotide has a role equivalent to that of the metazoan consensus AAUAAA polyadenylation signal.

Trichomonas vaginalis is a parasitic protozoan that causes trichomoniasis, one of the most common sexually transmitted infections in humans (34). Phylogenetic analyses based on small-subunit (SSU) rRNA have placed this organism among the earliest-diverging eukaryotic branches (27) and, more recently, in the Excavata supergroup according to the phylogenetic egalitarian model (7). Synthesis of functional mRNA in eukaryotes involves polyadenylation of the precursor transcripts. *T. vaginalis* mRNAs are polyadenylated, but this process has not been experimentally addressed. This situation contrasts with the considerable effort directed toward the analysis of transcription by RNA polymerase II in this organism: core promoters have been found to consist solely of a highly conserved initiator element (Inr), IBP39 is a protein involved in promoter recognition, and no homologues have been identified in other species (14). Regarding RNA splicing, relatively few genes (about 65) in the *T. vaginalis* genome appear to contain an intron (4), and the ones recognized are short and characterized by a conserved 12-nucleotide (nt) sequence (30); spliceosomal components also have been identified (26).

The 3' poly(A) tail of mRNA is involved in regulatory mechanisms such as mRNA stability, export of mRNA from the nucleus to the cytoplasm, and initiation of translation (22, 33). Recognition of the polyadenylation site is coordinated by interactions between processing factors and RNA sequence elements (*cis* elements) located both upstream and downstream from the poly(A) addition site. The consensus polyadenylation signal in metazoa is the hexanucleotide AAUAAA, which is located within the transcribed 3' untranslated region (3' UTR) of the mRNA, 10 to 30 nucleotides upstream of the actual cleavage site for polyadenylation. This motif is recognized by the cleavage/polyadenylation specificity factor 160 subunit (CPSF-160). Additional sequence motifs have been described as being relevant for the polyadenylation process; the dinucleotide CA has been identified as the cleavage site, recognized by the CPSF-73 endonuclease. A U-rich or

GU-rich region located up to 30 bases downstream from the cleavage site is also relevant in metazoans. This U-rich element is recognized by cleavage stimulation factor (CstF), which interacts with CPSF to cleave the mRNA and start the addition of the poly(A) tail using the poly(A) polymerase in a template-independent manner (5, 17, 33). In mammals, UGUA elements are generally present within 40 to 100 nucleotides upstream of the poly(A) site (12, 31). In the yeast *Saccharomyces cerevisiae*, sequences directing this process are less conserved than in metazoans. Nonetheless, four sequence elements have been identified: the AU-rich element, the A-rich positioning element, the cleavage site, and the U-rich elements flanking the cleavage site (11, 20). The motif UA(A/U)UU in *Entamoeba histolytica* and the sequence AGURAAAY in *Giardia* spp. have been proposed as the equivalent polyadenylation sequences (1, 15), although no experimental evidence that demonstrates their function has been provided so far.

Based on a comparative sequence analysis of the 3' UTRs of several *T. vaginalis* cDNAs, we have previously proposed the UAAA tetranucleotide to be the polyadenylation signal in *T. vaginalis* (9). It was shown there that the mRNA in this organism poses a very short 3' UTR and that the consensus eukaryotic polyadenylation signal (AAUAAA) is rarely found. The single identified consensus sequence related to this motif was the UAAA tetranucleotide in the 3' UTR, which may overlap the UAA translation stop codon. Additional *cis* elements were also proposed:

Received 4 January 2012 Accepted 26 March 2012

Published ahead of print 30 March 2012

Address correspondence to Imelda López-Villaseñor, imelda@biomedicas.unam.mx.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.
doi:10.1128/EC.05322-11

TABLE 1 Plasmids used in this study

Plasmid name	Sequence surrounding UAAA motif ^a	Description
pcat-UAAA	...CACU <u>U</u> AAACAAU...	Parental plasmid described in Fig. 1
pcat-GCCC	...CACU <u>G</u> CCCCAAU...	Change of the four bases of the UAAA motif
pcat-GAAA	...CACU <u>G</u> AAACAAU...	Change of the first base of the UAAA motif
pcat-UCAA	...CACU <u>C</u> AACAAU...	Change of the second base of the UAAA motif
pcat-UACA	...CACU <u>A</u> CACAAU...	Change of the third base of the UAAA motif
pcat-UAAC	...CACU <u>A</u> ACCAAU...	Change of the fourth base of the UAAA motif
pcat-GUAAA	...CAC <u>G</u> UAAACAAU...	Change of the upstream base adjacent to the UAAA motif
pcat-UAAAG	...CACU <u>A</u> AAAGAAU...	Change of the downstream nucleotide adjacent to the UAAA motif
pcat-AAUAAA	.CUUCA <u>AA</u> UAAACAAU...	Introduction of two A nucleotides to generate the eukaryotic polyadenylation signal AAUAAA
pcat-revUAAA	...CACU <u>U</u> AAACAAU...	Revertant construct from pcat-UCAA
pcat-UAAA74	...UG <u>U</u> AAAUACACUUCAUUAAUUAUUGUGUUUAAUUG...	Relocation of the UAAA motif 74 nt downstream from the stop codon in pcat-UCAA
pcat-UAAA118	...AC <u>U</u> AAACGGAUGUACCCUGAACAAUUCUUUUCCA...	Relocation of the UAAA motif 118 nt downstream from the stop codon in pcat-UCAA
pcat-UAAA118CS	...ACUAAAACGGAUGUACCCUGAACAAUUC <u>AAU</u> UCCA...	Addition of a cleavage site in the pcat-UAAA118 plasmid

^aThe UAAA motif is in bold. Underlining denotes a nucleotide changed from that in the parental plasmid UAAA.

Y ↓ (A)₂₋₅UU as the site for cleavage and a tract of Us downstream from the cleavage site (9). Here, the recognition of the UAAA motif as the polyadenylation signal is experimentally addressed. Based on the evidence presented, we propose that the UAAA tetranucleotide in *T. vaginalis* is functionally equivalent to the canonical metazoan AAUAAA polyadenylation signal.

MATERIALS AND METHODS

In silico analysis. Sixty *T. vaginalis* cDNA sequences were identified and retrieved from the NCBI database. Additionally, we searched for the presence of the UAAA motif in the 3' noncoding regions of genes that have evidence of expression. These genes were obtained from the expressed sequence tag (EST) database available at www.trichdb.org. Only genes with a 100% match to an EST were included in the analysis (780 genes). The region downstream from the translation stop codon was obtained using the tool retrieve sequence by gene IDs available at www.trichdb.org and analyzed using Excel (Microsoft Inc.). Sequence logo representations of multiple-sequence alignments were prepared with the WebLogo tool, version 2.8.2 (<http://weblogo.berkeley.edu/>) (6, 24).

Parasites and culture conditions. The *T. vaginalis* isolate CNC147 was used throughout the experimental work. Parasites were grown in Trypticase-yeast extract-maltose (TYM) medium supplemented with 10% (vol/vol) heat-inactivated horse serum (GIBCO) as previously described (8).

Construction of a transient-expression vector for *T. vaginalis*. An *Spe*I-*Eco*RI-digested fragment (2,827 bp) from a *T. vaginalis* genomic actin clone (GenBank access number AF257734) was cloned in pBlue-script II KS(-) (Stratagene) (see Fig. 2A). The actin-coding sequence was substituted for the chloramphenicol acetyltransferase (CAT) gene (*cat*) coding sequence by PCR amplification, which also included the addition of *Rca*I and *Sac*I restriction sites. In this construct, the UAG stop codon was used for the *cat* gene, while the original actin 3' UTR (which included the UAA stop codon) was moved 9 bp downstream (see Fig. 2B). The purpose of this strategy was to separate the stop codon for the *cat* gene from the analyzed UAAA motif because this tetranucleotide overlapped the actin UAA stop codon in the original genomic clone. The resulting pcat-UAAA plasmid was used as the reference plasmid from which all derived mutant plasmids were constructed (Table 1). From the mutant pcat-UCAA plasmid, three plasmids were derived: the revertant pcat-revUAAA, in which the UCAA motif was reverted to UAAA by site-directed mutagenesis, and two plasmids containing the relocated UAAA motif, pcat-UAAA74 (proximal) and pcat-UAAA118 (distal); in

the pcat-UAAA118 plasmid a putative cleavage site was added 20 nt downstream from the relocated UAAA motif, thus rendering the pcat-UAAA118CS plasmid (Table 1; see Fig. 3 and Fig. 5). The QuikChange site-directed mutagenesis kit (Stratagene) was used for these constructs following the manufacturer's directions, and the specific oligonucleotides that were used for mutagenesis are listed in Table 2. All plasmids were verified by sequencing.

Transfection of *T. vaginalis* cells. *T. vaginalis* cultured cells were transfected by electroporation (28). Briefly, *T. vaginalis* cells were grown to a density of 1.5×10^6 cells/ml in TYM medium. The cells were collected and washed in cold Zimmerman solution [1.5 mM KH_2PO_4 , 132 mM NaCl, 8 mM KCl, 8 mM Na_2HPO_4 , 0.5 mM $(\text{CH}_3\text{COO})_2\text{Mg}$, 90 mM

TABLE 2 Oligonucleotides used in this study

Oligonucleotide	Sequence
CATSTOP 5	5'GCGTAGGAGCTCACTGAAACAAATTCGGTAA3'
CATSTOP 6	5'TACCGAAATGTTTCAGTGAGCTCCTACGCS3'
CATSTOP 7	5'CGTAGGAGCTCACTCAACAAATTCGGTAA3'
CATSTOP 8	5'ATACCGAAATGTTGAAGTGAGCTCCTACG3'
CATSTOP 9	5'GTAGGAGCTCACTACACAATTCGGTAA3'
CATSTOP 10	5'TATACCGAAATGTTGAAGTGAGCTCCTAC3'
CATSTOP 11	5'TAGGAGCTCACTAACCAATTCGGTAA3'
CATSTOP 12	5'TTATACCGAAATGTTGAAGTGAGCTCCTAA3'
CATSTOP 13	5'TAGGAGCTCACTGCCCAATTCGGTAA3'
CATSTOP 14	5'TTATACCGAAATGCGGCAAGTGAGCTCCTAA3'
CATSTOP15	5'AGGAGCTCACTAAAGAATTCGGTAA3'
CATSTOP16	5'ATTATACCGAAATTCCTTAAGTGAGCTCCT3'
GTA	5'GCGTAGGAGCTCACGTAACAAATTCGGTAA3'
RC-GTAAA	5'TATTATACCGAAATGTTTACGTCAGCTCCTACGCS3'
INS TAAA	5'GCGTAGGAGCTCACTTCAATAAACAAATTCGGTAA3'
RC-INS	5'TATTATACCGAAATGTTTATTGAAGTGAGCTCCTAAGCS3'
F-REGSP	5'GCGTAGGAGCTCACTTAAACAAATTCGGTAA3'
RC-REGSP	5'TATTATACCGAAATGTTTAAAGTGAGCTCCTAAGCS3'
PAS	5'TGGTTTATATGTCATAAAAGATGTAACCTGAA3'
RC-PASN	5'GTTACGGGTACATCGTTTATGACATATAAAAC3'
CS	5'TACCTGAACAATTCATTCACATATGATTTTAC3'
RC-CS	5'GTAAAAATCATATTGGAATGAAATGTTACGGGTAA3'
F-NSP	5'GTGAACCTATTTGTAATCACTCATTAAATATG3'
RC-NSP	5'CAATATAATGAAGTATTTACAAAATAGGTTAC3'
AP	5'GGCCACGGCTGACTAGTACTTTTTTTTTTTTTTT3'
AUAP-B	5'GGATCCGGCCACGGCTGACTAGTAC3'
CAT10	5'CACATCTTGCCCGCTG3'
CAT11	5'GAGCTGTGATATGGGAT3'
CAT 20	5'AAGTTGTCCATATTGGCC3'
For-imit	5'TAATACGACTCACTATAGGGAATTACACAGTACTGCGAT3'
Rev-imit	5'ATTTCAAAAATAGGTTCACT3'
For-CATqPCR	5'GGCCCGGTTTTACCATGG3'
For-GFPqPCR	5'CCGTGCTGCTGCCCGACA3'

(CH_3COO)₂Ca] and brought to 6×10^8 cells/ml. From this suspension, 400 μl of cells was mixed with 50 μl of plasmid DNA (1 $\mu\text{g}/\mu\text{l}$ in Tris-EDTA [TE] buffer [10 mM Tris, 0.2 mM EDTA, pH 7.4]) and incubated for 5 min on ice. Mock-transfected cells with 50 μl of TE buffer were used as controls. Cells were electroporated at 360 V, 1,500 μF , and 13 Ω using a BTX600 Electro Cell Manipulator (BTX). Following electroporation, transfected cells were incubated in 30 ml of serum-enriched TYM medium at 37°C for 22 h. Extracts from transfected cells were prepared as described below and were used to measure the biochemical activity of CAT and to obtain the RNA poly(A)⁺ fraction for reverse transcription-PCR (RT-PCR) assays. Transfection experiments, isolation of RNA poly(A)⁺ fractions for RT-PCR analyses, and measurements of CAT activity were repeated at least three times for each plasmid construct. For all transfection experiments, both a positive control (the pcat-UAAA plasmid) and a negative control (cells transfected with TE buffer) were included. For the cotransfection experiments, 25 μg of each plasmid (pgfp-UAAA plasmid as the internal control and the pcat-derived tested plasmid) were used under the same conditions as described above. The highest CAT activity was consistently obtained in cell extracts at 22 h after transfection.

Analysis of cat transcripts from transfected *T. vaginalis* cells. Trichomonads from 22 h after transfection were collected by centrifugation and washed with phosphate-buffered saline (PBS). Total RNA was prepared using TRIzol (Invitrogen). The RNA poly(A)⁺ fraction was isolated from total RNA using the QuickPrep Micro mRNA purification kit (Amersham Biosciences) according to the manufacturer's instructions. The 3' rapid amplification of cDNA ends (3' RACE) method was used to identify the 3' end of the *cat* mRNA and, consequently, the polyadenylation site (10). Briefly, for the synthesis of cDNA, 500 ng of poly(A)⁺ RNA and 18 pmol of AP adapter primer (Table 2) were heat denatured and used for the reverse transcriptase reaction (1 \times First Strand buffer, 0.5 mM each deoxynucleoside triphosphate [dNTP], 5 mM dithiothreitol [DTT], 100 units of SuperScript III reverse transcriptase [Invitrogen]). The reaction was carried out for 50 min at 42°C, followed by heat inactivation of the enzyme and degradation of the RNA with RNase H. The amplification of *cat* cDNA 3' ends from the total cDNA pool was performed using the CAT10 primer and the AUAP-B adapter primer (Table 2) in a 100- μl PCR mixture (40 pmol of each primer, 1 \times PCR buffer, 0.6 mM each dNTP, 3 mM MgCl₂, 5 U of Taq DNA polymerase [Invitrogen]). The amplification reaction consisted of 30 cycles of denaturing at 94°C for 1 min, annealing at 45°C for 2 min, and extension at 72°C for 3 min, followed by a last extension cycle of 5 min at 72°C. To increase the specificity of the amplicons, a second (seminested) PCR amplification was performed on the initial PCR product using the CAT11 primer and the AUAP-B adapter primer (Table 2) under similar conditions. PCR products were cloned in the Zero Blunt TOPO PCR cloning system (Invitrogen) and sequenced. Plasmid preparations and other common molecular biology techniques were performed essentially as previously described (23).

