



Universidad Nacional Autónoma de México.

---

---

Facultad de Estudios Superiores Iztacala.

Instituto de Investigaciones Biomédicas.

Clonación del factor de corte y  
poliadenilación CPSF-160 de  
*Trichomonas vaginalis*.

TESIS

QUE PARA OBTENER EL TÍTULO DE

BIÓLOGO

PRESENTA:

ERIK ALEJANDRO ALVAREZ ARCE.



Directora de Tesis: Dra. María Imelda López Villaseñor.

México D. F.

2012.



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.

ESTE TRABAJO DE TESIS SE DESARROLLO EN EL DEPARTAMENTO DE BIOLOGÍA MOLECULAR Y BIOTECNOLOGÍA DEL INSTITUTO DE INVESTIGACIONES BIOMÉDICAS DE LA UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO. EL TRABAJO FUE DESARROLLADO BAJO LA ASESORÍA DE LA DOCTORA MARÍA IMELDA LÓPEZ VILLASEÑOR.

# *Dedicatoria.*

---

---

## *A mi familia.*

A mis padres, por darme la vida, por darme su vida; por ser mis amigos, cómplices y defensores de mi sinsentido. Gracias mamá y gracias papá, porque sin su esfuerzo nada de lo que soy sería ni remotamente posible, gracias por hacerme lo que soy. Gracias por demostrarme que a pesar de creerme el muy fuerte, no aguantaría ni un instante luchando como ustedes; que a pesar de creerme el muy sabio, hoy me doy cuenta que no sé nada comparándome a ustedes.

A mi hermano, por tu ejemplo y tus regaños, por tu tiempo y tu apoyo, porque eres el mejor amigo y compañero que la vida pudo escoger para mí.

## *A mis amigos de Willi's Bar.*

A Will por ser la mejor mala influencia que he tenido, por los días ligeros, por tu compañía y tus demencias que han cooperado a ser parte de mi cordura. Gracias Almendro porque me has permitido pensar que algún día, de seguir tus pasos, puedo madurar y ser responsable aunque no vaya conmigo. Gracias Enano porque del regaño a tiempo, conseguí esto en donde ahora escribo, porque tu consejo me llevó a encontrar los pasos que había perdido. Gracias Lucio por ser mi colega y mostrarme el camino a seguir. Gracias Angelito por compartir cada momento, porque fuiste el primero que creyó en el otro sueño. Gracias Paco hasta halla en tu rincón del cielo, porque a veces me gusta pensar que pasas a visitarme y ver cómo me ha ido, hoy te escribo aquí, que no me va tan mal.

### *A mis amigos de la FES Izacala.*

Gracias a mis queridos compañeros de viaje, Ángel Ongay, Isaac, Adrián y Paulina, que compartimos días, tardes y noches de trabajo. Que me conceden el privilegio de regalarme su amistad, gracias por tantas horas compartidas de trabajo y de futbol.

Gracias Ángel por ser mi amigo, porque cada una de mis pisadas en la arena, para fortuna mía, siempre ha estado acompañada de las tuyas. Gracias por el esfuerzo compartido. Gracias Isaac por tus palabras, por tus consejos y porque de alguna forma demente queremos mejorar las cosas desde nuestra trinchera.

### *A mis amigos del laboratorio.*

Sin su sabio consejo, no hubiera llegado a ningún lado, gracias a Bernardo y Vane, sin ustedes seguramente hubiera destruido el laboratorio en un par de ocasiones, gracias por su tiempo, su esfuerzo, su paciencia y su ejemplo. Quizá si me esfuerzo, algún día pueda ser la mitad de lo maravillosos que son ustedes.

# *Agradecimientos.*

---

---

A la Universidad Nacional Autónoma de México, por darme la oportunidad de ser uno de sus hijos, por darme las herramientas necesarias para considerarme un mejor profesional y un mejor ser humano.

A la Dra. María Imelda López Villaseñor, por enseñarme a disfrutar de mi trabajo, por sus consejos, por haberme permitido formar parte de su grupo de trabajo.

Al Dr. Roberto Hernández y su grupo de trabajo, por sus consejos cada seminario.

A la QFB. Vanesa Fuentes Morales.

A los miembros del jurado:

Dr. Elías Piedra Ibarra.	Presidente.
Dr. Santiago Martínez Calvillo.	Vocal.
Dra. María Imelda López Villaseñor.	Secretario.
Dr. Ignacio Peñaloza Castro.	Suplente.
Dra. Emma Bertha Gutiérrez Cirlos Madrid.	Suplente.

## LISTA DE ABREVIATURAS.

°C: Grados Celsius.

A: Adenina.

ATP: Adenosin Trifosfato.

C: Citosina.

cDNA: Cadenas de DNA complementarias al RNA.

CF: Factor de corte.

CPSF: Factor de corte y poliadenilación.

CPSF-160: Factor de Corte y Poliadenilación (160KDa).

CstF: Factor de estimulación de corte.

DEPC: Dietilpircarbonato (inhibidor de RNAasa).

DNA: Ácido desoxirribonucleico.

G: Guanina.

Hrs: Horas.

Icr: Región de control interna.

Kb: Kilobases.

mins: Minutos.

µm: Micrómetros.

mRNA: RNA mensajero.

nm: Nanómetros.

OD: Densidad Óptica.

PABP: Proteína de unión a la polimerasa.

PAS: Señal de poliadenilación.

pH: Potencial de hidrógeno.

RNA: Ácido ribonucleico.

RNApol II: RNA polimerasa II.

rpm: Revoluciones por minuto.

rRNA: RNA ribosomal.

SDS: Dodecil Sulfato de Sodio.

seg: Segundos.

T: Timina.

tRNA: RNA de transferencia.

U: Uracilo.

## CLONACIÓN DEL FACTOR DE CORTE Y POLIADENILACIÓN CPSF160 DE *Trichomonas vaginalis*.

<b>1. INTRODUCCIÓN.</b>	<b>1</b>
1.1. Clasificación taxonómica de <i>Trichomonas vaginalis</i> .	1
1.2. Generalidades de <i>T. vaginalis</i> .	2
1.2.1. Morfología.	2
1.3. Organización celular.	4
1.3.1. Citoplasma.	4
1.3.2. Núcleo.	4
1.3.3. Metabolismo.	5
1.4. Ciclo de vida.	6
1.5. Tricomonosis.	6
1.5.1. Manifestaciones clínicas.	6
1.5.2. Epidemiología.	7
1.6. Aspectos moleculares de <i>T. vaginalis</i> .	8
1.6.1. Importancia biológica.	8
1.6.2. Organización genética.	8
1.6.3. Regulación de la expresión génica.	9
<b>2. ANTECEDENTES.</b>	<b>11</b>
2.1. Transcripción en procariontes.	11
2.2. Transcripción en eucariontes.	12
2.2.1. Polimerasas.	13
2.3. Proceso de poliadenilación en eucariontes superiores.	13
2.3.1. Relevancia de la cola de poli (A) en los mensajeros	14
2.3.2. Elementos en la secuencia del premRNA en mamíferos.	17
2.3.3. Factores protéicos involucrados en el proceso de poliadenilación en mamíferos.	18
2.3.4. Modelo de procesamiento y poliadenilación en mamíferos.	21
2.4. Proceso de poliadenilación en levaduras.	22
2.4.1. Elementos en la secuencia del premRNA en levaduras.	22
2.4.2. Factores protéicos involucrados en el proceso de poliadenilación en levaduras.	24
2.5. Señales de poliadenilación en <i>T. vaginalis</i> .	25

<b>3. HIPÓTESIS.</b>	<b>26</b>
<b>4. JUSTIFICACIÓN.</b>	<b>26</b>
<b>5. OBJETIVOS.</b>	<b>26</b>
5.1. Objetivo General.	26
5.2. Objetivos Particulares.	26
<b>6. METODOLOGÍA.</b>	<b>27</b>
6.1. Diseño experimental.	27
6.2. Alineamientos de secuencias.	28
6.3. Cultivo de <i>T. vaginalis</i> .	28
6.4. Extracción de DNA.	29
6.5. Condiciones del PCR.	30
6.6. Purificación de fragmentos de DNA.	31
6.7. Clonación del producto de PCR en el vector pCR-Blunt II TOPO.	31
6.8. Obtención del DNA plasmídico.	32
6.9. Extracción de RNA total.	33
6.10. Purificación de RNA Poli(A)+ a partir de RNA total.	34
6.11. Northern blot.	35
6.12. Preparación de la sonda radiactiva.	37
6.13. Pre-hibridación de la membrana.	38
6.14. Hibridación de la membrana.	38
6.15. Lavado de la membrana.	38
6.16. Subclonación de CPSF-160 en el vector de expresión pQE30.	39
6.17. Transformación de <i>E. coli</i> con el vector pQE30CPSF160.	40
6.18. Expresión del vector pQE30CPSF160.	40
6.19. Modelado <i>in silico</i> de CPSF-160.	41
<b>7. RESULTADOS.</b>	<b>42</b>
7.1. Alineamiento de secuencias.	42

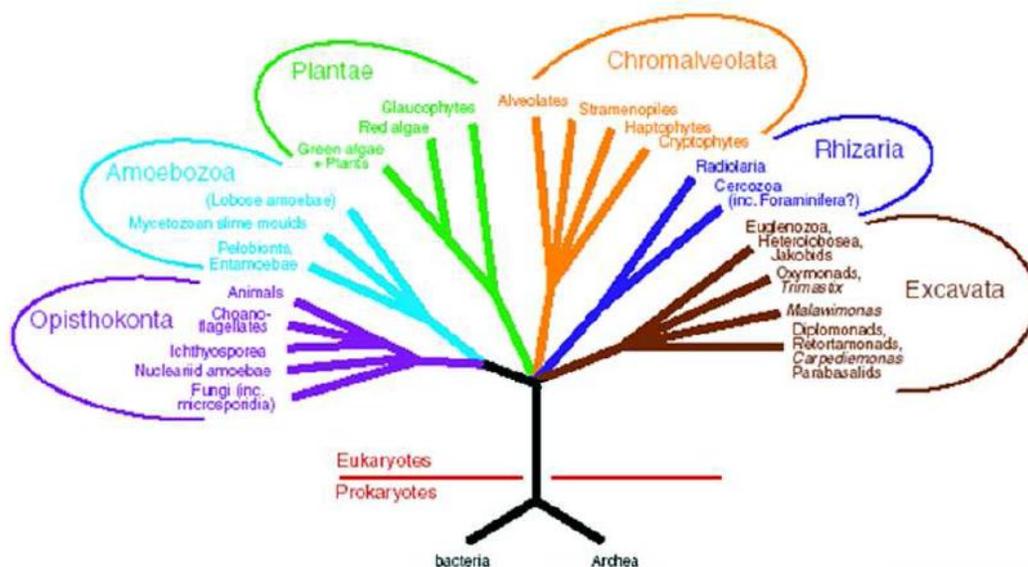
7.2. Análisis de la secuencia de TVAG-077030.	44
7.3. Extracción de DNA genómico y amplificación del gen de CPSF-160 DE <i>T. vaginalis</i> .	46
7.4. Análisis del plásmido TOPOCPSF-160.	47
7.5. Subclonación de CPSF-160 en el vector de expresión pQE30.	49
7.6. Secuenciación y análisis de la clona 4pQE30CPSF160.	52
7.7. Northern blot.	59
7.8. Expresión de la proteína TvCPSF-160 a partir del vector pQE30CPSF160.	61
7.9. Análisis de uso de codones.	62
<b>8. DISCUSIÓN.</b>	<b>64</b>
<b>9. CONCLUSIONES Y PERSPECTIVAS.</b>	<b>66</b>
<b>10. REFERENCIAS BIBLIOGRÁFICAS.</b>	<b>67</b>
<b>11. APÉNDICES.</b>	<b>79</b>

# 1 INTRODUCCIÓN.

## 1.1 CLASIFICACIÓN TAXONÓMICA DE *Trichomonas vaginalis*.

*Trichomonas vaginalis* es un organismo asexual, protozooario anaerobio y flagelado (Klenia P. y col. 2004). Es parásito del tracto urogenital humano (Vanacova S. y col. 2003) además de ser un miembro de las ramificaciones más tempranas del linaje evolutivo de los eucariontes, la clase de los parabasálidos (Keeling P. J. y col. 2000).

La clasificación de *T. vaginalis* se muestra según el modelo de clasificación de eucariontes más aceptado en la actualidad que es el de los seis súper-grupos: Amoebozoa, Opisthokonta, Excavata, Chromalveolata, Plantae y Rhizaria. Este modelo está basado en la comprensión actual de la relación evolutiva de los organismos y los conocimientos en filogenética molecular (Fig. 1.) (Simpson A.G. y Roger A. J. 2004).



**Fig. 1.** Diagrama que describe la organización de la mayoría de los eucariontes en seis súper-grupos. (Tomado de Simpson A.G. y Roger A.J. 2004.).

Dentro del grupo Excavata se encuentra el Orden Trichomonadida compuesto de organismos unicelulares (Levine N.D. 1985). A este Orden pertenece el parásito *T. vaginalis* dentro de la Familia Trichomonadidae. La característica de la Familia Trichomonadidae es que todos los miembros poseen hidrogenosomas, organelos de doble membrana que realizan una función similar a la de las mitocondrias, produciendo energía en forma de ATP a través del metabolismo del piruvato (Hernández P. A. 2010).

*T. vaginalis* pertenece al Subfilo de los Sarcomastigóforos que tiene como característica la presencia de uno o más flagelos a lo largo de su ciclo de vida; pertenece a la Clase de los Zoomastigóforos cuya principal característica es carecer de cromatóforos. Su nutrición es parasitaria y su reproducción de tipo asexual por fisión binaria longitudinal (Hernández P. A. 2010).

Súper Reino: Eukaryota  
Reino: Protista  
Filo: Protozoa  
Sin clasif: Excavata  
Subfilo: Sarcomastigophora  
Clase: Parabasalia  
Orden: Trichomonadida  
Familia: Trichomonadidae  
Género: Trichomonas  
Especie: *Trichomonas vaginalis*

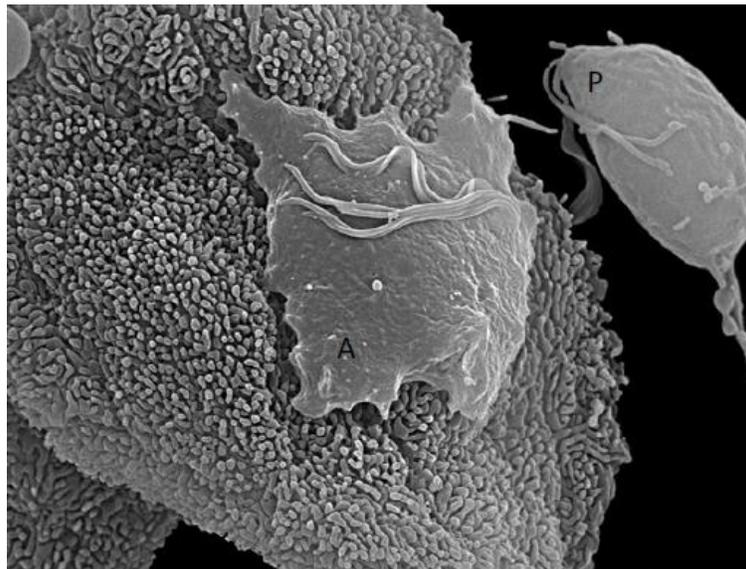
## **1.2 GENERALIDADES DE *T. vaginalis*.**

### **1.2.1 MORFOLOGÍA.**

*T. vaginalis* está rodeado por una membrana plasmática de 9-10nm de grosor, semejante a las de eucariontes superiores (Nadler S. A. y Honigberg B. M. 1988). El parásito presenta cuatro flagelos ubicados anteriormente y un quinto flagelo adosado a la membrana ondulante. También posee

una estructura hialina llamada axostilo que comienza en el núcleo y recorre longitudinalmente a la célula. El tamaño de este parásito puede variar en condiciones silvestres o en cultivos, pero generalmente sus dimensiones van de 10-20µm de largo por 7µm de diámetro (Petrin y col. 1998). Es posible observar dos formas diferentes de estos parásitos: las formas ovoides y redondas, y las formas ameboides aplanadas (Fig. 2). Se puede encontrar a los organismos ovoides en cultivo líquido y secreciones vaginales, mientras que las formas aplanadas se encuentran a menudo en frotis y biopsias. Cuando las células de *T. vaginalis* de tipo ovoide entran en contacto con tejido del epitelio vaginal (*in vitro*) las células adoptan la forma ameboide (Arroyo R. y col. 1993).

En su interacción con células del epitelio vaginal *in vitro*, los microorganismos de forma ovoide se transforman morfológicamente en formas ameboides en menos de tres minutos: forman lamelipodios y son capaces de emitir pseudópodos que se interdigitan en diferentes sitios con la membrana plasmática de las células blanco, reduciendo el espacio intercelular en las áreas de contacto (Arroyo R. y col. 1993; Arroyo R. 2000). El cambio en la forma de ovoide a ameboide le confiere al parásito un aumento en la superficie de unión a la célula del epitelio vaginal, lo cual puede ser necesario para llevar a cabo una parasitosis más estable y eficiente, además puede ser un factor de virulencia (Arroyo R. y col. 1993).



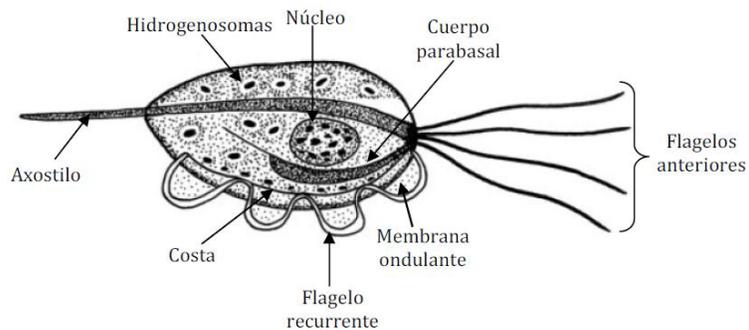
**Fig. 2. Adhesión de *T. vaginalis* a células del epitelio vaginal. Se observa la diferencia entre la forma ameboide (A) y la ovoide (P) (Pereira-Neves A. Benchimol M. 2008).**

### 1.3 ORGANIZACIÓN CELULAR.

#### 1.3.1 CITOPLASMA.

En el citoplasma de *T. vaginalis* se han identificado las siguientes estructuras: aparato de Golgi, retículo endoplásmico, sistema vacuolar rico en fagosomas y fagolisosomas, gránulos paraxostilares y paracostales (hidrogenosomas), gránulos de glucógeno, ribosomas libres y asociados al retículo endoplásmico y un citoesqueleto muy complejo. Este organismo no posee mitocondrias ni peroxisomas, en su lugar posee un hidrogenosoma en el cuál se lleva a cabo la generación de ATP (Petrin y col. 1998).

El citoesqueleto de *T. vaginalis* consiste en un complicado sistema de microtúbulos que se clasifican de acuerdo a su posición como axostilares, flagelares o mitóticos que forman estructuras como el axostilo, la costa y la pelta. Cuatro flagelos anteriores se originan en el complejo cinetosomal anterior y un flagelo recurrente adosado a la membrana ondulante también parte del complejo cinetosomal anterior. El axostilo, los flagelos anteriores y el flagelo recurrente participan en la motilidad del parásito (Fig. 3.) (Hernández P. A. 2010).



**Fig. 3. Representación esquemática de *T. vaginalis* (Tomado de Hernández P. A. 2010).**

#### 1.3.2 NÚCLEO.

El núcleo en *T. vaginalis* se encuentra localizado en la porción anterior del organismo. Es piriforme, es decir, presenta una forma de pera alargada y está rodeado de poros nucleares. Mide entre 4 y

5µm, se encuentra limitado por una envoltura nuclear y se pueden observar gránulos diseminados de cromatina. El núcleo contiene un nucléolo esférico pequeño, rodeado de un área libre de cromatina (Barrera G. 2006).

### **1.3.3 METABOLISMO.**

Uno de los aspectos más notorios de *T. vaginalis* es que posee un metabolismo anaerobio, debido a que carece de mitocondrias. *T. vaginalis* posee un organelo llamado hidrogenosoma donde se realiza el metabolismo de carbohidratos. Los hidrogenosomas se dividen de manera autónoma por fisión binaria, y entre sus características está la de producir hidrógeno molecular y ATP. A diferencia de las mitocondrias, los hidrogenosomas no poseen material genético, no presentan citocromos, tampoco hay un ciclo de ácidos tricarbónicos ni reacciones de fosforilación oxidativa; en su lugar utilizan enzimas (piruvato-ferredoxina oxidoreductasa e hidrogenasa) típicamente restringidas a organismos anaerobios (Müller. 1993).

En el hidrogenosoma se identificó y analizó la Beta-succinil-CoA sintetasa ( $\beta$ -SCS), una proteína típicamente localizada en la mitocondria de eucariontes superiores (Lahti C. J. y col. 1992). Otros estudios realizados en este parásito han mostrado que en este organelo existen proteínas relacionadas con la mitocondria, lo cual sugiere que *T. vaginalis* y otros protistas amitocondriados, se separaron del linaje eucariota antes de la endosimbiosis que dio origen a las mitocondrias (Roger A. J. y col. 1996).

La energía metabólica se deriva del catabolismo de la glucosa a piruvato o succinato a través de la vía glicolítica. Se puede generar energía adicional cuando el piruvato se convierte a acetato como resultado de una reacción ligada a la coenzima A; esto ocurre en el hidrogenosoma, que contiene a las enzimas responsables del metabolismo del piruvato (Lindmark y Muller. 1973). *T. vaginalis* es un parásito obligado debido a que es incapaz de sintetizar algunas macromoléculas esenciales, tal es el caso de algunas purinas, pirimidinas y lípidos. Estas macromoléculas son obtenidas del medio, tanto de las secreciones vaginales como de la fagocitosis tanto de células del hospedero y células bacterianas. Es por eso que los medios de cultivo son ricos en lípidos, ácidos grasos, aminoácidos y trazas de metales (Petrin y col. 1998).

## **1.4 CICLO DE VIDA.**

El ciclo celular de *T. vaginalis* presenta las fases M, G1, S y G2 con una predominancia hacia la fase G2 (Riley D. E. y col. 1993). El tiempo de duración aproximado de la mitosis en este organismo es de 45 minutos y puede ser dividido en cinco fases: profase, metafase, anafase temprana y tardía, telofase temprana y tardía y citocinesis (Gómez-Conde E. y col. 2000). El ciclo de vida de este parásito incluye una sola fase, el trofozoito, y es transmitido sexualmente por contacto directo de persona a persona ya que reside en el tracto urogenital humano. *In vitro*, *T. vaginalis* crece óptimamente a pH de 6 a 6.3 (Fuentes V. 2005), condiciones de pH que se presentan en la vagina.

## **1.5 TRICOMONOSIS.**

### **1.5.1 MANIFESTACIONES CLÍNICAS.**

*T. vaginalis* es el agente infeccioso responsable de la tricomonosis, una de las enfermedades más comúnmente transmitida por contacto sexual (WHO, 2001). La infección una vez establecida persiste durante largos periodos en mujeres, en los hombres sólo está presente durante cortos periodos. El periodo de incubación varía de 4 a 28 días (Fuentes V. 2005).

En las mujeres la tricomonosis produce daño en los tejidos del tracto urogenital provocando embarazos de alto riesgo (Cotch M. F. y col. 1997), mayor susceptibilidad a la infección por el virus de la inmunodeficiencia adquirida (Sorvillo F. Kerndt P. 1998), y predisposición de algunas pacientes con tricomonosis al cáncer cérvico-uterino (Vikki M. y col. 2000). Las complicaciones asociadas a la tricomonosis en hombres son: prostatitis, balanopostitis (inflamación del pene), epididimitis, enfermedad uretral e infertilidad (Meysick K. C. y col. 1992; Lehker M. y col. 2000).

El tratamiento primario utilizado para la tricomonosis es el 5-nitroimidazol (metronidazol), el cual es activado en los hidrogenosomas del parásito en donde ocurre la reducción del grupo nitrito del fármaco por la piruvato:ferredoxin oxidoreductasa resultando en la producción de intermediarios citotóxicos que reaccionen con el DNA del parásito. Esto provoca que la división celular y la movilidad cesen y ocurra la muerte de la célula (Petrin y col. 1998; Meysick K. C. y col. 1992; Lehker M y col. 2000).

### 1.5.2 EPIDEMIOLOGÍA.

A nivel mundial se estima que aparecen 174 millones de casos de tricomonosis al año, comparado con los 92 millones de casos de *Chlamydia trachomatis* y los 62 millones de casos de *Neisseria gonorrhoeae* (Johnsotn V. J. Mabey D. C. 2008). Según el CENAVECE (Centro Nacional de Vigilancia Epidemiológica y Control de Enfermedades) en México la Tricomonosis o Tricomoniiasis se encontró en el lugar número 17 de entre las principales 20 causas de morbilidad en el año 2008 (Fig. 4.) (CENAVECE. 2009). Los estados que reportan las mayores incidencias de tricomonosis en la población general son Veracruz con 12,811 casos, el Estado de México con 7,412 casos y Puebla con 7,310 casos registrados para 2009 (CENAVECE. 2009).

Número	Padecimiento	Número	Tasa*
1	Infecciones respiratorias agudas	24 120 252	22 609.376
2	Infecciones intestinales por otros organismos y las mal definidas	4 701 671	4 407.162
3	Infección de vías urinarias	3 244 994	3 041.73
4	Úlceras, gastritis y duodenitis	1 523 116	1 767.067
5	Otitis media aguda	666 493	624.7
6	Amebiasis intestinal	531 818	445.7
7	Hipertensión arterial	526 484	698.6
8	Gingivitis y enfermedades periodontales	513 456	481.3
9	Diabetes mellitus no insulino dependiente (Tipo II)	396 374	371.5
10	Conjuntivitis	376 485	352.9
11	Otras helmintiasis	325 728	305.3
12	Varicela	323 981	303.7
13	Asma y estado asmático	298 652	279.9
14	Candidiasis urogenital	295 912	277.4
15	Intoxicación por picadura de alacrán	271 976	254.9
16	Neumonías y bronconeumonías	146 454	137.3
17	Tricomoniiasis urogenital	136 002	127.5
18	Desnutrición leve	129 041	121.0
19	Paratifoidea y otras salmonelosis	122 422	114.8
20	Quemaduras	117 435	110.1
Total 20 principales causas		38 768 746	36 340.29
Otras causas		1 244 325	1 166.381
TOTAL GLOBAL		40 013 071	37 506.7

**Fig. 4. Veinte principales causas de morbilidad en los Estados Unidos Mexicanos en el año 2008.** (Tomado de CENAVECE, Epidemiología 2009). Tasa de 100,000 habitantes.

## **1.6 ASPECTOS MOLECULARES DE *T. vaginalis*.**

### **1.6.1 IMPORTANCIA BIOLÓGICA.**

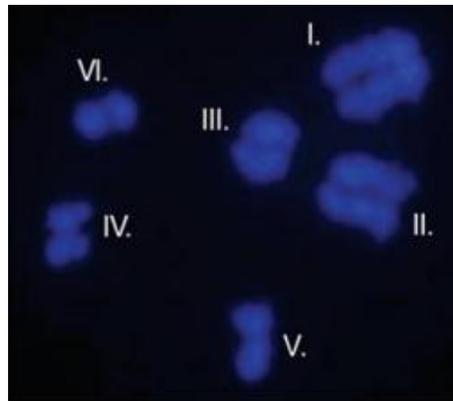
Los eucariontes unicelulares tienen un papel muy importante en la historia evolutiva, juegan un papel como intermediarios entre procariontes y eucariontes por un lado, y eucariontes unicelulares, plantas y animales por otro lado. Muchos de los actuales protistas, presumiblemente corresponden a descendientes de alguno de los grupos que proliferaron durante estas dos transiciones evolutivas importantes (Baroin A. y col. 1988). *T. vaginalis* es un protozoo flagelado de evolución temprana (Barrera G. 2006). Este parásito representa uno de los primeros organismos que divergieron en los linajes eucarióticos según estudios comparativos de secuencias de la subunidad grande y la subunidad pequeña del RNA ribosomal (rRNA) (Baroin A. y col. 1988).

A principios del siglo XXI se hicieron análisis filogenéticos más completos y complejos utilizando diversas fuentes de información, como son las secuencias de DNA de varios genes (no sólo los ribosomales) y características morfológicas y fisiológicas de los organismos como de secuencias de la subunidad pequeña de rRNA (Sogin M.L. 1991). De estos estudios es que surgió el modelo de los 6 súper-grupos (Amoebozoa, Opisthokonta, Excavata, Chromalveolata, Plantae y Rhizaria) (Fig. 1). *T. vaginalis* puede proporcionarnos información que nos permita comprender las relaciones filogenéticas dentro del grupo Excavata al cual pertenece el Orden Trichomonadida. El conocimiento de la biología de *T. vaginalis* podría proporcionarnos también información acerca de la evolución molecular y celular en organismos de evolución temprana (Barrera G. 2006).

### **1.6.2 ORGANIZACIÓN GENÉTICA.**

*T. vaginalis* es un organismo diploide que posee seis pares de cromosomas monocéntricos (Fig. 5) (Yeong-Seng y col. 1997; Yeong-Seng y col. 1998). Con la secuenciación del genoma de *T. vaginalis* (Carlton J. M. y col. 2007) se encontró que este organismo haploide tiene un genoma de alrededor de 176 millones de pares de bases, aunque es difícil calcular el tamaño exacto debido a la gran cantidad de transposones y elementos repetidos que posee y que reflejan una reciente expansión. En el año de 1985 se estimó el tamaño total del genoma utilizando métodos cuantitativos con DNA

aislado de *T. vaginalis* y se concluyó que el tamaño es cerca de la mitad de lo que se considera por la secuencia del genoma (Wang y Wang. 1985). Con estimaciones experimentales midiendo el contenido de DNA por citometría de flujo de varios tricomonádidos, se encontró que en general son organismos con genomas de gran tamaño (86-177 Mb) (Zubáková Z. y col. 2008). En comparación con los genomas secuenciados de otros protozoarios parásitos (cuyos genomas oscilan en su tamaño entre 2.9 y 60 millones de bases), el de *T. vaginalis* es un genoma muy grande (TrichDB, 2012). Se han encontrado muy pocos intrones en los genes de *T. vaginalis*, únicamente en 65 del total de genes anotados. Según lo descrito por Carlton J. M y col, este es uno de los organismos eucariontes con mayor capacidad codificadora, con aproximadamente 99 mil genes codificantes para proteínas (Carlton J. M. y col. 2007; TrichDB. 2012). Se han identificado genes para componentes de la vía de RNA de interferencia (RNAi), lo que podría ser una poderosa herramienta para manipular la expresión génica en *T. vaginalis*. Se han identificado genes con un posible origen procarionte que podrían haber sido adquiridos por transferencia horizontal, la mayoría de ellos relacionados con enzimas metabólicas. Entre los genes identificados también llaman la atención aquellos relacionados con la meiosis. Si bien este organismo no presenta un ciclo sexual de reproducción, es posible que los mecanismos de recombinación de DNA le permitan llevar a cabo eventos de recombinación homóloga e intercambio genético (Malik S. B. y col. 2007).



**Fig. 5.** Cariotipo determinado mediante hibridación de fluorescencia in situ (Fluorescent In Situ Hybridization FISH), tomado de Carlton J. M. y col. 2007. Es posible observar los 6 cromosomas de *T. vaginalis* (I – VI).

### 1.6.3 REGULACIÓN DE LA EXPRESIÓN GÉNICA.

En eucariontes la transcripción de los genes que codifican para proteínas está dirigida por la RNA polimerasa II (pol II). En *T. vaginalis* pol II presenta algunas características diferentes a las RNA polimerasas de eucariontes superiores. Estudios de transcripción *in vitro* han mostrado que la pol II de *T. vaginalis* es resistente a altas concentraciones de  $\alpha$ -amanitina (50% inhibición con 250  $\mu\text{g/ml}$ ), a diferencia de lo observado con otras RNA pol II que son inhibidas con 2-20  $\mu\text{g/ml}$  de este compuesto. Esta resistencia parece deberse a cambios en el sitio probable de unión a  $\alpha$ -amanitina encontrados en la subunidad grande de la pol II de *T. vaginalis* (Liston D.R. y Johnson P. J. 1998).

