

UACM

Universidad Autónoma
de la Ciudad de México

Nada humano me es ajeno

UNIVERSIDAD AUTÓNOMA DE LA CIUDAD DE MÉXICO
POSGRADO EN CIENCIAS GENÓMICAS

IDENTIFICACIÓN DE LA MAQUINARIA DE PROCESAMIENTO DEL EXTREMO 3'UTR DE LOS RNAm DE *Trichomonas vaginalis*

T E S I S

QUE PARA OBTENER EL GRADO DE

MAESTRO EN CIENCIAS GENÓMICAS

P R E S E N T A

M.C. MIGUEL ÁNGEL DEL MORAL STEVENEL

TUTORA:

Dra. María Elizbeth Álvarez Sánchez.

COMITÉ ASESOR:

Dra. Mavil López Casamichana.

Dr. César López Camarillo.

Dr. Máximo Berto Benítez Martínez

COMITÉ TUTORIAL

SISTEMA BIBLIOTECARIO DE INFORMACIÓN Y DOCUMENTACIÓN



UNIVERSIDAD AUTÓNOMA DE LA CIUDAD DE MÉXICO COORDINACIÓN ACADÉMICA

RESTRICCIONES DE USO PARA LAS TESIS DIGITALES

DERECHOS RESERVADOS[©]

La presente obra y cada uno de sus elementos está protegido por la Ley Federal del Derecho de Autor; por la Ley de la Universidad Autónoma de la Ciudad de México, así como lo dispuesto por el Estatuto General Orgánico de la Universidad Autónoma de la Ciudad de México; del mismo modo por lo establecido en el Acuerdo por el cual se aprueba la Norma mediante la que se Modifican, Adicionan y Derogan Diversas Disposiciones del Estatuto Orgánico de la Universidad de la Ciudad de México, aprobado por el Consejo de Gobierno el 29 de enero de 2002, con el objeto de definir las atribuciones de las diferentes unidades que forman la estructura de la Universidad Autónoma de la Ciudad de México como organismo público autónomo y lo establecido en el Reglamento de Titulación de la Universidad Autónoma de la Ciudad de México.

Por lo que el uso de su contenido, así como cada una de las partes que lo integran y que están bajo la tutela de la Ley Federal de Derecho de Autor, obliga a quien haga uso de la presente obra a considerar que solo lo realizará si es para fines educativos, académicos, de investigación o informativos y se compromete a citar esta fuente, así como a su autor ó autores. Por lo tanto, queda prohibida su reproducción total o parcial y cualquier uso diferente a los ya mencionados, los cuales serán reclamados por el titular de los derechos y sancionados conforme a la legislación aplicable.

DIRECTOR

Dra. María Elizbeth Álvarez Sánchez.

Profesora Titular. Posgrado en Ciencias Genómicas

Universidad Autónoma de la Ciudad de México

GRUPO ASESOR

Dr. Máximo Berto Martínez Benítez.

Director del Centro de Diagnóstico y Vigilancia Epidemiológico del ICyT

Distrito Federal.

Dra. Mavil López Casamichana.

Profesora. Posgrado en Ciencias Genómicas

Universidad Autónoma de la Ciudad de México.

Dr. César López Camarillo.

Profesor Titular. Posgrado en Ciencias Genómicas.

Universidad Autónoma de la Ciudad de México.

El presente proyecto de investigación se realizó en el laboratorio 1 del Posgrado en Ciencias Genómicas de la Universidad Autónoma de la Ciudad de México, bajo la dirección de la Dra. María Elizabeth Alvarez Sánchez y bajo la asesoría del Dr. Máximo Berto Martínez Benítez, la Dra. Mavil López Casamichana y el Dr. César López Camarillo. Agradezco a mis lectores de tesis Dra. Elisa Azuara Liceaga y Dra. Berenice Parra Ortega. El proyecto fue financiado con fondos del Posgrado en Ciencias Genómicas de la Universidad Autónoma de la Ciudad de México y del CONACYT con número de proyecto 83808.

El estudiante agradece el apoyo de la Universidad Autónoma de la Ciudad de México, al Posgrado en Ciencias Genómicas, al CONACYT y al ICyTDF como becario.

Agradecimientos...

Agradezco a mi Esposa Xóchitl y mis Hijos Natalia, Sebastián y Andrea por el tiempo que no estuve con ellos por estar en el laboratorio y por su apoyo, sin ustedes nunca se hubiera podido realizar este sueño. Ustedes son mi razón para seguir adelante gracias.

Agradezco a mi toda mi Familia por su apoyo en este proceso de superación en mi vida.

Mi más sincero y eterno agradecimiento a la Dra. Elizabeth Álvarez por aceptarme como miembro de su equipo de investigación, por su confianza, por todos sus consejos y guía, por su infinita paciencia para explicarme; por ser un ejemplo de vida para mi... y por su amistad.

Agradezco a mi amiga, Alma, por su asesoría en el laboratorio, por compartir conmigo buenos y no tan buenos momentos en el laboratorio, por las risas, por sus palabras de ánimo, por su apoyo y sobre todo por su amistad incondicional... gracias...

¡¡¡sí se pudo sensei!!!

Agradezco a Itzel Quintas, por su apoyo en la parte final de la tesis y sus tips.

Agradezco a José Luis por su amistad y apoyo en el laboratorio y en clases.

Agradezco a Laura por su apoyo en el cuarto de cultivo y sus llamadas de atención.

Agradezco a mis amigos y compañeros de laboratorio: José Alí, Berenice, Daniel, Marcos, Diana, Roberto, Miguel Fonseca, Itzel López, Olga, Helios, Rubí, Lidia, Karen, Fernando, Payán, Mendieta, Areli, Marco, Alma Olivarez, Alma Genis, María Elena, Manuel, Zefene, Karla, Elba, Yaqui, Brenda por todos esos momentos que compartimos y su apoyo, porque sin ellos, la estancia en el posgrado no hubiera sido lo mismo.

A todos los profesores del Posgrado en Ciencias Genómicas.

A mis compañeros y amigos del Consejo Universitario, en especial a la Rectora Dra. Esther Orozco Orozco por impulsar siempre la superación académica de todos, por su ejemplo de fortaleza, dignidad y convicción.

A Caty, Lulú y Viviana por su eficiencia y apoyo en todos los trámites que se necesitaron.

A aquel en quien todo lo puedo y es el alfa y el omega.

Dedicatoria

A mi Esposa.

A mi Hijos.