The relative homogeneity of *cat* mRNA was estimated by qualitative RT-PCR from cotransfected cells. The cDNA was synthesized from total RNA (5 μg) and 18 pmol of AP adapter primer, heat denatured, and used for the reverse transcriptase reaction (1 \times First Strand buffer, 0.5 mM each dNTP, 5 mM DTT, 100 units of SuperScript III reverse transcriptase [Invitrogen]). The reaction was carried out for 50 min at 42°C, followed by heat inactivation of the enzyme and degradation of the RNA with RNase H. The cDNA was digested overnight with appropriate restriction endonucleases to remove any possible contaminating DNA. The cDNA was extracted with phenol-chloroform (1:1) and precipitated overnight at -20°C with 2-propanol, 10 μg of tRNA, and 0.3 M sodium acetate. The cDNA was suspended in water. A multiplex PCR was used to simultaneously amplify *cat* and green fluorescent protein (GFP) cDNA 3' ends from the total cDNA pool using the For-CATqPCR primer, For-GFPqPCR, and the AUAP-B adapter primer (Table 2). PCRs were performed in 25 μl (20 pmol of each primer, 1 \times PCR Buffer, 0.6 mM each dNTP, 3 mM MgCl₂, 1.25 U of Taq DNA polymerase [Invitrogen]) and consisted of 16, 20, 24,

and 30 cycles of denaturing at 92°C for 1 min, annealing at 56°C for 1 min, and extension at 72°C for 2 min, followed by a last extension cycle of 10 min at 72°C. Five microliters of the PCR products was electrophoresed in 5% polyacrylamide gels stained with ethidium bromide and visualized by UV light. The size and homogeneity of the amplified products were evaluated with this approach. As an additional way to determine the sites where the analyzed mRNAs were polyadenylated, these PCR products were cloned in the pGEM-T Easy vector system (Promega) and some of the clones were sequenced.

Preparation of RNA probes by *in vitro* transcription. Single-stranded RNA probes (121 bases long) were used for RNA gel mobility shift assays. These RNA probes were synthesized by *in vitro* transcription and contained 39 nt upstream and 79 nt downstream of the *cat* stop codon. The DNA templates for the transcription were PCR products that were amplified from the pcat-UAAA and the pcat-GCCC plasmids with the For-*in vitro* and Rev-*in vitro* primers (Table 2). The forward primer includes the T7 RNA polymerase promoter sequence (TAATAAGACTCA CTATAGG) for *in vitro* transcription. The PCR products were cloned in the Zero Blunt TOPO cloning system (Invitrogen), and plasmids from single colonies were sequenced. The inserts were purified from the TOPO clones and were used as templates for *in vitro* RNA synthesis using T7 RNA polymerase with the MAXiScript kit (Ambion), following the manufacturer's instructions. The transcription reaction mixture was incubated at 37°C for 10 min and included 20 μCi [α -³²P]UTP (3,000 Ci/mmol; Perkin-Elmer). The RNA was precipitated with 2-propanol and dissolved in nuclease-free water. The size and integrity of the transcripts were verified on a 4% polyacrylamide-1 \times Tris-borate-EDTA (TBE)-8 M urea gel.

RNA mobility shift assays. *T. vaginalis* proteins were prepared essentially as described for the binding assays (16, 18, 19) with minor modifications as follows. *T. vaginalis* cells were grown to a density of 1.5×10^8 cells/ml in TYM medium. The cells (4×10^9 cells/ml) were collected by centrifugation, washed two times with PBS, suspended in 4 packed cell volumes (PCV) of hypotonic buffer (10 mM HEPES-KOH [pH 7.9], 1 mM EDTA, 5 mM DTT, 0.02 mM leupeptin, 0.2 mM *N*-*p*-tosyl-L-lysine chloromethyl ketone [TLCK], 1 mM EGTA, 1 mM phenylmethylsulfonyl fluoride [PMSF]), and incubated for 20 min on ice. Cells were gently lysed with a Dounce homogenizer, 4 PCV of lysis buffer (50 mM HEPES-KOH [pH 7.9], 10 mM MgCl₂, 2 mM DTT, 25% sucrose, 50% glycerol, 0.02 mM leupeptin, 0.2 mM TLCK, 1 mM PMSF) was added, 1 PCV of 4 M ammonium sulfate was slowly added with gentle stirring, and the solution was stirred for another 30 min. The cell lysate was centrifuged for 3 h at 190,000 \times g. Proteins from the supernatant were precipitated with 60% ammonium sulfate and were collected by centrifugation for 20 min at 10,000 \times g. Proteins were dissolved in 5% of the original volume of F buffer (25 mM HEPES-KOH [pH 7.9], 75 mM KCl, 12 mM MgCl₂, 0.25 mM EDTA, 2.5 mM DTT, 10% glycerol, 10 $\mu\text{g}/\text{ml}$ leupeptin, 50 $\mu\text{g}/\text{ml}$ TLCK, 0.2 mM EGTA, 1 mM PMSF) and were dialyzed overnight with 100 volumes of F buffer. A sample from the protein extracts was analyzed by electrophoresis and quantified by the Bradford method (3). Aliquots were frozen with dry ice-ethanol and stored at -70°C. Binding reaction mixtures contained 5 μg of *T. vaginalis* protein extract, 5,000 cpm of RNA labeled probe, 1.6 mM ATP, 20 mM phosphocreatine, 0.5 mM MgCl₂, 10 mM HEPES-KOH (pH 7.9), 50 mM KCl, 0.05 mM EDTA, 2 mM DTT, 0.1 mg/ml tRNA, 40 U RNase inhibitor (RNase Out; Invitrogen), and 10% glycerol in a final volume of 30 μl . The reaction mixtures were incubated for 10 min at 4°C followed by 10 min at 30°C. Samples were heated prior to loading on a 4% native polyacrylamide 0.5 \times TBE gel. Gels were dried, and the bands were detected by autoradiography. The assays were independently repeated at least three times with similar results.

Evaluation of CAT activity. Protein extracts were prepared from 15 ml of the transfected *T. vaginalis* cultures described above. Cells were collected by centrifugation, washed twice in 1 ml TEN buffer (40 mM Tris-HCl [pH 8], 1 mM EDTA [pH 8], 15 mM NaCl) and resuspended in 75 μl of 25 mM Tris-HCl (pH 8). Cells were lysed by freezing in liquid

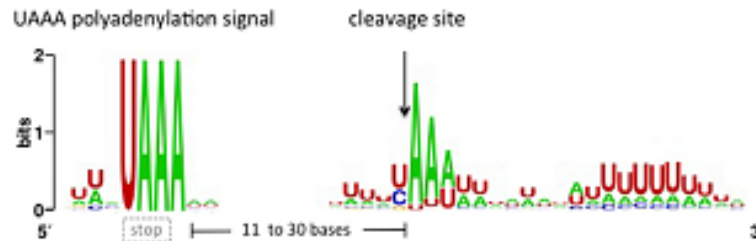


FIG 1 Sequence logo obtained by alignment of the 3' UTRs from 67 cDNAs and their corresponding genes. The overall height of the stack indicates the sequence conservation at that position, while the height of symbols within the stack indicates the relative frequency of each nucleotide at that position. The dashed gray box indicates that the UAAA motif may overlap the UAA translation stop codon.

nitrogen followed by thawing at 37°C in a water bath three times, followed by 10 min of incubation at 65°C. The soluble material was recovered from the supernatant of a 10-min centrifugation at 14,000 rpm and 4°C in a microcentrifuge. The supernatant was used to test for the enzymatic activity of the reporter protein (chloramphenicol acetyltransferase [CAT]). The protein concentration in the extracts was determined colorimetrically by the Bradford method (3) with the aid of a Bio-Rad protein assay solution. The reactions to measure the CAT activity were carried out with 5 µl of the extracted cellular protein fraction, which was incubated for 2.5 h at 37°C in the presence of 0.2 µCi [³H]chloramphenicol (New England Nuclear) and 250 µM *N*-butyryl coenzyme A in a final volume of 100 µl, as previously described (25). Thereafter, the samples were extracted with xylene, and the radioactivity in the xylene phase was measured with a liquid scintillation counter. The radioactivity values were adjusted for the protein concentration. For comparison among experiments, the CAT activity is reported as a percentage of activity compared to that of the positive control. Transfection assays, followed by the determination of CAT activity, were independently repeated at least four times for each plasmid construct, with consistent results.

RESULTS AND DISCUSSION

Presence of the UAAA motif in the 3' UTRs of *T. vaginalis* mRNAs and expressed genes. Based on the sequence comparison of the 3' UTRs from 13 *T. vaginalis* cDNA clones, the UAAA tetranucleotide was previously proposed as the potential polyadenylation signal in this organism (9). Since then, many more mRNA sequences have been made available, as is the case for EST database and the complete *T. vaginalis* genome (4), which allowed for an extended bioinformatics search of this sequence element. For this analysis we considered that if UAAA is the polyadenylation signal in *T. vaginalis*, then this motif should be present in all the mRNA 3' UTRs.

In a first approach, we searched for the UAAA motif in 67 *T. vaginalis* cDNA sequences accessed in the NCBI database. These sequences show the site where the poly(A) tail is added to the mRNA. It was found that all the cDNAs analyzed have a UAAA tetranucleotide in the 3' UTR within 11 to 30 nucleotides upstream from the poly(A) tail. Interestingly, in about one-third of the cDNAs studied, the UAAA overlapped the UAA stop codon. A particular case is that of the *T. vaginalis* syntaxin 16/TLG2-like protein cDNA (accession number AY344235), where the UAAA was not found in the short (21-nt) 3' UTR sequence; nevertheless, this motif is comprised within the last two codons preceding the stop codon (UCU AAA UAA UUU). A sequence logo diagram (<http://weblogo.berkeley.edu/logo.cgi>) was generated for the UAAA motifs and the adjacent bases in the 67 cDNA sequences; it shows that no additional bases are conserved apart from this tetranucleotide, and therefore there is no preference for penta- or

hexanucleotide variants. This alignment also revealed a preference for Y ↓ (A)₂₋₅UU as the cleavage site and the presence of a downstream U-rich region (Fig. 1). These findings are in good agreement with our previous proposal (9).

As a second genomic approach, a similar presumption was made: the UAAA element should be present in essentially all *T. vaginalis* genes that are expressed, as inferred from the trichdb EST database (<http://www.trichdb.org/trichdb/>). However, for most of the genes that are annotated here, the polyadenylation site has not been experimentally determined. We therefore considered that in these genes the UAAA motif should be present in the 3' intergenic region (including the stop codon). A total of 780 *T. vaginalis* genes with a 100% match with an EST were selected for analysis, and it was found that 99.5% of them do possess a UAAA motif either overlapping or downstream from the translation stop codon; Table 3 shows the position of the first tetranucleotide found in the 3' intergenic region. UAA is by far the most common stop codon found in *T. vaginalis*; it is used in 695 (89%) of the analyzed genes, and among these, 455 are followed by an A, thus forming a UAAA motif that might be used as the polyadenylation signal. When this is not the case or when UAG or UGA is used as a stop codon, one or more UAAA motifs can be found nearby. Again, no preference for pentanucleotides or hexanucleotides was found. This finding agrees with the known observation that *T. vaginalis* mRNAs have short 3' UTR sequences (29). It is important to point out that in this organism the single occurrence of a UAAA tetranucleotide is not sufficient to direct polyadenylation in a precise manner and that additional downstream sequences appear to be important for this process (see below). Relative to these findings, a short polyadenylation signal has been identified in *Spironucleus salmonicida*, a diplomonad that infects salmon. The position of the poly(A) tail in relation to the termination codon was mapped in 134 EST clones, revealing very short 3' UTRs. The bioinformatics analyses did not identify any strong consensus sequence outside the conserved unique termination

TABLE 3 Position of the first UAAA motif in the 3' intergenic regions of 780 *T. vaginalis* expressed genes

Position of UAAA motif downstream from stop codon	% of genes
Overlapping UAA stop codon	58.3
1–25 bases	18.5
26–50 bases	14.8
51–100 bases	5.5
101–400 bases	2.4

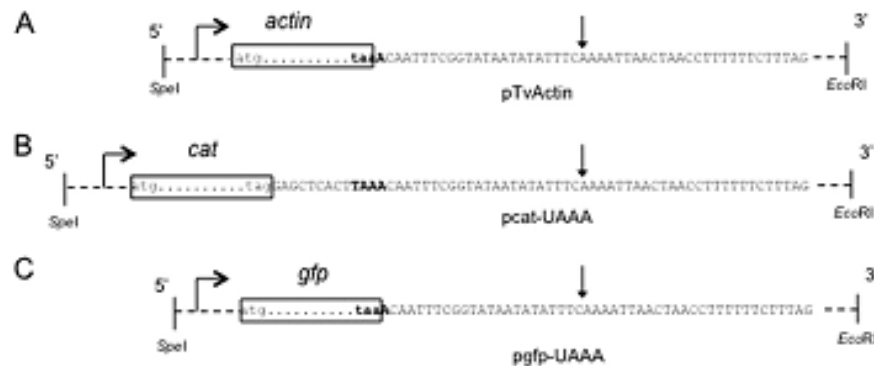


FIG 2 The reporter system to study RNA motifs involved in polyadenylation. The actin gene coding region from a genomic *T. vaginalis* clone (pTvActin) (A) was replaced with the *cat* coding region (pcat-UAAA) (B) or with the GFP gene coding region (pgfp-UAAA) (C). The coding regions are represented as dots between the initiation (atg) and the stop (taa or tag) codons inside boxes. Dashed lines represent upstream (934-bp) and downstream (762-bp) noncoding regions of the *T. vaginalis* actin gene. The DNA sequence is shown. The TAAA motif is in bold. In the pcat-UAAA plasmid, the TAAA motif was placed 9 nucleotides downstream from the stop codon, and TAG was introduced as the stop codon. Horizontal arrows represent the start of transcription. The vertical arrow marks the site where mRNAs are polyadenylated.

codon UGA, which is also the only motif conserved as a polyadenylation signal. Furthermore, a preference for a U-rich region close to the poly(A) tail is also apparent (2). Similar very short 3' UTRs have also been observed in *Giardia lamblia*, where AGURAAAY has been identified as a consensus polyadenylation signal. Interestingly, the three most conserved bases are the central URA nucleotides, which may overlap the translation stop codons of *Giardia* mRNAs (1, 21).