A pesar del progreso alcanzado en la comprensión de los mecanismos bioquímicos involucrados en la transcripción en eucariontes superiores, se conoce poco acerca de los mecanismos de regulación de la expresión génica en *T. vaginalis*. Por ejemplo en la región 5' de los genes que codifican proteínas se ha caracterizado una secuencia con actividad promotora de la transcripción. Este elemento rico en pirimidinas, conocido como elemento iniciador (Inr), rodea al sitio de inicio de la transcripción de todos los genes de *T. vaginalis* examinados, se compone de 13 nucleótidos cuya secuencia consenso es TCAYTWYTCATTA (Y= T/C, W= T/A) y se localiza de 6 a 20 nucleótidos río arriba del codón de inicio de la traducción (ATG). El análisis del elemento Inr de *T. vaginalis* ha mostrado que es esencial para la transcripción. No se han encontrado cajas TATA típicas de eucariontes a 25-30 nucleótidos río arriba a los sitios de inicio de la transcripción (Quon V. K. D. y col. 1994; Liston D. R. y Johnson P. J. 1999). Se sabe que el iniciador es reconocido por una proteína de 39Kda (IBP39: Initiator binding protein). La región carboxilo terminal de IBP39 interactúa con el CTD (Carboxil Terminal Domain) de la subunidad mayor de la pol II, reclutándola para el inicio de la transcripción (Lau A. O. y col. 2006).

En cuanto a la región 3' de los genes transcritos por pol II, en nuestro grupo de investigación se ha demostrado que en *T. vaginalis*, la secuencia UAAA es la señal de poliadenilación en los precursores de mRNA y la secuencia Y $\downarrow$ (A)<sub>2-5</sub>UU es el sitio de corte (Fuentes V. y col. 2012).

## **2 ANTECEDENTES.**

---

### **2.1 Transcripción en procariontes.**

La definición más simple de un gen es una “unidad de DNA que contiene la información para especificar la síntesis de una única cadena polipeptídica o RNA funcional (tal como un tRNA)” (Lodish H. 2005). Para estudiar la expresión génica en procariontes uno de los modelos más recurrentes es el de la bacteria *Escherichia coli*, se trata de una bacteria que posee un único cromosoma circular (Watson J. D. y col. 2003).

La disposición más frecuente de genes codificadores de proteínas en todos los procariontes tiene un orden notable: los genes dedicados a un único objetivo metabólico se suelen encontrar en una disposición contigua en el DNA, tal disposición de genes en un grupo funcional se denomina operón, porque opera como una única unidad a partir de un único promotor (Lodish H. 2005). La transcripción de un operón produce una hebra continua de mRNA. En el DNA procarionte los genes están estrechamente empaquetados pues hay pocas regiones no codificadoras, carecen de intrones y el DNA se transcribe directamente en RNA colineal para ser traducido a proteínas (Watson J. D. y col. 2003; Lewin B. 2004). Los organismos procariontes poseen una sola RNA polimerasa (Watson J. D. y col. 2003).

Los promotores de procariontes son muy heterogéneos, puede haber promotores fuertes y débiles, esto en función del número de transcritos que se generan en un momento dado, esto se ve afectado en principio debido a la unión de la polimerasa al promotor, la isomerización y la facilidad de la polimerasa para liberarse (Watson J. D. y col. 2003; Lewin B. 2004). En los procariontes no existe una secuencia consenso a partir de la cual se detenga la transcripción, sino que la terminación de la transcripción se da gracias a secuencias ya transcritas de RNA. Se trata de regiones ricas en G y C, con lo que la polimerasa irá más lenta. Además, en esas regiones, las bases están colocadas de manera que al irse transcribiendo se de complementariedad en el RNA y se formen bucles o “hairpins”. (Watson J. D. y col. 2003; Lewin B. 2004). Las secuencias de finalización de la transcripción denominadas terminadores desencadenan la disociación de la polimerasa y la liberación de la hebra de RNA. En las bacterias los terminadores son de dos tipos, independientes de Rho y dependientes de Rho (Watson J. D. y col. 2003).

**Terminadores Intrínsecos o independientes de rho:** Estos terminadores están compuestos por dos elementos en su secuencia: Una repetición invertida corta de más o menos 20 nucleótidos seguida por un segmento de unos 8pb A:U. Cuando la polimerasa transcribe una secuencia repetida e invertida, el RNA resultante puede formar una estructura tallo-asa (horquilla) por apareamiento entre sus propias bases, se cree que esta estructura causa la terminación porque distorsiona el complejo de alargamiento (Watson J. D. y col. 2003).

**Terminadores Dependientes de rho:** Los terminadores dependientes de Rho poseen elementos de RNA que no están tan bien caracterizados y para que funcionen se necesita la acción del factor protéico Rho. La proteína Rho se une al RNA monocatenario conforme sale de la polimerasa. En un primer nivel de especificidad Rho se une a regiones ricas en C. En un segundo nivel de especificidad Rho no se une a transcritos en proceso de traducción. En las bacterias la transcripción y la traducción están acopladas. En consecuencia lo típico es que Rho termine sólo los transcritos que todavía están en proceso de transcripción más allá del extremo de un gen o un operón (Watson J. D. y col. 2003).

## **2.2 TRANSCRIPCIÓN EN EUKARIOTES.**

A diferencia de los procariontes, los genes en el DNA eucarionte están distribuidos en diversos cromosomas. A menudo los genes dedicados a una única vía metabólica están separados físicamente en el DNA. Cada gen eucarionte se transcribe a partir de su propio promotor, produciendo un mRNA, que generalmente produce un solo polipéptido (Lodish H. 2005).

En el sistema génico de eucariontes no hay presencia de operones. Entre los diferentes fragmentos del genoma encontramos regiones codificadoras y regiones no codificadoras. Mientras que en los procariontes la transcripción y traducción pueden suceder al mismo tiempo, en los eucariontes no, ya que el núcleo está separado del citoplasma que es donde ocurre la traducción. Los transcritos primarios de los genes codificadores de proteínas son precursores de los mRNA (pre-mRNA), que deben atravesar varias modificaciones, denominadas procesamiento-post transcripcional del RNA, para producir un RNA funcional. Este RNA debe ser transportado al citoplasma antes de ser traducido, por tanto transcripción y traducción no pueden ocurrir de

manera simultánea (Watson J. D. y col. 2003). El RNA que se forma como resultado de la transcripción se denomina transcrito primario. Se ha de tener en cuenta que existe tres tipos principales de RNA, dos de los cuales son productos finales, como el rRNA y el tRNA, mientras que el mRNA deberá llegar a los ribosomas, donde se traduce para dar lugar a las proteínas (Lewin B. 2004). Existen también los miRNAs' de un tamaño de 20-25 nucleótidos, estos regulan diversos procesos tales como la especificación del destino celular, la apoptosis y el metabolismo (Eisenmann D. M. 2005).

### **2.2.1 POLIMERASAS.**

La transcripción en eucariontes es llevada a cabo por tres distintas RNA polimerasas, son clasificadas como I, II o III. La polimerasa con la mayor actividad es Pol I, esta reside en el nucleólo y es responsable de la transcripción de los genes que codifican para rRNA, excepto el 5S. La enzima que sigue en actividad es Pol II localizada en el nucleoplasma y es responsable de la síntesis de los precursores de mRNA. La polimerasa III es la de menor actividad, esta enzima nucleoplásmica sintetiza los tRNAs y los RNAs pequeños nucleares además del rRNA 5S (Lewin B. 2004; Watson J. D. y col. 2003). Además de las tres RNA polimerasas nucleares, hay que añadir las que podemos encontrar en cloroplastos y mitocondrias.

### **2.3 PROCESO DE POLIADENILACIÓN EN EUKARIONTES SUPERIORES.**

La presencia de la cola de poli (A) en el extremo 3' es una de las características más reconocibles de los mRNA en eucariontes (Dominski Z. Marzluff W. F. 2007), con excepción de los mRNAs que codifican para las proteínas histonas en metazoos (Keller W. y Minvielle-Sebastiat L. 1997). En eucariontes, los precursores de mRNA (pre-mRNAs) se transcriben en el núcleo, a partir del DNA genómico por pol II. Estos pre-mRNAs deben someterse a un extenso procesamiento co-transcripcional antes de que puedan ser transportados al citoplasma para su traducción en proteínas. Los eventos de procesamiento son: Adición del CAP (7 -metil-guanosina), donde la guanosina del extremo 5' terminal es metilada en el carbón número 7; el *splicing* para remover secuencias intrónicas del pre-mRNA y finalmente el corte y poliadenilación en el extremo 3' terminal (Mandel C. R. y col. 2008; Proudfoot y col. 2002).

Un paso crítico común para todos los transcritos primarios es la generación de los extremos 3' maduros poliadenilados. Se requieren señales de poliadenilación en el transcrito para la terminación de la transcripción y para la adición de residuos de adenilato en el extremo 3' (Whitelaw E. Proudfoot N. 1986; Proudfoot N. 1989; McCracken S. y col. 1997).

Varios estudios han demostrado que los transcritos primarios de RNA se extienden más allá del sitio de poliadenilación, lo que indica que el procesamiento del extremo 3' de los pre-mRNA en eucariotas requiere del corte del transcrito primario de RNA y la poliadenilación (Nevins J. 1978). El proceso de poliadenilación requiere entonces de dos reacciones: en la primera el transcrito es cortado endonucleolíticamente generándose dos moléculas; en la segunda reacción el sitio 3'OH generado en el corte es poliadenilado, es decir se agregan de 150 a 200 residuos de adenilato de manera independiente del templado. Estas observaciones contrastan con la forma en que se procesa el extremo 3' en bacterias, donde el extremo 3'-terminal del mRNA se forma por la terminación de la transcripción (Mandel C. R. y col. 2008).

Se ha demostrado que se requiere de una gran colección de factores proteicos (más de 14 moléculas para mamíferos y más de 20 moléculas para levadura) necesarios para el procesamiento final del extremo 3'. Esta maquinaria es dirigida al sitio de corte por elementos que se encuentran dentro de la secuencia del extremo 3' del pre-mRNA (Mandel C. R. y col. 2008). La cola de poli (A) de los mRNA se sintetiza en el núcleo de la célula y tiene una longitud de ~250 nucleótidos en células de mamíferos (Khun U. 2009). Todos estos eventos son necesarios para la maduración de los pre-mRNA (Mandel C. y col. 2008).

### **2.3.1 RELEVANCIA DE LA COLA DE POLI (A) EN LOS MENSAJEROS.**

La formación de la cola de poli (A) en eucariontes depende de tres secuencias presentes en los pre-mRNAs: la secuencia consenso del hexanucleótido AAUAAA, altamente conservado, el sitio de corte definido por la secuencia consenso CA y una región río abajo al sitio de corte rica en guanosinas y uracilo (Colgan D.F. Manley J.L. 1997; Wahle E. Ruegsegger U. 1999; Zhao y col. 1999).

El procesamiento del pre-mRNA en el extremo 3' es de crucial importancia funcional en las células eucariotas, y la interrupción de este proceso tiene efectos catastróficos en el crecimiento de las células y su viabilidad (Mendel C. y col. 2008).

En primer lugar, el procesamiento y poliadenilación del extremo 3' promueven el transporte de los mRNA del núcleo hacia el citoplasma (Vinciguerra P. Stutz F. 2004). La sustitución del sitio de poliadenilación de los pre-mRNA en el extremo 3' por un sitio de corte de rRNA, produce mRNA que se corta, pero no se poliadenila. Esta sustitución disminuye la proporción de mRNA nuclear que migra al citoplasma en diez veces, lo que indica una reducción en el transporte de mRNA y en consecuencia una reducción en la expresión de proteínas (Huang Y. Carmichael G. 1996).

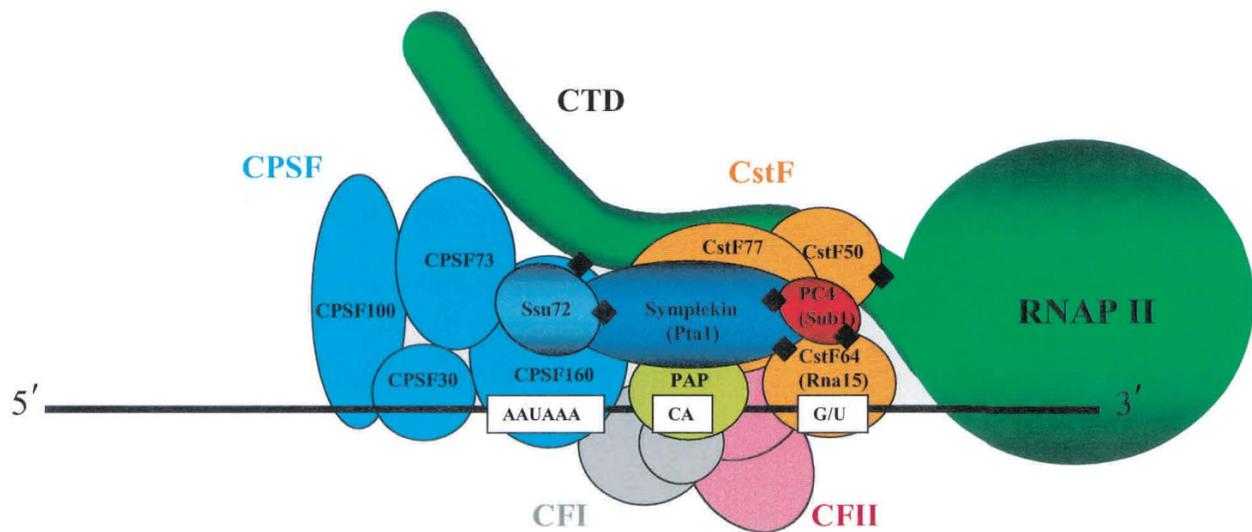
En segundo lugar, el procesamiento del extremo 3' promueve la estabilidad del mRNA. En el citoplasma, los mRNAs son degradados por el extremo 3' principalmente, lo que indica la importancia de proteger el extremo 3' (Wickens M y col. 1997). La adición de la cola de poli (A) y la posterior unión de la Proteína de Unión a la Cola de Poli (A) (Poly A Binding Protein PABP), ha demostrado prevenir la degradación de los mRNAs en células de mamíferos (Ford L. 1997) y ovocitos de *Xenopus* (Wormington M. 1996). De hecho, sólo la presencia de PABP sin la cola de poli (A) puede ser suficiente para prevenir la degradación del mRNA (Coller J. 1998).

En tercer lugar, se ha observado que la cola de poli (A) promueve el inicio de la traducción. Las PABPs asociadas con la cola de poli (A) promueven la traducción mediante la interacción con el factor de iniciación IF4G, probablemente, a través de mecanismos adicionales (Sachs A. 2002; Amrani N. y col. 2008). Estudios han demostrado que en la levadura la sola presencia de la cola de poli (A) es suficiente para iniciar la traducción de manera eficiente; pero la presencia tanto de la cola de poli (A) y el CAP del extremo 5' es óptima para la traducción (Preiss T. Hentze M. 1998).

La longitud de la cola de poli (A) también puede ser importante. Por ejemplo, la regulación génica durante la maduración de los ovocitos y el desarrollo embrionario temprano de los animales depende de la regulación de la traducción de los mRNA de la madre, y los cambios en la longitud de la cola de poli (A) de mRNA específicos juegan un papel importante en esta regulación de la traducción. Las colas de poli (A) largas favorecen la traducción, mientras que una reducción en el tamaño de la cola de poli (A) promueve la disminución de la traducción del mensajero (Wickens M.

y col. 2000; Thompson B. 2007). Mecanismos similares parecen operar en las neuronas (Richter J. Klann E. 2007; Kwak J. y col. 2008) y posiblemente en otras células somáticas (Fig. 6) (Saitoh S. y col. 2002).

Finalmente, el procesamiento del extremo 3' está íntimamente ligado a la transcripción y a la maquinaria del *splicing* (Calvo O. Manley J. 2003; Maniatis T. Reed R. 2002; Proudfoot N. y col. 2002.).



**Fig. 6. Modelo del procesamiento de mRNA en eucariotes (Tomado de Olga C. 2003).**

Poco después de que la RNA pol II inicia transcripción el extremo 5' del RNA naciente se modifica con el 7 metil-guanilato. La terminación de la transcripción está acoplada al proceso de poliadenilación, de tal manera que el corte endonucleolítico para la poliadenilación del mensajero genera también un extremo 5'OH en el RNA que queda río abajo; este RNA es degradado hasta alcanzar a la polimerasa que sigue transcribiendo y así terminar la transcripción. Una serie de residuos de adenosina (A) son añadidos en el extremo 3'. La cola de poli (A) contiene aproximadamente 250 residuos de (A) en mamíferos, 150 en insectos y 100 en levaduras (Lodis H. y col. 2005).

### 2.3.2 ELEMENTOS EN LA SECUENCIA DEL PRE-mRNA EN MAMÍFEROS.

Los pre-mRNAs' de mamíferos contienen 3 elementos de secuencia primarios que definen el sitio de poliadenilación (Zhao J. 1999) que son: La señal de poliadenilación (PAS, por sus siglas en inglés) que es el hexanucleótido AAUAAA (y algunas variantes), el sitio de corte determinado por el dinucleótido CA y una región rica en U o G/U río abajo del sitio de corte, llamado elemento río abajo (DSE, por sus siglas en inglés) (Fig. 7) (Zhao J. 1999; Colgan D.F. Manley J.L. 1997).

La señal de poliadenilación localizada en la mayoría de los transcritos a una distancia de 10 a 30 nucleótidos río arriba al sitio de procesamiento, es una señal altamente conservada para la poliadenilación en mamíferos y es esencial para el procesamiento de los extremos 3' de los mRNAs (Sheets, M. D., 1990). Se ha encontrado una variación del hexámero, en la segunda posición se presenta un U creando el hexámero AUUAAA que está presente en aproximadamente 10% de los mRNAs analizados. Un estudio de más de 400 ESTs en humanos (expressed sequence tags, por sus siglas en inglés); encontró que el 77% presentaba la señal de poliadenilación, de estos, el 75% presentaba la señal AAUAAA y el 20% presentaba el hexámero AUUAAA (Mandel C. y col. 2008). Se ha reportado que la eliminación del elemento AAUAAA suprime el procesamiento y la poliadenilación de los transcritos, mientras que las mutaciones puntuales en dicha secuencia consenso reducen en gran medida la eficiencia del procesamiento en el extremo 3' (Sheets, M. 1990).

Otro elemento es la región río abajo del sitio de procesamiento rica en G/U (DSE). Esta región es menos conservada que la señal de poliadenilación (Zhao J. 1999). La presencia de este elemento fue sugerida por la observación de que la eliminación de más de 10 nucleótidos río abajo del sitio de corte (CA) reducía los niveles de poliadenilación en tres veces (Sittler A. 1994). La DSE se ha observado en dos formas: un elemento rico en G/U que tiene una secuencia de YGUGUUY (Y = pirimidina) y una secuencia rica en U (UUUUU). Se ha visto que en algunos casos solo uno de estos elementos es suficiente para la poliadenilación; sin embargo, también se ha observado que si están presentes los dos tipos de elementos, se presenta un efecto de sinergismo (Chou Z. 1994; Gil A. Proudfoot N. 1987; McLauchlan J. 1985). Mutaciones puntuales en la DSE tienen un efecto menor en la actividad del corte, mientras que la eliminación de segmentos cortos del DSE tiene un impacto significativamente mayor (McDevitt, M. y col. 1986; Zarkower D. Wickens M. 1988).

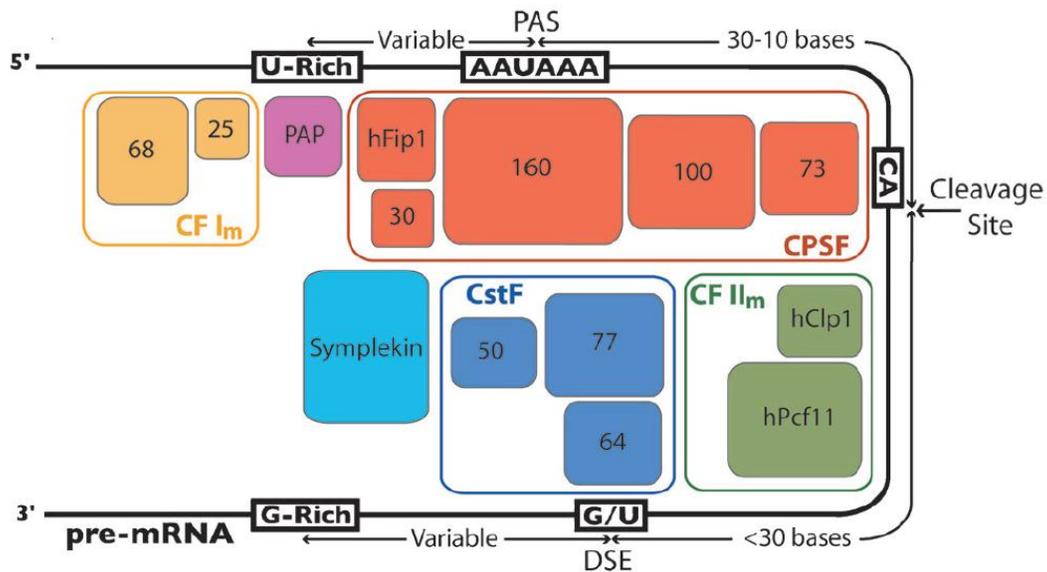
El sitio de corte del pre-mRNA se ubica entre la señal de poliadenilación y la DSE, dentro de una región de 13 nucleótidos (Chen F. y col. 1995). La secuencia de nucleótidos que rodea el sitio de corte no está estrictamente conservada. En vertebrados 70% de los sitios de corte en pre-mRNAs, se encuentra en el lado 3' de un residuo de adenosina, con una preferencia de nucleótidos de A>U>C>>G frecuentemente después del dinucleótido CA (Fig. 7) (Chen F. y col. 1995; Mandel C. y col. 2008). El nucleótido que precede al sitio de corte es citosina en el 59% de los casos de 269 secuencias de pre-mRNA analizadas (Takagaki Y. Manley J. 1997).

Existen dos elementos de secuencia secundarios o auxiliares que mejoran el procesamiento del extremo 3': el elemento auxiliar río arriba y el DSE auxiliar (Zhao J. 1999). El elemento auxiliar río arriba a la señal de poliadenilación, no tiene una secuencia consenso pero a menudo se compone de una secuencia rica en U (UUUU) o secuencias similares (UGUA, UAU) (Hu J. y col. 2005). La eficiencia en el corte y la poliadenilación es aumentada por la presencia de este elemento, pues permite la unión de factores de poliadenilación hacia el sitio de corte (Huang Y, y col. 1999; Le Hir H. y col. 2003). Son pocos los elementos de secuencia auxiliares que se conocen río abajo del sitio de corte; son generalmente ricos en G, pero carecen de una secuencia conservada y una distancia determinada hacia el sitio de corte. Además, puede haber más de un auxiliar presente en un gen (Mandel C. y col. 2008).

#### **2.3.4 FACTORES PROTEICOS INVOLUCRADOS EN EL PROCESO DE POLIADENILACIÓN EN MAMÍFEROS.**

Los estudios con extractos nucleares de células de mamíferos, así como los estudios por otros métodos bioquímicos y genéticos, sugieren que un gran número de factores proteicos son necesarios para el procesamiento del extremo 3' (Chou Z. y col. 1994). En la actualidad, más de 14 proteínas han sido identificadas como la maquinaria de procesamiento del extremo 3' en mamíferos (Mandel C. y col. 2008). El procesamiento del extremo 3' en mamíferos necesita de algunos sub-complejos (Fig. 7) que incluyen: El Factor específico de corte y poliadenilación (CPSF, *Cleavage and Polyadenylation Specificity Factor*), factor de estimulación de corte (CstF, *Cleavage Stimulation Factor*), factor de corte I (CFI, *Cleavage Factor I*), factor de corte II (CFII, *Cleavage Factor II*). Además son requeridas algunas otras proteínas como la Poli (A) polimerasa (PAP),

Proteína de unión a la Poli (A) polimerasa (PABP, *Poly (A) Binding Protein*), Simplekina y el dominio carboxilo terminal (CTD) de la polimerasa II (Pol II) (Mandel C. y col. 2008). Todas las proteínas excepto la PABP son requeridas para la reacción de corte *in vitro*, mientras que solamente CPSF, PAP, y PABP son requeridas para la reacción de poliadenilación *in vitro* (Fig. 7) (Zhao J. y col. 1999).



**Fig. 7. Esquema que representa la maquinaria de procesamiento del extremo 3' en mamíferos.** La subunidad CPSF-160 reconoce la señal de poliadenilación, mientras que la subunidad CstF-64 reconoce el elemento DSE (Tomado de Mandel C. y col. 2008).

El CPSF en mamíferos, está conformado por 5 subunidades: CPSF-160, CPSF-30, CPSF-73, CPSF-100, y hFip1. Todas las subunidades son necesarias para la división eficiente y la poliadenilación del pre-mRNA (Mandel C. y col. 2008). La subunidad CPSF-160 es la más grande del complejo CPSF e interactúa directamente con el pre-mRNA, reconoce el motivo AAUAAA para dirigir la reacción de corte y también es requerida para la poliadenilación. La unión de CPSF-160 con la señal de poliadenilación es de alta afinidad por la secuencia AAUAAA y de menor afinidad por otras secuencias de la señal de poliadenilación (Dantonel J. y col. 1997; Zhao y col. 1999). Además de su papel en el procesamiento del extremo 3', CPSF-160 se asocia con factores involucrados en el inicio de la transcripción (TFIID) y con el dominio CTD de la subunidad mayor de Pol II (McCracken S. y col. 1997).

La subunidad CPSF-30 es necesaria para el corte y la poliadenilación (Barabino S. M. L. y col. 2000). CPSF-30 posee 5 motivos de dedos de zinc (CCCH) en eucariontes, con la secuencia conservada CX<sub>8</sub>CX<sub>5</sub>CX<sub>3</sub>H; en metazoarios existe un dedo de zinc más (CCHC) con la secuencia conservada CX<sub>2</sub>CX<sub>4</sub>HX<sub>4</sub>C; ambos motivos han demostrado que se unen a RNA (Barabino S. M. L. y col. 1997; Colgan y col. 1997). La subunidad CPSF-73 es una proteína miembro de la familia de las metalo-β-lactamasas, en su mayoría son enzimas hidrolíticas que utilizan los iones de zinc durante la catálisis, CPSF-73 posee un dominio metalo β-lactamasa en su dominio amino terminal (Aravind L. 1999; Callebaut I. 2002). Un número cada vez mayor de estudios genéticos y bioquímicos indican que el corte endonucleolítico que antecede a la poliadenilación se lleva a cabo por CPSF-73 (Mandel C. y col. 2006; Calvo R. Manley J. 2004). La subunidad CPSF-100 interactúa con CPSF-30, sin embargo no se conoce su función con claridad actualmente (Mandel C. y col. 2008; Rao S. y col. 2009). La subunidad Fip1p (Factor de Interacción con Pap1p) fue identificado por primera vez en levadura, se identificó posteriormente por análisis de secuencias de cDNA en células HeLa y se confirmó que un miembro del complejo CPSF (Mandel C. y col. 2008).

El Factor de estimulación de Corte (CstF) está conformado por tres subunidades: CstF-64, CstF-77 y CstF-50. CstF es requerido para procesar de manera eficiente los pre-mRNAs y para que se lleve a cabo la poliadenilación, reconoce regiones ricas en U y G/U en la secuencia de RNA y ayuda a estabilizar el factor específico de corte y poliadenilación (CPSF) (MacDonald CC. y col. 1994; Moreno-Morcillo M. 2011). La subunidad CstF-64 posee un motivo de unión a RNA (RRM, RNA Recognition Motif), este dominio es suficiente por si solo para reconocer las regiones ricas en G/U (DSE) (Dai W. y col. 2011). La subunidad CstF-77 mantiene una interacción con CstF-64 para que ocurra el correcto corte por parte de CPSF. La interacción de CstF-64 y CstF-77 regula los niveles estequiométricos nucleares de los componentes del complejo CstF, regula también su localización intracelular, estabilidad y función (Juge F. y col. 2000; Ruepp M. y col. 2011). La subunidad CstF-50 es requerida para el corte del pre-mRNA *in vitro* (Takagaki Y. Manley J. L. 1994). CstF-50 y CstF-77 se unen al dominio CTD de la subunidad mayor de Pol II durante la transcripción (McCracken S. y col. 1997). Los factores de corte (CFI y CFII) son requeridos para que ocurra la reacción de corte del pre-mRNA *in vitro*, CFI parece ser específico de eucariontes superiores. La función principal del CFI es proporcionar reconocimiento adicional del pre-mRNA y ayudar con el reconocimiento del sitio adecuado para la poliadenilación (Ruegsegger U. y col. 1996; Brown P. y col., 1991).

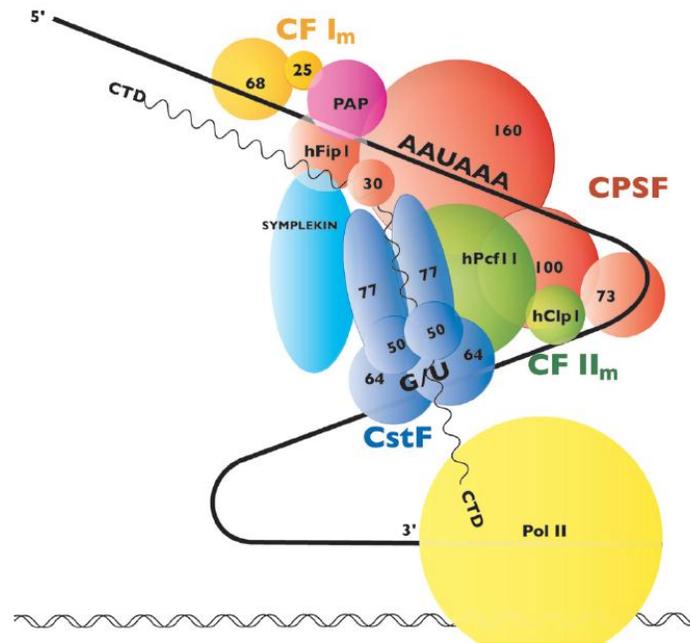
La enzima poli (A) polimerasa (PAP) es requerida en los mamíferos para el corte y la poliadenilación. La PAP interactúa con otras proteínas del complejo de procesamiento del extremo 3', importantes para la poliadenilación y la definición de la longitud adecuada de la cola de poli (A) (Kuhn U. y col. 2009; Mandel C. y col. 2008). De la misma forma la proteína de unión a la poli (A) polimerasa (PABP) es necesaria para la correcta y eficiente poliadenilación; la reacción de poliadenilación puede ocurrir en ausencia de la PABP, sin embargo esta proteína es necesaria para determinar el tamaño de la cola de poli (A) (Sachs A. B. 1989; Kuhn U. y col. 2009). Además, la unión de la PABP al pre-mRNA y la PAP aumenta la eficiencia de la reacción de poliadenilación hasta en 80 veces (Kerwitz Y. y col. 2003). La Simplekina ha sido identificada como una proteína capaz de formar un complejo estable entre CPSF y CstF, a través de su interacción con ambas. También puede funcionar como una proteína de andamiaje en el procesamiento del extremo 3' del pre-mRNA de histonas (Mandel C. y col. 2008).

#### **2.3.4 MODELO DE PROCESAMIENTO Y POLIADENILACIÓN EN MAMÍFEROS.**

Para que se pueda iniciar el procesamiento del extremo 3' del pre-mRNA es necesario que se lleve a cabo un reconocimiento de señales; en primer lugar, CPSF debe reconocer la señal de poliadenilación, la subunidad CPSF-160 es la encargada de reconocer el hexanucleótido AAUAAA, en una reacción asistida por el factor de corte I (CFI); es probable que en este momento la subunidad CPSF-160 esté unida a proteínas de transcripción como TFIID y el CTD de Pol II. Por otro lado la subunidad Cstf-64 debe reconocer las secuencias ricas en G/U (DSE) del pre-mRNA, de esta forma el complejo CstF ayuda a estabilizar el factor específico de corte y poliadenilación (CPSF) (Fig. 8) (Mandel C. y col. 2008; Barrera G. 2006).