A mis Padres

A mis Hermanos,

INDICE GENERAL

LISTA DE ABREVIATURAS

LISTA DE FIGURAS

LISTA DE TABLAS

I. INTRODUCCIÓN	1
1. ANTECEDENTES GENERALES	1
1.1. Tricomonosis	1
1.2 Epidemiología	5
1.3 Etiología	5
1.3.1 Características del parásito	5
1.3.1.1 Ciclo de vida	11
1.3.1.2 Genoma de <i>T. vaginalis</i>	14
1.3.1.3 Requerimientos nutricionales	16
1.4 Patogenia	19
1.4.1 Adhesión	20
1.4.2 Citotoxicidad	21
1.4.3 Fagocitosis	23
1.4.4 Evasión de la respuesta inmune	23
1.4.5 Apoptosis	26
1.4.6 Hemólisis	27
2. ESTADO DEL ARTE	27
2.1 Expresión de genes	27

2.2 Generalidades de la maquinaria de procesamiento en eucariontes.	29
2.3 Mecanismos alternativos de poliadenilación	31
2.4 Elementos del procesamiento de los 3'UTR de los RNAm de humanos.	35
2.4.1 Elementos de secuencia.	35
2.4.2 Maquinaria de procesamiento del extremo 3'UTR del RNAm del humano y sus complejos.	36
2.4.2.1 Complejo CPSF	36
2.4.2.2 Complejo CstF	39
2.4.2.3 Complejo CFIm	39
2.4.2.4 Complejo CFIlm	40
2.4.2.5 Elementos proteicos que no forman parte de complejos	40
2.5. Elementos en <i>cis</i> del procesamiento del extremo 3'UTR del RNAm en otros organismos.	41
2.6. Maquinaria de procesamiento del extremo 3'UTR en otros organismos.	46
2.6.1 <i>Saccharomyces cerevisiae</i>	46
2.6.2 <i>Arabidopsis thaliana</i>	47
2.6.3 <i>Entamoeba histolytica</i>	47
2.7 Procesamiento del extremo 3'UTR de los RNAm de <i>Trichomonas vaginalis</i> .	50
III. JUSTIFICACIÓN	51

IV. HIPÓTESIS	51
V. OBJETIVOS	51
5.1 OBJETIVO GENERAL	51
5.2 OBJETIVOS PARTICULARES	51
VI. ESTRATEGIA EXPERIMENTAL	52
VII. MATERIAL Y MÉTODOS	53
7.1 Detección <i>in silico</i>	53
7.2 Cultivo de <i>T. vaginalis</i>	53
7.3 Condición de cultivo en Zn ²⁺	54
7.4 Extracción de DNA	54
7.5 Extracción de RNA	55
7.6 Diseño de oligos	55
7.7 RT-PCR	56
7.8 Extracción de proteínas totales de <i>T. vaginalis</i>	57
7.9 Electroforesis en doble dimensión	57
7.10 Western blot	58
VIII. RESULTADOS	59
8.1 Detección <i>in silico</i> de las posibles proteínas que integran la maquinaria del procesamiento del extremo 3'UTR de los RNAm de <i>T. vaginalis</i> .	59
8.2 Expresión del RNAm de algunos de los genes que codifican para las proteínas candidatas de <i>T. vaginalis</i> involucradas en procesamiento del 3'UTR de los RNAm .	69

8.3 Inmunodetección de PC4 en extractos de <i>T. vaginalis</i>	74
IX. DISCUSIÓN	85
X. CONCLUSIONES	89
XI. REFERENCIAS	93
XII. APÉNDICE	101
ANEXO II	107

LISTA DE ABREVIATURAS

- Å Amstrongs □ A Adenina.
- C Citosina.
- c.b.p Cuanto baste para.
- cDNA Ácido desoxirribonucleico complementario.
- CEVs Células del epitelio vaginal.
- CFI A Factor de corte 1-A, hace referencia a levadura.
- CFIm Factor de corte 1 de mamífero.
- CFIlm Factor de corte 2 de mamífero.
- CNCD147 Aislado de tricomonas de una paciente mujer.
- CO₂ Dióxido de carbono.
- CPF Factor de corte y poliadenilación (levadura).
- CPs Cisteín proteinasas.
- CPSF Factor de especificidad de corte y poliadenilación.
- CstF Factor estimulante de corte.
- CTD-RNAPolIII Dominio carboxilo-terminal de la RNA polimerasa II.
- DEPC Dietilpirocarbonato.
- DNA Ácido desoxirribonucleico.
- DNAg DNA genómico.
- dNTPs Deoxinucléotido tri-fosfatados.
- DSE Elemento río abajo (down stream element).
- dsRNA RNA de doble cadena.
- DTT Ditioneitol.
- DU-145 Células de metástasis cerebral de carcinoma de próstata.
- ECM Matriz extracelular.
- EE Elemento de eficiencia (USE de levadura) □ ESTs
Blanco de secuencias de expresión.
- ETS Enfermedad de transmisión sexual.
- Fe²⁺ Hierro.
- FUE Elemento lejano corriente arriba (USE de Arabidopsis)
- G Guanina.

- g Gramos.
- h Hora(s).
- HCl Acido clorhídrico.
- HGMN01 Aislado de tricomonas de paciente masculino.
- H/M Proporción Hombre/Mujer.
- kDa Kilodaltones.
- µg Microgramos.
- min Minutos.
- µl Microlitros.
- ml Mililitros.
- mM Milimolar.
- MS Espectrometría de masas.
- NUE Elemento cercano río arriba (PAS de Arabidopsis) □ O.N.
Toda la noche (de over nigh).
- OMS Organización Mundial de la Salud.
- PABP Proteína de unión a poli-A.
- PAGE Electroforesis en gel de poliacrilamida.
- PAP Polimerasa de poli-A.
- PAS Señal de poliadenilación.
- pb Pares de base.
- PBS Amortiguador de sales de fosfato.
- PC4 Coactivador positivo 4 (Positive coactivator 4).
- PCR Reacción en cadena de la polimerasa.
- PE Elemento de posicionamiento (PAS de levadura).
- PFO Piruvato ferredoxina óxido-reductasa.
- pH Potencial de hidrogeno.
- pI Punto isoeléctrico.
- preRNAm RNA recién transcrito ó RNAm inmaduro.
- PSA Persulfato de amonio.
- RNA Acido ribonucleico.
- RNAm Acido ribonucleico mensajero.
- RNAi Acido ribonucleico de interferencia.
- RNPs Ribonucleoproteínas.

- ROS Especies reactivas de oxígeno.
- rpm Revoluciones por minuto.
- rRNA RNA ribosomal.
- RT-PCR Reacción en cadena de la polimerasa retrotranscriptasa □ SC
 Suero de caballo.
- SDS Dodecil sulfato de sodio.
- snRNPs Pequeñas ribonucleoproteínas nucleares.
- SPI Suero preinmune.
- T Timina.
- TA Temperatura ambiente.
- TAE Amortiguador Tris-Acetato-EDTA.
- TCA Ácido tricloro acético.
- TEMED N,N,N', N'tetrametilen-diamina.
- tRNA RNA de transferencia.
- TYM Medio de cultivo tripticasa-extracto de levadura-maltosa.
- U Uracilo.
- USE Elemento río arriba.
- UTR Región no traducida.
- V Volts.
- WB Western blot.
- Zn₂₊ Zinc.
- 2-DE Electroforesis en doble dimensión.
- 2D-WB Western blot en doble dimensión.