Experimental system designed to analyze polyadenylation sequences. As an approach to dissect the components of the proposed UAAA motif, a transient-expression vector was designed based on a bacterial coding sequence (*cat*) flanked by the 5' and 3' noncoding regions of a *T. vaginalis* actin gene known to be expressed. The 5' region contains the consensus sequence for the initiation of mRNA transcription (13), and the 3' region contains the UAA stop codon and 762 bp downstream (Fig. 2A). For the construction of the parental pcat-UAAA expression vector, the UAG stop codon was used for the *cat* reporter gene, followed by the 3' region of the actin gene (Fig. 2B). This strategy was used to separate the *cat* stop codon from the UAAA motif, which originally overlapped the actin UAA stop codon. From the pcat-UAAA parental plasmid, several derivatives were constructed by site-directed mutagenesis and are referred to here as pcat vectors. The purpose of these constructs was to define and delimit the role of the UAAA motif in the polyadenylation process by evaluating the effects of (i) the change of the four nucleotides, (ii) the single mutation of each of the four bases, (iii) the specific mutation of either the upstream or the downstream nucleotide adjacent to the UAAA motif, (iv) the insertion of two adenines to introduce the consensus metazoan AAUAAA polyadenylation signal, and (v) a revertant construct (UCAAA to UAAA) to recover the original UAAA motif. These plasmids are listed and described in Table 1. The ability of the UAAA motif to direct polyadenylation was evaluated by identifying the polyadenylation sites of the *cat* transcripts derived from the pcat vectors.

Four phenotypes associated with changes to the UAAA motif were additionally analyzed by (i) evaluating the *cat* mRNA abundance and size homogeneity relative to a second tran-

scriptional reporter, (ii) repositioning the UAAA motif further downstream from the intergenic region and evaluating the effect of its relocation upon mRNA polyadenylation, (iii) assessing the ability of an RNA fragment (containing the UAAA motif) to interact with a protein extract from *T. vaginalis*, and (iv) estimating the levels of CAT activity in cells transfected with the different pcat vectors.

Identification of the polyadenylation site in the reporter *cat* mRNA. The correlation between the UAAA motif and the selection of the polyadenylation site of the reporter *cat* mRNA was evaluated by cloning and sequencing 22 *cat* cDNA clones derived from the parental construct pcat-UAAA. These clones were readily obtained, and in all cases polyadenylation occurred 22 nt downstream the UAAA motif, at the same position as the endogenous actin mRNA from which the reporter vector was derived (Fig. 3A). The processing site occurred after the C in the CAAAAUU sequence, which is one of the proposed cleavage sites in *T. vaginalis* [Y ↓ (A)₂₋₅UU].

In contrast, any change within the UAAA motif altered the polyadenylation site of the *cat* mRNA, as shown in the schematic representation in Fig. 3B. The identified processing sites are indicated in the genomic sequence in Fig. 3D. Transcripts were found to be polyadenylated at several positions downstream from the parental site, giving rise to mRNAs with longer and heterogeneous 3' UTRs. These cDNA clones were difficult to obtain, likely because of the scarcity of poly(A) transcripts derived from these mutant constructs. The occurrence of potential nonpolyadenylated transcripts cannot be ruled out, but these would escape our detection system based on oligo(dT) initial priming. The effect of changing any of the four nucleotides is similar, with no apparent correlation between the mutated position and the site of polyadenylation. It is important to point out that several UAAA motifs are found in the 3' noncoding region, downstream from the mutated one, but none is able to unambiguously direct polyadenylation to a single site in the mutant mRNAs. However, more cDNA clones (16 out of 22) were found to be polyadenylated close to and downstream from UAAA motifs than otherwise. Particularly, one cleavage site seemed to be preferred, with seven cDNA clones identified. This site is in good agreement with the proposed cleavage site Y ↓ (A)₂₋₅UU for *T. vaginalis* (9). Nonetheless, the pro-

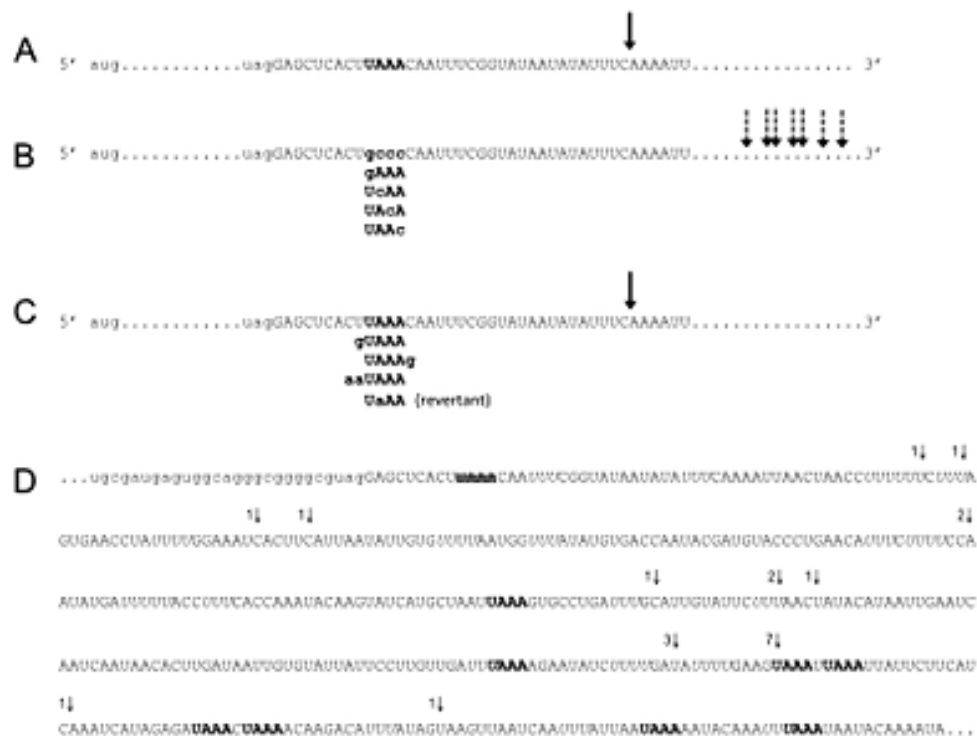


FIG 3 Schematic representation of the polyadenylation site in mRNAs derived from parental and mutant *pcatt* plasmids. (A) Partial sequence of the 3' UTR of *cat* transcripts from *T. vaginalis* cells that were transfected with the *pcatt*-UAAA plasmid. The UAAA motif is in bold, and the arrow denotes the site at which all mRNAs were found to be polyadenylated. (B) Schematic representation of the 3' UTR of *cat* transcripts from cells transfected with the different *pcatt* mutant plasmids. In these plasmids, the UAAA motif was changed to the indicated sequences. The arrows represent the sites at which mRNAs were found to be polyadenylated (fully described in panel D). The mutated nucleotides are shown in bold lowercase letters. (C) Sequence of the 3' UTR of *cat* transcripts from cells transfected with *pcatt* mutant plasmids in which the UAAA motif remained unaltered but the adjacent bases were changed as indicated. The UaAA (revertant) is a revertant construct derived from *pcatt*-UCAA by site-directed mutagenesis. The arrow denotes the site at which all the mRNAs analyzed were found to be polyadenylated and is the same position as in the wild type (A). The mutated nucleotides are shown in bold lowercase letters. (D) RNA sequence from the *pcatt* mutant plasmids in panel B, showing the actual positions where mRNAs were found to be polyadenylated. Part of the *cat* coding sequence is in lowercase letters. The 3' region is in uppercase. The parental UAAA motif is not present in the *pcatt* mutant plasmids and is therefore crossed out; this tetranucleotide was replaced by the bold motifs depicted in panel B. The last nucleotides identified before the poly(A) tail in *cat* cDNA clones are identified by arrows. The number of clones found for each position is shown next to the vertical arrows. Additional UAAA motifs in the 3' noncoding region are shown in bold, and U-rich elements are in italic. In all cases, the sites where the mRNAs were polyadenylated were deduced from sequencing oligo(dT)-derived cDNA clones.

posed U-rich region downstream from the processing site is not present (Fig. 3D).

The function of the UAAA motif is restricted to the tetranucleotide, since changing the contiguous bases showed no difference in the polyadenylation site selection compared to that of the parental control. This was also the case for the *pcatt*-AAUAAA plasmid, which contained the consensus metazoan polyadenylation signal (AAUAAA) (Fig. 3C). This result was expected since the UAAA motif is contained within.

Finally, to corroborate that changes in the polyadenylation site of the transcripts derived from the mutant plasmids were directly related to the UAAA motif, a single-nucleotide revertant plasmid was constructed from the mutant *pcatt*-UCAA plasmid to reconstruct the original UAAA sequence (*pcatt*-revUAAA). This revertant plasmid completely recovered the original phenotype (Fig. 3C).

The above results suggest that although the UAAA motif plays an essential role in defining the polyadenylation site, additional sequences must be involved that help to identify the site for precise

and efficient mRNA processing. These sequences could be the cleavage site and a downstream U-rich region.

Evaluation of *cat* mRNA size homogeneity. The size homogeneity of *cat* mRNAs in the transfected cells was evaluated relative to a cotransfected nonrelated gene as a control, transcribed from plasmid *pgfp*-UAAA. This plasmid is similar to the parental *pcatt*-UAAA, except that the *gfp* gene was cloned in place of the *cat* gene (Fig. 2C). Qualitative RT-PCR was used in a multiplex reaction where the 3' ends of *cat* and *gfp* cDNAs were specifically amplified from the total cDNA pool [oligo(dT)-derived]. Since the expected sizes of individual PCR products differ by 35 bp, it was then possible to analyze the size homogeneity and estimate the relative abundance of *cat* and control *gfp* mRNAs directly from acrylamide gels. The identity of the amplified products was corroborated by Southern hybridization (data not shown). In all cases the *gfp* RT-PCR amplification band obtained from the cotransfected control plasmid is well defined (Fig. 4). Also, *cat* mRNA is expressed as an apparently homogeneous population of transcripts when it is de-

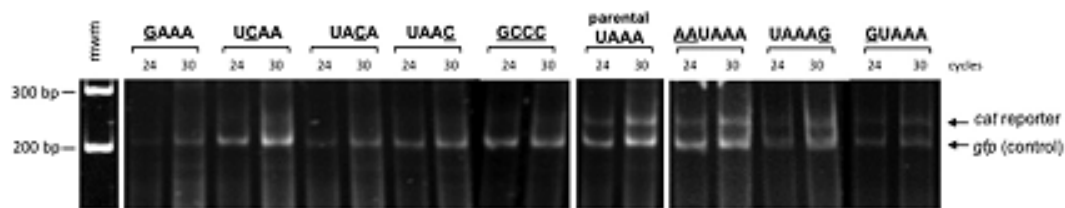


FIG 4 Size homogeneity of *cat* mRNAs. *T. vaginalis* cells were cotransfected with the different *pcat* plasmids and the *pgfp*-UAAA plasmid, which was used as an internal control for transfection and transcription. An RT-PCR approach was used, and samples from 24 and 30 cycles of amplification are shown. Amplification products were electrophoresed in a 5% acrylamide gel and stained with ethidium bromide. The lower band of about 200 bp corresponds to the RT-PCR product derived from the *gfp* mRNA control. This band migrates at the expected size, and it is constant and well defined throughout all the assays. The upper band of about 235 bp corresponds to the RT-PCR product derived from the *cat* mRNAs. This band is well defined for the cases where the transfected plasmids contain the UAAA motif (parental UAAA, AAUAAA, UAAAG, and GUAAA). RT-PCR products derived from plasmids without the UAAA motif appear as ill-defined bands compared with the *gfp* control in each lane, thus suggesting mRNAs of heterogeneous size and low abundance (GAAA, UCAA, UACA, UAAC, and GCCC). mwm, molecular size markers.

rived from the parental *pcat*-UAAA vector, as deduced from a well-defined RT-PCR amplification band (Fig. 4, parental UAAA). This is also the case for transcripts derived from all mutant plasmids where the UAAA motif is present (Fig. 4, AAUAAA, UAAAG, and GUAAA). In contrast, *cat* RT-PCR products derived from plasmids without the UAAA motif appear heterogeneous in size as observed from the blurred amplification products (Fig. 4, GAAA, UCAA, UACA, UAAC, and GCCC). These ill-defined products would reflect differences in the 3' UTR sizes of mutant mRNAs, which is in good agreement with the variability in the polyadenylation sites of the mutants described earlier (Fig. 2D). In addition, the apparent scarcity of these mRNAs may be due to the presence of nonpolyadenylated transcripts that would not be amplified and would escape our detection system. This result shows that a single change in the UAAA tetranucleotide causes an apparent decrease in the relative abundances of the derived mRNAs, and it reflects the poorly defined site for processing and polyadenylation of these transcripts.

Repositioning of the UAAA motif relocates the polyadenylation site of transcripts. To further support the directing role of the UAAA motif in polyadenylation, derivative plasmids in which the UAAA was relocated either 74 or 118 nucleotides downstream from the *cat* stop codon in the mutant *pcat*-UCAA plasmid were constructed (Fig. 5A). These two sites were chosen since runs of several Us were present but the UAAA motif was not found close by. The plasmids (*pcat*-UAAA74 and *pcat*-UAAA118) were transfected in *T. vaginalis* cells, and the derived mRNA was characterized by RT-PCR. The expected sizes of the amplified products were about 280 and 320 bp, respectively, if polyadenylation was directed by the inserted UAAA motifs. Figure 5B shows well-defined RT-PCR bands that correspond to the expected sizes, although very small differences in the amplification products would not be appreciated with this approach. By cDNA cloning and sequencing, we determined that mRNAs were processed downstream from the introduced UAAA motifs, but the new processing site was still not well defined since cleavage took place at several close-by positions (Fig. 5A). When the proposed (9) minimal cleavage motif (CAAUU) was introduced downstream the UAAA motif (plasmid *pcat*-UAAA118CS), all the analyzed clones were polyadenylated at the predicted cytosine (Fig. 5A and B).

It can therefore be concluded that although the UAAA motif is an essential sequence for polyadenylation, it is not enough to direct this process by itself and needs additional downstream se-

quences to attain its specificity. One of these signals would be the site for endonuclease cleavage (CAAUU and related sequences). Moreover, since the introduced motifs were located upstream from U-rich regions, these sequences might also aid in positioning the polyadenylation machinery, as has been proposed earlier (9, 17, 32). Experiments to further characterize these downstream elements are currently in progress in our laboratory.