Las interacciones individuales de CPSF y CstF son débiles pero se estabilizan por la interacción de Simplekina, este correcto ensamblaje puede favorecer la actividad de endonucleasa de la subunidad CPSF-73 y entonces permitir que se posicione sobre el sitio de corte y realice el corte catalítico en el pre-mRNA. Es entonces que la PAP puede catalizar la formación de la cola de poli (A), en una reacción que en conjunto con la PABP determina el tamaño de la cola de poli (A) y mejora la eficiencia de poliadenilación. Además de la unión y el procesamiento de los pre-mRNA, algunos de los factores protéicos en este mecanismo también son necesarios para coordinar la

iniciación y terminación de la transcripción, la adición del 7-metil guanosina en el extremo 5' y el *splicing* (Fig. 8) (Mandel C. y col. 2008; Barrera G. 2006).



**Fig. 8. Modelo esquemático de la maquinaria requerida para el procesamiento del extremo 3' de pre-mRNA de mamíferos (Tomado de Mandel C. y col. 2008).**

## 2.4 PROCESO DE POLIADENILACIÓN EN LEVADURAS.

### 2.4.1 ELEMENTOS EN LA SECUENCIA DEL PRE-mRNA EN LEVADURAS.

Las señales de procesamiento del extremo 3' de pre-mRNA en la levadura *Saccharomyces cerevisiae* son algo diferentes de las señales utilizadas en eucariontes superiores, tanto en secuencia como en organización. Las señales de poliadenilación en levaduras son menos conservadas que las de mamíferos (Zhao J. 1999). El sitio de poliadenilación en levaduras es definido por cuatro elementos de secuencia: El elemento de eficiencia rico en A/U (EE, Efficiency Element), el elemento de posicionamiento rico en A (PE, Positioning Element), el sitio de corte y los elementos ricos en U alrededor del sitio de corte (UUE o DUE) (Mandel C. y col. 2008).



## 2.4.2 FACTORES PROTÉICOS INVOLUCRADOS EN EL PROCESO DE POLIADENILACIÓN EN LEVADURAS.

El procesamiento del extremo 3' en levaduras incluye también los procesos de corte y poliadenilación del pre-mRNA. Algunos de los sub-complejos de la maquinaria de poliadenilación en levaduras son:

El factor de corte y poliadenilación (CPF, *Cleavage and Polyadenylation Factor*), encargado de regular el corte y la poliadenilación, además de reconocer las secuencias ricas en U (DUE y UEE) que flanquean al sitio de corte. El CPF está formado por el factor de corte II (CF II, *Cleavage Factor II*) y el factor de poliadenilación I (PF I, *Polyadenylation Factor I*) (Mandel C. y col. 2008). El CF II y el PF I están formados por varias subunidades. La subunidad Cft1p/Yhh1p que conserva un 24% de identidad y un 51% de similitud con la subunidad CPSF-160 de mamíferos Cft1p/Yhh1p se une directamente al RNA, pero no se une a la señal de PE (el equivalente de la señal de poliadenilación para mamíferos), en cambio, la unión se observó principalmente en la región rica en A cerca del sitio de corte (Fig. 9) (Dantonel J. y col. 1997; Dichtl B. y col. 2002).

La reacción de corte del extremo 3' está mediada por diferentes subunidades. La subunidad Yth1p, al ser mutada en sus motivos de dedos de zinc reduce la actividad de corte (Barabino S. M. L. y col. 2000). La subunidad Brr5p/Ysh1p y el CPSF-73 de mamíferos poseen una secuencia altamente conservada, debido a los dominios metalo  $\beta$ -lactamasa presentes en cada una de las subunidades, estos dominios pertenecen a proteínas de unión a RNA y endonucleasas (Dominski Z. y Marzluff W. F. 2007; Ma Y. y col. 2002). Esta evidencia estructural y bioquímica ha sido encontrada recientemente y se propone que Brr5p/Ysh1p es la endonucleasa de procesamiento del extremo 3' en levaduras (Ryan K. y col. 2004).

El factor de corte IA (CF IA, *Cleavage Factor IA*) y el factor de corte IB (CF IB, *Cleavage Factor IB*), reconocen las secuencias ricas en A (PE) y las secuencias ricas en U (EE) respectivamente. La subunidad Rna15p del complejo CF IA, posee motivos de reconocimiento de RNA altamente específicos, y esta especificidad se alcanza cuando Rna15p interactúa con otras proteínas como Rna14p y Hrp1p. Mutaciones en Rna15p eliminan la interacción con RNA (Gross S. Moore C. L. 2001; Hammell C. M. y col. 2002; Kessler M. M. y col. 1997).

Hrp1p es el único miembro que se conoce del factor de procesamiento CF IB, es requerido para la correcta reacción de corte en levaduras y no presenta un homólogo en mamíferos. Hrp1p posee dos motivos de unión a RNA y presenta una alta especificidad por la región rica en A/U (EE) (Perez-Canadillas J.-M. 2006).

La maquinaria de procesamiento del extremo 3' en levaduras, contiene también algunos factores proteicos adicionales en CPF. Aunque algunos de ellos contienen homólogos en mamíferos, la función de estos no ha sido establecida con claridad. Incluso la reconstitución de la reacción de corte *in vitro* requiere únicamente de: CF IA, CF IB y CF II, mientras que la poliadenilación *in vitro* requiere de: CPSF, CF IA, CF IB, y Pab1p. Los factores adicionales desempeñan generalmente un papel de moduladores. A pesar de la identificación de esta gran colección de factores proteicos, la identidad de la nucleasa que cataliza la reacción de escisión no se conoce (Mandel C. y col. 2008).

## **2.5 SEÑALES DE POLIADENILACIÓN EN *T. vaginalis*.**

En el laboratorio de la Dra. Imelda López Villaseñor se logró postular tras el análisis de secuencias de cDNAs, el tetranucleótido UAAA, como la señal de poliadenilación (Espinosa N. y col. 2002), actualmente se ha demostrado mediante el uso de un sistema transitorio de un gen reportero (cat, flanqueado por secuencias no codificantes de actina de *T. vaginalis*) que el tetranucleótido UAAA (que puede incluir el codón de paro UAA) es la señal de poliadenilación en *T. vaginalis* (Fuentes V. y col. 2012).

Al realizar mutaciones en la secuencia del tetranucleótido UAAA se producen cambios en la poliadenilación de los mRNAs reporteros, cuando los cambios ocurren en las bases a cada lado del tetranucleótido no se produce ningún cambio en la poliadenilación, incluso después de restablecer la secuencia del tetranucleótido se logra restablecer el procesamiento y poliadenilación de los mRNAs reporteros, además se logró identificar una señal de corte en los pre-mRNAs,  $Y^{\downarrow}(A)_{2-5}UU$ . Se demostró también la interacción de la señal de poliadenilación con proteínas, proporcionando evidencia que sugiere que el tetranucleótido UAAA cumple funciones similares en *T. vaginalis*, que las funciones que cumple el hexanucleótido AAUAAA en metazoarios (Fuentes V. y col. 2012).

### **3. HIPÓTESIS.**

---

CPSF-160 es la proteína que reconoce la señal de poliadenilación en metazoarios. Dado que los mRNAs de *T. vaginalis* están poliadenilados, entonces *T. vaginalis* posee una proteína de tipo CPSF-160, homóloga a la reportada para metazoarios capaz de reconocer la señal atípica de poliadenilación UAAA en este organismo.

### **4. JUSTIFICACIÓN.**

---

No se conocen los factores proteicos involucrados en el procesamiento del extremo 3' de los mRNAs en *T. vaginalis*. La identificación de estos factores es relevante dadas las señales cortas y atípicas identificadas para la poliadenilación de los mRNAs en este organismo.

### **5. OBJETIVOS.**

---

#### **5.1 OBJETIVO GENERAL.**

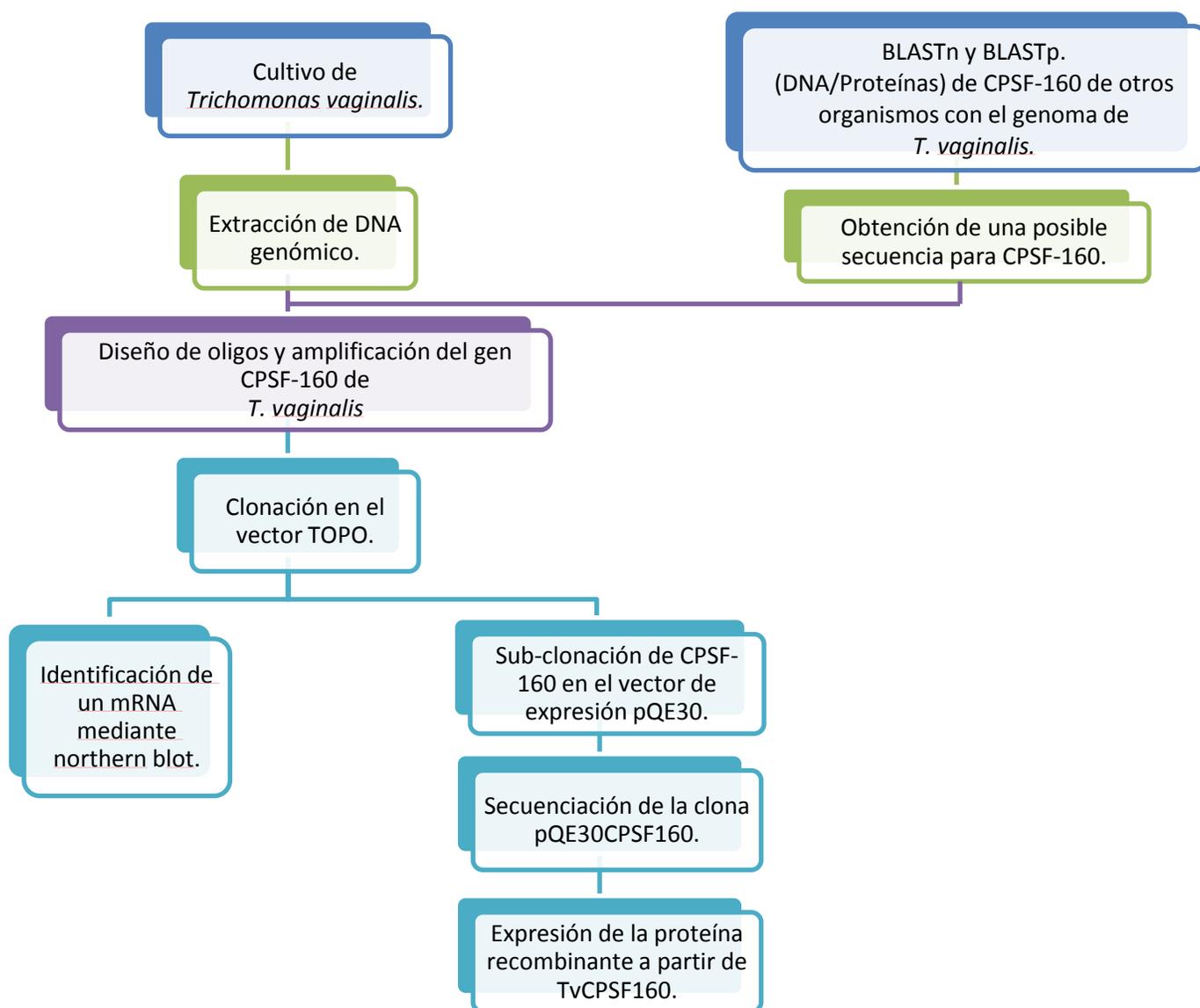
Identificar, clonar y expresar el gen que codifica para la proteína de *T. vaginalis* homóloga a CPSF-160.

#### **5.2 OBJETIVOS PARTICULARES.**

- Identificar la probable secuencia de DNA que codifica para el CPSF-160 de *T vaginalis*, a partir del análisis bioinformático del genoma accesado de este organismo.
- Amplificar la secuencia de CPSF-160 mediante PCR, a partir de DNA genómico.
- Clonar el gen de CPSF-160 en el vector TOPO.
- Corroborar la expresión del mRNA mediante northern blot.
- Sub-clonar el gen de CPSF-160 en el vector pQE30.
- Expresar la proteína recombinante de CPSF-160 de *T. vaginalis*.

## 6. METODOLOGÍA.

### 6.1 DISEÑO EXPERIMENTAL.



**Fig. 10.** Esquema que sintetiza las metodologías utilizadas durante el proceso experimental.

## 6.2 ALINEAMIENTOS DE SECUENCIAS.

Se realizó un ejercicio bioinformático de alineamiento de secuencias de DNA y de proteínas, comparando las secuencias de los genes CPSF-160 reportados para otros organismos. Se hizo una búsqueda en el genoma de *T. vaginalis*, con la finalidad de encontrar un gen candidato para ser el gen de CPSF-160 en *T. vaginalis*. El proceso se realizó de la siguiente manera:

1. Se hizo la búsqueda de cDNAs de CPSF-160 reportados para los siguientes organismos: *Arabidopsis thaliana* (NM\_124545), *Drosophila melanogaster* (NM\_206111), *Homo sapiens* (BC017232), *Mus musculus* (AF322193), *Trypanosoma cruzi* y *Saccharomyces cerevisiae* (NM\_001180609).

2. Los genes reportados se compararon con el genoma de *T. vaginalis* (BLASTn). Se alineó la secuencia codificante de los genes reportados con la base de datos de los transcritos del genoma de *T. vaginalis*, para identificar un gen que pudiera ser el CPSF-160 de *T. vaginalis*.

4. Las secuencias proteicas reportadas para CPSF-160 se compararon con las proteínas de *T. vaginalis* (BLASTp) para identificar una proteína que pudiera ser el CPSF-160 de *T. vaginalis*.

5. Se identificaron dos secuencias consenso una de cDNA y otra proteica, ambas del mismo gen.

Para realizar los alineamientos se utilizó el software disponible en línea (<http://TrichDB.org>) (Aurrecoechea C. y col. 2009), que proporciona la Herramienta de Búsqueda de Alineamiento Local Básico (BLAST, por sus siglas en ingles), dicho software utiliza la versión del genoma liberado de *T. vaginalis* (Carlton J. M. y col. 2009). Para los alineamientos de nucleótidos se utilizó la herramienta tblastx que usa la secuencia de nucleótidos y analiza los 6 diferentes marcos de lectura y lo compara con la base de datos de transcritos. Para los alineamientos de proteínas se usó la herramienta blastp, que permite el uso de gaps y compara una secuencia de aminoácidos con otra.

## 6.3 CULTIVO DE *T. vaginalis*.

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 A) 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 cada dos días.

#### 6.4 EXTRACCIÓN DE DNA.

Para la extracción del DNA genómico, se utilizó un cultivo en caja de 50ml de *T. vaginalis* en fase logarítmica ( $1.6 - 2 \times 10^6$  cel/ml). El volumen total del cultivo se dividió en dos tubos de 50ml y se procesaron independientemente de la siguiente manera:

1. Cada tubo Nalgene se centrifugó a 3,000rpm durante 10min a 4°C en un rotor Sorvall SS34. Se decantó el sobrenadante cuidadosamente que fue esterilizado posteriormente. El botón de células se lavó con 20ml de PBS pH 7.4 (Apéndice A), se resuspendió nuevamente en PBS y se centrifugó de nuevo, de la misma forma que en el punto número 1.
2. Se decantó la solución de lavado y el botón de células se resuspendió en 2ml de PBS. En dos tubos eppendorf de 1.5ml se colocó 1ml de células resuspendidas del inciso anterior, y se centrifugó a 14,000rpm durante 1min en una microcentrifuga. Se decantó el sobrenadante y se resuspendió el botón con el líquido que sobró (20µl).
3. Se agregaron 300µl de buffer de extracción (0.2% SDS, 100mM NaCl, 10mM EDTA y 10mM Tris-HCl, pH 8), y 300µl de fenol. Se agitó en vortex durante 30seg a una velocidad de 3, se centrifugó a velocidad máxima durante 3min. A partir de este paso fue muy importante que la extracción se realizara de forma rápida, ya que el parásito tiene muchas nucleasas que degradan rápidamente el DNA.
4. Se colocó la fase superior cuidadosamente (sin tomar la interfase) en otro tubo eppendorf con 300µl de fenol y 300µl de cloroformo (SIGMA), se agitó y se centrifugó igual que en el paso anterior. Se transfirió la fase superior cuidadosamente a otro tubo eppendorf con 500µl de cloroformo, se agitó y centrifugó de la misma forma. Se colocó la fase superior cuidadosamente en otro tubo eppendorf limpio y se precipitó con 1ml de etanol absoluto y 40µl de acetato de sodio 3M. Se dejó a -20°C durante 2Hrs (Se puede dejar a -70°C por una hora).
5. Se centrifugó el tubo eppendorf con el DNA precipitado en una microcentrifuga a velocidad máxima durante 15min. Se eliminó el sobrenadante y el botón se lavó con 1ml de etanol al 70% frío, se decantó y centrifugó durante 15min a velocidad máxima. Se quitó el exceso de etanol con una micropipeta y se dejó secar el botón de DNA durante 15min, hasta que el botón quedó transparente. El botón de DNA se resuspendió en 50µl de TE (Apéndice A).
6. Se agregaron 0.5µl de RNAsa (GIBCO) (20µg/ml) y se dejó a 37°C durante 30min. El DNA se guardó a -20°C.

## 6.5 CONDICIONES DEL PCR.

Para amplificar el gen CPSF-160 de *T. vaginalis* se realizó un PCR bajo las siguientes condiciones:

Paso	Tiempo (minutos)	Temperatura (°C)
1	10'	92°
2	1'	92°
3	2'	43°
4	10'	72°
5	30 Ciclos desde el paso 2	
6	10'	72°
7	24Hrs	4°
8	END	

La reacción de PCR se llevó a cabo en un volumen total de 50µl para la muestra con DNA genómico y 25µl para el control de amplificación de la reacción, los reactivos se adicionaron en el siguiente orden y en las cantidades indicadas.

Solución	Muestra	Control
H <sub>2</sub> O	26µl	16.25µl
Amortiguador para PCR 10x	5µl	2.5µl
MgCl <sub>2</sub> 50mM (4.5mM final)	4.5µl	1.5µl
d'NTPS 100mM (10mM c/u)	3µl	1.5µl
Oligo 5' CPSF-BH1 (200ng/µl)	3µl	1.5µl
Oligo 3' CPSF-K1 (200ng/µl)	3µl	1.5µl
DNA genómico	5µl	-
Total	49.5µl	24.5µl

En el minuto 9 del primer ciclo se adicionaron 0.5µl de la enzima Taq Polimerasa a la muestra con DNA genómico y 0.25µl a la muestra sin DNA genómico. A cada una de las muestras se le adicionó 100µl de Aceite Mineral con la finalidad de que las muestras no se evaporaran.

#### **6.6 PURIFICACIÓN DE FRAGMENTOS DE DNA.**

El producto de PCR se corrió en un gel preparativo de agarosa al 0.8% en TAE 1X. La banda de interés se cortó y se purificó de la siguiente forma:

1. La banda de interés se cortó con un bisturí desinfectado y se colocó en una minicolumna. Para prepara la minicolumna se perforó la punta de un tubo eppendorf de 0.5ml con una aguja caliente, se colocó al fondo del tubo fibra de vidrio siliconizada y estéril (un fragmento de 2mm). Este tubo se colocó dentro de un tubo eppendorf de 2ml.
- 3.- El fragmento del gel que contenía el DNA se colocó en la minicolumna y se centrifugó a 6,00rpm durante 5min. El DNA se colectó y se centrifugó de la misma forma durante 5 ocasiones más.
- 4.- El DNA extraído se limpió dos veces con un volumen de fenol cloroformo 1:1 y dos veces con un volumen de cloroformo. Entre cada extracción se vortexeó a una velocidad de 5 durante 30segundos y después se centrifugó a 12,000rpm durante 3min.
- 5.- Se precipitó con 2.5 volúmenes de etanol al 100% y 1/10 de volumen de 3M acetato de sodio pH6 durante dos horas a -20°C.
- 6.- Se centrifugó a 14,000rpm durante 20min, el pellet se lavó con 1ml de 70% etanol frío, se dejó secar a temperatura ambiente y se resuspendió en 25µl de TE.

#### **6.7 CLONACIÓN DEL PRODUCTO DE PCR EN EL VECTOR PCR-BLUNT II TOPO Y ELECTROPORACIÓN.**

Para clonar el fragmento de DNA purificado se utilizó el sistema pCR-Blunt II TOPO (Invitrogen):

1. Se preparó una mezcla de reacción para ligar el fragmento al vector pCR-Blunt II TOPO de la siguiente manera:

	Cantidad
<b>DNA</b>	4µl
<b>Solución salina (Invitrogen)</b>	1µl
<b>Vector pCR-Blunt II TOPO (Invitrogen).</b>	1µl

2. La reacción de ligación se incubó 20min a temperatura ambiente.
3. Se tomaron 4µl de la reacción de ligación y se añadieron a un vial de células electrocompetentes *E.coli* TOP10 (Invitrogen) descongeladas en hielo; se mezcló de forma suave.
4. Se electroporaron las células a 2.46kV en una cubeta de electroporación (GIBCO) a 4°C. Inmediatamente después de electroporar se adicionaron 250µl de medio SOC (Invitrogen).
5. Las células (300µl) se colocaron en un tubo falcon estéril. El tubo se colocó en agitación a 37°C durante 1Hr.
6. Se plateó 1/10 y 9/10 del volumen total de células en cajas de LB + Kanamicina (50µg/ml).
7. Las clonas obtenidas se cultivaron en tubos de vidrio con 5ml de LB + Kanamicina (50µg/ml).

## 6.8 OBTENCIÓN DEL DNA PLASMÍDICO.

Para obtener el DNA plasmídico de las clonas de *Escherichia coli*, se realizó lo siguiente:

1. Se cultivaron las clonas toda la noche en 5ml de LB + Kanamicina (50µg/ml) a 37°C. Una parte del cultivo (600µl) se mezcló junto con glicerol (600µl) y se almacenaron a -70°C.
2. Se colocaron en hielo las soluciones I y III (Apéndice A) el fenol y el cloroformo; se preparó la solución II (Apéndice A) que se usó fresca y a temperatura ambiente.
3. Se colocaron 1.5ml de cultivo en un tubo eppendorf de 1.5ml. Se centrifugó 30seg a 12,000rpm, y se desechó el sobrenadante. Se Colocó 1.5ml del cultivo nuevamente y se centrifugó del mismo modo, se desechó el sobrenadante.
4. Se resuspendió el pellet con el sobrenadante que quedó en el tubo, se mezcló en el vortex a una velocidad de 4. Se agregaron 100µl de la solución I fría y se mezcló en el vortex a una velocidad de 4.
5. Se agregaron 200µl de la solución II invirtiendo el tubo de 10 veces.

6. Se colocaron los tubos en hielo, ahí se agregaron 150µl de la solución III fría, se mezcló invirtiendo el tubo de 10 veces, se incubó en hielo x 5min.
7. Se centrifugó a 12,000rpm durante 5min, recuperando el sobrenadante en otro tubo eppendorf y desechando el pellet.
8. Al tubo con el sobrenadante se agregaron 250µl de fenol frío y 250µl cloroformo frío, se mezcló en el vortex a velocidad de 5 durante 30seg, se centrifugó a 12,000rpm durante 3min.
9. Se recuperó la fase de arriba y se agregaron 500µl de cloroformo frío, se mezcló en el vortex a velocidad de 5 durante 30seg, y se centrifugó a 12,000rpm durante 3min.
10. Se recuperó la fase superior y se agregaron 1000µl de etanol al 100% a temperatura ambiente, se mezcló en el vortex a velocidad máxima durante 30seg. Se Incubó a temperatura ambiente durante 2Hrs.
11. Se centrifugó 10min a velocidad máxima, se desechó el sobrenadante y se lavó con etanol 70% frío rápidamente, se dejó secar 10min, se centrifugó 10seg a máxima velocidad y se retiró el etanol con una pipeta.
12. Se resuspendió el pellet en 50µl de RNAsa (20µg/µl), se mezcló en el vortex a velocidad máxima y luego se centrifugó durante 30seg a velocidad máxima.
13. Se incubaron los tubos a 37°C durante 20min.
14. Los plásmidos se corrieron en un gel preparativo de agarosa 0.8% en TAE 1X para comprobar el tamaño de los plásmidos (Se comprobaron las clonas mediante digestiones).
15. La clona identificada se envió para ser secuenciada a la unidad de secuenciación del Instituto de Investigaciones Biomédicas.

## **6.9 EXTRACCIÓN DE RNA TOTAL.**

Para la extracción de RNA total se siguió la siguiente metodología:

1. Se realizó un cultivo de 15ml de la cepa TOPOCPSF160, el cultivo se colocó en un tubo cónico de 15ml y se centrifugó a 3,000 rpm durante 10 min a 4°C.
2. Se desechó el sobrenadante. El pellet se resuspendió en 1ml de PBS/H<sub>2</sub>O DEPC estéril y se transfirió a un tubo eppendorf de 1.5ml, se centrifugó a 14,000rpm durante 1min a 4°C, se desechó la solución de lavado.

3. El pellet se resuspendió con 1ml de trizol (INVITROGEN), y por pipeteo repetitivo se lisaron las células.
4. Las muestras se incubaron durante 5min a temperatura ambiente para permitir una disociación completa de los complejos nucleoproteicos.
5. Se agregaron 200µl de cloroformo y se mezcló durante 15 segundos por medio de inversión, se incubó 3 minutos a temperatura ambiente.
6. Se centrifugaron las muestras 15 minutos a 14,000rpm a 4°C.
7. Se transfirió la fase acuosa a un tubo EPPENDORF de 1.5ml y se precipito el RNA con 0.5ml de alcohol isoprophilico, se incubó la muestra 10min a temperatura ambiente.
8. Se centrifugó la muestra precipitada a 14,000rpm durante 10min a 4°C. El sobrenadante se desechó y se lavó el pellet de RNA con 1ml de Etanol 75% (H<sub>2</sub>O DEPC).
9. Se mezcló ligeramente la muestra en el vortex y se centrifugó a 10,000rpm durante 5min a 4°C.
10. Se dejó secar el botón de RNA. Se disolvió el RNA en 10µl AGUA/DEPC.
11. El RNA se corroboró en un gel de agarosa 0.8% TBE.
11. Se almacenó a -70°C.

#### **6.10 PURIFICACIÓN DE RNA POLI(A)+ A PARTIR DE RNA TOTAL.**

Para purificar la fracción poli(A)+ se utilizó el kit Quickprep micro RNA purification. (Amersham Pharmacia Biotech), que cuenta con una resina oligo (dT) celulosa. Se realizaron los siguientes pasos:

1. 20 o 30min antes de realizar la extracción se dejó el kit a temperatura ambiente, el buffer de extracción se colocó a 37°C hasta que los cristales se disolvieron.
2. La oligo (dT) celulosa se mezcló por inversión e inmediatamente se tomó 1ml y se colocó en un tubo eppendorf de 1.5 ml.
3. En otro tubo eppendorf se colocaron 30µl de RNA total, se le agregan 400µl de buffer de extracción, también se agregan 800µl del buffer de elusión y se mezcló suavemente.
4. En un tubo eppendorf nuevo se colocó 1ml de buffer de elusión y se mantuvo a 65°C hasta que se utilizó para eluir el RNA poli A+.
5. El tubo que contiene la oligo (dt) celulosa se centrifugó por un minuto a velocidad máxima, y se eliminó el sobrenadante.

6. Al tubo anterior se le agregó la mezcla realizada en el paso 3. Se mezcló suavemente por inversión 3 minutos y se centrifugó a máxima velocidad durante 10 segundos.
7. El sobrenadante se transfirió a un tubo nuevo ya que en este se encuentra la población RNA poli (A)-.
8. La resina se lavó cinco veces de la siguiente forma: se resuspendió en 1ml de amortiguador High-salt por inversión, se centrifugó a 14,000rpm por 10 segundos y se eliminó el sobrenadante. Posteriormente la resina se resuspendió en 1ml 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.
9. La resina se resuspendió en 300µl de amortiguador Low-salt, y se transfirió a una microcolumna que se colocó dentro de un microtubo de 1.5ml y se centrifugó a 14,000rpm durante 5 segundos.
10. Se cambió la columna a un microtubo nuevo y se agregaron 500µl de amortiguador Low-salt y se centrifugó 5 segundos a 14,000rpm, la columna se cambió a un tubo nuevo y se repitió el paso anterior.
11. El RNA poli (A)+ se eluyó de la resina de la siguiente manera: La columna se colocó en un microtubo de 1.5ml y se le agregaron 200µl del amortiguador de elusión a 65°C, la columna se centrifugó a 14,000rpm durante 5 segundos, y la elusión se guardó ya que contiene al RNA poli (A)+.
12. La columna se lavó con 200µl de amortiguador de elusión a 65°C y se repitió el paso anterior.
13. El RNA poli (A)+ se cuantificó por espectrofotometría de la siguiente manera: el RNA se diluyó 1:100 en amortiguador de elusión, en un volumen total de 1ml. En una celda de cuarzo pre tratada con H<sub>2</sub>O DEPC se leyó la absorbancia a A=260nm y A=280nm.

### **6.11 NORTHERN BLOT.**

Para realizar el northern blot se utilizaron guantes durante todo el experimento, y se siguieron estos pasos:

1. Se prepararon 4 litros de H<sub>2</sub>O DEPC 4 litros (Apéndice A). Se dividió en 5 botellas de 1 Litro con ~800ml cada una.
2. Se esterilizó el siguiente material: Tubos eppendorf de 0.5ml solo para RNA. Puntas de 200µl y 1000µl nuevas y estériles, solo para RNA.

3. El material de vidrio que se utilizó se lavó muy bien, se esterilizó por autoclave y se horneó a 80°C toda la noche, cubriéndolo con papel aluminio. (1 probeta de 100ml, 1 probeta de 50ml, 1 vaso de pp. de 100ml, 1 pipetero con pipetas de 10ml, 2 espátulas, 1 frasco de 1l y 1 molde pyrex.
4. El gel se corrió en una cámara de electroforesis de 11x14 cm. La agarosa se preparó de la siguiente forma: 1g de agarosa, 16.2ml de formaldehído, 73ml de H<sub>2</sub>O DEPC y 10ml de buffer MOPS/EDTA 10X (Apéndice A) a temperatura ambiente. El gel se dejó solidificar 3Hrs dentro de la campana de extracción debido al formaldehído.
5. El buffer de electroforesis MOPS/EDTA 1X se preparó a partir del buffer MOPS/EDTA 10X. (80ml de MOPS/EDTA 10X + 720ml de H<sub>2</sub>O DEPC).
6. La preparación de muestras se hizo 30min antes de cargar el gel de la siguiente manera (por cada tubo):
  - 0.78μl de MOPS/EDTA 10X.
  - 0.78μl de H<sub>2</sub>O DEPC.
  - 2.72μl de formaldehído.
  - 6.8μl de formamida (desionizada)
  - 3μl de RNA
  - 1.5μl de H<sub>2</sub>O DEPC
7. Se calentó la mezcla y también 4μl de cada uno de los marcadores de peso molecular: bajo (0.16-1.77 Kb) y alto (0.24-9.5 Kb) a 65°C por 15 min. Se pusieron en hielo por 2 min.
8. Se agregó a la muestra 4.5μl de buffer de carga para RNA 6X y 4μl a los marcadores de peso molecular. Se cargaron en el gel las muestras que fueron 8μl de cada uno de los pesos moleculares y 20μl de cada muestra. Se hace por duplicado ya que la mitad del gel se tiñe y la otra mitad será para hacer la transferencia.
9. El gel se corre durante 3 horas a 80 volts con buffer MOPS/EDTA 1X.
10. Una vez que terminó de correr el gel, se colocó en el refractario pirex estéril para ahí poder cortarlo. La mitad del gel se tiñó con bromuro de etidio durante 5min y se puso a desteñir con H<sub>2</sub>O DEPC toda la noche a 4°C.
11. Al día siguiente en la membrana se marcaron los pozos con lápiz. Se lavó con H<sub>2</sub>O DEPC (~10 min.), se dejó secar por una hora sobre papel whattman a temperatura ambiente y luego se fijó en una cámara UV a 12,000 μJoules, dejando la membrana lista para ser hibridada con una sonda radiactiva.