LISTA DE FIGURAS

Figura 1.- Epidemiología de <i>T. vaginalis</i> .	6-7
Figura 2.- Morfología de <i>T. vaginalis</i> .	9-10
Figura 3.- Ciclo de vida de <i>T. vaginalis</i> .	12-13
Figura 4.- Mecanismos de fagocitosis de <i>T. vaginalis</i> .	24-25
Figura 5.- Poliadenilación alternativa.	33-34
Figura 6.- Maquinaria de poliadenilación del humano.	37-38
Figura 7.- Elementos <i>cis</i> para la poliadenilación de los RNAm en diferentes organismos	44-45
Figura 8. Estudio comparativo de estructuras tridimensionales	66,67,68
Figura 9.- Amplicones de los transcritos de los genes candidatos del aislado HGMN01 que participan en la maquinaria de procesamiento de <i>T. vaginalis</i> .	72-73
Figura 10.- Inmunodetección de la proteína TvPC4 en <i>T. vaginalis</i> HGMN01 en 1-D	76-77
Figura 11.- Inmunoreconocimiento de PC4 en el aislado HGMN01 de <i>T. vaginalis</i> en 2-D	78-79
Figura 12.- Análisis por Melanie de la expresión de TvPC4 de <i>T. vaginalis</i> en presencia y ausencia de Zn^{2+}	81-82
Figura 13.- Predicción de modificaciones postraduccionales de TvPC4 de <i>T. vaginalis</i> .	83-84
Figura 14.- Modelo de la maquinaria de procesamiento del extremo 3'UTR de los RNAm de <i>T. vaginalis</i> .	91-92

LISTA DE TABLAS

Tabla 1.- Taxonomía de <i>Trichomonas vaginalis</i> .	8
Tabla 2.- Características del genoma de <i>T. vaginalis</i> .	16
Tabla 3.- Homología de las maquinarias de procesamiento del extremo 3'UTR de los preRNAm de <i>Homo Sapiens</i> , <i>Saccharomyces cerevisiae</i> , <i>Arabidopsis thaliana</i> y <i>Entamoeba histolytica</i> .	49
Tabla 4.- Elementos de la maquinaria de corte y poliadenilación de los RNAm en <i>H. sapiens</i> y secuencias relacionadas en <i>T. vaginalis</i> .	60
Tabla 5.- Elementos de la maquinaria de corte y poliadenilación de los RNAm en <i>S. cerevisiae</i> y secuencias relacionadas en <i>T. vaginalis</i> .	61
Tabla 6.- Elementos de la maquinaria de corte y poliadenilación de los RNAm en <i>A. thaliana</i> y secuencias relacionadas en <i>T. vaginalis</i> .	62
Tabla 7.- Elementos de la maquinaria de corte y poliadenilación de los RNAm en <i>E. histolytica</i> y secuencias relacionadas en <i>T. vaginalis</i>	63
Tabla 8.- Proteínas detectadas <i>in silico</i> como candidatas a formar parte de la maquinaria de procesamiento del extremo 3'UTR de los RNAm de <i>Trichomonas vaginalis</i> .	64-65
Tabla 9.- Características de los oligos iniciadores utilizados para la RT-PCR.	70-71

I.- INTRODUCCIÓN:

1. ANTECEDENTES GENERALES:

1.1. TRICOMONOSIS

La tricomonosis es una enfermedad de transmisión sexual (ETS) ocasionada por el parásito *Trichomonas vaginalis* que infecta el tracto genitourinario. En la mujer, la infección por *T. vaginalis* puede presentar un cuadro agudo que se caracteriza por vulvovaginitis, la cual, a su vez se manifiesta por eritema de las paredes vaginales (22-37%) y la presencia hasta en el 2% de los casos de cérvix en fresa (*colpitis macularis*); prurito, con un flujo verde-amarillento de olor fétido (40-50% de los casos), el cual, puede cambiar sus características si existe una infección concomitante (25-38% de los casos). Además puede manifestarse dispareunia (dolor durante el coito) y distensión abdominal (Schwebke y Burgess 2004). También puede presentar síntomas urinarios como disuria y/o polaquiuria que corresponden a un cuadro de uretritis y/o cistitis. La infección crónica por *T. vaginalis* es causa de enfermedad inflamatoria pélvica (Moodley, Wilkinson y col.; 2002) y puede condicionar infertilidad. Otras complicaciones de importancia son abortos, ruptura prematura de membranas y recién nacidos con bajo peso, así como aumento en el riesgo de contraer infecciones postquirúrgicas (Petrin, Delgaty y col.; 1998; Swygard, Sena y col.; 2004; Fichorova 2009; Shafir, Sorvillo y col.; 2009). Asimismo, la tricomonosis se ha asociado con el aumento en el riesgo de contraer el virus de la inmunodeficiencia humana (VIH), como resultado del proceso inflamatorio crónico, lo cual aumenta la cantidad de células blanco para el VIH al haber un aumento de linfocitos T CD4+ en la región (Guenther, Secor y col.; 2005). También se propone que la infección por *T. vaginalis* degrada las proteasas secretoras de los leucocitos que bloquean el anclaje del virus sobre sus células blanco, además de inhibir la producción de células inmunes (Thurman y Doncel 2011). La destrucción del epitelio urogenital por parte de *T. vaginalis* condiciona una entrada más fácil para los virus (Guenther, Secor y col.; 2005) como el virus del papiloma humano asociado al Cáncer Cervicouterino (Fichorova 2009).

En el caso del hombre, la infección es asintomática hasta en el 60% de los casos, por lo cual, se le ha atribuido un papel de portador para este parásito. Los reportes de prevalencia de tricomonosis masculina varían de acuerdo a la población y al método de diagnóstico utilizado, esta infección puede eliminarse en un lapso de 2 semanas o autolimitarse (Borchardt, al-Haraci y col.; 1995), como se ha demostrado en otros trabajos, esto puede deberse a que las concentraciones de Zn^{2+} en las secreciones prostáticas son tricomonicidas (Langley, Goldsmid y col.; 1987). En los fluidos prostáticos del hombre, la concentración de Zn^{2+} varía de 4.5 hasta 7 mM, *in vitro* se ha demostrado que la concentración tricomonicida del Zn^{2+} es igual o mayor a 1.6 mM. A esta concentración el parásito disminuye su crecimiento y reduce la citotoxicidad hacia las células DU-145 comparado con parásitos crecidos en ausencia de Zn^{2+} , además se encontró que el parásito no cambia a su forma ovoide como sucede cuando se une a las células del epitelio vaginal, sino que muestra proyecciones flagelares hacia las células prostáticas (Vázquez-Carrillo, Granados y col.; 2011).