RNA mobility shift assays support protein-RNA complex formation mediated by UAAA. To determine whether the UAAA motif, in the context of the mRNA analyzed here, is recognized by proteins from *T. vaginalis*, the interaction with an RNA fragment was tested by electrophoretic mobility shift assay (EMSA). For this purpose, 121-nt RNA transcripts were used as probes. These consisted of the last 39 nt of the *cat* coding region and 79 nt downstream from the *cat* stop codon; the UAAA probe contained the original UAAA motif, while the GCCC mutant probe was identical except for a four-nucleotide change of UAAA to GCCC (Fig. 6). A high-molecular-weight complex is formed with the probe containing the UAAA motif, while a much less intense shift of the probe is detected with the mutant probe containing the GCCC motif. The finding that the GCCC probe is able to interact with *T. vaginalis* proteins to a much lesser extent is not surprising given that proposed additional *cis* elements that could be recognized by protein factors are present in this RNA fragment. The formation of the complex was partially competed with the UAAA unlabeled RNA fragment, whereas no competition was observed with the mutant GCCC competitor (Fig. 6). However, when single point mutations in the UAAA motif were analyzed, an apparent mobility shift profile similar to that for the parental probe was observed (not shown).

Regarding the occurrence of potential orthologous proteins for the polyadenylation machinery, a comparative search was done in the *T. vaginalis* genome database. Several genes with more similarity to the mammalian than to the yeast counterparts were identified; among them were the ones coding for CPSF160, CstF, and symplekin. Of particular relevance is the CPSF160 subunit, since it recognizes the AAUAAA polyadenylation signal in metazoans. Taken together, the EMSAs and the identification of genes coding for putative polyadenylation proteins suggest that *T. vaginalis* possesses a polyadenylation machinery similar to the metazoan apparatus that can recognize the UAAA motif in a 3' UTR context. These protein factors are a current topic of study in our laboratory.

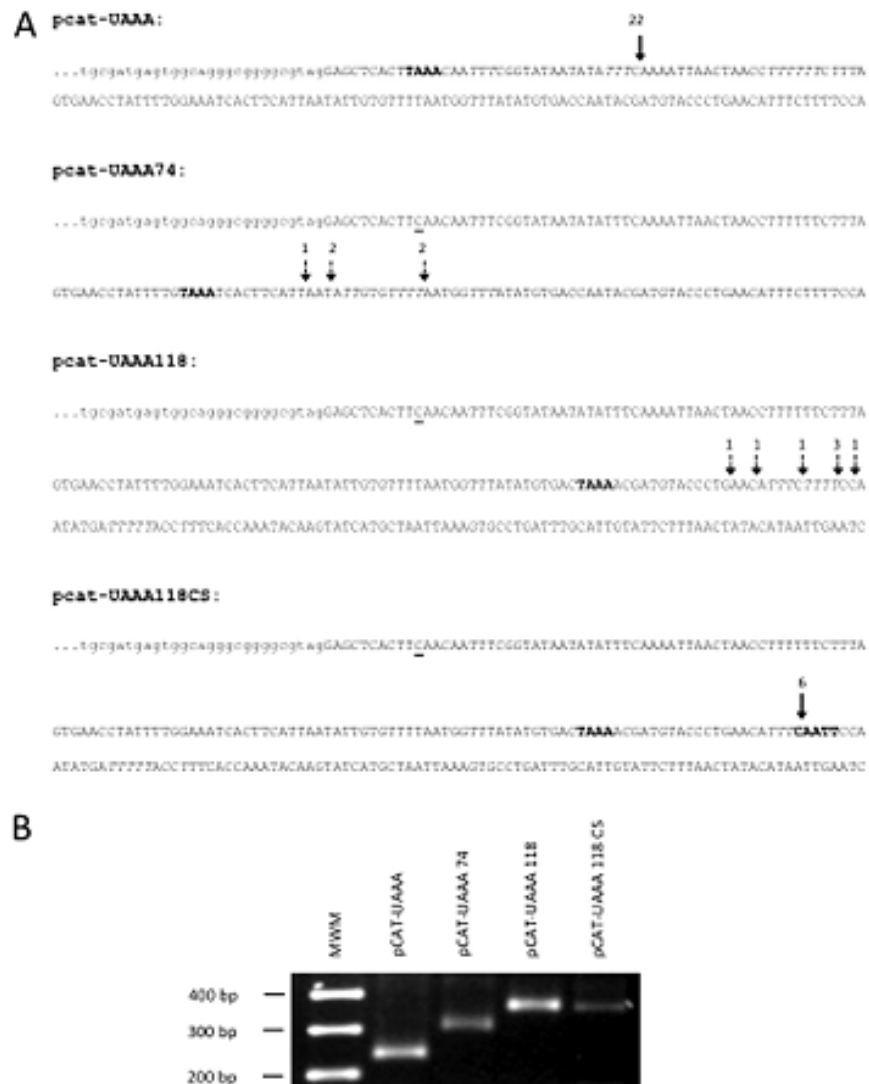


FIG 5 Relocation of the UAAA motif redirects the polyadenylation site. (A) DNA sequences of the plasmid constructs derived from the pcat-UCAA mutant, in which the UAAA motif was relocated downstream from the parental site. pcat-UAAA, parental sequence. The arrow denotes the site where all mRNAs were found to be polyadenylated. pcat-UAAA74, the UAAA motif (bold) was placed 74 nt downstream from the stop codon, and the parental UAAA motif was mutated (underlined C). The arrows represent the close-by sites where mRNAs were found to be polyadenylated. pcat-UAAA118, the UAAA motif (bold) was placed 118 nt downstream from the stop codon, and the parental UAAA motif was mutated (underlined C). The arrows represent close-by sites where mRNAs were found to be polyadenylated. pcat-UAAA118CS, the CAATT motif (bold), proposed as one of the cleavage site consensus motifs, was placed 20 nt downstream from the UAAA sequence in the pcat-UAAA118 plasmid. The arrow represents the site where all mRNAs were polyadenylated. The numbers above the arrows represent the size and homogeneity of the cat mRNAs by RT-PCR. Bands of the expected size were obtained, as predicted from the position where the UAAA motifs were introduced. Amplification products were analyzed on a 2.2% agarose gel and stained with ethidium bromide.

Reduced CAT activity from plasmids lacking the UAAA motif. As an indirect way to estimate the effect of changes in the polyadenylation process of the reporter transcript, the biochemical activity of CAT in the extracts of *T. vaginalis* cells that were transformed with the pcat vectors was evaluated. In all transfections, both a positive control (the plasmid pcat-UAAA) and a negative control (cells transfected with TE buffer) were included.

Table 4 shows that any change to the UAAA motif resulted in a lower (15 to 20%) level of CAT activity than that in the UAAA control. In contrast, cells that were transformed with plasmids bearing changes in the bases adjacent to the UAAA motif or with the revertant plasmid showed CAT activity levels similar to those in cells transformed with the parental pcat-UAAA. The same was observed with the plasmid bearing the consensus metazoan poly-

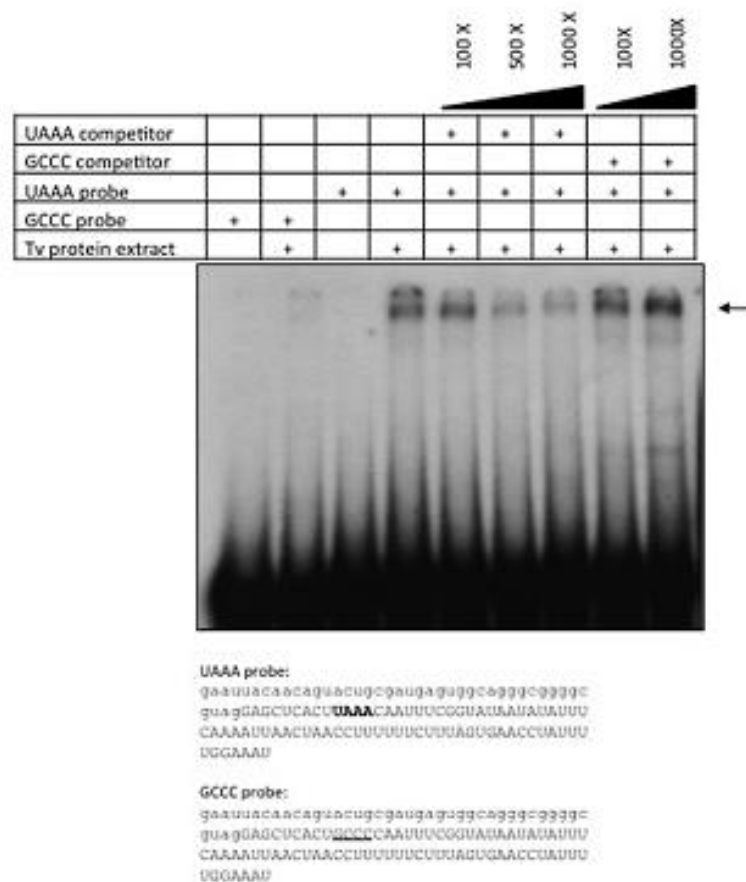


FIG 6 UAAA-specific binding activity is present in *T. vaginalis* protein extracts. A 121-nt RNA probe (corresponding to positions -51 to +66 relative to the UAAA motif in the *pcat*-UAAA plasmid) was used in RNA mobility shift assays with 5 μ g of *T. vaginalis* protein extracts. A similar RNA probe in which UAAA was replaced by GCCC was used as a mutant control. The sequences of both probes are shown. The *cat* coding region is denoted in lowercase letters, the 3' downstream region is shown in uppercase, the UAAA motif is in bold, and the GCCC motif is underlined. Competition assays were done with 100 \times , 500 \times , and 1,000 \times molar ratios of unlabeled UAAA and GCCC probes. The complexes were analyzed on a 4% acrylamide gel. The arrow shows the migration of the complex.

adenylation signal (Table 4). Given that all *pcat* plasmids provide the same promoter and 5' flanking sequence, we assume that transcription initiation is similar for all *cat* mRNA. The difference among transcripts would be the polyadenylation site, with aberrantly positioned poly(A) tails or nonpolyadenylated transcripts in the mutants within the UAAA motif. Our proposal is that these mutants would render fewer translatable mRNAs and, in turn, a lower abundance of CAT protein and lower measured activity in the transfected cell extracts.

TABLE 4 CAT activities in *T. vaginalis* cells transfected with the *pcat* vectors

Plasmid*	Mean (SD) % CAT activity
<i>pcat</i> -UAAA	100
<i>pcat</i> -GCCC	11.7 (4.05)
<i>pcat</i> -GAAA	10.2 (8.5)
<i>pcat</i> -UCAA	9.9 (3.1)
<i>pcat</i> -UACA	10.3 (5.8)
<i>pcat</i> -UAAC	14.5 (9.3)
<i>pcat</i> -GUAAA	146.9 (7.9)
<i>pcat</i> -UAAAAG	101.6 (7.8)
<i>pcat</i> -AAUAAA	127.4 (22.4)
<i>pcat</i> -revUAAA	122.3 (31.5)

*The UAAA motif is in bold.

Although UAAA is an apparently strict and essential element for defining the polyadenylation site of mRNA in *T. vaginalis*, it is not sufficient to direct this process by itself. Our data indicate that at least two additional elements (U-rich regions and a consensus cleavage site) enhance the specificity and direct the polyadenylation process in this early-divergent organism. Further work will address the identification of *trans*-acting factors of the polyadenylation complex in *T. vaginalis*.

ACKNOWLEDGMENTS

This work was supported by grant P45037-Q from Consejo Nacional de Ciencia y Tecnología (CONACYT), México, and grant IN220210 from Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIT), Universidad Nacional Autónoma de México (UNAM). We acknowledge scholarships to Vanessa Fuentes and Guadalupe Barrera from CONACYT, México.

This work was done by Vanessa Fuentes in partial fulfillment of the requirements for a Ph.D. degree from Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México.

We thank Jessica Guzmán for initial work on the construction of the reporter vector and Patricia de la Torre for help with DNA sequencing.