## 6.12 PREPARACIÓN DE LA SONDA RADIOACTIVA.

El DNA que se utilizó como sonda para marcarlo radiactivamente provino de dos plásmidos, el plásmido pQE30CPSF160 y como control se utilizó un plásmido de actina, pCRactinaNorma, de la Dra. Norma Espinoza Sánchez. El plásmido pQE30CPSF160 se digirió con la enzima *HindIII*, y se usó un fragmento de 1904pb como sonda.

La sonda se preparó con el kit comercial Rediprime II Random Prime Labelling System (PROMEGA). Se siguieron los siguientes pasos.

1. Se diluyó el DNA a ser marcado a una concentración ~2.5 a 25ng en 45µL (final) de Tris-HCl pH=8 10 mM/EDTA 1mM (Buffer T.E).
2. Se desnaturalizó la muestra de DNA por calentamiento a 95-100°C por 5 minutos en baño de agua.
3. Transcurrido el tiempo, se pasaron las muestras a hielo durante 5 min.
4. Se centrifugó unos segundos para mantener la reacción en el fondo del tubo.
5. Se adiciono el contenido del tubo (45µL) al tubo de reacción (PROMEGA).
6. Se adicionaron 1µL [<sup>32</sup>P] dCTP (~10 µCi) y se mezclaron por medio de pipeteo lento de 5 a 10 veces.
7. Se Incubó 2Hrs a temperatura ambiente.
8. Transcurrido el tiempo, el protocolo del Kit indica detener la reacción con 5µL de EDTA 0.2M, pero se hizo de la siguiente forma:  
Se adicionaron 2µL tRNA (5µg/µl), 2.5 partes de Etanol absoluto (125µL), 1/10 parte de Acetato de Sodio 3M (5µL) y se dejó toda la noche precipitando a -20°C.

Para desnaturalizar la sonda se realizó lo siguiente:

1. Se centrifugó 10 min a velocidad máxima. Se descartar el sobrenadante y se resuspendió en 100µL de buffer TE.
2. Se desnaturalizó el DNA marcado por adición de 1/10 vol de NaOH 3N (10µl) y se incubó 5 min a temperatura ambiente.
3. Se adicionaron 10µl de HCl 3N (para neutralizar) y 10µl de Tris 1M pH=7.4 (para amortiguar).
4. Se agregó la sonda radiactiva a la solución de hibridación.

### **6.13 PRE-HIBRIDACIÓN DE LA MEMBRANA.**

Después de fijada la membrana, esta se pre-hibridó durante toda la noche en 20ml de solución de pre-hibridación (Apéndice A), a 37°C.

### **6.14 HIBRIDACIÓN DE LA MEMBRANA.**

1. Se agregó al tubo de hibridación con la membrana 10ml de solución de hibridación fresca conteniendo la sonda radiactiva desnaturalizada.
2. La sonda se hibridó durante  $2 Cot \frac{1}{2}$  (Carletti E. y col. 2006) a 42°C, calculados con base en la siguiente ecuación:

$$Cot \frac{1}{2} = 2 (1/x)(y/5)(z/10)$$

Donde:

x= Peso de la sonda en  $\mu\text{g}$  (0.08 $\mu\text{g}$ ).

y= Tamaño de la sonda en Kb (3.914Kb).

z= Volumen de la hibridación (10ml).

### **6.15 LAVADO DE LA MEMBRANA.**

1. La membrana se lavó con 2X SSC/0.5% SDS. Los primeros lavados se hicieron a temperatura ambiente y después se aumenta la temperatura paulatinamente.
2. La membrana se envolvió en plástico Egga Pack y se expuso a película de rayos X (Hyperfilm, Amersham) con pantalla intensificadora.

### 6.16 SUBLOCACIÓN DE CPSF-160 EN EL VECTOR DE EXPRESIÓN pQE30.

Después de clonar el gen CPSF-160 de *T. vaginalis* en el vector pCR-Blunt IITOP, se subclonó dicho gen en el vector pQE30 para la expresión de la proteína CPSF-160 de *T. vaginalis*. La estrategia fue la siguiente:

1. Se digirió el vector TOPOCPSF160 con las enzimas de restricción BamHI y KpnI, se purificó un fragmento de aproximadamente 3914pb.

TOPOCPSF160 (50µl)	
Buffer 4 1X (Invitrogen)	5µl
DNA	10µl
KpnI	1µl
BamHI	1µl
H <sub>2</sub> O	33µl

2. Por otra parte el vector pQE30 fue digerido con las enzimas de restricción BamHI y KpnI, se purificó el plásmido linearizado.

pQE30 (50µl)	
Buffer 4 1X (Invitrogen)	5µl
DNA	30µl
KpnI	1µl
BamHI	1µl
H <sub>2</sub> O	13µl

3. Los fragmentos se corrieron en un gel preparativo de agarosa 0.8% y se purificaron bajo las mismas condiciones que en el punto 6.5 de esta sección.

4. Después de purificados, se corroboró la integridad de los fragmentos en un gel de agarosa 0.8%.

5. Para realizar la ligación se utilizó la siguiente reacción:

Ligación pQE30+CPSF160 (10.5µl)	
Buffer de ligación (Invitrogen)	2µl
CPSF160	4µl

pQE30	4 $\mu$ l
T4 Ligasa	0.5 $\mu$ l

---

La ligación se dejó durante 2Hrs a 16°C y toda la noche a temperatura ambiente.

### 6.17 TRANSFORMACIÓN DE *E. coli* CON EL VECTOR pQE30CPSF160.

Para llevar a cabo la transformación con células químicamente competentes se realizó la siguiente metodología:

1. La ligación se colocó en hielo.
2. Se descongeló el stock de células JM16 y se agregó un volumen de 2 $\mu$ l de la ligación, se incubó inmediatamente en hielo a 4°C durante 30min. La ligación agregó a las células bajo campana.
3. Se dio un shock térmico de 43°C durante 30seg.
4. Se adicionaron 0.8ml de medio LB al tubo eppendorf y se dejó durante 1Hr a 37°C en agitación.
5. Se plateó 1/10 y 9/10 del volumen total de células en cajas de LB + Ampicilina (100 $\mu$ g/ml), y se dejaron a 37°C durante toda la noche.

### 6.18. EXPRESIÓN DEL VECTOR pQE30CPSF160.

Se utilizaron células de *E. coli* para realizar la expresión.

1. Se Inocularon 20mL de medio de cultivo LB, conteniendo ampicilina (100 $\mu$ g/mL), se inocularon 2 matraces con la cepa BL21Star y dos matraces con la cepa XL1Blue, de cada una de las cepas se tomó un matraz y se colocó a 37°C, los otros dos matraces se colocaron a temperatura ambiente; todos en agitación constante.
2. Los matraces de 37° se inocularon con 200 $\mu$ l del cultivo crecido toda la noche y los matraces de temperatura ambiente se inocularon con 400 $\mu$ l.
3. Los cultivos crecieron hasta alcanzar una densidad óptica a 600nm de 0.5 (3Hrs después) para cada cepa.

4. Cuando alcanzaron la DO se tomaron 10ml del cultivo, y se colocaron en un tubo falco. Esta muestra fue el control no inducido, se dejó en agitación de la misma forma que el cultivo inducido.
5. Se indujo la expresión mediante la adición de IPTG a una concentración de 1mM durante 24Hrs para las dos cepas. 1µl de IPTG 1M por cada ml de cultivo que quedó. Se dejó en agitación a temperatura ambiente.
5. Se prepararon los geles de poliacrilamida unas horas antes.
6. Se colectaron las muestras inducidas y no inducidas de ambas cepas. Se cosecharon las células por centrifugación a 4000xg durante 20 min. Se resuspendieron en 100µl 5XSDS-PAGE sample buffer.
7. Se calentaron las muestras a 95°C por 5 min.
8. Se centrifugaron a 15,000xg por 10 min.
9. Se cargaron 20µl de las muestras sin inducir e inducidas en un gel SDS-PAGE 12 % (Apéndice A).

#### **6.19. MODELADO *IN SILICO* DE CPSF-160.**

El modelado de la proteína CPSF-160 de *T. vaginalis*, se realizó utilizando una herramienta disponible en línea conocida como PHYRE2 (Protein Homology analogY Recognition Engine V 2.0). Esta herramienta utiliza como base el modelado de homología o modelado comparativo. El modelado de homología se basa en la identificación de uno o más dominios estructurales de proteínas conocidas, que se asemejan a la estructura de la secuencia de la consulta (Kelley L.A. Sternberg M.J.E. 2009).

## 7 Resultados.

### 7.1 ALINEAMIENTO DE SECUENCIAS.

El primer objetivo del proyecto era determinar una secuencia de nucleótidos que compartiera similitud con las secuencias reportadas para CPSF-160. Al realizar el ejercicio bioinformático de alineamiento de secuencias (BLASTn) de los cDNAs de CPSF-160 reportados, con el genoma de *T. vaginalis*, se obtuvo un gen (TVAG\_077030) que presentaba los mayores grados de identidad (Tabla. 1).

Organismo	Gen concordante	High Score	P	Tamaño	Identidad
<i>Arabidopsis thaliana</i>	TVAG_077030	198	$1.2e^{-21}$	3910pb	36.92%
<i>Drosophila melanogaster</i>	TVAG_077030	188	$1.8e^{-18}$	3910pb	35.86%
	TVAG_094790	113	0.043	3060pb	28.86%
<i>Homo sapiens</i>	TVAG_077030	196	$8.8e^{-24}$	3910pb	34.86%
	TVAG_283250	91	0.992	1582pb	14.7%
<i>Mus musculus</i>	TVAG_0.77030	206	$1.4e^{-24}$	3910pb	35.22%
<i>Trypanosoma cruzi</i>	TVAG_0.77030	95	$2.6e^{-5}$	3910pb	36.22%
	TVAG_058680	94	0.99	2181pb	14.27%
<i>Saccharomyces cerevisiae</i>	TVAG_077030	95	0.00045	3910pb	39.5%
	TVAG_013920	94	0.97	1209pb	12.81%

**Tabla. 1. Resultados del alineamiento de secuencias de cDNAs de CPSF-160 reportados, con el genoma de *T. vaginalis*.** El gen TVAG\_077030, presenta los porcentajes de identidad más altos.

También se realizó un alineamiento de las secuencias proteicas de CPSF-160 reportadas, con las secuencias de proteínas de *T. vaginalis*. Los resultados fueron los siguientes:

Organismo	Proteína concordante	High Score	P	Tamaño	Identidad
<i>Arabidopsis thaliana</i>	TVAG_077030	198	1.5e <sup>-22</sup>	147KDa	66.71%
	TVAG_326480	58	0.97	7.7KDa	4.02%
	TVAG_166610	80	0.98	40KDa	16.91%
	TVAG_228910	54	0.999	5.7KDa	3.26%
<i>Drosophila melanogaster</i>	TVAG_077030	188	1.1e <sup>-21</sup>	147KDa	66.21%
	TVAG_094790	113	0.0050	113KDa	51.56%
	TVAG_067770	88	0.52	52KDa	25.29%
	TVAG_319010	87	0.996	58KDa	27.2%
<i>Homo sapiens</i>	TVAG_077030	196	2.7e <sup>-25</sup>	147KDa	63.96%
	TVAG_283250	70	0.16	58KDa	4.37%
<i>Mus musculus</i>	TVAG_077030	206	1.8e <sup>-26</sup>	147KDa	64.41%
	TVAG_013050	87	0.56	128KDa	57.42%
	TVAG_129760	77	0.96	23KDa	11.1%
	TVAG_094790	74	0.97	113KDa	51.28%
	TVAG_248400	56	0.996	8.5KDa	4.37%
<i>Trypanosoma cruzi</i>	TVAG_077030	95	4.2e <sup>-7</sup>	147KDa	63.91%
	TVAG_058680	94	0.26	80KDa	33.86%
	TVAG_243970	56	0.996	8KDa	3.9%
	TVAG_408410	55	0.9992	7.5KDa	3.76%
<i>Saccharomyces cerevisiae</i>	TVAG_077030	95	0.00012	147KDa	71.91%
	TVAG_013920	94	0.11	8.5KDa	4.64%
	TVAG_458590	62	0.69	42KDa	20.56%

**Tabla. 2. Resultados del alineamiento de secuencias de proteínas de CPSF-160 reportadas, con las proteínas de *T. vaginalis*.** La secuencia proteica derivada del gen TVAG\_077030, aparece en todas las búsquedas realizadas, y presenta los porcentajes de identidad más altos.

Nuevamente la secuencia proteica deducida del gen TVAG\_077030, resultó con el mayor porcentaje de identidad para los CPSF-160 analizados, por lo tanto se decidió asignar la secuencia del gen TVAG\_077030 (Fig. 11) del genoma liberado de *T. vaginalis*, como la secuencia que codifica para CPSF-160. A partir de ella se diseñaron oligonucleótidos que permitieran amplificar y clonar este gen.

## 7.2 ANÁLISIS DE LA SECUENCIA DE TVAG-077030.

La secuencia del gen TVAG\_077030 se utilizó en este estudio, debido a que cumplía con los criterios de mayor identidad comparada con los genes de CPSF-160 reportados. El tamaño del cDNA se estimó en 3894pb y se lograron identificar algunos elementos en la secuencia tales como: la región promotora (Inr) de tipo (A/C) AAAAAT (G/T) (Smith A.J. y col. 2011), el codón de inicio en fase con el codón de termino y la señal de poliadenilación (TAAA) propuesta en el laboratorio de la Dra. Imelda López Villaseñor.

```
1      ATATTGGTGA GTCCAAAACA TAGGCAAAAA TTGCAAAATT TATTAGAAGC TCCTAAAAAA
61      ATGAATGTTA CGCATCGAAC AATATGTGAT TCCCCAGCCC CGGAATCAGC AATTCGATGC
121     CACCTTCCGT TTTCAAAGGC AGATTGCATT GTTACTGCCA ATGGAAACAA GCTCCAAGTC
181     TATTCAACGA AGGAAGATGA TTTAAGATTA GTATGGGAAA AGAAGTTTTG GGGTGA AATT
241     TTCGGTGTAT ACAGGCATAA AAGTGGTGGG GAATACGATT CAATCATTGT TGGCTGCGAT
301     ACAAGTAAAG TTATTGTTCT ACAGGTTATA GATAATGATT TGAAAGAAAC AGAATATCAC
361     GAATTTAACA GACCTGGCCC ACCAGAGCCA GATCCACCAA AACCAGAACG TCCTTTCGAT
421     ATTAGTACAA GATTACGCAA TAAAACAATT ATGGATGCCG ACCCAACAGG TACATGCCTT
481     GCATTACTTC TTGCTCAAAA TATTCTTTAC GTTTTACCTC TTGCCAACAA GATTA A AATC
541     GAAAGTACCG AAAAAGCTGG TGATGAATAC CATTCCAGCT GGAAAGTAAT TAAAGATGCG
601     TTCGCGTATG ATGTCCACAC AGATTTTAAA TCGCCGTTAT ACAGAATTCG AGATATGGTT
661     TTCTTGACG GATATAAAAA TCCAACCTC GCTATTATTC ACGAATTAAT TCCTACATGG
721     AGTGTCAGAT TACCACTTCA GAAATCAACT GTTGCGGTTT CAATTGTTTC TCCTCCATTA
781     AAAAAAAGG AAAGTGT TTTT AATTTT CAGCA TCAATCGATA AAGTTACAAT GTGGACATCC
841     CGTGCATTAC CACATAACTC ATTTGGCTTA GTTCATGTTC CAGATCCAAT TGGAGGCTTC
901     CTGTTCTCT CAAAGAACGC AATTATTTAC ATGGACCATA CAAACATTGT TGCTTTAGCA
961     TTAACAAGC TCGCATACTT AGATGATGAA GTTCCAGTCG ATATTACAGC AAATGGACCT
1021    GGTGTGTCAG AATTATATTC AAAAGTTGGT ACAGCTATCG ATAAATCACA TATTTTATTG
1081    ACCGTTGACC AACATTATTT GTCCATTTTG ACTTTACATT ACAATGGTGT CAAAGTAACA
1141    AATCTTTCTT TAAATGTCAA CTTAAATCTT GAATTTTATC CATCATGCTT CCTTTCTCTT
1201    AACTACACGA ACAACAGGAG CCTTGTTTTT ATGGGATCTA CAACTCATGA CTCAACTTTG
1261    TCCGAAATAA TTTTAGAAAT TGAAGAAGAA GAACTTGCAA GCTTTTTGGT CGACCATGTC
1321    ATGACAGAAA CACAAAATGA TCTTACCAG AAGTTCTTTA AGTCATTACC AAGGCCTAAA
1381    ACCCTTTCAA AAGTCAAATC ATTTTCTGTC AACAACATTT CTTACATCAG ACAGCTTGGA
1441    ACTGTTGCCA ATGCCACTCC TTGCCTCAAG CTGAGCCAGA TAGATGAAAA TGAAGAACAG
1501    ATCGCGATGG CTCTCGCATG TGGCTTCAA AAATACGGTT GTTTGCAATTT CGTTCGTAGT
1561    GGTATCGATC CAAACCTTGT CAGTGAATTT AATATCCGCG GAGTTACTGG AATTTATTCA
1621    TCAGAAAACT TCTCATATAT CTTGATCTCA ACAGATGAAG GAACATCCGT ACTTTACGCC
1681    GATAGCCAGA ACCTTCCAGA GATCAACGAT GAGTTTACGA GCCAAGTAAG TAAGGAAAAG
1741    ACTCTTTTTG CAGGAGATTT CAACAACCTT TTCATCCAGA TCACTAAAAC AGTCGTAAGA
1801    GCGATCAAAT ATTCACAAGA AGATGGTTTC TCCCACACTG ACCTCTATTC TGGCTCTTCT
1861    CCAGCTGTTT TGGATGCAAG AATTTCTTCA AATGCCCTTG TTGTTCTATT TGACAACATAT
1921    ACAATATGCG CATATGACAG AGCGTTGGAA CCAATGACAC CTCTCACTAA CAAATATTTT
1981    AGCCGAATCG CAATTTTCGG TAAATATTTA TTTGTTTATT GTACAAACGG CGTTTTGAGA
2041    ATGTTTTCCG TAGAAACGAT GACGGAAGTT GCCACATTTG ATAATTTCAA GGCATTTCTT
```

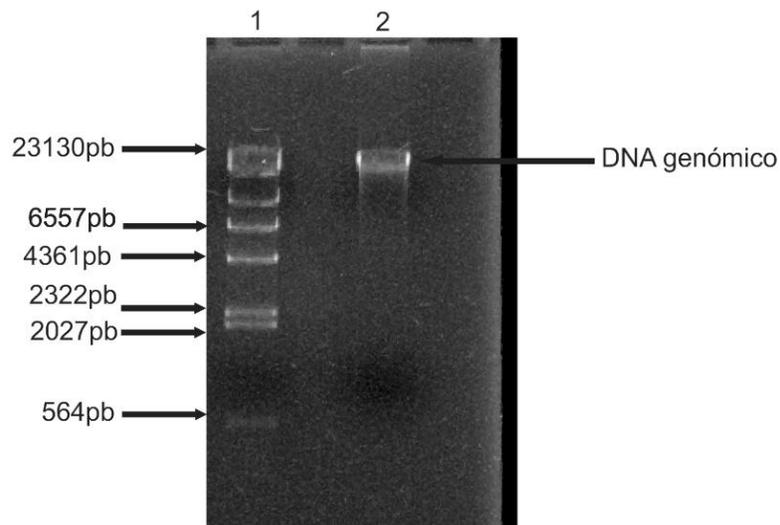
2101 GATTTTCATTG CTCCTCTCGG AAAAGATGAG ATTCCATTGC AGCCATCTTC AATTGGTATC  
 2161 GTTGATATCT CGTTCGTTGT CTTCGAATCA GTTGTCTTC TTGTTCTTAC ACCTCGCGAT  
 2221 GGCCCTGCTT TCTTCTATCA ATGGGTTGAA TCTAAATTAT TCTTCAGAAG GATCAAAAATG  
 2281 CACAGATTTC CAGTTACTTC ACGAAAAGAA ATTGATATCA TCCCATTCAA CGGCATTTCG  
 2341 GATGTTTTTCG ATGGTTGTTT CGTCACAGAT CCATCAAATC CTTTCTTTAT CCTCTCGGAA  
 2401 AACGGTTACC CACGTATCAT CCCACTCTCC GAGTCATTCA ACTGCAAACC ACGTGTCTTT  
 2461 GCCCCTGTAA TCGGTGGTCC TTACAAGAA CACTTTTTAA TTGCGGACGA AGACCAAATC  
 2521 AGGCTCTGCA ATCTCGAGAA TATTAAGCCA GAACACAATT TCTTCATCAT TGATGGCTGT  
 2581 ATTGTGCGAGA GAATCCCTGT TGGAAAGACT GTAAGAAGAA TTGCCTACTG CCAGAACCCG  
 2641 AATTGCGTCG CTTTCATCGC TTCCCATCCA GAACCATTCA CAACAGAGAA CGAAAAGAAG  
 2701 ATCGATGTCG AAGTTTATGA GAATCTTCAG GTTCATTACC AAGAACCACC TTCTCCAGCC  
 2761 AAAGTTTACC CAGATGAGGA CTACGAAACA ATTCCTAAGT GGAATGAAGA GCGATACTCC  
 2821 CTGTTTCTCT ATACAAAAGA TGGTTTGCAA CAAATGGTCG ATTACGCAAA CCACGAAATC  
 2881 GTCAATACAG TGCAATTTGT ACATACTACT CCGATGCCAG AGGATGGAAT AACCCCTCTT  
 2941 AACACGTATT TGGCCGTAGG AAGTGGATT CTTGTCTCAAC CGGAAAAAAT GATGAGAGGA  
 3001 GTCCTTTATA TTTACCAAAT CAGGTACATG CAGAACGATG AGGGATTTAA CGAGATCAGC  
 3061 TTAAGACCTT TGTACAACGA AACAAACAAA ATTTACAAA ATCCGATTAT CGAAATCAGC  
 3121 GATAACTCAG GATACATGGC TATTTTCTGC GGTAACCTT TGTATCTCAT GAGGTCTTC  
 3181 AACGAAAACA CCGTAAAGAT CGAAGCTTTC CTTGTTGGAA GATTCTTTGC ATCATCTATC  
 3241 GTTTCCTCA AGAATTATCT GCTCTACGCT GATTCTTACG AAGGCTTTGA GGTGCGCAGA  
 3301 TGGCGTAAAT ACGGAAAGAA GCTCATTTC AATGCAAGAG ATACGATGAC TAAATTACCT  
 3361 CTCAGTGCAG CATTCTACA ATACGAGGAT TGCTTGGGAG GTGTCGTATT TGATGATGAC  
 3421 GGCAACGCTC ATATCTTTGA CGTTGACGAA TACGCGATTC CAGCAGATGC AGTTGTCAGA  
 3481 AAGTCAATTT TCTACATTGG AGGCCGAGCA ATCAGTTCTG GTCAGTTCCC AATCAAGGCT  
 3541 GTGACCCAGG CAACTCAGCA AAATCCAAAT GAAGAAATG ATGAAGAATT ATTGCAGCTG  
 3601 CAGACAAAGA TTGGAGGCCA CATCGCATGG TATGTTACTA CACATGGTAA GATCGGTGCA  
 3661 TTCACTCAA TTGATGAAAA CGATCGACAT AAATTGGTCG GAGTCAATC CGCATAAGAA  
 3721 AAATCTCTTT GCGGACTGAG CCACTTAGAA TACAGGAGTG GAAAGTCAA GAATATGATT  
 3781 GAACAGGATA TTTTCAACCA ATCTCCAAAG AATGTCATTG ATTGCGACAT GTTGATTGAT  
 3841 TTGATTGAAG ATATGCCTGA TCACTTGAAG TTTGCAACAA AGGGATTAAG AACACAGGAC  
 3901 TTCTTATCTG AATTAAGAAA GATTTACAAT AATGGTGTCA ATGTATTCCA **TAA**ATAACT  
  
 3961 ATATATAGAT TTTTATTATT TATTATATTA AAATGATTTA ATATTTTTCT GAATAAAGCA  
 4021 AATTTTCATT AGAAATCTCA CCTTTTTTTC TGAATTAAG GTGCACACGA GAGGATCATG  
 4081 TAAACTTCAG AATTTAGAAC CTTATTTTTA GGCATTATCT TGGCTACCTA TTTAGGCTCG  
 4141 AAAAGCAAAA AATATCAAAA AATACAACCT GACTCCATTA TGACCCTTTC ATGGAGGTAG  
 4201 TTTTATCCG CTAGTATTTT CCCTTGAAT

**Fig. 11. Secuencia del gen TVAG\_077030 del genoma liberado de *T. vaginalis* 5'-3'. En un cuadro verde se observa la región promotora (Inr) (A/C) AAAAAT (G/T), en cuadros en color rojo se muestran los codones de inicio y de término, subrayado en naranja se muestra la posible señal de poliadenilación.**

La localización de estos elementos permitió establecer que la secuencia del gen TVAG\_077030 del genoma liberado de *T. vaginalis* tiene los elementos de un gen que se expresa en *T. vaginalis*. A partir de ella se diseñaron oligonucleótidos para clonar esta secuencia.

### 7.3 EXTRACCIÓN DE DNA GENÓMICO Y AMPLIFICACIÓN DEL GEN DE CPSF-160 DE *T. vaginalis*.

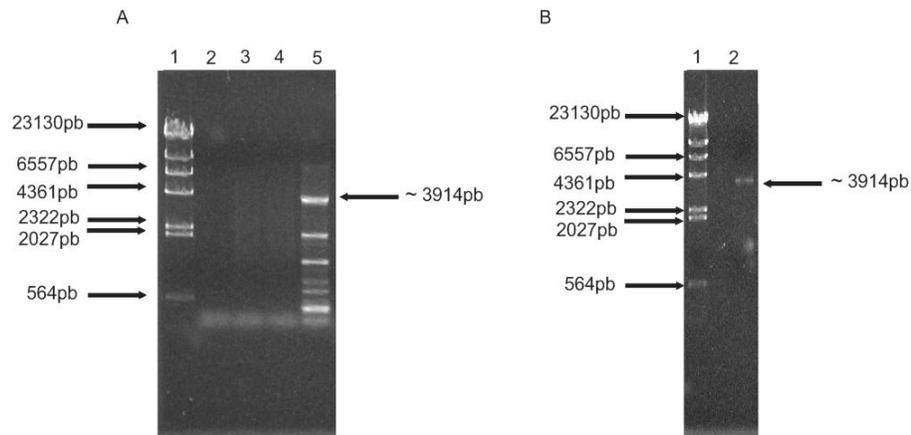
El DNA genómico purificado de *T. vaginalis*, se corrió en un gel de agarosa para verificar su integridad (Fig. 12). Se utilizó el DNA genómico purificado como templado para realizar una reacción PCR que permitiera amplificar el gen de CPSF-160.



**Fig. 12. DNA genómico de *T. vaginalis*.** El DNA genómico se corrió en un gel de agarosa 0.8% TBE, se utilizó como templado para amplificar el gen de CPSF-160. (1) Marcador de peso molecular (DNA  $\lambda$ .digerido con HindIII) (2) DNA genómico de *T. vaginalis*.

Para amplificar el gen de CPSF-160 de *T. vaginalis*, se diseñaron dos oligonucleótidos: CPSF-K1 (5' GGTGTCAATGTATTCCAATAAGGTACCGATG3') y CPSF-BH1 (5' CATCGGATCCATGAATGTTACGCATCG AAC 3'), basados en la secuencia de TVAG\_077030 (Fig. 11) del genoma liberado de *T. vaginalis* (Carlton J.M. y col. 2009). Como templado se utilizó el DNA genómico purificado de *T. vaginalis*. El

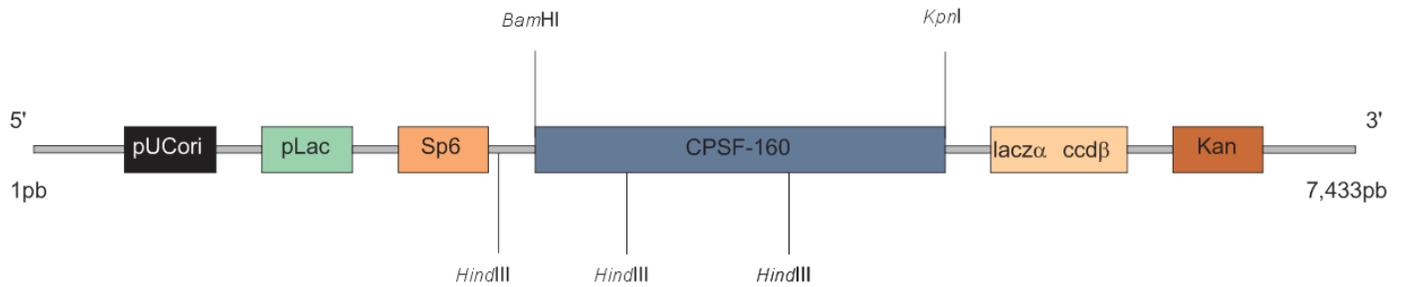
producto de PCR se analizó en un gel de agarosa al 0.8% (Fig. 13A) Se consiguió amplificar un fragmento del tamaño esperado (3914pb), sin embargo, también se amplificaron fragmentos de menor tamaño. Se purificó el fragmento más grande (Fig. 13B) y se clonó en el vector para productos de PCR, pCR-Blunt II TOPO (INVITROGEN) para generar el plásmido TOPOCPSF160.



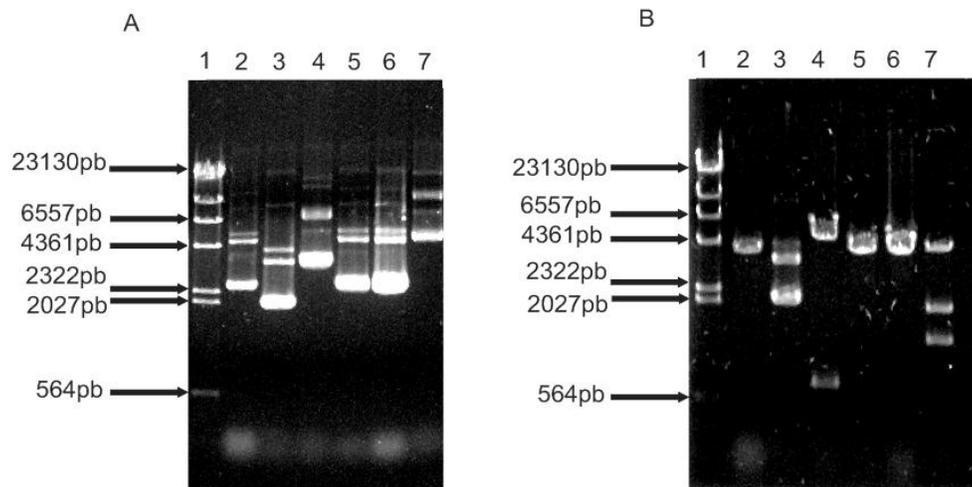
**Fig. 13. Productos de PCR analizados en geles de agarosa 0.8% TBE . Muestra la amplificación de un fragmento de DNA de un tamaño aproximado al esperado (3914pb). (A) 1 Marcador de peso molecular l. 2 Control de PCR sin Taq Polimerasa. 3 Reacción de PCR con baja sal, 1.5mM MgCl<sub>2</sub>. 4 Reacción de PCR con mediana sal, 3mM MgCl<sub>2</sub>. 5 Reacción de PCR con alta sal, 4.5mM MgCl<sub>2</sub>. (B) Fragmento de ~3914pb purificado.**

#### 7.4 ANÁLISIS DEL PLÁSMIDO TOPOCPSF160.

Se seleccionaron 6 clonas transformadas de TOPOCPSF160 (Apéndice B), se purificaron los plásmidos y se analizaron en un gel de agarosa 0.8% para evaluar el tamaño de los plásmidos, asumiendo que el plásmido más grande tendría el inserto (Fig. 14).



**Fig. 14. Mapa del plásmido TOPOCPSF160 5'-3'.** El gen CPSF-160 de *T. vaginalis* se observa en color morado, bajo el promotor Sp6. Se muestra la posición de los sitios de restricción BamHI y KpnI agregados con los oligonucleótidos y tres sitios de restricción HindIII utilizados para evaluar el tamaño del plásmido. Se muestran los sitios de origen de la transcripción y el gen de kanamicina.