Estas observaciones sugieren que pacientes masculinos que presenten concentraciones inferiores a 1.6 mM de Zn^{2+} en sus fluidos prostáticos pueden no autolimitar la infección. Por otra parte, también el cuadro clínico de la tricomonosis en el sexo masculino depende de la región afectada y los signos y síntomas pueden variar, pudiendo presentar cuadros de balanitis, que se caracteriza por eritema balanoprepucial; uretritis con síntomas de disuria y secreción uretral; epididimitis con dolor e inflamación del epidídimo, y prostatitis que puede cursar con secreción fétida, retardo en la micción y disminución del chorro miccional. Las complicaciones que se dan se deben básicamente a la cronicidad de la infección y pueden ser: infertilidad debido a la disminución de la movilidad y viabilidad del espermatozoide; predisposición al VIH y al cáncer de próstata, para este último se proponen mecanismos como daño al epitelio urogenital, inflamación crónica, inhibición de la apoptosis y alteración de los niveles locales de poliaminas, que condicionan metaplasia y displasia en el tejido prostático (Hobbs, Lapple y col.; 2006; Sutcliffe,

Giovannucci y col.; 2006; Abdolrasouli, Amin y col.; 2007; Sutcliffe, Alderete y col.; 2009).

El diagnóstico de la tricomonosis puede ser hecho por la presencia de los signos y síntomas clínicos ya mencionados. Un estudio en trabajadoras sexuales de Camerún mostró que el valor predictivo de los médicos para diagnosticar tricomonosis únicamente basados en los hallazgos clínicos es del 47%, y que los síntomas que pueden orientar el diagnóstico es el “cérvix en fresa”, el flujo vaginal y la disuria, a excepción del primero que es patognomónico, los otros síntomas son inespecíficos pero son más frecuentes, por lo cual se hace necesario un método de diagnóstico sensible y específico que sea usado sistemáticamente. Actualmente para el diagnóstico se pueden obtener diferentes muestras del paciente como son: orina, fluido vaginal, semen, frotis endocervical y frotis uretral. Con estas muestras se pueden hacer diferentes estudios (Swygard, Sena y col.; 2004). Entre estos estudios se encuentra la visualización directa del parásito y se logra por diferentes métodos; el primero es mediante una “preparación húmeda” en la cual, una muestra es suspendida en una solución salina al 0.85% e inmediatamente visualizada al microscopio, aunque es un método rápido y de bajo costo, es poco sensible. Otra forma de visualizar al parásito es extendiendo una muestra de moco cervical en una laminilla, la cual, también se observa al microscopio. El estándar de oro por visualización del parásito es el cultivo de la muestra, con medio modificado de Diamond, este método es mucho más sensible que una preparación húmeda pero el resultado tarda entre 3 y 7 días por lo que lo hace poco factible. Diferentes variantes de medios de cultivo y de tinciones para observar al microscopio se han utilizado con resultados similares (Swygard, Sena y col.; 2004). Debido a la problemática de las preparaciones húmedas y del cultivo, el uso de métodos como la PCR ó prueba de amplificación de ácido nucleico (NAAT de nucleic acid amplification test) ha ido incrementándose. Para este estudio se necesita DNA ya sea de parásitos vivos o muertos en la muestra y en concentraciones tan bajas como un solo microorganismo en la muestra, de la cual se amplifica un gen específico para tricomonas. Aunque este método tiene una sensibilidad del 96.7% y una

especificidad del 97.5% tiene como inconveniente el alto costo y la poca disponibilidad comercial. Un ejemplo de estas pruebas es el Kit Aptima (Gen-Probe, San Diego, CA) (Campbell, Woods y col.; 2008; Pattullo, Griffeth y col.; 2009). Otra opción para identificar al parásito es realizando pruebas antigénicas, mediante se trata de un artefacto con una tira que detecta proteínas de membrana del parásito y puede utilizar las muestras de una preparación húmeda o un frotis que hubiera salido negativo, por lo cual es un excelente método como una segunda línea de diagnóstico considerando su rapidez y el bajo costo. Se ha reportado con una sensibilidad del 85 al 90% y una especificidad del 100%. Su desventaja actual es su baja disponibilidad, un ejemplo es el OSOM Trich Rapid Test (Genzyme Diagnostics, Cambridge, MA) (Swygard, Sena y col.; 2004; Campbell, Woods y col.; 2008; Pattullo, Griffeth y col.; 2009).

El tratamiento definitivo de la tricomonosis es a base de nitroimidazoles como el metronidazol, tinidazol, ornidazol, cernidazol y nimorazol; los primero dos son los más ampliamente utilizados. Estos medicamentos son reducidos en los hidrogenosomas provocando desestabilidad del DNA en el parásito. El metronidazol es el medicamento más utilizado para el tratamiento de la tricomonosis, su dosis vía oral es de 2 g en una sola dosis, o bien, 500mg cada 8 hrs por 5 a 7 días. La aplicación tópica en forma de óvulos que se aplican diariamente hasta por 10 días, no es tan efectiva pero es considerada en el caso de mujeres embarazadas, debido al riesgo de que el metronidazol induzca teratogénesis en el feto. Los efectos secundarios del tratamiento pueden ser un sabor metálico, náusea, vómitos y gastritis medicamentosa, síntomas que son más frecuentes con la monodosis de 2 g. Una reacción adversa es el efecto antabús o disulfiram que se da al consumir concomitantemente alcohol. Se ha reportado hasta un 10% de resistencia al tratamiento con metronidazol y dicha resistencia puede predecir la resistencia a otros nitroimidazoles como el tinidazol (Swygard, Sena y col.; 2004).

1.2 EPIDEMIOLOGÍA

A nivel mundial se reporta una incidencia anual de 250 millones de casos (reporte OMS 2010), 8% en hombres y 92% en mujeres, con una prevalencia del 5-74% para mujeres y del 5-29% para hombres (Swygard, Sena y col.; 2004). En la República Mexicana durante el 2011 se reportaron 113,334 casos, de los cuales el 97.3% (110,174) corresponde a mujeres y 2.7% (3060) a hombres. Los estados más afectados en orden decreciente fueron Veracruz, Puebla, Estado de México, Chiapas, Oaxaca y Michoacán (**Fig. 1**). En el Distrito Federal se han reportado en lo que va del 2012, 817 casos, de los cuales 812 (99.4%) corresponden a mujeres y 5 casos (0.6%) a hombres, con una tasa de incidencia de 43 por cada 100,000 habitantes. A nivel nacional esta incidencia se estima en 118.2/100,000 habitantes, ocupando el 2º lugar dentro de las ETS. Es importante remarcar que las mujeres son las más afectadas con una relación Hombre/Mujer (H/M) de 1:37 y el grupo etario donde se reportan más incidencias es el de 25-44 años, lo cual se enmarca en la edad laboral y reproductiva del ser humano (<http://www.dgepi.salud.gob.mx>).