REFERENCES

- Adam RD. 1991. The biology of *Giardia* spp. *Microbiol. Rev.* 55:706–732.
- Andersson JO, et al. 2007. A genomic survey of the fish parasite *Spiro-nucleus salmonicida* indicates genomic plasticity among diplomonads and significant lateral gene transfer in eukaryote genome evolution. *BMC Genomics* 8:51. doi:10.1186/1471-2164-8-51.
- Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72:248–254.
- Carlton JM, et al. 2007. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. *Science* 315:207–212.
- Colgan DF, Manley JL. 1997. Mechanism and regulation of mRNA polyadenylation. *Genes Dev.* 11:2755–2766.
- Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Res.* 14:1188–1190.
- Dacks JB, Walker G, Field MC. 2008. Implications of the new eukaryotic systematics for parasitologists. *Parasitol. Int.* 57:97–104.
- Diamond LS. 1957. The establishment of various trichomonads of animals and man in axenic cultures. *J. Parasitol.* 43:488–490.
- Espinosa N, Hernandez R, Lopez-Griego L, Lopez-Villasenor I. 2002. Separable putative polyadenylation and cleavage motifs in *Trichomonas vaginalis* mRNAs. *Gene* 289:81–86.
- Frohman MA, Dush MK, Martin GR. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. U. S. A.* 85:8998–9002.
- Guo Z, Sherman F. 1995. 3'-end-forming signals of yeast mRNA. *Mol. Cell. Biol.* 15:5983–5990.
- Hu J, Lutz CS, Willusz J, Tian B. 2005. Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. *RNA* 11:1485–1493.
- Liston DR, Lau AO, Ortiz D, Smale ST, Johnson PJ. 2001. Initiator recognition in a primitive eukaryote: IBP39, an initiator-binding protein from *Trichomonas vaginalis*. *Mol. Cell. Biol.* 21:7872–7882.
- Liston DR, Johnson PJ. 1998. Gene transcription in *Trichomonas vaginalis*. *Parasitol. Today* 14:261–265.
- Lopez-Camarillo C, Orozco E, Marchat LA. 2005. *Entamoeba histolytica*: comparative genomics of the pre-mRNA 3' end processing machinery. *Exp. Parasitol.* 110:184–190.
- Lovrien RE, Matulis D. 2001. Selective precipitation of proteins. *Curr. Protoc. Protein Sci.* 4:4.5.
- Mandel CR, Bai Y, Tong L. 2008. Protein factors in pre-mRNA 3'-end processing. *Cell. Mol. Life Sci.* 65:1099–1122.
- Manley JL, Fire A, Cano A, Sharp PA, Geffer ML. 1980. DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract. *Proc. Natl. Acad. Sci. U. S. A.* 77:3855–3859.
- Manley JL, Fire A, Samuels M, Sharp PA. 1983. In vitro transcription: whole-cell extract. *Methods Enzymol.* 101:568–582.
- Millevoi S, Vagner S. 2010. Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. *Nucleic Acids Res.* 38:2757–2774.
- Morrison HG, et al. 2007. Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia*. *Science* 317:1921–1926.
- Proudfoot NJ, Furger A, Dye MJ. 2002. Integrating mRNA processing with transcription. *Cell* 108:501–512.
- Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Schneider TD, Stephens RM. 1990. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* 18:6097–6100.
- Seed B, Sheen JY. 1988. A simple phase-extraction assay for chloramphenicol acyltransferase activity. *Gene* 67:271–277.
- Simoës-Barbosa A, Meloni D, Wohlschlegel JA, Konarska MM, Johnson PJ. 2008. Spliceosomal snRNAs in the unicellular eukaryote *Trichomonas vaginalis* are structurally conserved but lack a 5'-cap structure. *RNA* 14:1617–1631.
- Sogin ML. 1991. Early evolution and the origin of eukaryotes. *Curr. Opin. Genet. Dev.* 1:457–463.
- Tsai CD, Liu HW, Tai JH. 2002. Characterization of an iron-responsive promoter in the protozoan pathogen *Trichomonas vaginalis*. *J. Biol. Chem.* 277:5153–5162.
- Vanacova S, Liston DR, Tachezy J, Johnson PJ. 2003. Molecular biology of the amitochondriate parasites, *Giardia intestinalis*, *Entamoeba histolytica* and *Trichomonas vaginalis*. *Int. J. Parasitol.* 33:235–255.
- Vanacova S, Yan W, Carlton JM, Johnson PJ. 2005. Spliceosomal introns in the deep-branching eukaryote *Trichomonas vaginalis*. *Proc. Natl. Acad. Sci. U. S. A.* 102:4430–4435.
- Venkataraman K, Brown KM, Gilmartin GM. 2005. Analysis of a non-canonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition. *Genes Dev.* 19:1315–1327.
- Wahle E, Keller W. 1996. The biochemistry of polyadenylation. *Trends Biochem. Sci.* 21:247–250.
- Wahle E, Rueggesser U. 1999. 3'-end processing of pre-mRNA in eukaryotes. *FEMS Microbiol. Rev.* 23:277–295.
- World Health Organization. 2001. Global prevalence and incidence of selected curable sexually transmitted infections. World Health Organization, Geneva, Switzerland.

9. METODOLOGÍA

Cultivo de *Trichomonas vaginalis*

El cultivo se realizó a partir de parásitos del aislado CNCD147 (donado por la Dra. Rossana Arroyo, CINVESTAV, IPN). Los parásitos se cultivaron en tubos con tapa de rosca estériles con 6 mL de medio TYM pH 6.2 (trypticase-yeast extract-maltose, apéndice I) suplementado con 0.6 mL (10% v/v) de suero de caballo inactivado (GIBCO) y se incubaron a 37 °C. Se utilizó como inóculo inicial aprox. 500,000 parásitos, para mantener los parásitos se realizaron subcultivos diariamente.

Mini-Preparación de DNA plasmídico (MINIPREPS)

Se tomó con un palillo estéril una colonia de bacterias transformantes de interés y se creció un cultivo en 5 mL de LB (apéndice I) (con el antibiótico adecuado) a 37°C toda la noche con agitación constante. Se colocó 1.5 mL del cultivo en un microtubo de 1.5 mL, se centrifugó durante 30 segundos a 12,000 rpm. El sobrenadante se desechó y se agregó en el mismo tubo 1.5 mL del mismo cultivo y se centrifugó durante 30 segundos a 12,000 rpm.

Se decantó el sobrenadante y se mezcló en el vortex para suavizar la pastilla. Se agregó 100 µL de la solución I fría (apéndice I) y se mezcló en el vortex, se agregó 200 µL de la solución II (apéndice I) recién preparada (a temperatura ambiente) y se mezcló por medio de inversión 6 veces.

Se colocaron los tubos en hielo y se agregó 150 µL de la solución III fría (apéndice I), se mezcló por inversión 6 veces y se incubaron en hielo durante 5 minutos.

Los tubos se centrifugaron por 5 minutos a 12,000 rpm, se recuperó el sobrenadante en otro microtubo de 1.5 mL y se desechó la pastilla. Al tubo con el sobrenadante se le añadió 250 µL de fenol frío y 250 µL de cloroformo frío, se mezcló en el vortex durante 30 segundos y se centrifugó por 3 minutos a 12,000 rpm. La fase superior se recuperó en otro microtubo de 1.5 mL y se le agregó 500 µL de cloroformo frío, se mezcló en el vortex durante 30 segundos y se centrifugó a 12,000 rpm durante 3 minutos. Se recuperó la fase superior en otro microtubo de 1.5 mL y se le agregó 1000 µL de etanol absoluto a temperatura ambiente, se mezcló en el vortex por 30 segundos. Se incubaron a temperatura ambiente al menos durante 10 minutos. Terminada la incubación se centrifugaron por 10 min. a 14,000 rpm, el sobrenadante se desechó y se lavó el botón de DNA con 700 µL de etanol al 70% frío e inmediatamente se desechó el líquido. El botón de DNA se dejó secar a temperatura ambiente.

Se resuspendió el botón de DNA con 50 μL de RNAsa (20 $\mu\text{g}/\text{mL}$), se mezcló en el vortex a velocidad máxima y se centrifugó por 30 segundos a 14,000 rpm.

Los tubos se incubaron a 37 °C al menos durante 20 min. Los minipreps se verificaron en un gel de agarosa 0.8% en TAE IX (apéndice I).

Mutagénesis Dirigida

Para la mutagénesis dirigida se utilizó el QuikChange site-directed mutagénesis kit (Stratagene). Este método de mutación consta de 4 pasos: 1) Diseño de los oligonucleótidos conteniendo la mutación deseada, 2) Replicación del DNA plasmídico a partir de los oligonucleótidos con la mutación, usando DNA polimerasa *PfuTurbo* 3) Digestión del DNA parental usando la endonucleasa *Dpn* I la cual es específica para DNA metilado y hemimetilado 4) Transformación de células XL1-Blue supercompetentes, con el DNA conteniendo la mutación deseada.

Para hacer la reacción de mutagénesis se puso en un microtubo lo siguiente:

5 μL de Amortiguador reacción 10X, 2.5 μL del DNA templado (20 ng/ μL), 1.25 μL Oligonucleotido # 1(100 ng/ μL), 1.25 μL Oligonucleotido # 2 (100 ng/ μL), 1 μL dNTP's mix, 38 μL H₂O bidestilada estéril.

Se añadió 1 μL DNA polimerasa *PfuTurbo*. Se utilizó el programa de PCR “CAT-CAAUU” (apéndice III).

Transcurrido el tiempo del programa de PCR se agregó 1 μL de la enzima *Dpn* I (10U/ μL). Se incubó la reacción a 37°C durante 1 hora. Terminada la incubación se llevó a cabo la transformación de las células XL1-Blue.

Se tomó un tubo con 50 μL de células XL1-Blue competentes y se agregaron 2 μL de la reacción de mutagénesis. Se mezcló e incubó en hielo durante 30 min. Posteriormente la reacción se incubó a 42°C durante 45 segundos y se puso en hielo durante 2 minutos.

Se añadieron 500 μL de medio NZY⁺ (apéndice I) precalentado a 42°C para la recuperación. Se pasaron las células transformantes a un tubo cónico de 15 mL y se incubó la reacción a 37°C durante 1 hora con agitación ~200-250 rpm.

Transcurrido el tiempo se platearon las células en dos cajas de LB+ Ampicilina, y se incubaron a 37°C durante por lo menos 16 horas.

Maxipreparación de DNA plasmídico (MAXIPREP)

Para la realización de esta técnica se utilizó el kit comercial de QIAGEN. Se inoculo medio LB+ antibiótico (100 mL para plásmidos de alta copia y 500 mL para plásmidos de baja copia) con una colonia de interés. Se crecieron a 37°C durante 12-16 h con agitación vigorosa (~300 rpm).

Las bacterias se bajaron por centrifugación a 6000 rpm en una centrífuga SORVALL con rotor GSA durante 15 min. a 4°C. El sobrenadante se desechó y se resuspendió la pastilla en 10 mL de amortiguador P1. Se añadió 10 mL de amortiguador P2, se mezcló ligeramente invirtiendo de 4-6 veces y se incubó a temperatura ambiente durante 5 min. Se añadió 10 mL de amortiguador P3 frío y se mezcló inmediatamente por inversión 4-6 veces y se incubó en hielo durante 20 min.

Se centrifugó a 13,000 rpm durante 30 min. a 4°C en una centrifuga SORVALL con rotor SS-34. Se decantó el sobrenadante en otro tubo estéril para rotor SS-34 y se recentrifugó durante 15 min. a 13,000 rpm a 4°C. Mientras tanto se equilibraron las columnas QIAGEN-tip 500 aplicando 10 mL de amortiguador QBT y dejando que se vaciara la columna por gravedad. Se aplicó el sobrenadante recentrifugado y se filtró por gravedad (guardar este filtrado (Sobrenadante)). Se lavó la columna con 2x30 mL de Amortiguador QC (lavado 1 y 2). El DNA se eluyó con 15 mL de Amortiguador QF y se recuperó en un tubo estéril para rotor SS-34. El DNA se precipitó añadiendo 10.5 mL de Isopropanol (0.7 vol.) a temperatura ambiente. Se mezcló y centrifugó inmediatamente a 11,000 rpm en una centrifuga SORVALL con rotor SS-34 durante 30 min. a 4°C.

Se lavó el botón de DNA con 5 mL de Etanol al 70% a temperatura ambiente y se centrifugó a 11,000 rpm durante 10 min. Se decantó cuidadosamente el sobrenadante, se dejó secar el botón de DNA y se resuspendió en 150 µL de TE. El plásmido se verificó en un gel de agarosa 0.8 % en TAE IX

Transfección de *T. vaginalis*

En un tubo de vidrio con tapa de rosca estéril se colocaron 6 mL de medio TYM pH 6.2 con 0.6 mL de suero de caballo inactivado (GIBCO) y se puso como inoculo inicial ~500,000 células; se incubó a 37°C durante 24 horas. En una caja de cultivo Nunc con capacidad de 260 mL se pusieron 50 mL de medio TYM pH 6.2 suplementado con 5 mL de suero de caballo inactivado (GIBCO) y se inocularon con 3.7×10^6 parásitos. Se incubó a 37 °C por 24 horas. Si se van a realizar varias transfecciones los cultivos se ponen en varias cajas de cultivo ya que para cada transfección se necesitan 240×10^6 parásitos.

Las células se resuspendieron y se tomó 100 µL de éstas para contar. Los cultivos se centrifugaron a 3,000 rpm durante 10 minutos a 4°C en una centrifuga SORVALL con rotor GSA. El sobrenadante se desechó y la pastilla se resuspendió suavemente en 20 mL de solución Zimmerman (apéndice I)

fría. Las células resuspendidas se transfirieron a un tubo para rotor SS-34, y se centrifugaron nuevamente a 3,000 rpm durante 10 minutos a 4°C.

Se desechó el sobrenadante y se adicionó la solución Zimmerman fría para obtener una concentración de 240×10^6 parásitos en un volumen de 400 μL para cada electroporación. Se adicionaron 200 μL más de solución Zimmerman fría al volumen final, se resuspendió la pastilla suavemente y se dejó en hielo. En cada celda de electroporación se puso la mezcla hecha de 400 μL de células + 50 μg del plásmido con el que se realizó la transfección (en un volumen máximo de 50 μL) y se incubaron en hielo por 5 minutos.

La electroporación se llevó a cabo en el BTX electroporation system, electro cell manipulator 600. Las condiciones iniciales del pulso fueron 1,500 μF , 360 Volts y 13 Ohms. Después de la electroporación de las células estas se dejaron a temperatura ambiente durante 5 minutos. Las células electroporadas se recuperaron en cajas de cultivo Nunc con capacidad de 260 mL con 30 mL de medio TYM pH 6.2 suplementado con 3 mL de suero de caballo inactivado (GIBCO) las cuales deben de encontrarse a 37°C, y se incubaron durante 22 horas.

Preparación del extracto proteico

Del volumen total del cultivo electroporado de *T. vaginalis* se transfirieron 15 mL a un tubo cónico de 15 mL. Este se centrifugó a 3,000 rpm durante 5 minutos a 4°C. Se eliminó el sobrenadante y la pastilla se resuspendió en 1 mL de solución TEN (apéndice I), subiendo y bajando con una micropipeta.

La pastilla resuspendida de cada muestra se transfirió a un microtubo nuevo de 1.5 mL y se centrifugó a 3,000 rpm durante 15 segundos a 4°C. Se retiró el sobrenadante con una pipeta y se resuspendió cada pastilla con 500 μL de solución TEN. Se centrifugó a 3,000 rpm durante 15 segundos a 4°C y se desechó el sobrenadante.

Se resuspendió la pastilla con 75 μL de Tris 25 mM pH 8, subiendo y bajando con una pipeta. Se lisaron las células con 3 ciclos de congelación-descongelación en nitrógeno líquido. Inmediatamente después de cada ciclo de congelación-descongelación los tubos se incubaron a 37°C durante 5 minutos. El lisado se incubó a 65°C durante 10 minutos y se mezcló con la punta de una pipeta. Se centrifugó a 14,000 rpm durante 10 minutos a 4°C. Se tomaron 50 μL totales del sobrenadante de cada muestra de extracto proteico y se alicuotaron en 2 tubos con 25 μL cada uno. Los extractos proteicos se guardaron a -70°C.

Extracción de RNA total a partir de células transfectadas

Cuidados para el manejo de RNA: Siempre utilizar guantes nuevos y cubrebocas, se utiliza material estéril y para la preparación de soluciones se utiliza H₂O DEPC.

15 mL del cultivo electroporado se colocaron en un tubo cónico de 15 mL y se centrifugó a 3,000 rpm durante 10 min. a 4°C. Se desechó el sobrenadante. La pastilla se resuspendió en 1 mL de PBS/DEPC estéril (apéndice I), se transfirió a un microtubo de 1.5 mL. Se centrifugó a 14,000 rpm durante 1 min. a 4°C y se desechó la solución de lavado. La pastilla se resuspendió con 1 mL de trizol (Invitrogen) y por pipeteo repetitivo se lisaron las células. Las muestras se incubaron durante 5 min. a temperatura ambiente para permitir una disociación completa de los complejos nucleoproteicos. Se agregaron 200 µL de cloroformo y se mezcló durante 15 segundos por medio de inversión, se incubó 3 minutos a temperatura ambiente. Se centrifugaron las muestras 15 minutos a 14,000 rpm a 4°C.