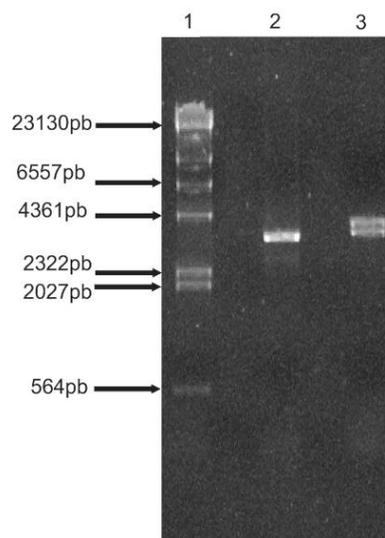


**Fig. 15. Geles de agarosa 0.8% para el análisis de las clonas TOPOCPSF160.** (A) Carril 1 Marcador de peso molecular  $\lambda$  digerido con HindIII. Carriles 2 – 7 plásmidos sin digerir de 6 de las 7 clonas de TOPOCPSF160. (B) Se muestra 6 clonas de TOPOCPSF160, digeridas con HindIII. La clona 7-TOPOCPSF160 mostró el patrón esperado (carril 7).

Las 6 clonas seleccionadas (Fig. 15A) se analizaron por restricción con la enzima *HindIII*, se esperaban fragmentos de 4220pb, 1904pb y 1309pb. Unicamente la clona 7-TOPOCPSF160 presentó el patrón de restricción esperado (Fig. 15B, carril 7). Dentro del inserto se encuentra la secuencia de dos sitios de restricción *HindIII* y uno más en el sitio múltiple de clonación del plásmido, por esta razón se realizó el análisis de restricción con dicha enzima. En la figura 16 podemos observar la posición los sitios de restricción *KpnI* y *BamHI*, agregados con los oligonucleótidos CPSF-K1 y CPSF-BH1, para poder sub-clonar el inserto en el vector de expresión pQE30.

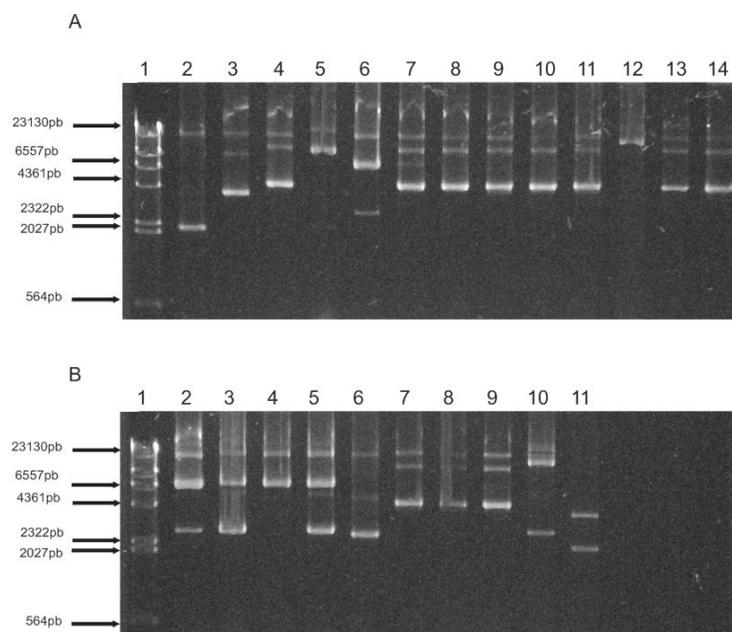
### 7.5 SUBCLONACIÓN DE CPSF-160 EN EL VECTOR DE EXPRESIÓN pQE30.

Para subclonar el inserto de CPSF-160 en pQE30, primero se liberó del plásmido pTOPOCPSF160 usando las enzimas *BamHI* y *KpnI*. Se esperaban dos fragmentos uno de 3904pb, que era el tamaño del inserto y otro de 3474pb que sería el vector. El vector pQE30 se digirió con las mismas enzimas, el fragmento esperado para el vector digerido era de 3439pb (Fig. 16).



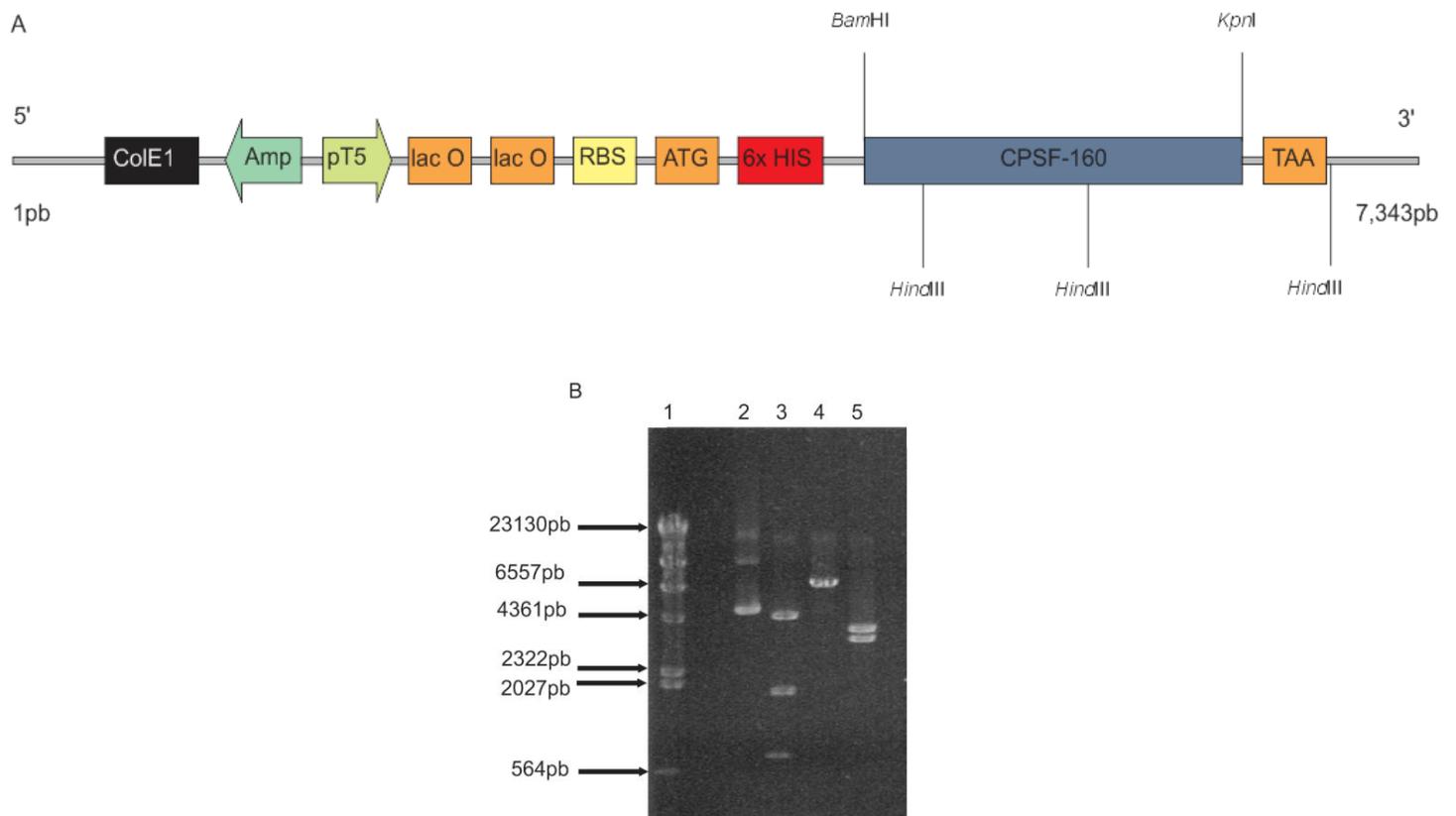
**Fig. 16. Gel de agarosa 0.8% TBE con las digestiones de TOPOCPSF160 y pQE30.** Se muestran los productos de restricción de los plásmidos pQE30 en el carril 2 y TOPOCPSF160 en el carril 3, ambos digeridos con las enzimas de restricción *BamHI* y *KpnI*. En el carril 1 se muestra el marcador de peso molecular  $\lambda$  digerido con *HindIII*.

Los fragmentos que coincidieron con el peso molecular esperado (3439pb del plásmido pQE30 linearizado y 3904pb del inserto TvCPSF-160) se purificaron y se colocaron en una reacción de ligación con la enzima T4 ligasa a 16°C. La ligación se transformó en células de *Escherichia coli* (JM109). Se seleccionaron 22 clonas de las cuales se obtuvo DNA plasmídico y se corrieron en un gel de agarosa 0.8% para analizar el tamaño de los plásmidos. Se usó como control de tamaño el plásmido pQE30 sin inserto. De las 22 muestras de DNA plasmídico se seleccionaron las muestras 4pQE30CPSF160, 5pQE30CPSF160 y 7pQE30CPSF160 (Fig. 17A, carriles 4, 5 y 7) pues eran muestras que representaban los diferentes tamaños de plásmidos encontrados



**Fig. 17. Geles de agarosa 0.8% para el análisis de DNA plasmídico de 22 clonas transformadas con la construcción pQE33CPSF160. En el carril 1 de ambos geles se observa el marcador de peso molecular  $\lambda$  digerido con *HindIII*. (B) En el carril 11 se observa el plásmido pQE30, (A y B) Se observa el DNA plasmídico de 22 clonas de la construcción pQE30CPSF160.**

Las 3 clonas seleccionadas se analizaron por restricción con la enzima *HindIII*, se esperaban fragmentos de 4663pb, 1904pb y 776pb; se realizó también un análisis de restricción con las enzimas *BamHI* y *KpnI*, se esperaban fragmentos de 3904pb y 3439pb y finalmente un análisis de restricción sólo con la enzima *BamHI* para linearizar el plásmido. Sólo la clona 4pQE30CPSF160 presentó el patrón de restricción esperado (Fig. 18B).

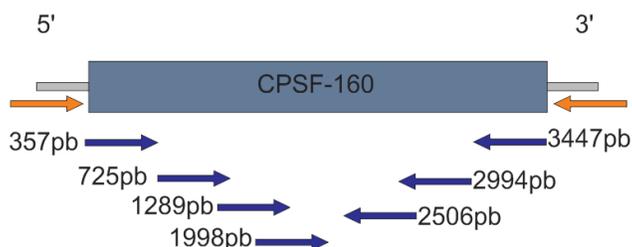


**Fig. 18. Mapa del plásmido TOPOCPSF160 5'-3' y análisis de las clonas pQE30CPSF160 en gel de agarosa 0.8%. (A) Mapa del plásmido pQE30CPSF160. (B) Carril 1 marcador de peso molecular  $\lambda$  digerido con HindIII. Carril 2: 4pQE30CPSF160 sin digerir. Carril 3: 4pQE30CPSF160 digerido con HindIII. Carril 4: 4pQE30CPSF160 digerido con BamHI. Carril 5 4pQE30CPSF160 digerido con BamHI y KpnI.**

El inserto de CPSF-160 al ser clonado en el vector pQE30 contendría un Tag de Histidinas en el extremo N-terminal. Bajo el promotor de T5 se encuentran las regiones del operador de lac, el sitio de unión al ribosoma (RBS, ribosome-binding site) y el ATG iniciador, el codón de paro está dentro de la secuencia del CPSF-160 (Fig. 18A). Éste plásmido se denominó como 4pQE30CPSF160 (Apéndice C).

## 7.6 SECUENCIACIÓN Y ANÁLISIS DE LA CLONA 4PQE30CPSF160.

La clona 4pQE30CPSF160 se secuenció en la unidad de secuenciación del Instituto de Investigaciones Biomédicas, de la Universidad Nacional Autónoma de México. Se utilizaron los oligonucleótidos comerciales (QUIAGEN) para la secuenciación de vectores de la familia pQE, para anclarse fuera de la región codificante; se diseñaron también 7 oligonucleótidos que se anclaban dentro de la región codificante para secuenciar completamente la clona (Fig. 19).



**Fig. 19. Sitios de anclaje de los oligonucleótidos utilizados para la secuenciación de de la clona TvCPSF-160.** El esquema muestra la región de anclaje de los oligonucleótidos, en rojo se esquematizan los oligonucleótidos comerciales para secuenciar fuera de la región codificante, en azul se muestran los oligonucleótidos diseñados y el sitio donde se anclan y el tamaño de los fragmentos secuenciados.

Tras analizar los electroferogramas se encontró que la secuencia presenta el ATG del vector pQE30 en fase con el ATG del inserto del gen TvCPSF-160, se observa también la secuencia del Tag de Histidinas, y la secuencia muestra al final el codón de término de la traducción (TAA), también en fase (Fig. 20).

```

1      ATGAGAGGAT CGCATCACCA TCACCATCAC GGATCCATGA ATGTTACGCA TCGAACAAATA
61     TGTGATTCCC CAGCCCCGGA ATCAGCAATT CGATGCCACC TTCCGTTTTC AAAGGCAGAT
121    TGCATTGTTA CTGCCAATGG AAACAAGCTC CAAGTCTATT CAACGAAGGA AGATGATTTA
181    AGATTAGTAT GGGAAAAGAA GTTTTGGGGT GAAATTTTCG GTGTATACAG GCATAAAAGT
241    GGTGGAGAAT ACGATTCAAT CATTGTTGGC TCGGATACAA GTAAAGTTAT TGTTCTACAG
301    GTTATAGATA ATGATTTGAA AGAAACAGAA TATCACGAAT TTAACAGACC TGGCCCACCA
361    GAGCCAGATC CACCAAAACC AGAACGTCCT TTCGATATTA GTACAAGATT ACGCAATAAA
421    ACAATTATGG ATGCCGACCC AACAGGTACA TGCCTTGCAT TATTTCTTGC TCAAAATATT
481    CTTTACGTTT TACCTCTTGC CAACAAGATT AAAATCGAAA GTACCGAAAA AGCTGGTGAT

```

541 GAATACCATT CCAGCTGGAA AGTAATTTAAA GATGCGTTCG CGTATGATGT CCACACAGAT  
601 TTAAATCGC CGTTATACAG AATTCGAGAT ATGGTTTTCT TGGACGGATA TAAAAATCCA  
661 ACTCTCGCTA TTATTCACGA ATTAATTCCT ACATGGAGTG TCAGATTACC ACTTCAGAAA  
721 TCAACTGTTG CGGTTTCAAT TGTTTCTCCT CCATTAAAAA AAAAGGAAAA TGTTTAAATT  
781 TCAGCATCAA TCGATAAAGT TACAATGTGG ACATCCCGTG CATTACCACA TAACTCATTT  
841 GGCTTAGTTC ATGTTCCAGA TCCAATTGGA GGCTTCCTTG TTCTCTCAA GAACGCAATT  
901 ATTTACATGG ACCATACAAA CATCGTTGCT TTAGCATTAA ACAAGCTCGC ATACTTAGAT  
961 GATGAAGTTC CAGTCGATAT TACAGCAAAT GGACCTGGTT GTCACGAATT ATATTCAAAA  
1021 GTTGGTACAG CTATCGATAA ATCACATATT TTATTGACCG TTGACCAACA TTATTGTCC  
1081 ATTTTGACTT TACATTACAA TGGTGTCAA GTAACAAATC TTCCTTTAAA TGTCACCTTA  
1141 AATCTTGAAT TTCATCCATC ATGCTTCCTT TCTCTTAACT ACACGAACAA CAGGAGCCTT  
1201 GTTTTCATGG GATCTACAAC TCATGACTCA ACTTTGTCCG AAATAATTTT AGAAATTGAA  
1261 GAAGAAGAAC TTGCAAGCTT TTTGGTCGAC CATGTATGA CAGAAACACA AAATGATCTC  
1321 TACCAGAAGT TCTTTAAGTC ATTACCAAGG CCTAAAACCC TTTCAAAAGT CAAATCATTT  
1381 TCTGTCAACA ACATTTCTTA CATCAGACAG CTTGGAAGT TGGCCAATGC CACTCCTTGC  
1441 CTCAAGCTGA GCCAGATAGA TGAAAATGAA GAACAGATCG CGATGGCTCT CGCATGTGGC  
1501 TTCAAAAAAT ACGGTTGTTT GCATTTGCTT CGTAGTGGTA TCGATCCAAA CCTTGTGAGT  
1561 GAATTCAATA TCCGCGGAGT TACTGGAATT TATTCATCAG AAAACTTCTC ATATATCTTG  
1621 ATCTCAACAG ATGAAGGAAC ATCCGTACTT TACGCCGATA GCCAGAACCT TCCAGAGATC  
1681 AACGATGAGT TTACGAGCCA AGTAAGTAAG GAAAAGACTC TTTTTGCAGG AGATTTCAAC  
1741 AACTTCTTCA TCCAGATCAC TAAAACAGTC GTAAGAGCGA TCAAATATC ACAAGAAGAT  
1801 GGTTCCTCCC AACTGACCT CTATTCTGGC TCTTCTCCAG CTGTTTTGGA TGCAAGAATT  
1861 TCTTCAAATG CCCTTGTTGT TCTATTTGAC AACTATACAA TATGCGCATA TGACAGAGCG  
1921 TTGGAACCAA TGACACCTCT CACTAACAAA TATTTAGTC GAATCGCAAT TTTCCGTAAA  
1981 TATTTATTTCG TTTATTGTAC AACCGGCGTT TTGAGAATGT TTTCCGTAGA AACGATGACG  
2041 GAAGTTGCCA CATTTGATAA TTTCAAGGCA TTTCTGATT TCATTGCTCC TCTCGGAAAA  
2101 GATGAGATTC CATTGCAGCC ATCTTCAATT GGTATCGTTG ATATCTCGTT CGTTGTCTTC  
2161 GAATCAGTTG TCCTTCTTGT TCTGACACCT CGCGATGGCC CTGCTTTCTT CTATCAATGG  
2221 GTTGAATCTA AATTATTCTT CAGAAGGATC AAAATGCACA GATTGCGAGT TACTTACCGA  
2281 AAGAATATTG ATATCATCCC ATTCAACGGC ATTTCCGATG TTTTCGATG TTGTTTCGTC  
2341 ACAGATCCAT CAAATCCTTT CTTTATCCTC TCGGAAAACG GTTACCCACG TATCATCCCA  
2401 CTCTCCGAGT CATTACCTG CAAACCACGT GCTTTTCCCC CTGTAATGCG TGGTCTTAC  
2461 AAGAATCACT TTTTAATTGC GGACGAAGAC CAAATCAGGC TCTGCAATCT CGAGAATATT  
2521 AAGCCAGAAC ACAATTTCTT CATCATTGAT GGCTGTATTG TCGAGAGAAT CCCTGTTGGA  
2581 ATGACTGTAA GAAGAATTGC CTAAGTCCAG AACCCGAATT GCGTCGCTTT CATCGCTTCC  
2641 CATCCAGAAC CATTACAAC AGAGAACGAA AAGAAGATCG ATGTCGAAGT TTATGAGAAT  
2701 CTTGAGTTC ATTACCAAGA ACCACCTTCT CCAGCCAAAG TTTACCCAGA TGAGGACTAC  
2761 GAAACAATTC CGAAGTGGAA TGAAGAGCGA TACTCCCTGT TTCTCTATAC AAAAGATGGT  
2821 TTGCAACAAA TGGTCGATTA CGCAAACCAC GAAATCGTCA ATACAGTGCA ATTTGTACAT  
2881 ACTACTCCGA TGCCAGAGGA TGAATAACC CTCCTTAAAC CGTATTTGGC CGTAGGAAGT  
2941 GGATTCTTGT CTCAACCGGA AAAAATGATG AGAGGAGTCC TTTATATTTA CCAAATCAGG  
3001 TACATGCAGA ACGATGAGGG ATTTAACGAG ATCACGTTAA GACCTTTGTA CAACGAAACA  
3061 AACAAAATTT ACAAATCC GATTATCGAA ATCACGGATA ACTCAGGATA CATGGCTATT  
3121 TTCTGCGGTA ACCTTTTGTA TCTCATGAGG TTCTTCAACG AAAACACCGT AAAGATCGAA  
3181 GCTTTCCTTG TTGGAAGATT CTTTGCATCA TCTATCGTTT CCCTCAAGAA TTATCTGCTC  
3241 TACGCTGATT CCTACGAAGG CTTTGGAGTC GCCAGATGGC GTAAATACGG AAAGAAGCTC  
3301 ATTTCAATGG CAAGAGATAC GATGACTAAA TTACCTCTCA GTGCAGCATT CCTACAATAC  
3361 GAGGATTGCT TGGGAGGTGT CGTATTTGAT GATGACGGCA ACGCTCATAT CTTTGACGTT  
3421 GACGAATACG CGATTCCAGC AGATGCAGTT GTCAGAAAGT CAATTTTCTA CATTGGAGGC

```

3481   CGAGCAATCA GTTCTGGTCA GTTCCCAATC AAGGCTGTGA CCCAGGCAAC TCAGCAAAAT
3541   CCAAATGAAG AAATTGATGA AGAATTATTG CAGCTGCAGA CAAAGATTGG AGGCCACATC
3601   GCATGGTATG TTACTIONACACA TGGTAAGATC GGTGCATTCA CTCCAATTGA TGAAAACGAT
3661   CGACATAAAT TGGTCGGAGT TCAATCCGCA TACGAAAAAT CTCTTTGCGG ACTGAGCCAC
3721   TTAGAATACA GGAGTGAAA GTTCAAGAAT ATGATTGAAC AGGATATTTT CAACCAATCT
3781   CCAAAGAATG TCATTGATTG CGACATGTTG ATTGATTGA TTGAAGATAT GCCTGATCAC
3841   TTGAAGTTG CAACAAAGG ATTAAGAACA CAGGACTTCT TATCTGAATT AAGAAAGATT
3901   TACAATAATG GTGTCAATGT ATTCCAA TAA

```

**Fig. 20.** Secuencia del gen TvCPSF-160 5'-3'. En un cuadro rojo se muestran los codones de inicio y de término de la traducción; en el cuadro negro se observa la secuencia del Tag de Histidinas; en un cuadro verde se muestra el ATG del inserto en fase con el ATG del plásmido pQE30.

La secuencia de TvCPSF-160 se alineó con la secuencia de DNA accesada en el genoma de *T. vaginalis*. Se encontraron 13 diferencias a nivel de nucleótidos, ninguna de estas genera un codon de término, y tampoco se encontraron inserciones o deleciones (Apéndice D). La secuencia también fue alineada con las secuencias de DNA de CPSF-160 de otros organismos. Se encontró un 35% de identidad con la secuencia de *Arabidopsis thaliana*, un 35.6% de identidad con la secuencia de *Drosophila melanogaster*, 33.56% con la secuencia de *Homo sapiens*, 34.56% con la secuencia de *Mus musculus*, 36.03% con la secuencia de *Trypanosoma cruzi*, y finalmente un 39.37% de identidad con la secuencia de DNA del CPSF-160 de *Saccharomyces cerevisiae* (Tabla. 1).

Conociendo la secuencia completa del gen TvCPSF-160 clonado, se dedujo la secuencia proteica. Esta secuencia se compone de 1297 aminoácidos y tiene un peso estimado de 146,935Da (147KDa) (Fig. 21).

```

1      MRGSHHHHHH GSMNVTHRTI CDSPAPESAI RCHLPFSKAD CIVTANGNKL QVYSTKEDDL
61     RLVWEKKFWG EIFGVYRHKS GGEYDSIIVG CDTSKVIVLQ VIDNDLKETE YHEFNRP GPP
121    EPDPPKPERP FDISTR LRNK TIMDADPTGT CLALFLAQNI LYVLPLANKI KIESTEKAGD
181    EYHSSWKVIK DAFAYDVHTD FKSPLY RIRD MVFLDGYRNF TLAIHELIF TWSVRLPLQR
241    STVAVSIVSP PLKKKENVLI SASIDKVTMW TSBALPHNSF GLVHVPDPIG GFLVLSKNAI
301    IYMDHTNIVA LALNKLAYLD DEVPVDITAN GPGCHELYSK VGT AIDKSHI LLTV DQHLYS
361    ILTLHYNGVK VTNLPLNVNL NLEFHPSCFL SLNYTNNRSL VFMGSTHDS TLSEIILEIE
421    EEELASFLVD HVMTETQNDL YQKFFKSLPR PKTLSKV KSF SVNNISYIRQ LGTVANATPC
481    LKLSQIDENE EQIAMALACG FKKYGCLHFV RSGIDPNLVS EFNIRGVTGI YSSENF SYIL

```

```

541      ISTDEGTSVL YADSQLPEI NDEFTSQVSK EKTLFAGDFN NFFIQITKTV VRAIKYSQED
601      GSSHTDLYSG SSPAVLDARI SSNALVLFED NYTICAYDRA LEPMTPLTNK YFSRIAIFGK
661      YLFVYCTTGV LRMFSVETMT EVATFDNFKA FPDFIAPL GK DEIPLQPSI GIVDISFVVF
721      ESVVLLVLT P RDGPAFFYQW VESKLFRRRI KMHRFAVTSR KNIDIIPFNG ISDVFDGCFV
781      TDPSNPFIL SENGYPRIIP LSESFTCKPR AFPPVMRGPY KNHFLIADED QIRLCNLENI
841      KPEHNFFIID GCIVERIPVG MTRRIAYCQ NPNCVAFIAS HPEPFTTENE KKIDVEVYEN
901      LQVHYQEPPS PAKVYPDEDY ETIPKWNEER YSLFLYTKDG LQQMVDYANH EIVNTVQFVH
961      TTPMPEDGIT LLNTYLAVGS GFLSQPEKMM RGVLYIYQIR YMQNDEGFNE ITLRPLYNET
1021     NKIYKNPIIE ITDNSGYMAI FCGNLLYLMR FFNENTVKIE AFLVGRFFAS SIVSLKNYLL
1081     YADSYEGFEV ARWRKYGKKL ISMARDTMTK LPLSAAFLQY EDCLGGVVFDD DGNNAHIFDV
1141     DEYAIPADAV VRKSIFYIGG RAISSGQFPI KAVTQATQQN PNEEIDEELL QLQTKIGGHI
1201     AWYVTTGKI GAFTPIDEND RHKLVGVQSA YEKSLCGLSH LEYRSGKFKN MIEQDIFNQS
1261     PKNVIDCDML IDLIEDMPDH LKFATKGLRT QDFLSELRKI YNNGVNVFQ*

```

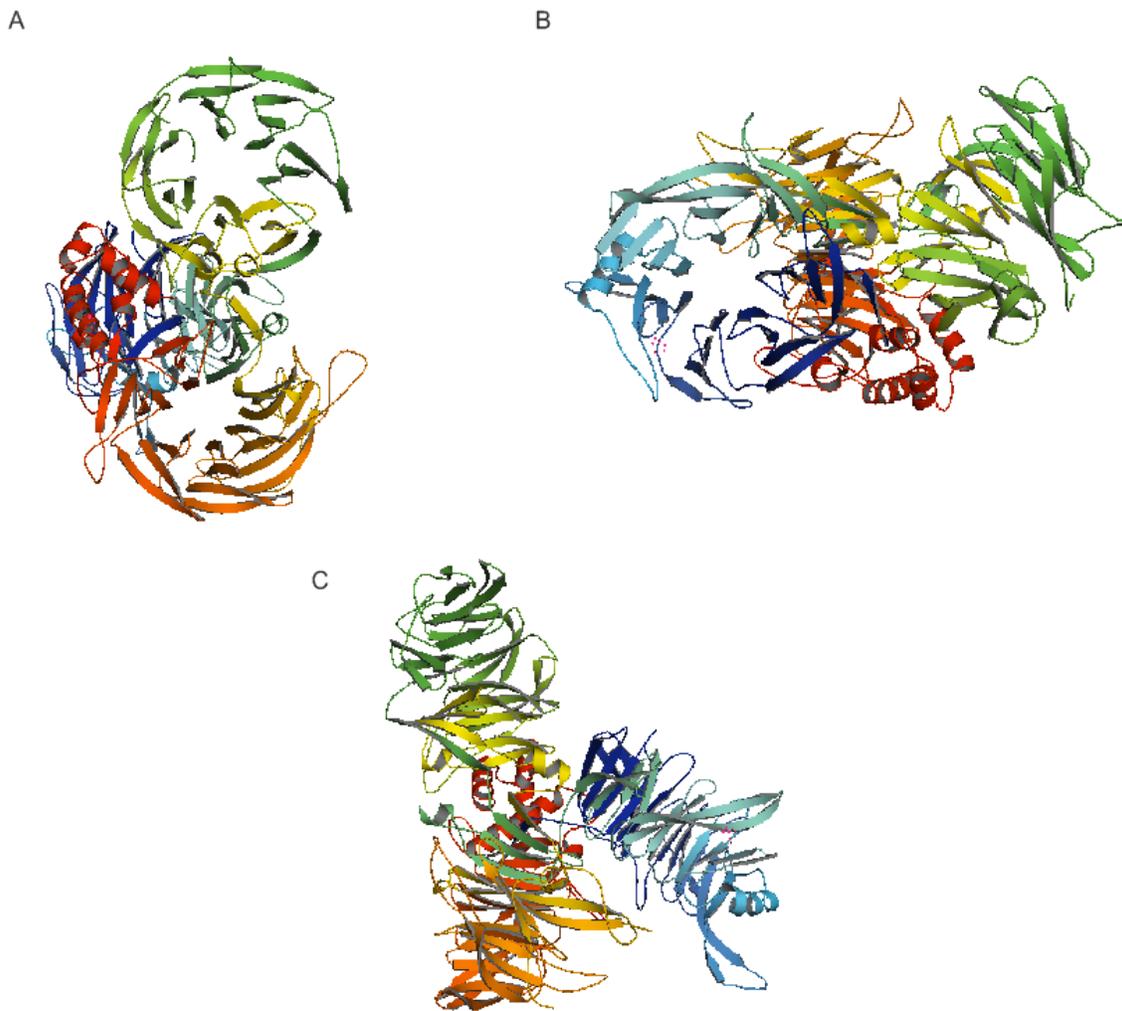
**Fig. 21. Secuencia proteica de TvCPSF-160.** Secuencia proteica deducida a partir de la secuencia de DNA del plásmido secuenciado 4pQE30CPSF160. En color verde se observa el dominio conservado de SFT1, en amarillo se observa el dominio conservado de CPSF160.

La secuencia proteica deducida de TvCPSF-160 (Fig. 21) se comparó con la secuencia proteica accesada en el genoma de *T. vaginalis* (TVAG\_077030), se encontraron 7 diferencias en aminoácidos. En la figura 24 se muestra el alineamiento de secuencias, la línea de arriba representa la secuencia proteica de TVA\_077030 y la línea de abajo representa la secuencia del gen TvCPSF-160. En la posición 143 se presenta un cambio de Fenilalanina (F) por Leucina (L), ambos son aminoácidos no polares. En la posición 245 se presenta un cambio de Asparagina (N) por Treonina (T) los dos son aminoácidos polares. En la posición 364 hay un cambio de Prolina (P) por Serina (S), en este caso se presenta un cambio de aminoácido no polar por uno polar. En la posición 590 hay un cambio de Serina (S) por Fenilalanina (F) esto representa otro cambio en la polaridad de un aminoácido polar por un no polar. Las posiciones 656 y 794 no presentan cambios en la polaridad hay un cambio de Treonina (T) por Asparagina (N) en ambos casos. La posición 801 muestra un cambio de Fenilalanina (F) por Alanina (A), esta modificación tampoco representa cambios en la polaridad (Fig. 22). Estos cambios pueden deberse a que la cepa utilizada en el proyecto de secuenciación del genoma es diferente a la cepa que se utiliza en nuestro laboratorio y con la que se desarrolló este trabajo.