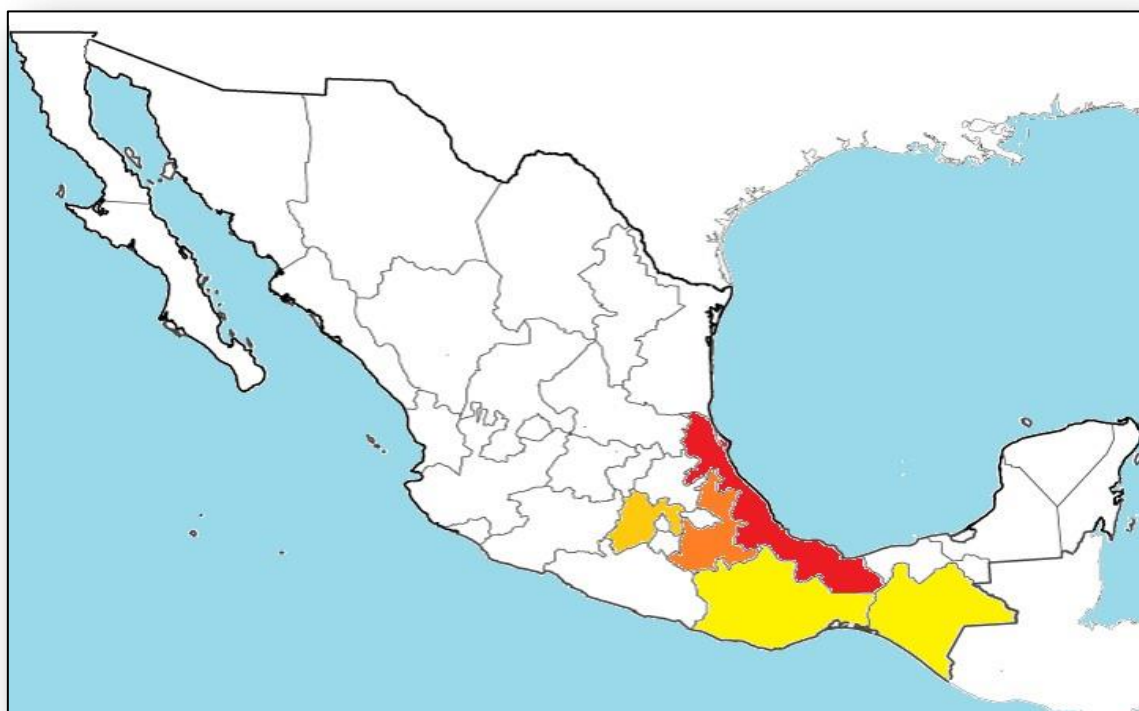
1.3 ETIOLOGÍA:

1.3.1 Características del parásito.

Trichomonas vaginalis; es un parásito extracelular del epitelio del tracto urogenital, cuya clasificación taxonómica se indica en la **Tabla 1**. Es un organismo microaerófilico o microaerotolerante, de 12-18 µm de largo por 5 µm de ancho, en su extremo anterior se distinguen 4 flagelos y se visualiza un quinto flagelo adherido al cuerpo del parásito por una membrana ondulante, estas estructuras le confieren una motilidad espasmódica que es una característica de este microorganismo. Cuando se observa al microscopio se aprecia un núcleo grande, ovalado, excéntrico hacia el extremo anterior (Schwebke, Desmond y col.; 2004; Swygard, Sena y col.; 2004; Mabey, Ackers y col.; 2006). El parásito adopta una

Fig. 1.- Epidemiología de *T. vaginalis*. A) Mapa de la República donde se señalan los cinco estados más afectados por la tricomoniasis durante el 2011. En orden decreciente, Rojo: Veracruz, Naranja: Puebla, Amarillo fuerte: Estado de México, Amarillo pálido: Oaxaca y Chiapas. B) Casos totales reportados por entidades en la República Mexicana durante el 2011 (Boletín Epidemiológico 2012, semana 1).

A)



B) CASOS EN EL 2011

Estado	Número de casos	Estado	Número de casos
Aguascalientes	883	Morelos	2,674
Baja California	884	Nayarit	1,611
Baja California Sur	346	Nuevo León	2,526
Campeche	590	Oaxaca	6,222
Coahuila	1219	Puebla	13,565
Colima	490	Querétaro	183
Chiapas	6,265	Quintana Roo	1,513

Chihuahua	2,358	San Luis Potosí	3,529
Distrito Federal	2,478	Sinaloa	1,993
Durango	2,539	Sonora	1,612
Guanajuato	5,040	Tabasco	3,173
Guerrero	4,727	Tamaulipas	5,076
Hidalgo	3,431	Tlaxcala	2,559
Jalisco	3,025	Veracruz	14,026
México	9,667	Yucatán	1,973
Michoacán	4,812	Zacatecas	2,612

forma ovoide cuando se encuentra en suspensión y una ameboide cuando se adhiere a sustratos, presentando morfología y tamaño variables (Honigberg, Gupta y col.; 1984; Rasmussen, Nielsen y col.; 1986; Arroyo, González-Robles y col.; 1993). Al entrar en contacto con las células del epitelio vaginal (VECs), se forman pseudópodos y filipodias adquiriendo su conformación ameboide, estas estructuras se interdigitan y se anclan a sitios específicos de la membrana de sus células blanco (Petrin, Delgaty y col.; 1998) (**Fig. 2**).

Tabla 1.- Taxonomía de *Trichomonas vaginalis*.

Taxonomía	Clasificación	Referencia
Grupo:	Excavata	Wegener y col., 2006
Subgrupo:	Arquezoa	Wegener y col., 2006
Clase:	Parabasalia	Wegener y col., 2006
Orden:	Trichomonadida	Honigberg., 1974
Familia:	Trichomonadidae	Wenyon., 1926
Género:	<i>Trichomonas</i>	
Especie:	<i>Trichomonas vaginalis</i>	Donne., 1836