La fase acuosa se transfirió a un microtubo de 1.5 mL y se precipitó el RNA con 0.5 mL de alcohol isopropílico, se incubó la muestra 10 minutos a temperatura ambiente. Se centrifugó la muestra precipitada a 14,000 rpm durante 10 minutos a 4°C. El sobrenadante se desechó y se lavó el botón de RNA con 1 mL de Etanol/H₂O DEPC 75%(v/v). Se mezcló ligeramente la muestra en el vortex y se centrifugó a 10,000 rpm durante 5 min. a 4°C.

Se dejó secar el botón de RNA y se resuspendió en 10 µL de H₂O DEPC y se guardó a -70 °C. La integridad del RNA se verificó en un gel de agarosa 0.8% en TBE IX (apéndice I). El RNA se descongeló y se calentó durante 10 min. a 60°C antes de tomar la alícuota para checarla en el gel. Se tomó 1 µL de la muestra de RNA+ 4 µL de H₂O DEPC + 2 µL de amortiguador de carga 6X para RNA, se calentaron a 75°C durante 10 minutos, se puso en hielo 5-10 min. y se cargó en el gel.

Ensayo para medir la actividad enzimática de la cloranfenicol acetil transferasa(CAT)

Para realizar este ensayo se siguieron los lineamientos para el manejo de material radiactivo. Se preparó una solución concentrada de [³H]-cloranfenicol 0.2µCi/µL (apéndice I) y se hizo una dilución de la solución concentrada para que quedará a una concentración de 0.01 µCi/µL. Se preparó el N-butil-CoA 2.5mM en acetato de sodio 10 mM pH 5 y se guardó a -70°C

Para realizar el ensayo CAT se utilizaron los extractos proteicos de cada electroporación (ver sección 5.6), poniendo los siguientes controles:

Tris 25 mM (control negativo)

T.E. 10.2 mM (control negativo de la transfección)

CAT 0.008U (control positivo)

CAT 0.016U (control positivo)

Todos los reactivos se pusieron en hielo y se preparó la mezcla maestra para 8 reacciones de la siguiente manera:

SOLUCIÓN:	[] FINAL:
160 μ L [3 H] Cloranfenicol/ 0.01 μ Ci/ μ L	0.2 μ Ci
80 μ L N-Butiril-CoA 2.5 mM	0.25 mM
80 μ L Tris 1M, pH 8	100 mM
<u>80 μL H₂O bidestilada estéril</u>	
400 μ L Volumen final	

En un microtubo de 1.5 mL se prepararon las siguientes reacciones:

50 μ L de Tris 25 mM pH 8. (control negativo)

b. 48.9 μ L de Tris 25 mM pH 8 + 1.12 μ L de la enzima CAT dilución 1:1000 en Tris 25 mM (CAT 0.008U).

c. 47.8 μ L de Tris 25 mM pH 8 + 2.24 μ L de la enzima CAT dilución 1:1000 en Tris 25 mM (CAT 0.016U).

d. 45 μ L de Tris 25 mM pH 8 + 5 μ L del extracto proteico de las células transfectadas con T.E.

e. 45 μ L de Tris 25 mM pH 8 + 5 μ L del extracto proteico (se repitió dependiendo del número de muestras).

Se añadió a cada tubo 50 μ L de la mezcla maestra para tener un volumen final de 100 μ L. Se incubaron los tubos a 37°C durante 2.5 horas.

Transcurrido el tiempo de incubación, se agregó a cada tubo 240 μ L de xileno, se mezclaron en el vortex y se centrifugaron a 14,000 rpm durante 2 minutos. Se tomaron 180 μ L de la fase superior y se pusieron en microtubos marcados. Se hizo una reextracción de la fase acuosa a los tubos originales agregando 210 μ L de xileno. Se vortexearon 30 segundos y se centrifugaron a 14,000 rpm por 2 minutos.

Se tomaron 200 μ L de la fase superior y se colocaron en los tubos marcados, para tener 380 μ L en total. A cada uno de los tubos marcados se les agregó 250 μ L de Tris 25 mM pH 8, se mezclaron en el vortex durante 30 segundos y se centrifugaron 2 minutos a 14,000 rpm.

Se tomaron 330 μL de la fase superior y se colocaron en otros microtubos de 1.5 μL . Se agregaron 250 μL de Tris 25 mM, pH 8. Se mezclaron en el vortex durante 30 segundos y se centrifugaron 2 minutos a 14,000 rpm.

Se tomaron 300 μL de la fase superior y se colocaron directamente en los viales que contienen 1.5 mL del líquido de centelleo. Antes de leer en el contador de centelleo líquido se normalizó, se leyó en el canal de tritio y se utilizó como blanco un vial con solo líquido de centelleo.

Cuantificación de proteínas

Para reportar la actividad de CAT/ μg de proteína se cuantificó la cantidad de proteína en los extractos mediante el método de Bradford (Bradford, M.M. 1976). Se utilizó una curva patrón de albúmina sérica bovina a diferentes concentraciones (0.5, 1, 1.5, 2, 2.5...10 $\mu\text{g}/\text{mL}$).

En un microtubo de 1.5 mL se colocaron 5 μL de cada extracto proteico a cuantificar y se agregó 795 μL de H_2O bidestilada estéril. Se agregaron 200 μL del reactivo Azul de Coomasie (BIORAD), se mezcló en el vortex durante 30 segundos y se incubó a temperatura ambiente por 10 minutos.

Las muestras se leyeron en el espectrofotómetro a $\lambda = 595 \text{ nm}$, antes de que transcurriera una hora de realizada la reacción. Los datos de la curva patrón se graficaron y las lecturas obtenidas de los extractos proteicos se interpolaron.

Purificación de RNA poly A+

A partir del RNA total de *T. vaginalis* se purificó el RNA poli A+ utilizando el Kit QuickPrep™ Micro mRNA Purification (Amersham Pharmacia Biotech), que cuenta con una resina oligo (dT) celulosa.

Al RNA total de *T. vaginalis* resuspendido en H_2O DEPC, se le agregó 0.4 mL de amortiguador de extracción y 0.8 mL del amortiguador de elusión. En un microtubo de 1.5 mL se colocó 1 mL de oligo (dT) celulosa, se centrifugó a 14,000 rpm durante 1 min. y se eliminó el sobrenadante. Se le agregó la mezcla hecha con el RNA total y se mezcló suavemente por inversión durante 3 minutos, se centrifugó a 14,000 rpm durante 10 segundos. El sobrenadante se transfirió a un tubo nuevo ya que en este se encuentra la población RNA poly A-. La resina se lavó cinco veces de la siguiente forma: se resuspendió en 1 mL de amortiguador High-salt por inversión, se centrifugó a 14,000 rpm por 10 segundos y se eliminó el sobrenadante.

Posteriormente la resina se resuspendió en 1 mL de amortiguador Low-salt por inversión, se centrifugó a velocidad máxima por 10 segundos y se eliminó el sobrenadante, este paso se repitió una vez más.

La resina se resuspendió en 300 μL de amortiguador Low-salt, y se transfirió a una microcolumna que se colocó dentro de un microtubo de 1.5 mL y se centrifugó a 14,000 rpm durante 5 segundos. Se cambió la columna a un microtubo nuevo y se agregaron 500 μL de amortiguador Low-salt y se centrifugó 5 segundos a 14,000 rpm, la columna se cambió a un tubo nuevo y se repitió el paso anterior.

El RNA poli A+ se eluyó de la resina de la siguiente manera: La columna se colocó en un microtubo de 1.5 mL y se le agregó 200 μL del amortiguador de elusión a 65°C, la columna se centrifugó a 14,000 rpm durante 5 segundos, y la elusión se guardó ya que contiene al RNA poly A+. La columna se lavó con 200 μL de amortiguador de elusión a 65°C y se repitió el paso anterior. El RNA poly A+ se cuantificó por espectrofotometría de la siguiente manera: el RNA se diluyó 1:100 en amortiguador de elusión, en un volumen total de 1 mL. En una celda de cuarzo pretratada con H₂O DEPC se leyó la absorbancia a $\lambda=260$ nm y $\lambda=280$ nm. La concentración de RNA se calculó utilizando la fórmula:

$$1 \text{ D.O.}_{260} = 40 \mu\text{g RNA/mL}$$

Amplificación rápida de los extremos 3' de cDNA (3' RACE)

La técnica de 3' RACE incluye los procesos de: a) síntesis de cDNA a partir de RNA poli A+, b) amplificación del extremo 3' del cDNA del gen mediante PCR y c) clonación del producto de PCR purificado.

Síntesis de cDNA de *T. vaginalis*.

Se utilizaron 10-500 ng de RNA poly A+ de *T. vaginalis*, 1 μL del oligonucleótido AP (adapter primer) (200 ng/ μL), 0.5 μL dNTP's mix 10 mM c/u (Invitrogen) y se agregó H₂O DEPC para tener un volumen de 7 μL .

Se calentó la mezcla a 65°C por 5 min. y se incubó 1 min. en hielo, el tubo se centrifugó para bajar y se agregó:

μL Amortiguador First-Strand 5X (Invitrogen), 0.5 μL 0.1 M DTT (Invitrogen), 0.5 μL SuperScript™ III RT (200U/ μL) (Invitrogen).

Se mezcló por medio de pipeteo subiendo y bajando, se incubó a 42°C durante 15 min. y a 50°C durante 30 min. Se inactivó la enzima calentando a 70°C durante 15 min. Se añadió 0.5 μL RNAsa H (Invitrogen) (2U/ μL). Se incubó a 37°C durante 35 minutos, se centrifugó para bajar y se guardó a -20°C.

b) Amplificación de los extremos 3' de cDNAs de CAT por medio de PCR.

En un microtubo se colocaron 1.5 μ L de cDNA de *T. vaginalis* (inciso a), 5 μ L de amortiguador PCR 10X (SIGMA), 6 μ L de $MgCl_2$ 25 mM (SIGMA), 3 μ L de dNTPs 10 mM (Invitrogen), 30 μ L H_2O bidestilada estéril, 2 μ L del oligonucleótido CAT 10 (20 pmol/ μ L), 2 μ L del oligonucleótido AUAP-B (20 pmol/ μ L) y 0.5 μ L de la enzima *Pfu*. Se utilizó el programa de PCR PROCE 1 (apéndice III).

Para realizar una segunda amplificación del cDNA de CAT (PCR anidado, o “nested”), se colocaron 0.5 μ L del producto de PCR de la primera amplificación, 5 μ L amortiguador PCR 10X (SIGMA), 6 μ L de $MgCl_2$ 25 mM (SIGMA), 3 μ L de dNTPs 10 mM (Invitrogen), 31 μ L H_2O bidestilada estéril, 2 μ L del oligonucleótido CAT11 (20 pmol/ μ L), 2 μ L del oligonucleótido AUAP-B (20 pmol/ μ L) y 0.5 μ L de la enzima *Pfu*. Se utilizó el programa de PCR PROSC2 (apéndice III).

Los productos de PCR se verificaron en un gel de agarosa al 1.5 % en TAE 1X

Ligación de los Productos del ensayo 3' RACE en el vector TOPO (Invitrogen)

La clonación en el vector TOPO (Zero Blunt TOPO PCR Cloning) se basa en la reacción que se lleva a cabo entre la enzima topoisomerasa I del virus de la *Vaccinia* y el DNA del vector plasmídico; la enzima corta después de la secuencia 5'- CCCTT de cada una de las cadenas del DNA (en este sitio se ligará el producto de PCR) y se forma un enlace covalente entre el grupo 3' fosfato del sitio de corte y el residuo de tirosina (Tyr 274) de la topoisomerasa I. Este enlace es atacado por el grupo 5' hidroxilo de la cadena de DNA liberando a la enzima, haciendo eficiente la clonación del producto de PCR con el vector. El vector TOPO tiene al gen *ccdB* que es un gen letal de *E. coli*, este gen esta fusionado a la región terminal del gen que codifica para el fragmento LacZa y si hay ligación del producto de PCR se inhibe la expresión de dicho gen, permitiendo crecer únicamente a las colonias recombinantes.

Se clonaron los productos del PCR anidado, de la siguiente manera: En un microtubo de 0.5 mL se colocaron 2 μ L del producto de PCR anidado, 0.5 μ L de la solución diluida de sal (1:4) (Invitrogen), y 0.5 μ L del vector TOPO. La mezcla se incubó por 30 minutos a temperatura ambiente para que se llevara a cabo la ligación. Se agregó 2 μ L de la mezcla anterior a un vial de células competentes (*E.coli*) mezclándose ligeramente. Se transfirieron las células a una celda para electroporación fría de 0.1 cm de diámetro, se electroporaron dando un pulso de 2.46 Kvolts, inmediatamente después se recuperaron agregando 250 μ L de medio SOC. La mezcla se transfirió a un tubo cónico de 15 mL y se incubó con agitación a 37 °C durante una hora.

Después de la incubación se plateó una dilución 1:10 (30 μL) y 9:10 (270 μL) de cada electroporación en cajas de medio LB sólido + kanamicina (50 ng/ μL), las cajas se incubaron a 37°C toda la noche.

Transcripción *in vitro* (Amplibio)

Se utilizó una cámara de electroforesis vertical marca Hoefer.

La transcripción se realizó de la siguiente manera:

A. Reacción de transcripción

1. Descongelar los reactivos congelados, colocar la RNA polimerasa en hielo, vortexear el buffer transcripción 10X y las soluciones de ribonucleótidos hasta que estén completamente en solución. Una vez descongelados, colocar los ribonucleótidos en hielo, pero mantener el **buffer transcripción 10X a temperatura ambiente**.

Se realizó la reacción de transcripción a temperatura ambiente ya que la espermidina en el buffer de transcripción 10X puede coprecipitar el templado de DNA si la reacción se realiza sobre el hielo.

Añadir el buffer de transcripción 10X después del agua y el DNA.

Isotopo a utilizar: UTP [α - ^{32}P]

Componente	Volumen para el transcrito
Agua libre nucleasa (kit)	2.05 μL
Banda del Transcrito (~200 ng/uL)	1 μL
Buffer Transcripción 10X	1 μL
10 mM ATP	0.5 μL
10 mM CTP	0.5 μL
10 mM GTP	0.5 μL
10 mM UTP	-----
120 μM UTP	0.25 μL
UTP [α-^{32}P]	3.2 μL
Mix Enzima (T7)	1 μL
Volumen total	10 μL

3. Agitar suavemente o pipetear la mezcla subiendo y bajando lentamente, centrifugar los tubos brevemente para recoger la mezcla de reacción en el fondo del tubo. Se incubó la reacción durante 10 min. a 37 °C. Posteriormente se añadió 1 µL TURBO DNAsa I (2U/µL) a los tubos, se mezcló e incubó a 37 °C por 15 min.