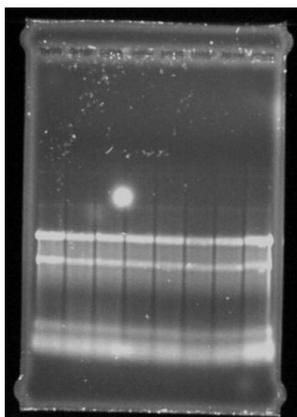
Se realizó un modelado *in silico* de la secuencia de TvCPSF-160, tomando como molde los primeros 1297 aminoácidos de la secuencia de TvCPSF-160. Se utilizó la herramienta bio-informática en línea denominada Protein structure prediction Phyre2, (Kelley L. A. Sternberg M. J. E. 2009). Se obtuvo una precisión de 82%, en el modelo es posible observar 3 estructuras probables de  $\beta$ -propelas o  $\beta$ -barriles, estas estructuras han sido reportados como elementos en la secuencia de Yhh1p/Cft1p, proteína homóloga de CPSF-160 en levaduras (Dichtl B. y col. 2002) (Fig. 23).



**Fig. 23. Modelo estructural teórico de la proteína TvCPSF-160 de *T. vaginalis*.** (A) En color verde y naranja es posible observar dos estructuras de tipo  $\beta$ -propelas. (B) En color azul se observa la tercera estructura  $\beta$ -propela. (C) Distribución espacial de las estructuras  $\beta$ -propelas.

## 7.7 NORTHERN BLOT.

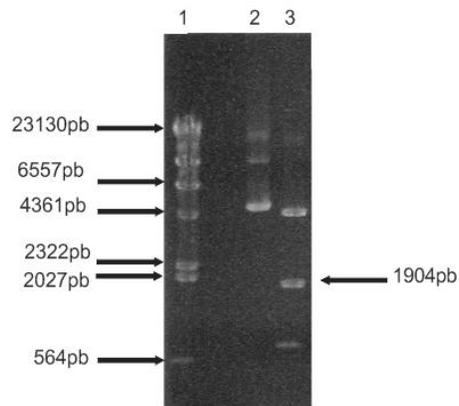
Para saber si el gen TVAG\_077030 clonado se expresaba en *T. vaginalis* se buscó la presencia del mRNA correspondiente mediante un análisis de northern blot. Se extrajo RNA total a partir de cultivos de *T.vaginalis*, se midió la absorbancia de las muestras a 260nm. La concentración de la muestra fue de 11.48µg/µl y se visualizó en un gel de agarosa H<sub>2</sub>O DEPC, se observó que el RNA se encontraba íntegro (Fig. 24).



**Fig. 24. Gel de agarosa al 0.8%.** Muestras de RNA total, extraído a partir de cultivos de *T. vaginalis*.

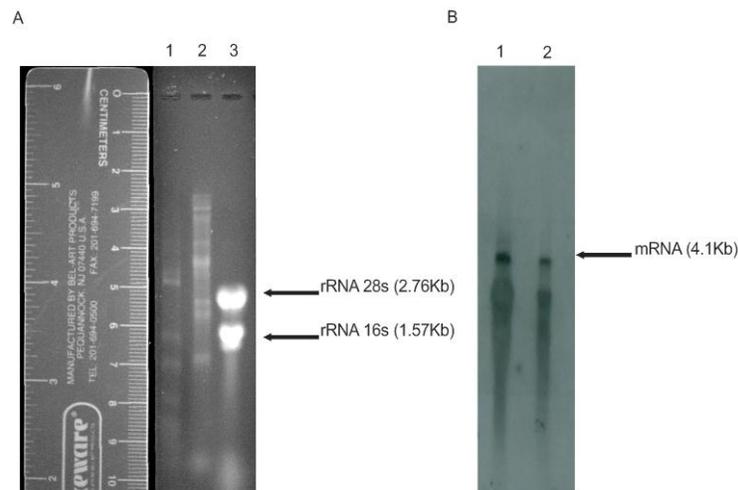
Para el northern se utilizó también mRNA, se realizó la extracción de RNA poly(A+) mediante el kit comercial Quickprep microRNA purification (Amersham Pharmacia Biotech, 2011). Se midió la absorbancia de la muestra a 260nm, se obtuvo una concentración de 0.6µg/µl (Fig. 24).

Se preparó una sonda radiactiva con un fragmento (1905pb) del plásmido pQE30CPSF160, como se describe en materiales y métodos (Fig. 25). La sonda radiactiva se preparó con el kit comercial Rediprime II Random Prime Labelling System (PROMEGA, 2011), como se describe en materiales y métodos. La sonda se hibridizó con la membrana del northern blot durante 39Hrs (2 Cot  $\frac{1}{2}$ ) (Carletti E. y col. 2006) a 42°C. La membrana se expuso a una placa fotosensible.



**Fig. 25. Gel de agarosa 0.8% TBE.** Se observa un fragmento de DNA de 1904pb utilizado para hacer la sonda radiactiva. El fragmento se obtuvo del plásmido pQE30CPSF160 digerido con *HidIII*.

Se logró observar una banda del tamaño esperado (4.1Kb), tanto en la muestra de RNA total, como en la muestra de RNA poly(A+); esto sugiere que el mRNA del gen clonado se expresa (Fig. 26). Se observan otras dos bandas de menor tamaño, esto puede ser un pegado inespecífico a los transcritos ribosomales debido a la baja astringencia con la que se lavó la membrana.

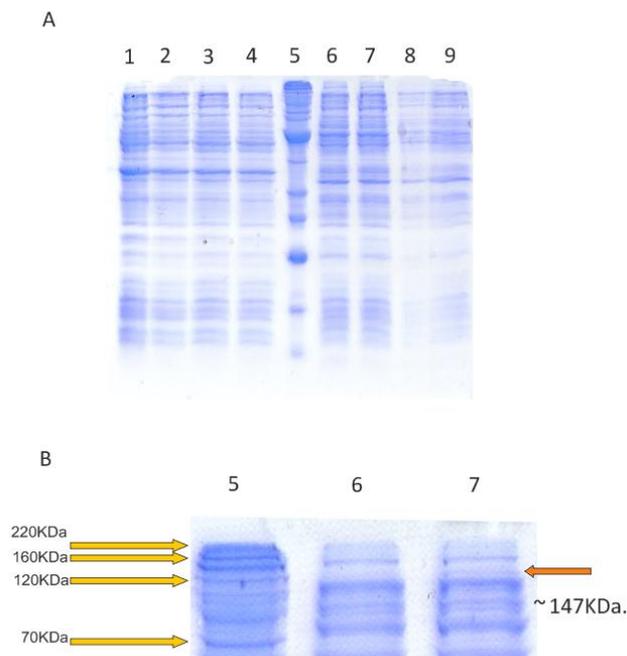


**Fig. 26. TvCPSF-160 se expresa como un transcrito de 4.1 Kb.** A) Gel de agarosa 0.8% con formaldehído utilizado para el análisis por northern blot; se muestran los marcadores de peso molecular (carriles 1 y 2) y el RNA total de *T. vaginalis* (carril 3). B) Autorradiografía del ensayo de

northern blot. Carril 1: RNA total de *T. vaginalis*. Carril 2: RNA poly(A+). La membrana se expuso durante 72Hrs. La flecha indica la banda esperada para el mRNA de TvCPSF-16, cuyo peso molecular se calculó de 4.1Kb

### 7.8 EXPRESIÓN DE LA PROTEÍNA TVCPSF-160 A PARTIR DEL VECTOR pQE30CPSF160.

Se realizaron ensayos de expresión para la proteína TVCPSF-160 utilizando el plásmido pQE30CPSF160. El vector se transformó en dos cepas de *E. coli*, XL1Blue y BL21Star. En un primer ensayo, se realizó un cultivo de ambas cepas en medio LB (ampicilina 100µg/ml), cada uno de los cultivos se incubó a 37°C y a temperatura ambiente. La inducción se llevó a cabo con IPTG 1mM cuando los cultivos alcanzaron una DO<sub>600</sub> 0.5, durante 4Hrs. Se esperaba observar una banda de aproximadamente 147Kda.



**Fig. 27. Análisis de la expresión de la proteína TvCPSF-160 por SDS-PAGE en poliacrilamida al 10% (Todas las células fueron transformadas con el vector pQE30CPSF160). Carriles 1,2: proteínas totales del cultivo BL21Star crecido a 37°C sin inducir e inducido respectivamente. Carriles 3,4: proteínas totales del cultivo BL21Star crecido a Temperatura Ambiente sin inducir e**

inducido respectivamente. Carriles 6,7: proteínas del cultivo XL1Blue crecido a 37°C sin inducir e inducido respectivamente. Carriles 8,9: proteínas totales del cultivo XL1Blue crecido a Temperatura Ambiente sin inducir e inducido respectivamente. Carril 5: se observa el marcador de peso molecular. **(B)** Imagen ampliada de los carriles 5, 6, 7, donde se observa una banda con el peso molecular esperado (147KDa), lo que puede ser una probable expresión de la proteína CPSF-160.

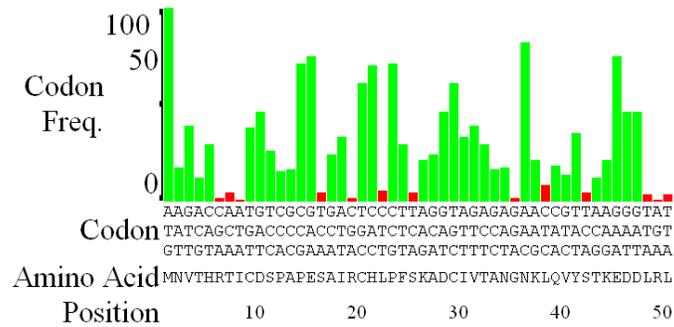
Se realizó un análisis SDS-PAGE en geles de poliacrilamida 10%. Fue posible observar una banda del tamaño esperado (147KDa) aunque muy tenue en el cultivo de XL1Blue crecido a 37°C (Fig. 27). La banda no se observa en el cultivo control, crecido en las mismas condiciones pero sin inducir con IPTG. Se realizaron varios ensayos, tratando de encontrar condiciones óptimas de expresión, se modificó tanto la temperatura de incubación como el tiempo de inducción, así como la concentración de poliacrilamida, para mejorar la definición de los geles, sin embargo, no se lograron encontrar condiciones en las que se observara una expresión más clara de la banda esperada.

## 7.9 ANÁLISIS DE USO DE CODONES.

Se determinó el uso de codones para las primeras 750pb de la secuencia codificante del DNA de CPSF-160, con la finalidad de compararlo con el uso de codones de *E. coli*, ya que las cepas XL1Blue y BL21Star son cepas de *E. coli*. Se utilizó la herramienta *E. coli* Codon Usage Analyzer 2.1, disponible online. La base de datos de este programa es de genes de clase II de análisis y predicciones en *E. coli* y *Salmonella* (Hénaut A. Danchin A. 1996).

Los resultados muestran que en la secuencia de *T. vaginalis* existen codones que son poco frecuentes en *E. coli*, es decir, que *E. coli* produce pocos tRNA para dichos codones. En la figura se puede observar la frecuencia en el uso de codones para los primeros 50 residuos de aminoácidos, codones como CGA, ACA, ATA, TCA, CTT, GGA, CTC, TTA y AGA presentan una frecuencia menor del 10% (Fig. 28). Es decir, del total de tRNAs producidos por *E. coli*, menos del 10% corresponde a los tRNAs requeridos para la traducción de los mRNAs de *T. vaginalis*. Por lo tanto, la poca

expresión de la proteína recombinante CPSF-160 de *T. vaginalis* podría deberse a que la maquinaria traduccional de *E. coli* es poco eficiente para expresar esta proteína.



**Fig. 28. Análisis de frecuencia en el uso de codones de las primeras 150pb de la secuencia de CPSF-160 de *T. vaginalis*, relativo al uso de codones de *E. coli*.**

## 8 DISCUSIÓN.

---

La expresión de genes codificadores de proteínas en las células de eucariontes implica múltiples procesos de transcripcionales y post-transcripcionales, los cuales, se reconoce cada vez con mayor frecuencia, están interconectados (Maniatis y Reed, 2002 ; Moore y Proudfoot, 2009). El proceso de poliadenilación, es decir, el corte del extremo 3' y el evento de poliadenilación es fundamental para la terminación de la transcripción (Zhao J. y col. 1999).

En 1991 se consiguió purificar el Factor específico de Corte y Poliadenilación (CPSF), a partir de células de timo de ternera y células HeLa (Bienroth S. 1991). Actualmente se sabe que el CPSF, en mamíferos, está conformado por 5 subunidades: CPSF-160, CPSF-30, CPSF-73, CPSF-100, y hFip1. Todas las subunidades son necesarias para la división eficiente y la poliadenilación del premRNA (Mandel C. y col. 2008).

Se ha conseguido purificar en modelos de *S. cerevisiae*, una proteína homóloga al CPSF-160 de mamíferos. Un componente del Factor de Corte CFII (Cleavage Factor II), designado CFT1 fue identificado por su secuencia similar a la subunidad CPSF de mamíferos (Stumpf G. Domdey H. 1996). El CPF de *S. cerevisiae* posee la subunidad Cft1p/Yhh1p que conserva un 24% de identidad y un 51% de similitud con la subunidad CPSF-160 de mamíferos (Dantonel J. y col. 1997; Dichtl B. y col. 2002).

La hipótesis de este trabajo es que *T. vaginalis* posee una proteína de tipo CPSF-160, homóloga al CPSF-160 de mamíferos y al Yhh1p /Cft1p de levaduras.

En este trabajo se logró identificar, clonar y secuenciar un gen que codifica para una proteína homóloga al CPSF-160 de mamíferos y al Cft1p/Yhh1p de levaduras. El TvCPSF-160 de *T. vaginalis* es homólogo al CPSF-160 reportado para mamíferos y al Cft1p/Yhh1p de levaduras, esto se demostró mediante análisis bioinformáticos, comparando las secuencias de DNA y las secuencias protéicas.

Se logró también demostrar la expresión de un mRNA en *T. vaginalis*, cuya secuencia es homóloga al DNA que codifica para la proteína tipo CPSF-160 de *T. vaginalis*. Sin embargo, no se logró

expresar en *E. coli* de manera clara a la proteína CPSF-160 según los resultados de los ensayos de SDS-PAGE. Para tratar de definir si la dificultad en la expresión estaba relacionada al uso de codones, se determinó el uso relativo de codones en *T. vaginalis* y *E. coli*. Resultó interesante el hecho de que muchos de los codones de la secuencia del CPSF-160 de *T. vaginalis*, son usados con poca frecuencia por *E. coli* (Hénaut A. Danchin A. 1996).

Finalmente se realizó un modelado de la secuencia proteica del CPSF-160 de *T. vaginalis*. La secuencia conserva algunos dominios que resultan importantes para la función teórica de la proteína, entre ellos los dominios de procesamiento y reconocimiento de RNA SFT1, y el dominio CPSF A, homólogo a las subunidades CPSF-160 reportadas; también se observaron similitudes estructurales, tal como las  $\beta$ -propelas, reportadas para otras proteínas de reconocimiento de RNA (Andreas J. Keller W. 1995; Salinas C. A y col. 1998; Li Y. y col. 2001; Dichtl, B. y col. 2002; Ivens A. C. y col. 2005). De esta forma es que el modelo obtenido comparte características tanto en secuencia como en estructura respecto a los modelos de CPSF-160 conocidos.

## **9. CONCLUSIONES Y PERSPECTIVAS.**

---

### **CONCLUSIONES.**

1. Se clonó un gen de *T. vaginalis* homólogo a los reportados para CPSF-160 de mamíferos.
2. Se obtuvo una secuencia consenso para CPSF-160 a partir del análisis del genoma accesado de *T. vaginalis*.
3. Se clonó el gen CPSF-160 de *T. vaginalis* en el vector pCR-Blunt II TOPO.
4. Se sub-clonó el gen CPSF-160 en el vector de expresión pQE30.
5. Se secuenció el gen CPSF-160 de *T. vaginalis*.
6. Se observó la expresión de un mRNA de un tamaño aproximado al esperado para un gen que codifica CPSF-160.
7. Se obtuvo un modelo teórico de la proteína a partir de la secuencia del gen CPSF-160 de *T. vaginalis*, que mostró similitudes estructurales y de secuencia, respecto a las proteínas CPSF-160 reportadas.
8. No se consiguió expresar la proteína CPSF-160, utilizando las cepas de expresión de *E. coli* (M15, XL1Blue, BL21Star y JM109).

### **PERSPECTIVAS.**

Para conseguir la expresión de la proteína CPSF-160, se recomienda utilizar la cepa comercial de expresión de *E. coli* denominada ROSSETA, que cuenta con plásmidos que sobre expresan los tRNAs que normalmente son poco abundantes en *E. coli*. Por otra parte, considerando que la proteína CPSF-160 tiene un tamaño grande (147KDa), se recomienda intentar expresar la proteína utilizando secciones de ésta y no la proteína completa.

Para aumentar la sensibilidad de los ensayos de expresión se deben realizar ensayos tipo western blot, utilizando la marca de histidinas (6His tag) que tiene la proteína. De esta manera se puede saber si la proteína se está expresando, pero lo hace en bajas cantidades.

Con la proteína recombinante purificada, mediante el tag de histidinas, será posible realizar ensayos de retardo para conocer el tipo de interacción de esta proteína con la señal de poliadenilación propuesta para *T. vaginalis*, y la interacción con algunos otros factores proteicos.

## 10 REFERENCIAS BIBLIOGRÁFICAS.

---

1. Adl SM, Simpson AG, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S, James TY, Karpov S, Kugrens P, Krug J, Lane CE, Lewis LA, Lodge J, Lynn DH, Mann DG, McCourt RM, Mendoza L, Moestrup O, Mozley-Standridge SE, Nerad TA, Shearer CA, Smirnov AV, Spiegel FW, Taylor MF. 2005. **The new higher level classification of eukaryotes with emphasis on the taxonomy of protists.** *J Eukaryot Microbiol.* 2005 Sep-Oct; 52(5):399-451.
2. Amrani N, Ghosh S, Mangus DA, Jacobson A. 2008. **Translation factors promote the formation of two states of the closed-loop mRNP.** *Nature.* 2008 Jun 26; 453(7199):1276-80.
3. Aravind L. 1999. **An evolutionary classification of the metallo-beta-lactamase fold proteins.** *In Silico Biol.* 1999; 1(2):69-91.
4. Arroyo R. 2000. **Biología molecular de la Trichomonosis.** *Genética y Biomedicina Molecular.* Edit. Limusa, Noriega. Cap. 19. P.p. 287-304.
5. Arroyo R, González-Robles A, Martínez-Palomo A, Alderete JF. 1993. **Signalling of *Trichomonas vaginalis* for amoeboid transformation and adhesion synthesis follows cytoadherence.** *Mol Microbiol.* 1993 Jan; 7(2):299-309.
6. Aurrecochea C, Brestelli J, Brunk BP, Carlton JM, Dommer J, Fischer S, Gajria B, Gao X, Gingle A, Grant G, Harb OS, Heiges M, Innamorato F, Iodice J, Kissinger JC, Kraemer E, Li W, Miller JA, Morrison HG, Nayak V, Pennington C, Pinney DF, Roos DS, Ross C, Stoeckert CJ Jr, Sullivan S, Treatman C, Wang H. 2009. **GiardiaDB and TrichDB: integrated genomic resources for the eukaryotic protist pathogens *Giardia lamblia* and *Trichomonas vaginalis*.** *Nucleic Acids Res.* 2009 Jan; 37(Database issue):D526-30.
7. Barabino SM, Hübner W, Jenny A, Minvielle-Sebastia L, Keller W. 1997. **The 30-kD subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are RNA-binding zinc finger proteins.** *Genes Dev.* 1997 Jul 1; 11(13):1703-16.
8. Barabino SM, Ohnacker M, Keller W. 2000. **Distinct roles of two Yth1p domains in 3'-end cleavage and polyadenylation of yeast pre-mRNAs.** *EMBO J.* 2000 Jul 17; 19(14):3778-87.
9. Baroin A, Perasso R, Qu LH, Brugerolle G, Bachellerie JP, Adoutte A. 1988. **Partial phylogeny of the unicellular eukaryotes based on rapid sequencing of a portion of 28S ribosomal RNA.** *Proc Natl Acad Sci U S A.* 1988 May; 85(10):3474-8.
10. Barrera G. 2006. **Análisis funcional de las señales de poliadenilación en *Trichomonas vaginalis*.** *Tesis Posgrado en Ciencias Biológicas.* Universidad Nacional Autónoma de México.

11. Bienroth S, Wahle E, Suter-Crazzolara C, Keller W. 1991. **Purification of the cleavage and polyadenylation factor involved in the 3'-processing of messenger RNA precursors.** *J Biol Chem.* 1991 Oct 15; 266(29):19768-76.
12. Brown PH, Tiley LS, Cullen BR. 1991. **Effect of RNA secondary structure on polyadenylation site selection.** *Genes Dev.* 1991 Jul; 5(7):1277-84.
13. Callebaut I, Moshous D, Mornon JP, de Villartay JP. 2002. **Metallo-beta-lactamase fold within nucleic acids processing enzymes: the beta-CASP family.** *Nucleic Acids Res.* 2002 Aug 15; 30(16):3592-601.
14. Calvo O, Manley JL. 2003. **Strange bedfellows: polyadenylation factors at the promoter.** *Genes Dev.* 2003 Jun 1; 17(11):1321-7.
15. Carletti E, Guerra E, Alberti S. 2006. **The forgotten variables of DNA array hybridization.** *Trends Biotechnol.* 2006 Oct; 24(10):443-8.
16. Carlton JM, Hirt RP, Silva JC, Delcher AL, Schatz M, Zhao Q, Wortman JR, Bidwell SL, Alsmark UC, Besteiro S, Sicheritz-Ponten T, Noel CJ, Dacks JB, Foster PG, Simillion C, Van de Peer Y, Miranda-Saavedra D, Barton GJ, Westrop GD, Müller S, Dessi D, Fiori PL, Ren Q, Paulsen I, Zhang H, Bastida-Corcuera FD, Simoes-Barbosa A, Brown MT, Hayes RD, Mukherjee M, Okumura CY, Schneider R, Smith AJ, Vanacova S, Villalvazo M, Haas BJ, Pertea M, Feldblyum TV, Utterback TR, Shu CL, Osoegawa K, de Jong PJ, Hrdy I, Horvathova L, Zubacova Z, Dolezal P, Malik SB, Logsdon JM Jr, Henze K, Gupta A, Wang CC, Dunne RL, Upcroft JA, Upcroft P, White O, Salzberg SL, Tang P, Chiu CH, Lee YS, Embley TM, Coombs GH, Mottram JC, Tachezy J, Fraser-Liggett CM, Johnson PJ. 2009. **Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*.** *Science.* 2007 Jan 12; 315(5809):207-12.
17. **Centro Nacional de Vigilancia Epidemiológica y Control de Enfermedades (CENAVECE).** 2009. Secretaria de Salud. *Boletín Epidemiología.* 35 vol. 21 Semana 35.
18. Chekanova JA, Belostotsky DA. 2006. **MicroRNAs and messenger RNA turnover.** *Methods Mol Biol.* 2006; 342:73-85.
19. Chen F, MacDonald CC, Wilusz J. 1995. **Cleavage site determinants in the mammalian polyadenylation signal.** *Nucleic Acids Res.* 1995 Jul 25; 23(14):2614-20.
20. Chou ZF, Chen F, Wilusz J. 1994. **Sequence and position requirements for uridylate-rich downstream elements of polyadenylation signals.** *Nucleic Acids Res.* 1994 Jul 11; 22(13):2525-31.
21. Colgan DF, Manley JL. 1997. **Mechanism and regulation of mRNA polyadenylation.** *Genes Dev.* 1997 Nov 1; 11(21):2755-66.
22. Collier JM, Gray NK, Wickens MP. 1998. **mRNA stabilization by poly(A) binding protein is independent of poly(A) and requires translation.** *Genes Dev.* 1998 Oct 15; 12(20):3226-35.

23. 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.* 1997 Jul; 24(6):353-60.
24. Crittenden SL, Bernstein DS, Bachorik JL, Thompson BE, Gallegos M, Petcherski AG, Moulder G, Barstead R, Wickens M, Kimble J. 2002. **A conserved RNA-binding protein controls germline stem cells in *Caenorhabditis elegans*.** *Nature.* 2002 Jun 6;417(6889):660-3.
25. Dai W, Zhang G, Makeyev EV. 2011. **RNA-binding protein HuR autoregulates its expression by promoting alternative polyadenylation site usage.** *Nucleic Acids Res.* 2011 Sep 24. [Epub ahead of print]
26. Dantoni JC, Murthy KG, Manley JL, Tora L. 1997. **Transcription factor TFIID recruits factor CPSF for formation of 3' end of mRNA.** *Nature.* 1997 Sep 25; 389(6649):399-402.
27. Dichtl B, Keller W. 2001. **Recognition of polyadenylation sites in yeast pre-mRNAs by cleavage and polyadenylation factor.** *EMBO J.* 2001 Jun 15; 20(12):3197-209.
28. Dichtl B, Blank D, Sadowski M, Hübner W, Weiser S, Keller W. 2002. **Yhh1p/Cft1p directly links poly(A) site recognition and RNA polymerase II transcription termination.** *EMBO J.* 2002 Aug 1; 21(15):4125-35.
29. Dominski Z, Marzluff WF. 2007. **Formation of the 3' end of histone mRNA: getting closer to the end.** *Gene.* 2007 Jul 15; 396(2):373-90.
30. Eisenmann, D. M. 2005. **WormBook.** The *C. elegans Research Community.* WormBook. <http://www.wormbook.org>.
31. Espinosa N, Hernández R, López-Griego L, López-Villaseñor I. 2002. **Separable putative polyadenylation and cleavage motifs in *Trichomonas vaginalis* mRNAs.** *Gene.* 2002 May 1; 289(1-2):81-6.
32. Ford LP, Bagga PS, Wilusz J. 1997. **The poly(A) tail inhibits the assembly of a 3'-to-5' exonuclease in an in vitro RNA stability system.** *Mol Cell Biol.* 1997 Jan; 17(1):398-406.
33. Fuentes V. 2005. **Análisis preliminar del procesamiento de transcritos en *Trichomonas vaginalis*.** Tesis para obtener el grado de Química Farmacéutica Bióloga. México D. F. Universidad Nacional Autónoma de México. Facultad de Química.
34. Fuentes V, Barrera G, Sánchez J, Hernández R, López-Villaseñor I. 2012. **Functional Analysis of Sequence Motifs Involved in the Polyadenylation of *Trichomonas vaginalis* mRNAs.** *Eukaryot Cell.* 2012 Mar 30.

35. Gil A, Proudfoot NJ. 1987. **Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit beta-globin mRNA 3' end formation.** *Cell*. 1987 May 8; 49(3):399-406.
36. Gómez-Conde E, Mena-López R, Hernández-Jaúregui P, González-Camacho M, Arroyo R. 2000. ***Trichomonas vaginalis*: chromatin and mitotic spindle during mitosis.** *Exp Parasitol*. 2000 Nov; 96(3):130-8.
37. Gross S, Moore CL. 2001. **Rna15 interaction with the A-rich yeast polyadenylation signal is an essential step in mRNA 3'-end formation.** *Mol Cell Biol*. 2001 Dec; 21(23):8045-55.
38. Hammell CM, Gross S, Zenklusen D, Heath CV, Stutz F, Moore C, Cole CN. 2002. **Coupling of termination, 3' processing, and mRNA export.** *Mol Cell Biol*. 2002 Sep; 22(18):6441-57.
39. Hénaut A, Danchin A. 1996. **Analysis and Predictions from Escherichia coli sequences in: Escherichia coli and Salmonella.** Vol. 2, Ch. 114:2047-2066, 1996, Neidhardt FC ed., ASM press, Washington, D.C.
40. Hernandez P. A. 2010. **Clonación y caracterización de una serín proteasa en Trichomonas vaginalis.** Tesis para obtener el grado de Doctor en Ciencias. México D. F. Universidad Nacional Autónoma de México. Instituto de Investigaciones Biomédicas.
41. Hu J, Lutz CS, Wilusz J, Tian B. 2005. **Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation.** *RNA*. 2005 Oct; 11(10):1485-93.
42. Huang Y, Carmichael GG. 1996. **Role of polyadenylation in nucleocytoplasmic transport of mRNA.** *Mol Cell Biol*. 1996 Apr; 16(4):1534-42.
43. Huang Y, Wimler KM, Carmichael GG. 1999. **Intronless mRNA transport elements may affect multiple steps of pre-mRNA processing.** *EMBO J*. 1999 Mar 15;18(6):1642-52.
44. Irniger S, Braus GH. 1994. **Saturation mutagenesis of a polyadenylation signal reveals a hexanucleotide element essential for mRNA 3' end formation in Saccharomyces cerevisiae.** *Proc Natl Acad Sci U S A*. 1994 Jan 4; 91(1):257-61.
45. Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, Sisk E, Rajandream MA, Adlem E, Aert R, Anupama A, Apostolou Z, Attipoe P, Bason N, Bauser C, Beck A, Beverley SM, Bianchetin G, Borzym K, Bothe G, Bruschi CV, Collins M, Cadag E, Ciarloni L, Clayton C, Coulson RM, Cronin A, Cruz AK, Davies RM, De Gaudenzi J, Dobson DE, Duesterhoeft A, Fazelina G, Fosker N, Frasch AC, Fraser A, Fuchs M, Gabel C, Goble A, Goffeau A, Harris D, Hertz-Fowler C, Hilbert H, Horn D, Huang Y, Klages S, Knights A, Kube M, Larke N, Litvin L, Lord A, Louie T, Marra M, Masuy D, Matthews K, Michaeli S, Mottram JC, Müller-Auer S, Munden H, Nelson S, Norbertczak H, Oliver K, O'neil S, Pentony M, Pohl TM, Price C, Purnelle B, Quail MA, Rabinowitsch E, Reinhardt R, Rieger M, Rinta J, Robben J, Robertson L, Ruiz JC, Rutter S, Saunders D, Schäfer M, Schein J, Schwartz DC, Seeger K, Seyler A, Sharp S, Shin H, Sivam D, Squares R, Squares S, Tosato V, Vogt C, Volckaert G, Wambutt R, Warren T, Wedler H, Woodward J, Zhou S, Zimmermann W, Smith DF,

- Blackwell JM, Stuart KD, Barrell B, Myler PJ. 2005. **The genome of the kinetoplastid parasite, *Leishmania major***. *Science*. 2005 Jul 15; 309(5733):436-42.
46. Jenny A, Keller W. 1995. **Cloning of cDNAs encoding the 160 kDa subunit of the bovine cleavage and polyadenylation specificity factor**. *Nucleic Acids Res*. 1995 Jul 25;23(14):2629-35.
47. Johnston VJ, Mabey DC. 2008. **Global epidemiology and control of *Trichomonas vaginalis***. *Curr Opin Infect Dis*. 2008 Feb; 21(1):56-64.
48. Juge F, Audibert A, Benoit B, Simonelig M. 2000. **Tissue-specific autoregulation of *Drosophila* suppressor of forked by alternative poly(A) site utilization leads to accumulation of the suppressor of forked protein in mitotically active cells**. *RNA*. 2000 Nov; 6(11):1529-38.
49. Keeling PJ, Palmer JD. 2000. **Parabasalian flagellates are ancient eukaryotes**. *Nature*. 2000 Jun 8; 405(6787):635-7.
50. Keller W, Minvielle-Sebastia L. 1997. **A comparison of mammalian and yeast pre-mRNA 3'-end processing**. *Curr Opin Cell Biol*. 1997 Jun; 9(3):329-36.
51. Kelley LA and Sternberg MJE. 2009. **Protein structure prediction on the web: a case study using the Phyre server**. *Nature Protocols* 4, 363 - 371 (2009)
52. Kerwitz Y, Kühn U, Lilie H, Knoth A, Scheuermann T, Friedrich H, Schwarz E, Wahle E. 2003. **Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA**. *EMBO J*. 2003 Jul 15;22(14):3705-14.
53. Kessler MM, Henry MF, Shen E, Zhao J, Gross S, Silver PA, Moore CL. 1997. **Hrp1, a sequence-specific RNA-binding protein that shuttles between the nucleus and the cytoplasm, is required for mRNA 3'-end formation in yeast**. *Genes Dev*. 1997 Oct 1; 11(19):2545-56.
54. Kleina P, Bettim-Bandinelli J, Bonatto SL, Benchimol M, Bogo MR. 2004. **Molecular phylogeny of Trichomonadidae family inferred from ITS-1, 5.8S rRNA and ITS-2 sequences**. *Int J Parasitol*. 2004 Jul; 34(8):963-70.
55. Kühn U, Gündel M, Knoth A, Kerwitz Y, Rüdell S, Wahle E. 2009. **Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor**. *J Biol Chem*. 2009 Aug 21; 284(34):22803-14.
56. Kwak JE, Drier E, Barbee SA, Ramaswami M, Yin JC, Wickens M. 2008. **GLD2 poly(A) polymerase is required for long-term memory**. *Proc Natl Acad Sci U S A*. 2008 Sep 23; 105(38):14644-9.