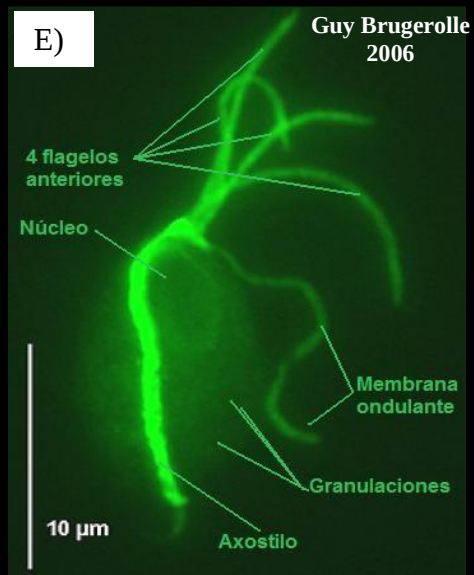
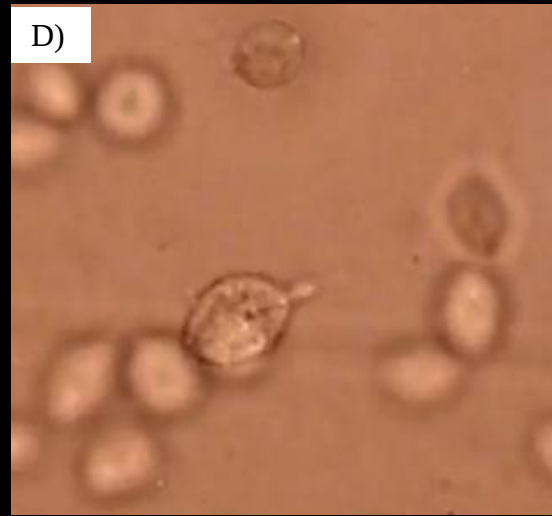
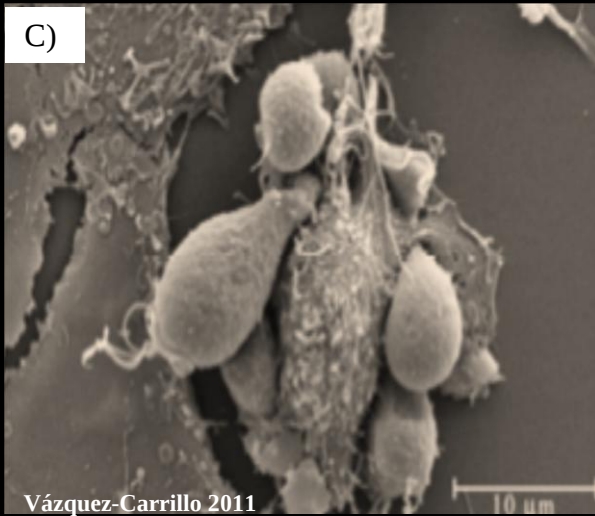
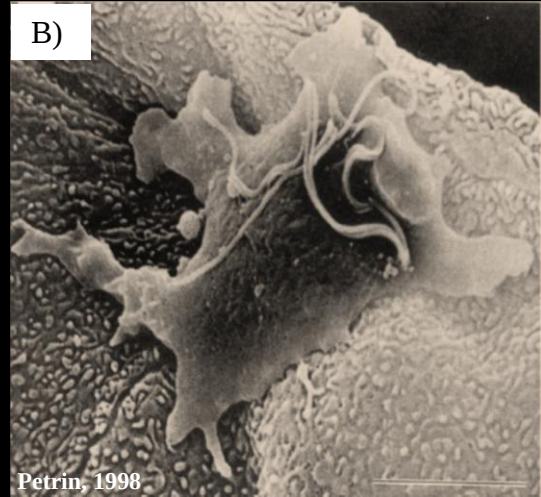
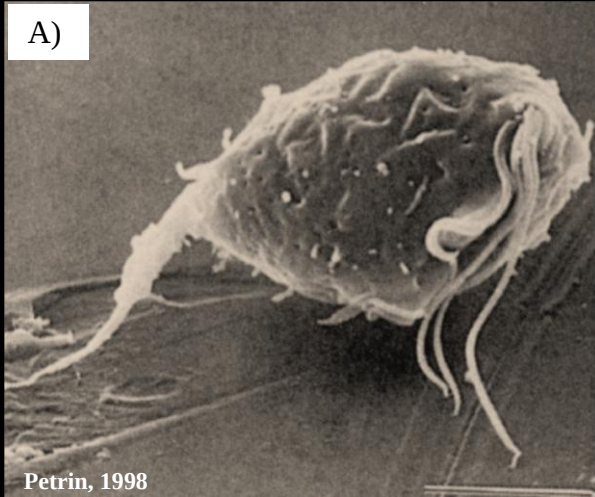
El citoesqueleto de *T. vaginalis* es altamente especializado compuesto por las proteínas: α -actina, β y γ tubulina, centrina, formina y profilina. Las fimbrinas son proteínas que se unen a la actina y forman retículas filamentosas. Las proteínas motoras son la dineína y la kinesina. El citoesqueleto posee una estructura fibrilar estriada denominada “costa” formada de proteínas contráctiles en el citoplasma

adyacente a la membrana ondulante que se conecta al cuerpo basal del flagelo recurrente y que es una característica distintiva de la familia Trichomonadidae. También posee una estructura rígida que surge a partir de los cuerpos parabasales llamada axostilo y cruza el cuerpo del parásito terminando en una espícula (**Fig. 2**) (Carlton, Hirt y col.; 2007).

El citoplasma de *T.vaginalis* es rico en carbohidratos, y contiene un gran número de vacuolas (incluyendo liposomas), carece de mitocondrias y peroxisomas pero posee una gran cantidad de hidrogenosomas (Schwebke y Burgess 2004; Carlton, Hirt y col.; 2007; Martincova, Voleman y col.; 2012), se ha sugerido que estos organelos sustituyen a las mitocondrias de los organismos aerobios y que ambos provienen de un organelo común. Los hidrogenosomas se caracterizan por carecer de ácido desoxirribonucleico (DNA) y por tener tres gránulos cromáticos (Schwebke y Burgess 2004), generan ATP a través de la vía enzimática de la piruvato ferredoxina oxido-reductasa (PFO), a diferencia de la mitocondria que utiliza la vía de la piruvato deshidrogenasa en organismos aerobios. La vía metabólica de PFO posee relevancia médica en tanto que esta enzima activa fármacos como los 5-nitroimidazoles, lo que hace susceptible a este parásito a los efectos del metronidazol, cuyo grupo nitro es reducido por la enzima y su producto desestabiliza la doble cadena de DNA (Leitsch, Kolarich y col.; 2009), lo mismo sucede con algunas bacterias que presentan esta vía, por lo cual se ha considerado a *T. vaginalis* como un eucarionte primitivo (Carlton, Hirt y col.; 2007).

Figura 2.- Morfología de *T. vaginalis*. A) *T. vaginalis* en su forma ovoide, se distinguen el axostilo, la membrana ondulante y los flagelos. B) Forma ameboide del parásito, adherida a una célula del epitelio vaginal, la parte que se adosa a la superficie celular es la contraria a la que se encuentra la membrana ondulante (Petrin y Col. 1998). C) Interacción de *T. vaginalis* con células prostáticas DU-145 donde se muestra que no cambia su forma ovoide (Carrilo y col. 2007). D) Aspecto de cultivo de *T. vaginalis*, por microscopía simple de luz, se observa la forma ovoide