6. Añadir 1 µL de 0.5 M EDTA DEPC para detener la reacción.

Es importante añadir EDTA a las reacciones que se van a calentar (para inactivar DNasa I, por ejemplo). El EDTA bloquea la degradación del RNA inducido por calor que puede ocurrir en el buffer de transcripción.

- **Análisis del Transcrito marcado en gel de acrilamida:**

Para la visualización del tamaño del transcrito se utilizó un gel de acrilamida al 4 % desnaturalizante.

- a. Mezclar 0.5 µL de los 10 µL de la reacción de transcripción con 3.5 µL Gel Loading Buffer II. Como marcador de peso molecular utilizar RNA total de *T. cruzi* (5 µL RNA total + 3.5 µL gel Loading Buffer II).
- b. Calentar las muestras a 95°C por 2 min.
- c. Cargar directamente en el gel o almacenar las muestras en hielo hasta que el gel sea cargado.
- d. Correr la electroforesis a 80 V durante 15 min y posteriormente a 40 V durante 1.5 h o hasta que el azul de bromofenol llegue al final del gel.
- e. Remover un vidrio y cubrir el gel con plástico (adherido al segundo vidrio). Exponer el gel.
- f. Tener cuidado con el buffer de la electroforesis pues está radiactivo. Desechar de manera apropiada.

Nota: Antes de cargar el gel es necesario lavar los pozos.

Terminado el corrimiento del gel, se cortó la parte del gel donde esta el RNA total de *T. cruzi*, se verificó que no se encuentre radiactivo y fue teñido con bromuro de etidio y se enjuago con H₂O DEPC.

Protocolo de obtención de extractos totales de *T. vaginalis*

1. Crecer de 2-3 litros de *T. vaginalis* a 2×10^6 cel/ml ($4-6 \times 10^9$ cel).
2. Centrifugarlas en la centrífuga Sorvall con el rotor GSA a 3000rpm (1500 x g) 10 minutos a 4°C.

-
3. Las células se resuspenden en 40 ml de PBS y se transfieren a un tubo falcon de 50 ml. Se centrifugan a 1500 x g 10'.
 4. El botón se lava nuevamente con 10ml de PBS, se centrifuga de igual forma que en el paso anterior y se determina el volumen celular empacado (PCV).
 5. El botón se resuspende en 4 PCV de buffer hipotónico y se incuba en hielo por 20 minutos.
 6. Se lisan las células con el homogenizador Dounce usando 20 golpes con el pistón Tight. La homogenización es lenta y suave.
 7. Se agregan 4 PCV de buffer de lisis. Se mezcla suavemente la suspensión y se pone a agitar lentamente, entonces se agrega 1 PCV de sulfato de amonio 4M de forma muy lenta (gotas) y se deja agitando 30 minutos (la concentración final de sulfato de amonio es 0.4M).
 8. El extracto se centrifuga a 45000 rpm (190 000 x g) en un rotor Beckman SW 50.1 por 3 horas.
 9. Se recupera el sobrenadante en Falcon de 50 mL y se determina su volumen. Se precipitan las proteínas agregando sulfato de amonio sólido (0.33g por ml de solución= 60% saturación).
 10. Después de la solubilización del sulfato de amonio, se agrega 1N de NaOH (0.1 mL/10g de sulfato de amonio sólido y se agita la suspensión por 30 minutos adicionales. Este paso neutraliza la acidez del sulfato de amonio. Agitar lentamente y sin formar espuma.
 11. Se colectó el precipitado (en tubo Falcon de 50 mL) centrifugando a 10 000 x g (8800 rpm en eppendorf 5804R) 20 minutos. Se decanta completamente el sobrenadante (y se almacena para precipitarse a saturación del 85% de sulfato de amonio) y el botón se disuelve en 5% del volumen original del sobrenadante con buffer de resuspensión (evitar la formación de espuma). Almacenar en hielo mientras se termina la siguiente precipitación.
 12. El sobrenadante se precipita con 0.14g/ml sobrenadante (80% saturación), se trata con NaOH 1N y se procesa como en el paso 11.
 13. Se dializa el extracto contra dos cambios de 100 (40ml) volúmenes del buffer de resuspensión toda la noche (esta vez los cambios fueron después de 2h y por la mañana para dejar la diálisis 3h más). El volumen de la solución incrementa del 30-45% durante la diálisis.
 14. Se centrifuga el dializado a 10 000 x g 10 minutos para remover los materiales insolubles. Revisar las proteínas en un gel de acrilamida y cuantificarlas por Bradford. Congelar alícuotas en hielo seco-alcohol.

- **Ensayos de Retardo**

Tratamiento de la cámara de electroforesis antes de utilizar (se utilizó una cámara Hoefer):

La cámara donde se corre el gel así como vidrios, pinzas, separadores y peines deben ser tratados con Etanol al 70 % por lo menos 20 min. Transcurrido el tiempo la cámara se enjuaga con H₂O DEPC y se deja secar. Los vidrios, pinzas, separadores y peines solo se les limpia el alcohol.

Mezcla de reacción:

Muestra	Contro	Muestra	Conc. Final
Reactivo	1		
Transcrito	X μ L	X μ L	5000 cpm*
Fosfocreatina 0.6 M	1 μ L	1 μ L	20 mM
MgCl₂ 50 mM	0.3 μ L	0.3 μ L	0.5 mM
Glicerol stock	3 μ L	3 μ L	10 %
HEPES- KOH (pH 7.9) 1M	0.3 μ L	0.3 μ L	10 mM
ATP 10 mM	4.8 μ L	4.8 μ L	1.6 mM
KCl 1 M	1.5 μ L	1.5 μ L	50 mM
EDTA 10 mM	0.15 μ L	0.15 μ L	0.05 mM
DTT 100 mM	0.6 μ L	0.6 μ L	2 mM
tRNA 10 μg/μL	0.6 μ L	0.6 μ L	0.2 mg/mL
RNaseOut 40 U/μL	1 μ L	1 μ L	40U
Extracto proteico	-----	X μ L	~ 14 μg
H₂O DEPC	X μ L	X μ L	-----
Volumen final	30 μL	30 μL	-----

* La cuantificación se realiza con el Geiger

-
1. Realizar la mezcla de reacción en hielo. Poner el buffer de interacción y el volumen de H₂O DEPC correspondiente a cada uno de los tubos.
 2. Agregar el extracto proteico a cada uno de los tubos.
 3. Agregar el transcrito marcado radiactivamente a todos los tubos. Mezclar cuidadosamente subiendo y bajando con la pipeta.
 4. Las reacciones se preincuban en hielo durante 10 min.
 5. Incubar las reacciones a 30 °C por 10 minutos.
 6. Cargar las muestras en un gel de poliacrilamida nativo 4 % con TBE 0.5 X. Azul de bromofenol y xilencianol se ponen en carriles vecinos y la electroforesis es llevada a cabo a temperatura ambiente.

Corrimiento del gel:

Voltaje: 120 V por 55 min.

7. Secar el gel a 70 °C durante 1 a 2 h. La temperatura de secado se va subiendo poco a poco, puede comenzar en 30 °C y se va subiendo de 5°C hasta llegar a 70 °C. Para poner a secar el gel se pega a un papel Whatman y se cubre con egapack. NOTA: Se le ponen dos whatman debajo del gel para que el líquido radiactivo no contamine el equipo.
8. Exponer el gel, el tiempo puede variar dependiendo del marcaje del transcrito.

10. SOLUCIONES Y OLIGONUCLEÓTIDOS UTILIZADOS

Ampicilina stock: 100mg/mL, concentración final en las cajas 100 µg/mL.

Medio SOC

- 10 mL de medio SOB
- 200 µL de glucosa 1M

Medio SOB

- Triptona 2 %
- Extracto de Levadura 0.5 %
- Cloruro de Sodio 10 mM
- Cloruro de Potasio 2.5 mM
- Cloruro de Magnesio 10mM
- Sulfato de Magnesio 10 mM

Medio LB sólido (para 1 Litro)

- 10 g Triptona
- 5 g Extracto de Levadura
- 5 g de NaCl
- 0.3 mL NaOH 10M
- 10 mL Tris 1M pH = 8
- 15 g de Bacto agar

Medio LB líquido (para 1 Litro):

- 10 g Triptona
- 5 g Extracto de Levadura
- 5 g de NaCl
- 0.3 mL NaOH 10M
- 10 mL Tris 1M pH = 8

Solución I: (Preparar 100ml).

-Glucosa-----50mM

-Tris/HCl pH 8-----25mM

-EDTA pH 8-----10mM

Nota: Preparar a partir de los stocks ya estériles y ponerlo en un frasco estéril.

Solución II: (Preparar 5ml)

-NaOH 0.2N

-SDS 1%

Nota: Preparar a partir de los stocks ya estériles y ponerlos en un matraz estéril.

Solución III: (Preparar 100mL)

-Acetato de potasio 5M-----60mL

-Ácido acético glacial concentrado -----11.5mL

-Agua bidestilada estéril-----28.5mL

Nota: Preparar a partir de los stocks ya estériles y ponerlos en un frasco estéril.

TAE 50X (para 500 mL)

- Tris base -----121g

- Acido acético glacial -----28.55 mL

- EDTA 0.5 M pH 8 ----- 50 mL

TAE 1X (para 1 Litro):

- 20 mL TAE 50 X

- 980 mL H₂O bidestilada

TE 10.2 mM

- Tris pH =7.4 ----- 10 mM

- EDTA ----- 0.2 mM

TBE 5X en H₂O DEPC (para 1 L):

- Tris ----- 54 g

- Acido bórico -----27.4 g

- EDTA 0.5 M pH = 8 -- 20 mL

TBE 1X/ H₂O DEPC

- 200 mL TBE 5X/ H₂O DEPC
- 800 mL H₂O DEPC

PBS/DEPC pH = 8 (para 1L):

- NaCl -----8 g
- KCl -----0.2 g
- Na₂HPO₄-----1.44 g
- KH₂PO₄ -----0.24 g

Medio TYM pH 6.2 (para 500 mL):

- Peptona ----- 10 g
- Extracto de Levadura --- 5 g
- Maltosa ----- 2.5 g
- L-cisteína ----- 0.5 g
- Acido ascórbico----- 0.1 g
- KH₂PO₄ ----- 0.4 g

Solución TEN

- Tris----- 40 mM pH 8
- EDTA----- 1mM pH 8
- NaCl -----15mM

Esterilizarlo por autoclave y conservarlo a temperatura ambiente

Solución desnaturalizante (para 500 mL)

- NaCl-----43.83 g
- NaOH 10N-----25 mL

Solución de neutralización (para 500 mL)

- Tris/HCl 2M pH 7.5-----125 mL
- NaCl-----87.66 g

Solución de transferencia:

- NH₄Ac -----1M

-
- NH₄OH-----1M
 - NaOH.-----0.02 N

Solución de hibridación

- Formamida----- 50%
- SSC----- 5X
- Tris -----50 mM pH 7.4
- Denhardt's -----5X
- SDS----- 1%

SSC 20 X (Sodium Chloridal Sodium Citrate) en H₂O DEPC

- NaCl-----3M
- Citrato de Sodio dihidratado-----0.3 M

Ajustar el pH 7 con 1M HCl

Stock de cloranfenicol tritiado 0.2μCi/μL

Tomar 125μL del vial original de cloranfenicol-H3 (250 μCi en 250μL de etanol). Agregar 480 μL de etanol absoluto y 40 μL de solución de cloranfenicol sin marcar 50 mg/mL para lo cual se pesan 100mg de cloranfenicol y se disuelven en 2 mL de etanol al 50%. Se mezclan y se guarda a -20°C.

- Buffer hipotónico [9ml]

10mM HEPES-KOH (pH 7.9) [90ul HEPES 1M]

1mM EDTA [18ul EDTA 0.5M]

5mM DTT [45ul DTT 1M]

Leupeptina 10ug/ml [19ul stock 10mM]

TLCK 50 ug/ml (0.2mM) [12.8ul stock 50mg/ml]

EGTA 1mM [18 ul EGTA 0.5M]

PMSF 1mM [90ul stock 100mM]

Agua DEPC [8.71ml]

- Buffer de lisis [9ml]

50mM HEPES-KOH (pH 7.9) [450ul HEPES 1M]

10mM MgCl₂ [90ul MgCl₂ 1M]

2mM DTT [18ul DTT 1M]

25% sucrosa [2.25g]
50% glicerol [4.5ml]
Leupeptina 10ug/ml [19ul stock 10mM]
TLCK 50 ug/ml [12.8ul stock 50mg/ml]
PMSF 1mM [90ul stock 100mM]
Agua DEPC [~2.40ml]

- 4M (NH₄)₂ SO₄ pH 7.9 (NaOH)
- Buffer de resuspensión y diálisis [120ml x 2 para poner por separado las diálisis]

25mM HEPES-KOH (pH 7.9) =20-50mM= [3ml HEPES 1M]

75mM KCl =50-100= [9ml KCl 1M]

12mM MgCl₂ [1.44ml MgCl₂ 1M]

0.25mM EDTA =0.1-0.5mM= [60ul EDTA 0.5M]

2.5mM DTT =1-5mM= [300ul DTT 1M]

10% glycerol =10-20%= [12ml]

Leupeptina 10ug/ml [252ul Stock 10mM]

TLCK 50 ug/ml [171ul stock 50mg/ml]

EGTA 0.2mM [48ul EGTA 0.5M]

PMSF 1mM [1.2ml stock 100mM]

Agua DEPC [92.53ml]

INHIBIDORES

1. Leupeptina 10mM: 0.0035g en 700ul agua
2. TLCK 140mM: 0.0206g en 400ul agua
3. PMSF 100mM: 0.052g en 3ml isopropanol
4. DTT 1M: 0.12 g en 750ul agua.