57. Lahti CJ, d'Oliveira CE, Johnson PJ. 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.* 1992 Nov; 174(21):6822-30.
58. Le Hir H, Nott A, Moore MJ. 2003. **How introns influence and enhance eukaryotic gene expression.** *Trends Biochem Sci.* 2003 Apr; 28(4):215-20.
59. Lau AO, Smith AJ, Brown MT, Johnson PJ. 2006. ***Trichomonas vaginalis* initiator binding protein (IBP39) and RNA polymerase II large subunit carboxy terminal domain interaction.** *Mol Biochem Parasitol.* 2006 Nov;150(1):56-62.
60. Lehker MW, Alderete JF. 1999. **Resolution of six chromosomes of *Trichomonas vaginalis* and conservation of size and number among isolates.** *J Parasitol.* 1999 Oct; 85(5):976-9.
61. Lehker MW, Alderete JF. 2000. **Biology of trichomonosis.** *Curr Opin Infect Dis.* 2000 Feb; 13(1):37-45.
62. Lewin B. 2004. **Genes VIII.** 6° ED. Pearson.
63. Lindmark DG, Müller M. 1973. **Hydrogenosome, a cytoplasmic organelle of the anaerobic flagellate *Tritrichomonas foetus*, and its role in pyruvate metabolism.** *J Biol Chem.* 1973 Nov 25; 248(22):7724-8.
64. Liston DR, Johnson PJ. 1998. **Gene Transcription in *Trichomonas vaginalis*.** *Parasitol Today.* 1998 Jul; 14(7):261-5.
65. Liston DR, Johnson PJ. 1999. **Analysis of a ubiquitous promoter element in a primitive eukaryote: early evolution of the initiator element.** *Mol Cell Biol.* 1999 Mar; 19(3):2380-8.
66. Li Y, Chen ZY, Wang W, Baker CC, Krug RM. 2001. **The 3'-end-processing factor CPSF is required for the splicing of single-intron pre-mRNAs in vivo.** *RNA.* 2001 Jun; 7(6):920-31.
67. Lodish H, Berk A, Matsudaira P, Kaiser C A, Krieger M, Scott M, P. Zipursky S, Darnell J. 2005. **Biología celular y molecular.** 5a Ed. Panamericana. 108,
68. Malik SB, Pightling AW, Stefaniak LM, Schurko AM, Logsdon JM Jr. **An expanded inventory of conserved meiotic genes provides evidence for sex in *Trichomonas vaginalis*.** *PLoS One.* 2007 Aug 6;3(8).
69. Mandel CR, Kaneko S, Zhang H, Gebauer D, Vethantham V, Manley JL, Tong L. 2006. **Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease.** *Nature.* 2006 Dec 14; 444(7121):953-6.
70. Mandel CR, Bai Y, Tong L. 2008. **Protein factors in pre-mRNA 3'-end processing.** *Cell Mol Life Sci.* 2008 Apr; 65(7-8):1099-122.
71. Maniatis T, Reed R. 2002. **An extensive network of coupling among gene expression machines.** *Nature.* 2002 Apr 4; 416(6880):499-506.

72. Ma Y, Pannicke U, Schwarz K, Lieber MR. 2002. **Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination.** *Cell*. 2002 Mar 22; 108(6):781-94.
73. McCracken S, Fong N, Yankulov K, Ballantyne S, Pan G, Greenblatt J, Patterson SD, Wickens M, Bentley DL. **The C-terminal domain of RNA polymerase II couples mRNA processing to transcription.** *Nature*. 1997 Jan 23; 385(6614):357-61.
74. McDevitt MA, Hart RP, Wong WW, Nevins JR. 1986. **Sequences capable of restoring poly(A) site function define two distinct downstream elements.** *EMBO J*. 1986 Nov; 5(11):2907-13.
75. MacDonald CC, Wilusz J, Shenk T. 1994. **The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location.** *Mol Cell Biol*. 1994 Oct; 14(10):6647-54.
76. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH. 2011. **CDD: a Conserved Domain Database for the functional annotation of proteins.** *Nucleic Acids Res*. 2011 Jan; 39(Database issue):D225-9.
77. Marchler-Bauer A, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Lu S, Marchler GH, Mullokandov M, Song JS, Tasneem A, Thanki N, Yamashita RA, Zhang D, Zhang N, Bryant SH. 2009. **CDD: specific functional annotation with the Conserved Domain Database.** *Nucleic Acids Res*. 2009 Jan; 37(Database issue):D205-10.
78. Marchler-Bauer A, Bryant SH. 2004. **CD-Search: protein domain annotations on the fly.** *Nucleic Acids Res*. 2004 Jul 1; 32(Web Server issue):W327-31.
79. McLauchlan J, Gaffney D, Whitton JL, Clements JB. 1985. **The consensus sequence YGTGTTY located downstream from the AATAAA signal is required for efficient formation of mRNA 3' termini.** *Nucleic Acids Res*. 1985 Feb 25; 13(4):1347-68.
80. Meysick KC, Garber GE. 1990. **Growth of *Trichomonas vaginalis* in a serum-free McCoy cell culture system.** *J Parasitol*. 1990 Dec; 76(6):926-8.
81. Meysick KC, Garber GE. 1992. **Interactions between *Trichomonas vaginalis* and vaginal flora in a mouse model.** *J Parasitol*. 1992 Feb; 78(1):157-60.
82. Moore MJ, Proudfoot NJ. 2009. **Pre-mRNA processing reaches back to transcription and ahead to translation.** *Cell*. 2009 Feb 20; 136(4):688-700.
83. Moreno-Morcillo M, Minvielle-Sébastien L, Mackereth C, Fribourg S. 2011. **Hexameric architecture of CstF supported by CstF-50 homodimerization domain structure.** *RNA*. 2011 Mar; 17(3):412-8.

84. Müller M. **The hydrogenosome.** *J Gen Microbiol.* 1993 Dec; 139(12):2879-89.
85. Nadler SA, Honigberg BM. 1988. **Genetic differentiation and biochemical polymorphism among trichomonads.** *J Parasitol.* 1988 Oct; 74(5):797-804.
86. Nevins JR, Darnell JE Jr. 1978. **Steps in the processing of Ad2 mRNA: poly(A)+ nuclear sequences are conserved and poly(A) addition precedes splicing.** *Cell.* 1978 Dec; 15(4):1477-93.
87. Olga C. 2003. **Strange bedfellows: polyadenylation factors at the promoter.** *Genes Dev.* 2003 17: 1321-1327.
88. Pérez-Cañadillas JM. 2006. **Grabbing the message: structural basis of mRNA 3'UTR recognition by Hrp1.** *EMBO J.* 2006 Jul 12; 25(13):3167-78.
89. Pereira-Neves A, Benchimol M. 2008. ***Trichomonas vaginalis*: in vitro survival in swimming pool water samples.** *Exp Parasitol.* 2008 Mar;118(3):438-41
90. Petrin D, Delgaty K, Bhatt R, Garber G. 1998. **Clinical and microbiological aspects of *Trichomonas vaginalis*.** *Clin Microbiol Rev.* 1998 Apr; 11(2):300-17.
91. Proudfoot NJ. 1989. **How RNA polymerase II terminates transcription in higher eukaryotes.** *Trends Biochem Sci.* 1989 Mar; 14(3):105-10.
92. Proudfoot NJ, Furger A, Dye MJ. 2002. **Integrating mRNA processing with transcription.** *Cell.* 2002 Feb 22; 108(4):501-12.
93. Proudfoot N. 2004. **New perspectives on connecting messenger RNA 3' end formation to transcription.** *Curr Opin Cell Biol.* 2004 Jun; 16(3):272-8.
94. 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.* 1994 May 10; 91(10):4579-83.
95. Quon DV, Delgadillo MG, Johnson PJ. 1996. **Transcription in the early diverging eukaryote *Trichomonas vaginalis*: an unusual RNA polymerase II and alpha-amanitin-resistant transcription of protein-coding genes.** *J Mol Evol.* 1996 Sep; 43(3):253-62.
96. Preiss T, Hentze MW. 1998. **Dual function of the messenger RNA cap structure in poly(A)-tail-promoted translation in yeast.** *Nature.* 1998 Apr 2; 392(6675):516-20.
97. Rao S, Dinkins RD, Hunt AG. 2009. **Distinctive interactions of the Arabidopsis homolog of the 30 kD subunit of the cleavage and polyadenylation specificity factor (AtCPSF30) with other polyadenylation factor subunits.** *BMC Cell Biol.* 2009 Jul 2; 10:51.
98. Richter JD, Klann E. 2007. **Making synaptic plasticity and memory last: mechanisms of translational regulation.** *Genes Dev.* 2009 Jan 1; 23(1):1-11.

99. Riley DE, Campbell LA, Puolakkainen M, Krieger JN. 1993. **Trichomonas vaginalis** and early evolving DNA and protein sequences of the CDC2/28 protein kinase family. *Mol Microbiol.* 1993 May; 8(3):517-9.
100. Roger AJ, Clark CG, Doolittle WF. 1996. A possible mitochondrial gene in the early-branching amitochondriate protist *Trichomonas vaginalis*. *Proc Natl Acad Sci USA.* 1996 Dec 10; 93(25):14618-22.
101. Ruepp MD, Schweingruber C, Kleinschmidt N, Schümperli D. 2011. **Interactions of CstF-64, CstF-77, and symplekin: implications on localisation and function.** *Mol Biol Cell.* 2011 Jan 1; 22(1):91-104.
102. Rügsegger U, Beyer K, Keller W. 1996. **Purification and characterization of human cleavage factor Im involved in the 3' end processing of messenger RNA precursors.** *J Biol Chem.* 1996 Mar 15; 271(11):6107-13.
103. Ryan K, Calvo O, Manley JL. 2004. **Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease.** *RNA.* 2004 Apr; 10(4):565-73.
104. Ryan K, Calvo O, Manley JL. 1989. **The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation.** *Cell.* 1989 Sep 8; 58(5):857-67.
105. Sachs AB, Sarnow P, Hentze MW. 1997. **Starting at the beginning, middle, and end: translation initiation in eukaryotes.** *Cell.* 1997 Jun 13; 89(6):831-8.
106. Schneider R, Agol VI, Andino R, Bayard F, Cavener DR, Chappell SA, Chen JJ, Darlix JL, Dasgupta A, Donzé O, Duncan R, Elroy-Stein O, Farabaugh PJ, Filipowicz W, Gale M Jr, Gehrke L, Goldman E, Groner Y, Harford JB, Hatzglou M, He B, Hellen CU, Hentze MW, Hershey J, Hershey P, Hohn T, Holcik M, Hunter CP, Igarashi K, Jackson R, Jagus R, Jefferson LS, Joshi B, Kaempfer R, Katze M, Kaufman RJ, Kiledjian M, Kimball SR, Kimchi A, Kirkegaard K, Koromilas AE, Krug RM, Kruijs V, Lamphear BJ, Lemon S, Lloyd RE, Maquat LE, Martinez-Salas E, Mathews MB, Mauro VP, Miyamoto S, Mohr I, Morris DR, Moss EG, Nakashima N, Palmenberg A, Parkin NT, Pe'ery T, Pelletier J, Peltz S, Pestova TV, Pilipenko EV, Prats AC, Racaniello V, Read GS, Rhoads RE, Richter JD, Rivera-Pomar R, Rouault T, Sachs A, Sarnow P, Scheper GC, Schiff L, Schoenberg DR, Semler BL, Siddiqui A, Skern T, Sonenberg N, Sossin W, Standart N, Tahara SM, Thomas AA, Toulmé JJ, Wilusz J, Wimmer E, Witherell G, Wormington M. **New ways of initiating translation in eukaryotes.** *Mol Cell Biol.* 2001 Dec; 21(23):8238-46.
107. Salinas CA, Sinclair DA, O'Hare K, Brock HW. 1998. **Characterization of a Drosophila homologue of the 160-kDa subunit of the cleavage and polyadenylation specificity factor CPSF.** *Mol Gen Genet.* 1998 Apr; 257(6):672-80.
108. Saitoh S, Chabes A, McDonald WH, Thelander L, Yates JR, Russell P. 2002. **Cid13 is a cytoplasmic poly(A) polymerase that regulates ribonucleotide reductase mRNA.** *Cell.* 2002 May 31; 109(5):563-73.

109. Sheets MD, Ogg SC, Wickens MP. 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.* 1990 Oct 11; 18(19):5799-805.
110. Simpson AG, Roger AJ. 2004. **The real 'kingdoms' of eukaryotes.** *Curr Biol.* 2004 Sep 7; 14(17):R693-6.
111. Sittler A, Gallinaro H, Jacob M. 1994. **Upstream and downstream cis-acting elements for cleavage at the L4 polyadenylation site of adenovirus-2.** *Nucleic Acids Res.* 1994 Jan 25; 22(2):222-31.
112. Smith AJ, Chudnovsky L, Simoes-Barbosa A, Delgadillo-Correa MG, Jonsson ZO, Wohlschlegel JA, Johnson PJ. 2011. **Novel core promoter elements and a cognate transcription factor in the divergent unicellular eukaryote *Trichomonas vaginalis*.** *Mol Cell Biol.* 2011 Apr;31(7):1444-58.
113. Sogin ML. 1991. **Early evolution and the origin of eukaryotes.** *Curr Opin Genet Dev.* 1991 Dec; 1(4):457-63.
114. Sorvillo F, Kerndt P. 1998. ***Trichomonas vaginalis* and amplification of HIV-1 transmission.** *Lancet.* 1998 Jan 17; 351(9097):213-4.
115. Stumpf G, Domdey H. 1996. **Dependence of yeast pre-mRNA 3'-end processing on CFT1: a sequence homolog of the mammalian AAUAAA binding factor.** *Science.* 1996 Nov 29; 274(5292):1517-20.
116. Takagaki Y, Ryner LC, Manley JL. 1988. **Separation and characterization of a poly(A) polymerase and a cleavage/specificity factor required for pre-mRNA polyadenylation.** *Cell.* 1988 Mar 11; 52(5):731-42.
117. Takagaki Y, Manley JL. 1994. **A polyadenylation factor subunit is the human homologue of the *Drosophila* suppressor of forked protein.** *Nature.* 1994 Dec 1; 372(6505):471-4.
118. Takagaki Y, Manley JL. 1997. **RNA recognition by the human polyadenylation factor CstF.** *Mol Cell Biol.* 1997 Jul; 17(7):3907-14.
119. Thompson SR, Goodwin EB, Wickens M. 2000. **Rapid deadenylation and Poly(A)-dependent translational repression mediated by the *Caenorhabditis elegans* tra-2 3' untranslated region in *Xenopus* embryos.** *Mol Cell Biol.* 2000 Mar; 20(6):2129-37.
120. 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.* 2003 Mar; 33(3):235-55.
121. Viikki M, Pukkala E, Nieminen P, Hakama M. 2000. **Gynaecological infections as risk determinants of subsequent cervical neoplasia.** *Acta Oncol.* 2000; 39(1):71-5.

122. Vinciguerra P, Stutz F. 2004. **mRNA export: an assembly line from genes to nuclear pores.** *Curr Opin Cell Biol.* 2004 Jun; 16(3):285-92.
123. Wahle E, Rügsegger U. 1999. **3'-End processing of pre-mRNA in eukaryotes.** *FEMS Microbiol Rev.* 1999 Jun; 23(3):277-95.
124. Wang AL, Wang CC. 1985. **Isolation and characterization of DNA from *Tritrichomonas foetus* and *Trichomonas vaginalis*.** *Mol Biochem Parasitol.* 1985 Mar; 14(3):323-35.
125. Watson J. D. Baker T. A. Bell S. P. Gann A. Levine M. Losick R. 2004. **Molecular Biology of the gene.** Fifth edition. Pearson.
126. Whitelaw E, Proudfoot N. 1986. **Alpha-thalassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3' end processing in the human alpha 2 globin gene.** *EMBO J.* 1986 Nov; 5(11):2915-22.
127. Wickens M, Anderson P, Jackson RJ. 1997. **Life and death in the cytoplasm: messages from the 3' end.** *Curr Opin Genet Dev.* 1997 Apr; 7(2):220-32.
128. Wilusz CJ, Wormington M, Peltz SW. 2001. **The cap-to-tail guide to mRNA turnover.** *Nat Rev Mol Cell Biol.* 2001 Apr; 2(4):237-46.
129. Wormington M, Searfoss AM, Hurney CA. 1996. **Overexpression of poly(A) binding protein prevents maturation-specific deadenylation and translational inactivation in *Xenopus oocytes*.** *EMBO J.* 1996 Feb 15; 15(4):900-9.
130. Yuh YS, Liu JY, Shaio MF. 1997. **Chromosome number of *Trichomonas vaginalis*.** *J Parasitol.* 1997 Jun; 83(3):551-3.
131. Zarkower D, Wickens M. 1988. **A functionally redundant downstream sequence in SV40 late pre-mRNA is required for mRNA 3'-end formation and for assembly of a precleavage complex in vitro.** *J Biol Chem.* 1988 Apr 25; 263(12):5780-8.
132. Zhao J, Hyman L, Moore C. 1999. **Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis.** *Microbiol Mol Biol Rev.* 1999 Jun; 63(2):405-45.
133. Zubáčová Z, Cimbůrek Z, Tachezy J. 2008. **Comparative analysis of trichomonad genome sizes and karyotypes.** *Mol Biochem Parasitol.* 2008 Sep;161(1):49-54.

## 11 APÉNDICES.

---

### 10.1 APÉNDICE A. SOLUCIONES.

Medio TYM.

Medio TYM (500ml)	
Peptona	10g
Extracto de levadura	5g
Maltosa	2.5g
L-cisteina	0.5g
Ac. Ascórbico	0.1g
KH <sub>2</sub> PO <sub>4</sub>	0.4g
K <sub>2</sub> HPO <sub>4</sub>	0.4g
H <sub>2</sub> O Bidestilada	Para 500ml
Ajustar pH 6.2 con HCl y esterilizar.	

Buffer PBS

PBS (1L)	
NaCl	8g
KCl	0.2g
Na <sub>2</sub> HPO <sub>4</sub>	1.44g
KH <sub>2</sub> PO <sub>4</sub>	0.24g
Ajustar pH 7.4 y esterilizar.	

Buffer TE.

TE (50ml)	
Tris pH 7.4 10mM	1ml
EDTA 0.2mM	20 $\mu$ l

SOLUCIONES PARA EXTRACCIÓN DE DNA PLASMÍDICO.

Solución I.

Glucosa	50mM
Tris-HCl pH8	25mM
EDTA pH8	10mM

Solución II.

NaOH	0.2N
SDS	1%

Solución III (100ml).

Acetato de potasio 5M	60ml
Ácido acético glacial	11.5ml
Agua bidestilada estéril	28.5ml

## SOLUCIONES PARA NORTHERN BLOT.

H<sub>2</sub>O DEPC.

En 4 litros de H<sub>2</sub>O bidestilada se agregaron 20 gotas de dietilpirocarbonato (DEPC) y se dejó en agitación toda la noche, al día siguiente se esterilizó el H<sub>2</sub>O antes de usarla.

Buffer MOPS/EDTA 10 X en H<sub>2</sub>O DEPC

- MOPS 0.5 M
- EDTA 0.01 M
- Ajustar a pH=7

SSC 20 X (sodium chloridal sodium citrate) en H<sub>2</sub>O DEPC

- 3M NaCl
- 0.3M Na<sub>3</sub>Citrato. 2H<sub>2</sub>O
- Ajustar a pH=7 com 1 M HCl

SSC 10X

- Preparar a partir de SSC 20X en H<sub>2</sub>O DEPC

Solución de Hibridación.

Concentración final.	Stock	Para 20ml
Tris 50mM pH7.4	1M	1ml
Denhardt's	100X	1ml
Formamida 50%	100%	10ml
SDS 1%	10%	2ml
SSC 5X	20X	5ml
H <sub>2</sub> O		1ml

Solución Desnaturalizante (500ml).

4.83g NaCl

25ml de NaOH 10N

Solución de Neutralización(500ml).

125ml de Tris-HCl 2M pH7.5

87.66g NaCl

## Soluciones para expresión y análisis SDS-PAGE.

### Gel inferior 10% (Separador)

- Acrilamida stock 3.5 mL
- Agua 4.5 mL
- Buffer inferior 2.7 mL
- Persulfato de amonio 10 % 60 uL
- TEMED 10 uL

### Gel Superior 4 % (Concentrador)

- Acrilamida stock 650 uL
- Agua 3.06 mL
- Buffer superior 1.25 mL
- Persulfato de amonio 10 % 20 uL
- TEMED 10 uL

### Acrilamida stock 30 % (100 mL) (19:1)

- acrilamida 30 %-----30 g
- Bis-acrilamida 0.8 %-----0.8 g

### Buffer inferior pH 8.8

- Tris 1.5 M
- SDS 0.4 %

Buffer superior pH 6.8

- Tris 0.5 M
- SDS 0.4 %

5x SDS-PAGE sample buffer:

- 0.225 M Tris-HCl, pH 6.8
- 50 % glicerol
- 5% SDS
- 0.05% azul de bromofenol
- 0.25 M DTT

Buffer de corrida 5X 1 L

- Tris 25 mMx5 = 15 g
- SDS 0.1% x 5 = 5 g
- Glicina 192 mM x 5 = 72 g

Se almacena a 4°C

Buffer de corrida 1X: Tomar 100 mL buffer corrida 5X + 400 mL de H<sub>2</sub>O

IPTG (1 M)

238 mg/mL en H<sub>2</sub>O, almacenar en alicuotas a -20 °C.

## 10.2 APÉNDICE B. SECUENCIA DEL PLÁSMIDO TOPOCPSF160.

```

1      AGCGCCAAT  ACGCAAACCG  CCTCTCCCCG  CGCGTTGGCC  GATTCATTAA  TGCAGCTGGC
61     ACGACAGGTT  TCCCAGACTGG  AAAGCGGGCA  GTGAGCGCAA  CGCAATTAAT  GTGAGTTAGC
121    TCACTCATT  GGCACCCCAG  GCTTTACACT  TTATGCTTCC  GGCTCGTATG  TTGTGTGGAA
181    TTGTGAGCGG  ATAACAATTT  CACACAGGAA  ACAGCTATGA  CCATGATTAC  GCCAAGCTAT
241    TTAGGTGACA  CTATAGAATA  CTCGAAGCTAT  GCATCAAGCT  TGGTACCGAG  CTCGGATCCA
301    CTAGTAACGG  CCGCCAGTGT  GCTGGAATTC  GCCCTTCATC  GGATCCATGA  ATGTTACGCA
361    TCGAACAATA  TGTGATTCCC  CAGCCCCGGA  ATCAGCAATT  CGATGCCACC  TTCGGTTTTC
421    AAAGGCAGAT  TGCATTGTTA  CTGCCAATGG  AAACAAGCTC  CAAGTCTATT  CAACGAAGGA
481    AGATGATTTA  AGATTAGTAT  GGGAAAAGAA  GTTTTGGGGT  GAAATTTTCG  GTGTATACAG
541    GCATAAAAGT  GGTGGAGAAT  ACGATTCAAT  CATTGTTGGC  TCGGATACAA  GTAAAGTTAT
601    TGTTCACAG  GTTATAGATA  ATGATTTGAA  AGAAACAGAA  TATCACGAAT  TTAACAGACC
661    TGGCCCACCA  GAGCCAGATC  CACCAAAACC  AGAACGTCCT  TTCGATATTA  GTACAAGATT
721    ACGCAATAAA  ACAATTATGG  ATGCCGACCC  AACAGGTACA  TGCCTTGCA  TACTTCTTGC
781    TCAAAATATT  CTTTACGTTT  TACCTCTTGC  CAACAAGATT  AAAATCGAAA  GTACCGAAAA
841    AGCTGGTGAT  GAATACCATT  CCAGCTGGAA  AGTAATTTAA  GATGCGTTCG  CGTATGATGT
901    CCACACAGAT  TTTAAATCGC  CGTTATACAG  AATTCGAGAT  ATGGTTTTCT  TGGACGGATA
961    TAAAAATCCA  ACTCTCGCTA  TTATTCACGA  ATTAATTCCT  ACATGGAGTG  TCAGATTACC
1021   ACTTCAGAAA  TCAACTGTTG  CGGTTTCAAT  TGTTCCTCCT  CCATTAAGAA  AAAAGGAAAC
1081   TGTTTTAATT  TCAGCATCAA  TCGATAAAGT  TACAATGTGG  ACATCCGTTG  CATTACACAA
1141   TAATCATT  GGCTTAGTTC  ATGTTCCAGA  TCCAATTGGA  GGCTTCCTTG  TTCTCTCAAA
1201   GAACGCAATT  ATTTACATGG  ACCATACAAA  CATTGTTGCT  TTAGCATTAA  ACAAGCTCGC
1261   ATACTTAGAT  GATGAAGTTC  CAGTCGATAT  TACAGCAAAT  GGACCTGGTT  GTCACGAATT
1321   ATATTCAAAA  GTTGGTACAG  CTATCGATAA  ATCACATATT  TTATTGACCG  TTGACCAACA
1381   TTATTTGTCC  ATTTTGACTT  TACATTACAA  TGGTGTCAA  GTAACAAATC  TTTCTTTAAA
1441   TGTCAACTTA  AATCTTGAAT  TTCATCCATC  ATGCTTCCTT  TCTCTTAAT  ACACGAACAA
1501   CAGGAGCCTT  GTTTTCATGG  GATCTACAAC  TCATGACTCA  ACTTTGTCCG  AAATAATTTT
1561   AGAAATTGAA  GAAGAAGAAC  TTGCAAGCTT  TTTGGTCGAC  CATGTCATGA  CAGAAACACA
1621   AAATGATCTC  TACCAGAAGT  TCTTTAAGTC  ATTACCAAGG  CCTAAAACCC  TTTCAAAAGT
1681   CAAATCATT  TCTGTCAACA  ACATTTCTTA  CATCAGACAG  CTTGGAAC TG  TTGCCAATGC
1741   CACTCCTTGC  CTCAAGCTGA  GCCAGATAGA  TGAAAATGAA  GAACAGATCG  CGATGGCTCT
1801   CGCATGTGGC  TTCAAAAAAT  ACGGTGTGTT  GCATTTCTGT  CGTAGTGGTA  TCGATCCAAA
1861   CCTTGTCACT  GAATTTAATA  TCCGCGGAGT  TACTGGAATT  TATTCATCAG  AAAACTTCTC
1921   ATATATCTTG  ATCTCAACAG  ATGAAGGAAC  ATCCGTACTT  TACGCCGATA  GCCAGAACCT
1981   TCCAGAGATC  AACGATGAGT  TTACGAGCCA  AGTAAGTAAG  GAAAAGACTC  TTTTTGCAGG
2041   AGATTTCAAC  AACTTCTTCA  TCCAGATCAC  TAAAACAGTC  GTAAGAGCGA  TCAAATATTC
2101   ACAAGAAGAT  GGTTCCTCCC  AACTGACCT  CTATTCTGGC  TCTTCTCCAG  CTGTTTTGGA
2161   TGCAAGAATT  TCTTCAAATG  CCCTTGTTGT  TCTATTTGAC  AACTATACAA  TATGCGCATA
2221   TGACAGAGCG  TTGGAACCAA  TGACACCTCT  CACTAACAAA  TATTTCAAGC  GAATCGCAAT
2281   TTTCCGGTAAA  TATTTATTTG  TTTATTTGAT  AAACGGCGTT  TTGAGAATGT  TTTCCGTAGA
2341   AACGATGACG  GAAGTTGCCA  CATTTGATAA  TTTCAAGGCA  TTTCTGATT  TCATTGCTCC
2401   TCTCGGAAAA  GATGAGATTC  CATTGCAGCC  ATCTTCAATT  GGTATCGTTG  ATATCTCGTT
2461   CGTTGTCTTC  GAATCAGTTG  TCCTTCTTGT  TCTTACACCT  CGCGATGGCC  CTGCTTTCTT
2521   CTATCAATGG  GTTGAATCTA  AATTATTCTT  CAGAAGGATC  AAAATGCACA  GATTCGCAGT
2581   TACTTCACGA  AAGAATATTG  ATATCATCCC  ATTCAACGGC  ATTTCCGATG  TTTTCGATGG
2641   TTGTTTCGTC  ACAGATCCAT  CAAATCCTTT  CTTTATCCTC  TCGGAAAACG  GTTACCCACG
2701   TATCATCCCA  CTCTCCGAGT  CATTCAACTG  CAAACCACGT  GCTTTTGCCC  CTGTAATGCG
2761   TGGTCCTTAC  AAGAATCACT  TTTTAATTGC  GGACGAAGAC  CAAATCAGGC  TCTGCAATCT
2821   CGAGAATATT  AAGCCAGAAC  ACAATTTCTT  CATCATTGAT  GGCTGTATTG  TCGAGAGAAT
2881   CCTGTGTTGA  ATGACTGTAA  GAAGAATTGC  CACTGCCAG  AACCCGAATT  GCGTCGCTTT
2941   CATCGCTTCC  CATCCAGAAC  CATTCAACAC  AGAGAACGAA  AAGAAGATCG  ATGTCGAAGT
3001   TTATGAGAAT  CTTCAAGTTC  ATTACCAAGA  ACCACCTTCT  CCAGCCAAAG  TTTACCCAGA
3061   TGAGGACTAC  GAAACAATTC  CTAAGTGAA  TGAAGAGCGA  TACTCCCTGT  TTCTCTATAC
3121   AAAAGATGGT  TTGCAACAAA  TGGTCGATTA  CGCAAACCAC  GAAATCGTCA  ATACAGTGCA
3181   ATTTGTACAT  ACTACTCCGA  TGCCAGAGGA  TGGAATAACC  CTCCTTAACA  CGTATTTGGC
3241   CGTAGGAAGT  GGATTCTTGT  CTCAACCGGA  AAAAATGATG  AGAGGAGTCC  TTTATATTTA
3301   CCAAATCAGG  TACATGCAGA  ACGATGAGGG  ATTTAACGAG  ATCACGTAA  GACCTTTGTA
3361   CAACGAAACA  AACAAAATTT  AAAAAATCC  GATTATCGAA  ATCACGGATA  ACTCAGGATA
3421   CATGGCTATT  TTCTGCGGTA  ACCTTTGTGA  TCTCATGAGG  TTCTTCAACG  AAAACACCGT
3481   AAAGATCGAA  GCTTTCCTTG  TTGGAAGATT  CTTTGCATCA  TCTATCGTTT  CCTCAGAA
3541   TTATCTGCTC  TACGCTGATT  CCTACGAAGG  CTTTGAGGTC  GCCAGATGGC  GTAAATACGG
3601   AAAGAAGCTC  ATTTCAATGG  CAAGAGATAC  GATGACTAAA  TTACCTCTCA  GTGCAGCATT