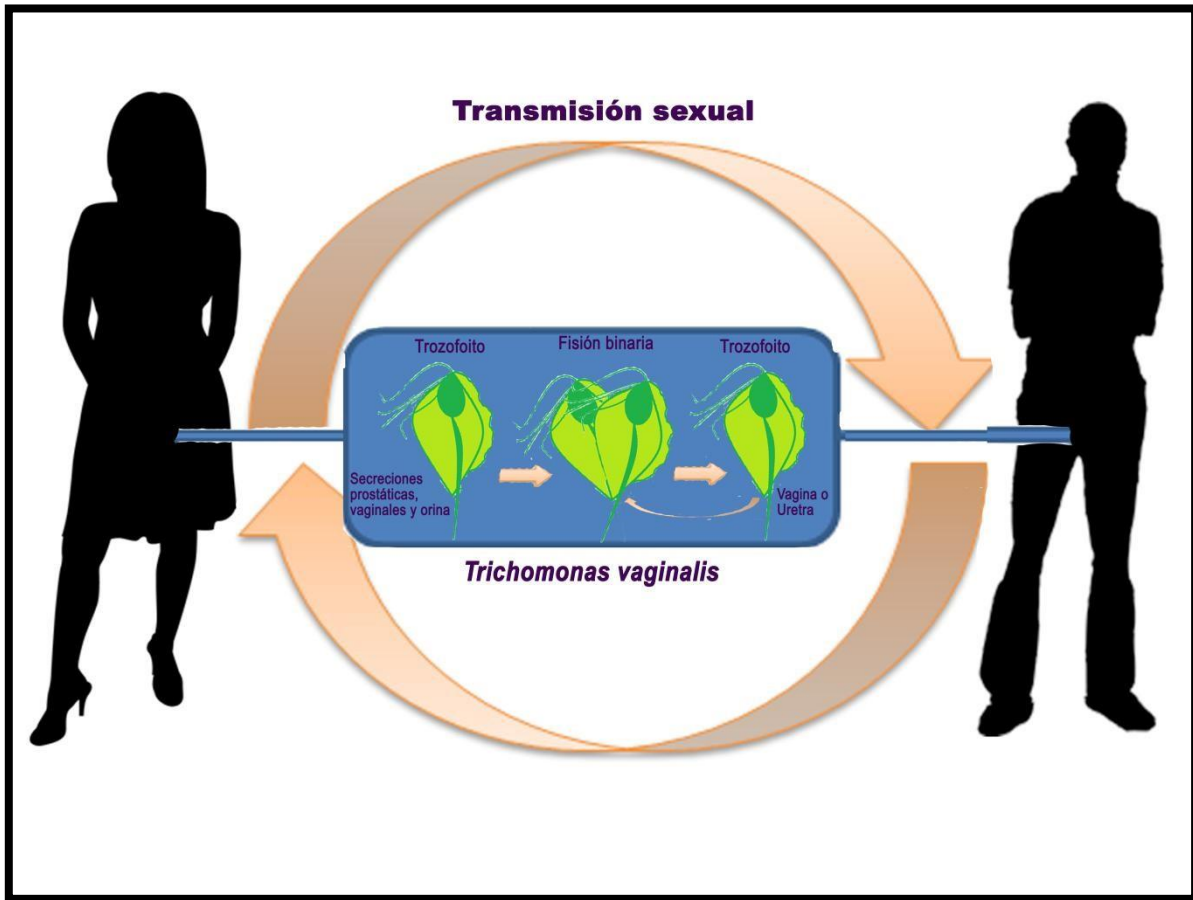
del parásito. E) Imagen de inmunofluorescencia con anticuerpo antitubulina donde se observan los cuatro flagelos anteriores, el flagelo recurrente con su membrana ondulante, el axostilo que cruza longitudinalmente el cuerpo del parásito y el borde del núcleo acéntrico, cercano a la región anterior (Guy Brugerolle).



1.3.1.1. Ciclo de vida de *Trichomonas vaginalis*.

T. vaginalis tiene un ciclo muy simple (**Fig. 3**), es un parásito obligado del epitelio urogenital humano, no existen formas libres de este microorganismo, no tiene huéspedes intermedios ni vectores, del mismo modo no se conocen fases quísticas. Se transmite exclusivamente por coito vaginal, aunque se considere también la transmisión vertical madre-hijo(a) durante el parto que puede afectar tanto las vías respiratorias como el tracto genital del recién nacido (Smith, Wang y col.; 2002; Carter y Whithaus 2008). Después de que el parásito se adhiere a las células del epitelio vaginal empieza multiplicarse por fisión binaria que *in vitro* sucede cada 4 hrs (Gómez-Conde, Mena-López y col.; 2000) hasta formar una monocapa. Pocos días después, los parásitos provocan degeneración y descamación del epitelio urogenital, con infiltración de leucocitos y aumento de las secreciones vaginales (Lehker y Alderete 2000). La infección puede ser tratada y resuelta con diferentes esquemas de antibióticos derivados del 5-nitroimidazol,

Figura 3.- Ciclo de vida de *T. vaginalis*. El parásito habita en el tracto urogenital femenino, así como en la uretra y próstata masculinas, donde se replica por fisión binaria cada 4 hrs. No se conocen formas quísticas del microorganismo y no sobrevive en el medio ambiente. *T. vaginalis* sólo se transmite entre humanos, que el único hospedero conocido, mediante contacto sexual directo o bien transmisión vertical madre-hijo o madre-hija, en los cuales se reportan casos de infección urogenital y en ocasiones afecciones pulmonares.



sin embargo en los casos en que los síntomas no son tan evidentes, o que no es tratada la infección o en aquellos en que se presente resistencia al antibiótico, se puede pasar a un estado de cronicidad con síntomas disminuidos.

1.3.1.2 Genoma de *Trichomonas vaginalis*.

El genoma de *T. vaginalis* fue secuenciado por el método “whole-genome shotgun” y publicado en el 2007. Presenta un tamaño de ~160 Mb, con 60,000 genes que codifican para proteínas. Se encontraron 65 genes con intrones, incluyendo los ~20 genes que presentan “splicing”, previamente documentados (Vanacova, Yan y col.; 2005; Carlton, Hirt y col.; 2007). El genoma

de *T. vaginalis* contiene los RNAs de transferencia (tRNA) que codifican para todos los aminoácidos y ~250 RNAs ribosomales (rRNA) localizados en un pequeño fragmento de secuencia continua dentro de uno de sus 6 cromosomas (Gunderson, Hinkle y col.; 1995; Carlton, Hirt y col.; 2007). Este genoma contiene 59 familias de repetidos que representan ~39 Mb del mismo, que están en cientos de copias localizados en pequeños fragmentos de secuencias continuas. Cada familia de repetidos es extraordinariamente homogénea con un promedio de polimorfismo de ~2.5%. El análisis de estas familias génicas indican que algunas sufrieron una expansión reciente, probablemente como adaptación del parásito al ambiente urogenital, puesto que las familias génicas que se expandieron corresponden a la maquinaria de tráfico de membranas, necesario para la secreción de proteínas patógenas y para la fagocitosis de células del hospedero y bacterias, otra familia génica expandida es la que codifica para alrededor de 880 proteínas-quinasas que constituyen el “kinoma” de *T. vaginalis* y que forman parte de diferentes vías de señalización. Además se encontraron 29 genes involucrados en el proceso de meiosis, muchos de ellos también participan en la reparación del DNA, pero 8 de ellos son específicos de recombinación meiótica que podrían participar en un ciclo sexual aún no descrito de *T. vaginalis* (Carlton, Hirt y col.; 2007).

En el 75% de las secuencias de las regiones 5' no traducidas (5'UTR) de los genes que codifican para diversas proteínas en *T. vaginalis*, se encontró el elemento promotor Inr común en otros eucariontes en forma conservada, con la secuencia TCA₊₁Y(T/C)W(T/A) que se encuentra entre 6 a 20 nucleótidos (nt) corriente arriba del inicio de la transcripción. Este elemento es reconocido por la proteína de unión a Inr (IBP39), la cual, interactúa con el dominio C-terminal de la RNA polimerasa II. En *T. vaginalis*, no se encuentra una secuencia TATA para el inicio de la transcripción, por lo cual este elemento Inr tiene un papel central en la expresión de genes (Carlton, Hirt y col.; 2007).