OLIGONUCLEÓTIDOS

Oligonucleótido	Secuencia
GTAAA	5' GCGTAGGAGCTCACGTAAACAATTTTCGGTATAATA 3'
RC GTAAA	5' TATTATACCGAAATTGTTTACGTGAGCTCCTACGC 3'
INS TAAA	5' GCGTAGGAGCTCACTTCAATAAACAATTTTCGGTATAATA 3'
RC INS	5' TATTATACCGAAATTGTTTATTGAAGTGAGCTCCTACGC 3'
F-REGSP	5' GCGTAGGAGCTCACTTAAACAATTTTCGGTATAATA 3'
RC-REGSP	5' TATTATACCGAAATTGTTTAAAGTGAGCTCCTACGC 3'
PAS	5' TGGTTTATATGTGACTAAAACGATGTACCCTGAAC 3'
RC-PASN	5' GTTCAGGGTACATCGTTTTAGTCACATATAAACCA 3'
CS	5' TACCCTGAACATTTCAATTCCAATATGATTTTTTAC 3'
RC-CS	5' GTAAAAATCATATGGAATTGAAATGTTTCAGGGTA 3'
F-NSP	5' GTGAACCTATTTTGTAAATCACTTCATTAATATTG 3'
RC-NSP	5' CAATATTAATGAAGTGATTTACAAAATAGGTTTAC 3'
AP	5' GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTTTT 3'
AUAP-B	5' GGATCCGGCCACGCGTCGACTAGTAC 3'
CAT10	5' CACATTCTTGCCCGCCTG 3'
CAT11	5' GAGCTGGTGATATGGGAT 3'
CAT 20	5' AAGTTGTCCATATGGCC 3'
For-invít	5' TAATACGACTCACTATAGGGAATTACAACAGTACTG CGAT 3'
Rev-invít	5' ATTTCCAAAATAGGTTCACT 3'
For-CATqPCR	5' CGCCCCGTTTTCCACCATGG 3'
For-GFPqPCR	5' CCGTGCTGCTGCCGACAA 3'

11. REFERENCIAS

- Adagu IS1, Nolder D, Warhurst DC, Rossignol JF. 2002. *In vitro* activity of nitazoxanide and related compounds against isolates of *Giardia intestinalis*, *Entamoeba histolytica* and *Trichomonas vaginalis*. *J Antimicrob Chemother.* **49(1)**:103-11.
- Adam RD.1991. The biology of *Giardia* spp. *Microbiol. Rev.***55**:706–732.
- Afzan MY, Sivanandam S, Kumar GS. 2010. Modified Field's staining--a rapid stain for *Trichomonas vaginalis*. *Diagn Microbiol Infect Dis.* **68(2)**:159-62
- Alderete JF, Garza GE. 1988. Identification and properties of *Trichomonas vaginalis* proteins involved in cytoadherence. *Infect Immun.* **56(1)**:28-33.
- Andersson JO, et al.2007. A genomic survey of the fish parasite *Spiroplasma salmonicida* indicates genomic plasticity among diplomonads and significant lateral gene transfer in eukaryote genome evolution. *BMC Genomics*8:51. doi:10.1186/1471-2164-8-51.
- Arroyo, R., Gonzalez-Robles, A., Martinez-Palomo, A. & Alderete, J. F. 1993. Signalling of *Trichomonas vaginalis* for amoeboid transformation and adhesion synthesis follows cytoadherence. *Mol Microbiol*, **7**:299-309.
- Bradford MM.1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*72:248–254
- Briselden AM, Hillier SL. 1994. Evaluation of affirm VP Microbial Identification Test for *Gardnerella vaginalis* and *Trichomonas vaginalis*. *J Clin Microbiol.* **32(1)**:148-52
- Carlton, J. M., Hirt, R. P., Silva, J. C., Delcher, A. L., Schatz, M., Zhao, Q., Wortman, J. R., Bidwell, S. L., Alsmark, U. C., Besteiro, S., Sicheritz-Ponten, T., Noel, C. J., Dacks, J. B., Foster, P. G., Simillion, C., Van de Peer, Y., Miranda-Saavedra, D., Barton, G. J., Westrop, G. D., Muller, S., Dessi, D., Fiori, P. L., Ren, Q., Paulsen, I., Zhang, H., Bastida-Corcuera, F. D., Simoes-Barbosa, A., Brown, M. T., Hayes, R. D., Mukherjee, M., Okumura, C. Y., Schneider, R., Smith, A. J., Vanacova, S., Villalvazo, M., Haas, B. J., Perlea, M., Feldblyum, T. V., Utterback, T. R., Shu, C. L., Osoegawa, K., de Jong, P. J., Hrdy, I., Horvathova, L., Zubacova, Z., Dolezal, P., Malik, S. B., Logsdon, J. M., Jr., Henze, K., Gupta, A., Wang, C. C., Dunne, R. L., Upcroft, J. A., Upcroft, P., White, O., Salzberg, S. L., Tang, P., Chiu, C. H., Lee, Y. S., Embley, T. M., Coombs, G. H., Mottram, J. C., Tachezy, J., Fraser-Liggett, C. M. & Johnson, P. J. 2007. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. *Science*, **315**:207-12.
- Cotch MF, Pastorek JG 2nd, Nugent RP, Hillier SL, Gibbs RS, Martin DH, Eschenbach DA, Edelman R, Carey JC, Regan JA, Krohn MA, Klebanoff MA, Rao AV, Rhoads GG. 1997. *Trichomonas vaginalis* associated with low birth weight and preterm delivery. The Vaginal Infections and Prematurity Study Group. *Sex Transm Dis.* **24(6)**:353-60.

-
- Cramer, P., Srebrow, A., Kadener, S., Werbach, S., de la Mata, M., Melen, G., Nogues, G. & Kornblihtt, A. R. 2001. Coordination between transcription and pre-mRNA processing. *FEBS Lett*, **498**:179-82.
- Cho, E. J., Takagi, T., Moore, C. R. & Buratowski, S. 1997. mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev*, **11**:3319-26.
- Espinosa, N., Hernandez, R., Lopez-Griego, L. & Lopez-Villasenor, I. 2002. Separable putative polyadenylation and cleavage motifs in *Trichomonas vaginalis* mRNAs. *Gene*, **289**:81-6.
- Fouts AC, Kraus SJ. 1980. *Trichomonas vaginalis*: reevaluation of its clinical presentation and laboratory diagnosis. *J Infect Dis*. **141**(2):137-143.
- Franco B, Hernández R, López-Villaseñor I. 2012. *Trichomonas vaginalis* ribosomal RNA: identification and characterisation of the transcription promoter and terminator sequences. *Mol Biochem Parasitol*. 185(1): 1-9.
- Fong, N. & Bentley, D. L. 2001. Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. *Genes Dev*, **15**:1783-95.
- Germot, A., Philippe, H. & Le Guyader, H. 1996. Presence of a mitochondrial-type 70-kDa heat shock protein in *Trichomonas vaginalis* suggests a very early mitochondrial endosymbiosis in eukaryotes. *Proc Natl Acad Sci U S A*, **93**:14614-7.
- Gomez-Conde, E., Mena-Lopez, R., Hernandez-Jauregui, P., Gonzalez-Camacho, M. & Arroyo, R. 2000. *Trichomonas vaginalis*: chromatin and mitotic spindle during mitosis. *Exp Parasitol*, **96**:130-8.
- Gomez, C., Esther Ramirez, M., Calixto-Galvez, M., Medel, O. & Rodriguez, M. A. Regulation of gene expression in protozoa parasites. *J Biomed Biotechnol*, **2010**:726045.
- Hahn, S. 2004. Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol*, **11**:394-403.
- Harp D. F. & Chowdhury, I. 2011. Trichomoniasis: evaluation to execution. *Eur J Obstet Gynecol Reprod Biol*, **157**:3-9.-associated, subtilisin-like serine protease in *Trichomonas vaginalis*. *Parasitology*. 137(11): 1621-35.
- Hernández-Romano P, Hernández R, Arroyo R, Alderete JF, López-Villaseñor I. 2010. Identification and characterization of a surface
- Howe, K. J. 2002. RNA polymerase II conducts a symphony of pre-mRNA processing activities. *Biochim Biophys Acta*, **1577**:308-24.
- Hussein EM, Atwa MM. 2008. Infectivity of *Trichomonas vaginalis* pseudocysts inoculated intravaginally in mice. *J Egypt Soc Parasitol*, **38**(3):749-62.
- Johnston VJ, Mabey DC. 2008. Global epidemiology and control of *Trichomonas vaginalis*. *Curr Opin Infect Dis*. **21**(1):56-64

-
- Lahti, C. J., d'Oliveira, C. E. & Johnson, P. J. 1992. Beta-succinyl-coenzyme A synthetase from *Trichomonas vaginalis* is a soluble hydrogenosomal protein with an amino-terminal sequence that resembles mitochondrial presequences. *J Bacteriol*, **174**:6822-30.
- Lau, A. O., Smith, A. J., Brown, M. T. & Johnson, P. J. 2006. *Trichomonas vaginalis* initiator binding protein (IBP39) and RNA polymerase II large subunit carboxy terminal domain interaction. *Mol Biochem Parasitol*, **150**:56-62.
- Lehker M. W. & Alderete, J. F. 2000. Biology of trichomonosis. *Curr Opin Infect Dis*, **13**:37-45.
- Leitsch D. 2016. Recent Advances in the *Trichomonas vaginalis* field. *F1000Res*. **11**;5. pii: F1000 Faculty Rev-162.
- Liston DR, Johnson PJ. 1999. Analysis of a ubiquitous promoter element in a primitive eukaryote: early evolution of the initiator element. *Mol Cell Biol*. **19**(3):2380-8.
- Martínez-Ruboco FW, Kohler R, Van der Waterbeemd M, Heck AJ, Hemann M, Herzog F, Stark H. & Cramer P. 2015. Molecular Basis of transcription-coupled pre-mRNA capping. *Mol Cell*, **58**(6):1079-89.
- Meyer, S., Temme, C. & Wahle, E. 2004. Messenger RNA turnover in eukaryotes: pathways and enzymes. *Crit Rev Biochem Mol Biol*, **39**:197-216.
- Paule, M. R. & White, R. J. 2000. Survey and summary: transcription by RNA polymerases I and III. *Nucleic Acids Res*, **28**:1283-98.
- Petrin, D., Delgaty, K., Bhatt, R. & Garber, G. 1998. Clinical and microbiological aspects of *Trichomonas vaginalis*. *Clin Microbiol Rev*, **11**:300-17.
- Ribeiro, K. C., Monteiro-Leal, L. H. & Benchimol, M. 2000. Contributions of the axostyle and flagella to closed mitosis in the protists *Tritrichomonas foetus* and *Trichomonas vaginalis*. *J Eukaryot Microbiol*, **47**:481-92.
- Quon DV, Delgadillo MG, Khachi A, Smale ST, Johnson PJ. 1994. Similarity between a ubiquitous promoter element in an ancient eukaryote and mammalian initiator elements. *Proc Natl Acad Sci U S A*. **91**(10):4579-83.
- Roger, A. J., Clark, C. G. & Doolittle, W. F. 1996. A possible mitochondrial gene in the early-branching amitochondriate protist *Trichomonas vaginalis*. *Proc Natl Acad Sci U S A*, **93**:14618-22.
- Schumacher, M. A., Lau, A. O. & Johnson, P. J. 2003. Structural basis of core promoter recognition in a primitive eukaryote. *Cell*, **115**:413-24.
- Schwebke, J. R. & Burgess, D. 2004. Trichomoniasis. *Clin Microbiol Rev*, **17**:794-803, table of contents.
- Seña AC, Miller WC, Hobbs MM, Schwebke JR, Leone PA, Swygard H, Atashili J, Cohen MS. 2007. *Trichomonas vaginalis* infection in male sexual partners: implications for diagnosis, treatment, and prevention. *Clin Infect Dis*. **44**(1):13-22

Sheets, M. D., Ogg, S. C. & Wickens, M. P. 1990. Point mutations in AAUAAA and the poly (A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation in vitro. *Nucleic Acids Res*, **18**:5799-805.

Simoës-Barbosa, A., Meloni, D., Wohlschlegel, J. A., Konarska, M. M. & Johnson, P. J. 2008a. Spliceosomal snRNAs in the unicellular eukaryote *Trichomonas vaginalis* are structurally conserved but lack a 5'-cap structure. *RNA*, **14**:1617-31.

Simoës-Barbosa, A., Louly, C., Franco, O. L., Rubio, M. A., Alfonzo, J. D. & Johnson, P. J. 2008b. The divergent eukaryote *Trichomonas vaginalis* has an m7G cap methyltransferase capable of a single N2 methylation. *Nucleic Acids Res*, **36**:6848-58.

Smith, A. J., Chudnovsky, L., Simoës-Barbosa, A., Delgadillo-Correa, M. G., Jonsson, Z. O., Wohlschlegel, J. A. & Johnson, P. J. Novel core promoter elements and a cognate transcription factor in the divergent unicellular eukaryote *Trichomonas vaginalis*. *Mol Cell Biol*, **31**:1444-58.

Sorvillo F, Kerndt P. 1998. *Trichomonas vaginalis* and amplification of HIV-1 transmission. *Lancet*. **351(9097)**:213-4.

Thorburn, AL. 1974. Alfred Francois Donné, 1801-1878, discoverer of *Trichomonas vaginalis* and of leukaemia. *Br J vener Dis*, **50(5)**:377-380.

Upcroft P. & Upcroft J. A. 2001. Drug targets and mechanisms of resistance in the anaerobic protozoa. *Clin Microbiol Rev*, **14**:150-64.

Van Grinsven KW1, Rosnowsky S, Van Weelden SW, Pütz S, Van der Giezen M, Martin W, Van Hellemond JJ, Tielens AG, Henze K. 2008. Acetate:succinate CoA-transferase in the hydrogenosomes of *Trichomonas vaginalis*: identification and characterization. *J Biol Chem*. **18;283(3)**:1411-8.

Vanacova, S., Liston, D. R., Tachezy, J. & Johnson, P. J. 2003. Molecular biology of the amitochondriate parasites, *Giardia intestinalis*, *Entamoeba histolytica* and *Trichomonas vaginalis*. *Int J Parasitol*, **33**:235-55.

Vanacova, S., Yan, W., Carlton, J. M. & Johnson, P. J. 2005. Spliceosomal introns in the deep-branching eukaryote *Trichomonas vaginalis*. *Proc Natl Acad Sci U S A*, **102**:4430-5.

Viikki M1, Pukkala E, Nieminen P, Hakama M. 2000. Gynaecological infections as risk determinants of subsequent cervical neoplasia. *Acta Oncol*. **39(1)**:71-5.

Woychik, N. A. & Hampsey, M. 2002. The RNA polymerase II machinery: structure illuminates function. *Cell*, **108**:453-63.

Yuh, Y. S., Liu, J. Y. & Shaio, M. F. 1997. Chromosome number of *Trichomonas vaginalis*. *J Parasitol*, **83**:551-3.

Yusof, A. & Kumar, S. 2012 Ultrastructural changes during asexual multiple reproduction in *Trichomonas vaginalis*. *Parasitol Res*, **110**:1823-8.

Zubáková Z, Cimbůrek Z, Tachezy J. 2008. Comparative analysis of trichomonad genome sizes and karyotypes. *Mol Biochem Parasitol.* **161(1)**:49-54