```

3661 CCTACAATAC GAGGATTGCT TGGGAGGTGT CGTATTTGAT GATGACGGCA ACGCTCATAT  
3721 CTTTGACGTT GACGAATACG CGATTCCAGC AGATGCAGTT GTCAGAAAGT CAATTTTCTA  
3781 CATTGGAGGC CGAGCAATCA GTTCTGGTCA GTTCCCAATC AAGGCTGTGA CCCAGGCAAC  
3841 TCAGCAAAAAT CCAAATGAAG AAATTGATGA AGAATTATTG CAGCTGCAGA CAAAGATTGG  
3901 AGGCCACATC GCATGGTATG TTACTACACA TGGTAAGATC GGTGCATTCA CTCCAATTGA  
3961 TGAAAACGAT CGACATAAAT TGGTCGGAGT TCAATCCGCA TACGAAAAAT CTCTTTGCGG  
4021 ACTGAGCCAC TTAGAATACA GGAGTGGAAA GTTCAAGAAT ATGATTGAAC AGGATATTTT  
4081 CAACCAATCT CCAAAGAATG TCATTGATTG CGACATGTTG ATTGATTGA TTGAAGATAT  
4141 GCCTGATCAC TTGAAGTTTG CAACAAAGGG ATTAAGAACA CAGGACTTCT TATCTGAATT  
4201 AAGAAAGATT TACAATAATG GTGTCAATGT ATTCCAATAA GGTACCGATG AAGGGCGAAT  
4261 TCTGCAGATA TCCATCACAC TGGCGGCCGC TCGAGCATGC ATCTAGAGAG CCCAATTCGC  
4321 CCTATAGTGA GTCGTATTAC AATTCACTGG CCGTCGTTTT ACAACGTCGT GACTGGGAAA  
4381 ACCCTGGCGT TACCCAACCT AATCGCCTTG CAGCACATCC CCCTTTTCGCC AGCTGGCGTA  
4441 ATAGCGAAGA GGCCCGCACC GATCGCCCTT CCCAACAGTT GCGCAGCCTA TACGTACGGC  
4501 AGTTTAAGGT TTACACCTAT AAAAGAGAGA GCCGTTATCG TCTGTTTTGTG GATGTACAGA  
4561 GTGATATTAT TGACACGCCG GGGCGACGGA TGGTGATCCC CCTGGCCAGT GCACGTCTGC  
4621 TGTCAGATAA AGTCTCCCGT GAACCTTACC CGGTGGTGA TATCGGGAT GAAAGCTGGC  
4681 GCATGATGAC CACCGATATG CCGAGTTGTC CGGTCTCCGT TATCGGGAA GAAAGTGGCTG  
4741 ATCTCAGCCA CCGCGAAAAT GACATCAAAA ACGCCATTAA CCTGATGTTT TGGGGAATAT  
4801 AAATGTCAGG CATGAGATTA TCAAAAAGGA TCTTCACCTA GATCCTTTTC ACGTAGAAAAG  
4861 CCAGTCCGCA GAAACGGTGC TGACCCCGGA TGAATGTCAG CTACTGGGCT ATCTGGACAA  
4921 GGGAAAACGC AAGCGCAAAG AGAAAAGCAG TAGCTTGACAG TGGGCTTACA TGGCGATAGC  
4981 TAGACTGGGC GGTTTTATGG ACAGCAAGCG AACCCGAAAT GCCAGCTGGG GCGCCCTCTG  
5041 GTAAGGTTGG GAAGCCCTGC AAAGTAAACT GGATGGCTTT CTCGCCGCA CCGATTGAT  
5101 GGCGCAGGGG ATCAAGCTCT GATCAAGAGA CAGGATGAGG ATCGTTTTGC ATGATTGAAC  
5161 AAGATGGATT GCACGCAGGT TCTCCGGCCG CTTGGGTGGA GAGGCTATTC GGCTATGACT  
5221 GGGCACAACA GACAATCGGC TGCTCTGATG CCGCCGTGTT CCGGCTGTCA GCGCAGGGGC  
5281 GCCCGGTTCT TTTTGTCAAG ACCGACCTGT CCGGTGCCCT GAATGAACTG CAAGACGAGG  
5341 CAGCGCGGCT ATCGTGGCTG GCCACGACGG GCGTTCCTTG CGCAGCTGTG CTCGACGTTG  
5401 TCACTGAAGC GGAAGGGAC TGGTGCTAT TGGGCGAAGT GCCGGGGCAG GATCTCCTGT  
5461 CATCTAACCT TGCTCCTGCC GAGAAAAGTAT CCATCATGGC TGATGCAATG CGCGCGCTGC  
5521 ATACGCTTGA TCCGGCTACC TGCCCATTCG ACCACCAAGC GAAACATCGC ATCGAGCGAG  
5581 CACGTAICTG GATGGAAGCC GGTCTTGTCG ATCAGGATGA TCTGGACGAA GAGCATCAGG  
5641 GGCTCGCGCC AGCCGAAC TGTCGCCAGG TCAAGGCGAG CATGCCCGAC GGCAGGATC  
5701 TCGTCGTGAC CCATGGCGAT GCCTGCTTGC CGAATATCAT GGTGGAATAA GGCCGCTTTT  
5761 CTGGATTTCAT CGACTGTGGC CGGCTGGGTG TGGCGGACCG CTATCAGGAC ATAGCGTTGG  
5821 CTACCCGTGA TATTGTGAA GAGCTTGGCG GCGAATGGGC TGACCGCTTC CTCGTCCTTT  
5881 ACGGTATCGC CGTCCCGAT TCGCAGCGCA TCGCCTTCTA TCGCCTTCTT GACGAGTCTT  
5941 TCTGAATTAT TAACGCTTAC AATTTCTTGA TCGGGTATTT TCTCCTTACG CATCTGTGCG  
6001 GTATTTTACA CCGCATAACG GTGGCACTTT TCGGGGAAAT GTGCGCGGAA CCCCTATTTG  
6061 TTTATTTTTT TAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC CCTGATAAAT  
6121 GCTTCAATAA TAGCACGTGA GGAGGGCCAC CATGGCCAAG TTGACCAGTG CCGTTCGGT  
6181 GCTCACCAGC CGCAGCTCG CCGGAGCGGT CGAGTTCTGG ACCGACCGGC TCGGTTCTC  
6241 CCGGACTTTC GTGGAGGAC ACTTCGCGG TGTGTCCGG GACGACGTGA CCGGTTCAT  
6301 CAGCGCGGTC CAGGACCAGG TGGTGCCGGA CAACACCCTG GCCTGGGTGT GGGTGCAGG  
6361 CCTGGACGAG CTGTACGCCG AGTGGTCCGGA GGTGCTGTCC ACGAACTTCC GGGACGCTC  
6421 CGGGCCGGCC ATGACCGAGA TCGGCGAGCA GCCGTGGGGG CGGGAGTTCG CCCTGCGCA  
6481 CCCGGCCGGC AACTGCGTGC ACTTCGTGGC CGAGGAGCAG GACTGACACG TGCTAAAAC  
6541 TCATTTTTAA TTTAAAAGGA TCTAGGTGAA GATCCTTTTT GATAATCTCA TGACCAAAAT  
6601 CCCTAACGTT GAGTTTTCGT TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC  
6661 TTCTTGAGAT CTTTTTTTTT TCGCGTAAAT CTGCTGCTTG CAAACAAAAA AACCACCGCT  
6721 ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTTCCGA AGGTAACCTG  
6781 CTTCAGCAGA GCGCAGATAC CAAATACTGT CTTTCTAGTG TAGCCGTAGT TAGGCCACCA  
6841 CTTCAGAAGC TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCAGTGGC  
6901 TGCTGCCAGT GGCATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCAGG  
6961 TAAGGCGCAG CGGTCCGGCT GAACGGGGGG TTCGTGCACA CAGCCCAGCT TGGAGCGAAC  
7021 GACCTACACC GAACTGAGAT ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGTTCCCGA  
7081 AGGGAGAAAAG GCGGACAGGT ATCCGGTAAG CCGCAGGGTC GGAACAGGAG AGCGCAGGAG  
7141 GGAGCTTCCA GGGGAAACG CCTGGTATCT TTATAGTCTT GTCGGGTTTC GCCACCTCTG  
7201 ACTTGAGCGT CGATTTTTGT GATGCTCGTC AGGGGGGCGG AGCCTATGGA AAAACGCCAG  
7261 CAACGCGGCC TTTTACGGT TCCTGGGCTT TTGCTGGCCT TTTGCTACA TGTTCTTCC  
7321 TGCGTTATCC CCTGATTCTG TGGATAACCG TATTACCGCC TTTGAGTGAG CTGATACCGC  
7381 TCGCCGACG CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAG

### 10.3 APÉNDICE C. SECUENCIA DEL PLÁSMIDO pQE30CPSF160.

```

1      CTCGAGAAAT CATAAAAAAT TTATTTGCTT TGTGAGCGGA TAACAATTAT AATAGATTCA
61     ATTGTGAGCG GATAACAATT TCACACAGAA TTCATTAAAG AGGAGAAATT AACTATGAGA
121    GGATCGCATC ACCATCACCA TCACGGATCC ATGAATGTTA CGCATCGAAC AATATGTGAT
181    TCCCCAGCCC CGGAATCAGC AATTCGATGC CACCTTCCGT TTTCAAAGGC AGATTGCATT
241    GTTACTGCCA ATGGAACAAA GTCCTCAAGT TATTCAACGA AGGAAGATGA TTTAAGATTA
301    GTATGGGAAA AGAAGTTTTG GGGTGAAATT TTCGGTGAT ACAGGCATAA AAGTGGTGGA
361    GAATACGATT CAATCATTGT TGGCTGCGAT ACAAGTAAAG TTATTGTTCT ACAGGTTATA
421    GATAATGATT TGAAAGAAAC AGAATATCAC GAATTTAACA GACCTGGCCC ACCAGAGCCA
481    GATCCACCAA AACCAGAACG TCCTTTTCGAT ATTAGTACAA GATTACGCAA TAAAACAATT
541    ATGGATGCCG ACCCAACAGG TACATGCCTT GCATTACTTC TTGCTCAAAA TATTCTTTAC
601    GTTTTACCTC TTGCCAACAA GATTTAAAAT GAAAGTACCG AAAAAGCTGG TGATGAATAC
661    CATTCCAGCT GGAAAGTAAT TAAAGATGCG TTCGCGTATG ATGTCCACAC AGATTTTAAA
721    TCGCCGTTAT ACAGAATTCG AGATATGGTT TTCTTGGACG GATATAAAAA TCCAACCTC
781    GCTATTATTC ACGAATTAAT TCCTACATGG AGTGTGATG TACCACTTCA GAAATCAACT
841    GTTGCAGTTT CAATTGTTTC TCCTCCATTA AAAAAAAGG AACTGTTTTT AATTTACAGC
901    TCAATCGATA AAGTTACAAT GTGGACATCC CGTGCATTAC CACATAACTC ATTTGGCTTA
961    GTTCATGTTT CAGATCCAAT TGGAGGCTTC CTGTCTCTCT CAAAGAACGC AATTATTTAC
1021   ATGGACCATA CAAACATTGT TGCTTTAGCA TTAACAAGC TCGCATACTT AGATGATGAA
1081   GTTCCAGTCG ATATTACAGC AAATGGACCT GGTGTGTCAG AATTATATTC AAAAGTTGGT
1141   ACAGCTATCG ATAAATCACA TATTTTATTG ACCGTTGACC AACATTATTT GTCCATTTTG
1201   ACTTTACATT ACAATGGTGT CAAAGTAACA AATCTTTCTT TAAATGTCAA CTTAAATCTT
1261   GAATTTCATC CATCATGCTT CCTTTCTCTT AACTACACGA ACAACAGGAG CCTTGTPTTC
1321   ATGGGATCTA CAACTCATGA CTCAACTTTG TCCGAAATAA TTTTAGAAAT TGAAGAAGAA
1381   GAACCTGCAA GCTTTTTGGT CGACCATGTC ATGACAGAAA CACAAAATGA TCTCTACCAG
1441   AAGTTCTTTA AGTCATTACC AAGGCCTAAA ACCCTTCAA AAGTCAAATC ATTTTCTGTC
1501   AACAACTTTT CTTACATCAG ACAGTTGGA ACTGTTGCCA ATGCCACTCC TTGCCTCAAG
1561   CTGAGCCAGA TAGATGAAAA TGAAGAACAG ATCGCGATGG CTCTCGCATG TGGCTTCAAA
1621   AAATACGGTT GTTTGCATTT CGTTCGTAGT GGTATCGATC CAAACCTTGT CAGTGAATTT
1681   AATATCCGCG GAGTTACTGG AATTTATTCA TCAGAAAAC TCTCATATAT CTTGATCTCA
1741   ACAGATGAAG GAACATCCGT ACTTTACGCG GATAGCCAGA ACCTTCCAGA GATCAACGAT
1801   GAGTTTACGA GCCAAGTAAG TAAGGAAAAG ACTCTTTTTG CAGGAGATTT CAAACACTTC
1861   TTCATCCAGA TCACTAAAAC AGTCGTAAGA GCGATCAAAT ATTCACAAGA AGATGGTTTC
1921   TCCCACACTG ACCTCTATTC TGGCTCTTCT CCAGCTGTTT TGGATGCAAG AATTTCTTCA
1981   AATGCCCTTG TTGTTCTATT TGACAACAT ACAATATGCG CATATGACAG AGCGTTGGAA
2041   CCAATGACAC CTCTCACTAA CAAATATTTT AGCCGAATCG CAATTTTCGG TAAATATTTA
2101   TTTGTTTATT GTACAAACGG CGTTTTGAGA ATGTTTTCCG TAGAAACGAT GACCGAAGTT
2161   GCCACATTTG ATAATTTCAA GGCATTTCTT GATTTTATTG CTCCTCTCGG AAAAGATGAG
2221   ATTCATTTGC AGCCATCTTC AATTGGTATC GTTGATAICT CGTTCTGTTT CTTGAAATCA
2281   GTTGTCTTTC TTGTTCTTAC ACCTCGCGAT GGCCCTGCTT TCTTCTATCA ATGGGTTGAA
2341   TCTAAATTAT TCTTCAGAAG GATCAAAATG CACAGATTCG CAGTTACTTC ACGAAAGAAT
2401   ATTGATATCA TCCCATTCAA CGGCATTTTC GATGTTTTTC ATGGTTGTTT CGTCACAGAT
2461   CCATCAAATC CTTTCTTTAT CCTCTCGGAA AACGGTTACC CACGTATCAT CCCACTCTCC
2521   GAGTCATTCA ACTGCAAACC ACGTGCTTTT GCCCCTGTAA TGCGTGGTCC TTACAAGAAT
2581   CACTTTTTAA TTGCGGACGA AGACCAAATC AGGCTCTGCA ATCTCGAGAA TATTAAGCCA
2641   GAACATAATT TCTTCATCAT TGATGGCTGT ATTGTGAGA GAATCCCTGT TGGAAATGACT
2701   GTAAGAAGAA TTGCTACTG TCAGAACCCG AATTGCGTCG CTTTCATCGC TTTCCATCCA
2761   GAACCATTCA CAACAGAGAA CGAAAAGAAG ATCGATGTCG AAGTTTATGA GAATCTTCAG
2821   GTTCATTACC AAGAACCACC TTCTCCAGCC AAAGTTTACC CAGATGAGGA CTACGAAACA
2881   ATTCCCTAAGT GGAATGAAGA GCGATACTCC CTGTTTCTCT ATACAAAAGA TGTTTTGCAA
2941   CAAATGGTCG ATTACGCAA CCACGAAATC GTCAATACAG TGCAATTTGT ACATACTACT
3001   CCGATGCCAG AGGATGGAAT AACCTCTCTT AACACGTATT TGGCCGTAGG AAGTGGATTTC
3061   TTGTCTCAAC CGGAAAAAAT GATGAGAGGA GTCCTTTATA TTTACCAAAT CAGGTACATG
3121   CAGAACGATG AGGGATTTAA CGAGATCAGC TTAAGACCTT TGTACAACGA AACAAACAAA
3181   ATTTACAAAA ATCCGATTAT CGAAATCAGC GATAACTCAG GATACATGGC TATTTTCTGC
3241   GGTAACCTTT TGTATCTCAT GAGGTTCTTC AACGAAAACA CCGTAAAGAT CGAAGCTTTC
3301   CTTGTTGGAA GATTCTTTGC ATCATCTATC GTTTCCCTCA AGAATTATCT GCTCTACGCT
3361   GATTCTACG AAGGCTTTGA GGTGCGCAGA TGGCGTAAAT ACGGAAAGAA GCTCATTTCA
3421   ATGGCAAGAG ATACGATGAC TAAATTACCT CTCAGTGCAG CATTCTTACA ATACGAGGAT
3481   TGCTTGGGAG GTGTCGTATT TGATGATGAC GGCAACGCTC ATATCTTTGA CGTTGACGAA
3541   TACGCGATTC CAGCAGATGC AGTTGTGACA AAGTCAATTT TCTACATTTG AGGCCGAGCA

```

3601 ATCAGTTCTG GTCAGTTCCC AATCAAGGCT GTGACCCAGG CAACTCAGCA AAATCCAAAT  
3661 GAAGAAATTG ATGAAGAATT ATTGCAGCTG CAGACAAAAGA TTGGAGGCCA CATCGCATGG  
3721 TATGTTACTA CACATGGTAA GATCGGTGCA TTCACTCCAA TTGATGAAAA CGATCGCAT  
3781 AAATTGGTCG GAGTTCAATC CGCATAACGAA AAATCTCTTT GCGGACTGAG CCACTTAGAA  
3841 TACAGGAGTG GAAAGTTCAA GAATATGATT GAACAGGATA TTTTCAACCA ATCTCCAAAG  
3901 AATGTCATTG ATTGCGACAT GTTGATTGAT TTGATTGAAG ATATGCCTGA TCACTTGAAG  
3961 TTTGCAACAA AGGGATTAAG AACACAGGAC TTCTTATCTG AATTAAGAAA GATTTACAAT  
4021 AATGGTGTCA ATGTATTCCA ATAAGGTACC CCGGGTCGAC CTGCAGCCAA GCTTAATTAG  
4081 CTGAGCTTGG ACTCCTGTTG ATAGATCCAG TAATGACCTC AGAACTCCAT CTGGATTGTG  
4141 TCAGAACGCT CGGTTGCCGC CGGGCGTTTT TTATTGGTGA GAATCCAAGC TAGCTTGGCG  
4201 AGATTTTCAG GAGCTAAGGA AGCTAAAAATG GAGAAAAAAA TCACTGGATA TACCACCGTT  
4261 GATATATCCC AATGGCATCG TAAAGAACAT TTTGAGGCAT TTCAGTCAGT TGCTCAATGT  
4321 ACCTATAACC AGACCGTTCA GCTGGATATT ACGGCCTTTT TAAAGACCGT AAAGAAAAAT  
4381 AAGCACAAGT TTTATCCGGC CTTTATTAC ATTCTTGCCC GCCTGATGAA TGCTCATCCG  
4441 GAATTTTCGTA TGGCAATGAA AGACGGTGAG CTGGTGATAT GGGATAGTGT TCACCCTTGT  
4501 TACACCGTTT TCCATGAGCA AACTGAAACG TTTTCATCGC TCTGGAGTGA ATACCACGAC  
4561 GATTTCCGGC AGTTTCTACA CATATATTCC CAAGATGTGG CGTGTACGG TGAAAACCTG  
4621 GCATTTTCC CTAAGGGTT TATTGAGAAT ATGTTTTTCG TCTCAGCCAA TCCCTGGGTG  
4681 AGTTTACCA GTTTTGATTT AAACGTGGCC AATATGGACA ACTTCTTCGC CCCCGTTTTT  
4741 ACCATGGGCA AATATTATAC GCAAGGCGAC AAGGTGCTGA TGCCGCTGGC GATTCAGGTT  
4801 CATCATGCCG TTTGTGATGG CTTCCATGTC GGCAGAATGC TTAATGAATT ACAACAGTAC  
4861 TGCGATGAGT GGCAGGGCGG GCGTAATTT TTTTAAGGCA GTTATTGGTG CCCTTAAACG  
4921 CCTGGGGTAA TGACTCTCTA GCTTGAGGCA TCAAATAAAA CGAAAGGCTC AGTCAAAAAG  
4981 CTGGGCCTTT CGTTTTATCT GTTGTTTGTC GGTGAACGCT CTCCTGAGTA GGCACAAATC  
5041 GCCCTCTAGA GCTGCCTCGC GCGTTTTCGG GATGACGGTG AAAACCTCTG ACACATGCAG  
5101 CTCCCGGAGA CGGTCACAGC TTGTCTGTAA GCGGATGCCG GGAGCAGACA AGCCCGTCAG  
5161 GCGCGTCAG CGGGTGTGG CGGGTGTGCG GCGCAGCCA TGACCCAGTC ACGTAGCGAT  
5221 AGCGGAGTGT ATACTGGCTT AACTATGCCG CATCAGAGCA GATTGTACTG AGAGTGACC  
5281 ATATGCGGTG TGAATACCG CACAGATGCG TAAGGAGAAA ATACCGCATC AGGCGCTCTT  
5341 CCGCTTCTC GCTCACTGAC TCGTGCCTG CGGTCGTTG GCTGCGGCGA GCGGTATCAG  
5401 CTCACTCAA GCGGTAATA GCGTTATCCA CAGAATCAGG GGATAACGCA GGAAGAACAA  
5461 TGTGAGCAA AGGCCAGCAA AAGGCCAGGA ACCGTAAAAA GGCCGCGTTG CTGGCGTTTT  
5521 TCCATAGGCT CCGCCCCCT GACGAGCATC ACAAATTCG ACGCTCAAGT CAGAGGTGGC  
5581 GAAACCCGAC AGGACTATAA AGATACCAGG CGTTTTCCCC TGGAAGCTCC CTCGTGCGCT  
5641 CTCTGTTCC GACCCTGCCG CTTACCGGAT ACCTGTCCGC CTTTCTCCCT TCGGGAAGCG  
5701 TGGCGCTTTC TCATAGCTCA CGCTGTAGGT ATCTCAGTTC GGTGTAGGTC GTTCGCTCCA  
5761 AGTGGGGCTG TGTGCACGAA CCCCCTGTC AGCCCGACCG CTGCGCCTTA TCCGGTAACT  
5821 ATCGTCTTGA GTCCAACCCG GTAAGACCG ACTTATCGCC ACTGGCAGCA CCGACTGTA  
5881 ACAGGATTAG CAGAGCGAGG TATGTAGGCG GTGCTACAGA GTTCTTGAAG TGGTGGCCTA  
5941 ACTACGGCTA CACTAGAAGG ACAGTATTTG GTATCTGCGC TCTGCTGAAG CCAGTTACCT  
6001 TCGGAAAAAG AGTTGGTAGC TCTTGATCCG GCAAACAAAC CACCGCTGGT AGCGGTGGTT  
6061 TTTTTGTTT CAAGCAGCAG ATTACGCGCA GAAAAAAGG ATCTCAAGAA GATCCTTTGA  
6121 TCTTTTCTAC GGGGTCTGAC GCTCAGTGA ACGAAAATC ACGTTAAGG ATTTTGGTCA  
6181 TGAGATTATC AAAAAGGATC TTCACCTAGA TCCTTTTTAA TTAATAATGA ATTTTAAAT  
6241 CAATCTAAAG TATATATGAG TAAACTTGGT CTGACAGTTA CCAATGCTTA ATCAGTGAGG  
6301 CACCTATCTC AGCGATCTGT CTATTTCTGT CATCCATAGT TGCCTGACTC CCCGTCGTGT  
6361 AGATAACTAC GATACGGGAG GGCTTACCAT CTGGCCCCAG TGCTGCAATG ATACCGGAG  
6421 ACCCAGCTC ACCGGCTCCA GATTTATCAG CAATAAACCA GCCAGCCGGA AGGGCCGAGC  
6481 GCAGAAGTGG TCCTGCAACT TTATCCGCCT CCATCCAGTC TATTAATTGT TGCCGGGAG  
6541 CTAGAGTAAG TAGTTCGCCA GTTAATAGTT TGCGCAACGT TGTGCCATT GCTACAGGCA  
6601 TCGTGGTGTG ACGCTCGTGG TTTGGTATGG CTTCAATCAG CTCCGGTCC CAACGATCAA  
6661 GGCGAGTTAC ATGATCCCC ATGTTGTGCA AAAAAGCGGT TAGCTCCTTC GGTCTCCGA  
6721 TCGTTGTGAG AAGTAAGTTG GCCGAGTGT TATCACTCAT GGTATGGCA GCACTGCATA  
6781 ATTCTCTTAC TGTCATGCCA TCCGTAAGAT GCTTTTCTGT GACTGGTGA TACTCAACCA  
6841 AGTCATTCTG AGAATAGTGT ATGCGGCGAC CGAGTTGCTC TTGCCCGCG TCAATACGGG  
6901 ATAATACCGC GCCACATAGC AGAAGTTTAA AAGTGCTCAT CATTGGAAAA CGTCTCTCGG  
6961 GCGAAAACT CTCAAGGATC TTACCTGTGT TGAGATCCAG TTCGATGTA CCCACTCGTG  
7021 CACCCAACTG ATCTTCAGCA TCTTTTACTT TCACCAGCGT TTCTGGGTGA CAAAAACAG  
7081 GAAGGCAAAA TGCCGCAAAA AAGGGAATAA GGGCGACACG GAAATGTTGA ATACTCATA  
7141 TCTTCTTTT TCAATATTAT TGAAGCATTT ATCAGGGTTA TTGTCTCATG AGCGGATACA  
7201 TATTTGAATG TATTTAGAAA AATAAACAAA TAGGGGTTC GCGCACATTT CCCGAAAAAG  
7261 TGCCACCTGA CGTCTAAGAA ACCATTATTA TCATGACATT AACCTATAAA AATAGGCGTA  
7321 TCACGAGGCC CTTTCGTCTT CAC

## 10. 4 APÉNDICE D. ALINEAMIENTO DE SECUENCIAS TvCPSF160 Y LA SECUENCIA ACCESADA

### TVAG\_077030.

Upper line: CPSF160 TVAG\_077030.seq, from 1 to 3894

Lower line: CPSF160 Secuenciado.seq, from 1 to 3894

99.67% (3881/3894) gap=0.00% (0/3894)

```
1   ATGAATGTTACGCATCGAACAAATATGTGATTCCCCAGCCCCGGAATCAGCAATTCGATGC
   |||
1   ATGAATGTTACGCATCGAACAAATATGTGATTCCCCAGCCCCGGAATCAGCAATTCGATGC

61  CACCTTCCGTTTTCAAAGGCAGATTGCATTGTTACTGCCAATGGAAACAAGCTCCAAGTC
   |||
61  CACCTTCCGTTTTCAAAGGCAGATTGCATTGTTACTGCCAATGGAAACAAGCTCCAAGTC

121 TATTCAACGAAGGAAGATGATTTAAGATTAGTATGGGAAAAGAAGTTTTGGGGTGAAATT
   |||
121 TATTCAACGAAGGAAGATGATTTAAGATTAGTATGGGAAAAGAAGTTTTGGGGTGAAATT

181 TTCGGTGTATACAGGCATAAAAAGTGGTGGAGAATACGATTCAATCATTGTTGGCTGCGAT
   |||
181 TTCGGTGTATACAGGCATAAAAAGTGGTGGAGAATACGATTCAATCATTGTTGGCTGCGAT

241 ACAAGTAAAGTTATTGTTCTACAGGTTATAGATAATGATTTGAAAGAAACAGAATATCAC
   |||
241 ACAAGTAAAGTTATTGTTCTACAGGTTATAGATAATGATTTGAAAGAAACAGAATATCAC

301 GAATTTAACAGACCTGGCCCACCAGAGCCAGATCCACCAAAACCAGAACGTCCTTTTCGAT
   |||
301 GAATTTAACAGACCTGGCCCACCAGAGCCAGATCCACCAAAACCAGAACGTCCTTTTCGAT

361 ATTAGTACAAGATTACGCAATAAAAACAATTATGGATGCCGACCAACAGGTACATGCCTT
   |||
361 ATTAGTACAAGATTACGCAATAAAAACAATTATGGATGCCGACCAACAGGTACATGCCTT

421 GCATTACTTCTTGCTCAAAAATATCTTTACGTTTTTACCTCTTGCCAACAAGATTAAAATC
   |||
421 GCATTACTTCTTGCTCAAAAATATCTTTACGTTTTTACCTCTTGCCAACAAGATTAAAATC

481 GAAAGTACCGAAAAAGCTGGTGATGAATACCATTCAGCTGGAAAAGTAATTAAAGATGCG
   |||
481 GAAAGTACCGAAAAAGCTGGTGATGAATACCATTCAGCTGGAAAAGTAATTAAAGATGCG

541 TTCGCGTATGATGTCCACACAGATTTTAAATCGCCGTTATACAGAATTCGAGATATGGTT
   |||
541 TTCGCGTATGATGTCCACACAGATTTTAAATCGCCGTTATACAGAATTCGAGATATGGTT

601 TTCTTGACGGATATAAAAAATCCAACCTCTCGCTATTATTACGAATTAATTCCTACATGG
   |||
601 TTCTTGACGGATATAAAAAATCCAACCTCTCGCTATTATTACGAATTAATTCCTACATGG
```



```

|||||
1501 GGTATCGATCCAAACCTTGTCTAGTGAATTCAATATCCGCGGAGTTACTGGAATTTATTCA
|||||
1561 TCAGAAAACCTTCTCATATATCTTGATCTCAACAGATGAAGGAACATCCGTACTTTACGCC
|||||
1561 TCAGAAAACCTTCTCATATATCTTGATCTCAACAGATGAAGGAACATCCGTACTTTACGCC
|||||
1621 GATAGCCAGAACCTTCCAGAGATCAACGATGAGTTTACGAGCCAAGTAAGTAAGGAAAAG
|||||
1621 GATAGCCAGAACCTTCCAGAGATCAACGATGAGTTTACGAGCCAAGTAAGTAAGGAAAAG
|||||
1681 ACTCTTTTTGCAGGAGATTTCAACAACCTTCTTCATCCAGATCACTAAAACAGTCGTAAGA
|||||
1681 ACTCTTTTTGCAGGAGATTTCAACAACCTTCTTCATCCAGATCACTAAAACAGTCGTAAGA
|||||
1741 GCGATCAAATATTCACAAGAAGATGGTTTTCTCCACACTGACCTCTATTCTGGCTCTTCT
|||||
1741 GCGATCAAATATTCACAAGAAGATGGTTCCTCCACACTGACCTCTATTCTGGCTCTTCT
|||||
1801 CCAGCTGTTTTGGATGCAAGAATTTCTTCAAATGCCCTTGTTGTTCTATTTGACAACAT
|||||
1801 CCAGCTGTTTTGGATGCAAGAATTTCTTCAAATGCCCTTGTTGTTCTATTTGACAACAT
|||||
1861 ACAATATGCGCATATGACAGAGCGTTGGAACCAATGACACCTCTCACTAACAAAATATTTT
|||||
1861 ACAATATGCGCATATGACAGAGCGTTGGAACCAATGACACCTCTCACTAACAAAATATTTT
|||||
1921 AGCCGAATCGCAATTTTCGGTAAATATTTATTTGTTTATTGTACAAACGGCGTTTTGAGA
|| |||||
1921 AGTCGAATCGCAATTTTCGGTAAATATTTATTCGTTTATTGTACAACGGCGTTTTGAGA
|||||
1981 ATGTTTTCCGTAGAAAACGATGACGGAAGTTGCCACATTTGATAATTTCAAGGCATTTCC
|||||
1981 ATGTTTTCCGTAGAAAACGATGACGGAAGTTGCCACATTTGATAATTTCAAGGCATTTCC
|||||
2041 GATTTTCATTGCTCCTCTCGGAAAAGATGAGATTCCATTGCAGCCATCTTCAATTGGTATC
|||||
2041 GATTTTCATTGCTCCTCTCGGAAAAGATGAGATTCCATTGCAGCCATCTTCAATTGGTATC
|||||
2101 GTTGATATCTCGTTCGTTGTCTTCGAATCAGTTGTCCCTTCTTGTCTTACACCTCGCGAT
|||||
2101 GTTGATATCTCGTTCGTTGTCTTCGAATCAGTTGTCCCTTCTTGTCTTACACCTCGCGAT
|||||
2161 GGCCCTGCTTTCTTCTATCAATGGGTTGAATCTAAATTATTCTTTCAGAAGGATCAAAATG
|||||
2161 GGCCCTGCTTTCTTCTATCAATGGGTTGAATCTAAATTATTCTTTCAGAAGGATCAAAATG
|||||
2221 CACAGATTCGCAGTTACTTTCACGAAAGAATATTGATATCATCCCATTCAACGGCATTTCG
|||||
2221 CACAGATTCGCAGTTACTTTCACGAAAGAATATTGATATCATCCCATTCAACGGCATTTCG
|||||
2281 GATGTTTTTCGATGGTTGTTTCGTACAGATCCATCAAATCCTTTCTTTATCCTCTCGGAA
|||||
2281 GATGTTTTTCGATGGTTGTTTCGTACAGATCCATCAAATCCTTTCTTTATCCTCTCGGAA
|||||
2341 AACGGTTACCCACGTATCATCCCCTCTCCGAGTCATTCAACTGCAAACCACGTGCTTTT
|||||

```