Los RNAm de *T. vaginalis* que contienen cap sugieren la existencia de la RNA polimerasa II, la cual es necesaria en el procesamiento del capping en el

extremo 5' al inicio de la transcripción (Liston y Johnson 1998; Liston, Carrero y col.; 1999; Liston y Johnson 1999). La RNA polimerasa II de *T. vaginalis* es inusual, pero se ha identificado por ser resistente a α -amanitina en un 50%, sugiriendo que también participa en la transcripción de varios genes en tricomonas (Liston y Johnson 1998). También ha sido descrita la RNA pol III, encargada de transcribir el gen que codifica para el rRNA 5Sr (Torres-Machorro, Hernandez y col.; 2006).

T. vaginalis es capaz de realizar "splicing"; proceso mediante el cual se eliminan los intrones de algunos genes. Reportes previos a la secuenciación del genoma de *T. vaginalis*, establecieron la presencia del gen que codifica para PRP8, considerada una proteína indispensable para este proceso (Fast y Doolittle 1999). Además, existen motivos conservados en el 5' y 3' de los intrones, que sirven como blanco para el complejo ribonucleoproteínico que conforma el "splicesoma". Estos motivos sirvieron para identificar por genómica comparativa de *T. vaginalis* con otras especies eucariontes, la presencia de intrones en genes ortólogos, los cuales tienen posiciones similares entre las diferentes especies. En *T. vaginalis* se encontró un arreglo altamente conservado de 12 nucleótidos en el 3' de dichos intrones (Vanacova, Yan y col.; 2005). Posteriormente, con la secuenciación del genoma de este parásito se demostró que varios genes de *T. vaginalis* presentan intrones (Carlton, Hirt y col.; 2007).

Con la secuenciación del genoma de este parásito, se hallaron genes "Dicer", "argonautas" y genes para helicasas, sugiriendo que *T. vaginalis* tiene la maquinaria para procesar RNA de interferencia (RNAi), aunque hasta el momento no existen reportes al respecto. Además, se encontraron 152 secuencias, que se especula, tienen su origen en la transferencia horizontal de genes, la mayoría participan en diferentes vías metabólicas y podrían tener su origen en las enterobacterias del tracto intestinal (Carlton, Hirt y col.; 2007). Algunas características generales del genoma se muestran en la **Tabla 2**.

Tabla 2.- Características del Genoma de *T. vaginalis*.

CARACTERÍSTICA	DATOS
Tamaño	160Mb
Número de cromosomas	6
Porcentaje de G+C	32.7%
Genes para proteínas	60,000
Porcentaje de secuencias repetidas	65%
Genes para Trna	20 aa
Genes con intrones	65
Secuencias con Inr en el extremo 5'UTR	75%
Casos de transferencia horizontal de genes	152
Vías localizadas:	RNAi Splicing

1.3.1.3 Requerimientos nutricionales de *Trichomonas vaginalis*.

T. vaginalis es capaz de fagocitar VECs, virus, bacterias, levaduras, eritrocitos y fragmentos celulares para la adquisición de nutrientes (López, Braga y col.; 2000; Pereira-Neves y Benchimol 2007). Los carbohidratos son la principal fuente de energía de *T. vaginalis* usando el metabolismo fermentativo tanto en condiciones aerobias como anaerobias; su fuente principal es la maltosa, aunque puede utilizar glucosa, galactosa, glucógeno o almidón. En el microambiente vaginal, utiliza los depósitos de glucógeno del epitelio vaginal como fuente de carbohidratos. Además, utiliza una variedad de aminoácidos como sustratos de energía, los cuales son catabolizados mediante las vías de la arginina hidrolasa, aminotransferasas y la glutamato deshidrogenasa, no obstante, estas reacciones pueden ser reversibles para sintetizar aminoácidos. *T. vaginalis* carece de las vías para sintetizar nucleótidos *de novo*, por lo cual depende de una fuente exógena como las células mencionadas, las cuales, al ser lisadas, liberan sus depósitos de nucleótidos y entran al parásito por difusión pasiva (Heyworth, Gutteridge y col.; 1982). El metabolismo de los lípidos está limitado a sintetizar *de novo* fosfatidiletanolamina, por lo que requiere de colesterol, fosfatidiletanolamina,

fosfatidilcolina y esfingomielina que también obtiene de fuentes exógenas (Beach, Holz y col.; 1990; Beach, Holz y col.; 1991).

T. vaginalis al fagocitar eritrocitos obtiene una fuente de Fe^{2+} , para el cual posee moléculas receptoras y transportadoras especializadas como la lactoferrina. El Fe^{2+} , es necesario para el parásito pues regula moléculas como la fibronectina, necesaria para la unión a su célula huésped (Alderete, Benchimol y col.; 2002), algunas moléculas se sobre-regulan en presencia de Fe^{2+} como las adhesinas AP23, AP33, AP51, AP65 AP120 (Alderete, Nguyen y col.; 2004; Moreno-Brito, Yáñez-Gómez y col.; 2005), TVLEGU1 (León-Félix, Ortega-López y col.; 2004); TVLip (Carvalho, Freitas y col.; 2005), NTPDasa 1, Ecto 5 nucleotidasa (Tasca, Bonan y col.; 2005) y TVCP4 (Solano-González, Burrola-Barraza y col.; 2007); mientras que otras se expresan más en ausencia de Fe^{2+} como los genes que codifican para proteínas parecidas a fibronectina 1 y 2 (Crouch y Alderete 2001); TVCP39 (Hernández-Gutierrez, Ortega-López y col.; 2003); TVCP12 (LeónSicairos, León-Felix y col.; 2004) y TVCP65 (Álvarez-Sánchez, Solano-González y col.; 2007). En estudios proteómicos se ha evidenciado la expresión diferencial de 45 proteínas en presencia de hierro las cuales participan entre otras funciones en el metabolismo energético, proteólisis y metabolismo hidrogenosomal de *T. vaginalis*, mientras que en ausencia de Fe^{2+} el parásito es capaz de internalizar los flagelos y el axostilo desarrollando una morfología de pseudoquiste (De Jesús, Cuervo y col.; 2007).

Otro elemento importante de dentro del metabolismo de *T. vaginalis* son las poliaminas, las cuales son aminas alifáticas de bajo peso molecular que se encuentran en todos los seres vivos. Las más estudiadas son la putrescina, espermidina y espermina, esta última sintetizada únicamente por organismos eucariontes. *T. vaginalis* no sintetiza poliaminas de *novo*, es por ello que necesita de su hospedero para obtener estas moléculas. La obtención de poliaminas la realiza mediante un sistema antiporte putrescina/espermina, en el cual, se excretan dos moléculas de putrescina para introducir una molécula de espermina. En *T.*

vaginalis la putrescina es sintetizada por la enzima ornitina descarboxilasa (ODC) mediante la vía de la arginina-hidrolasa. La espermina es procesada dentro de los hidrogenosomas y convertida a espermidina vía espermina N¹ acetiltransferasa (SSAT) acoplado a la vía de la poliamina oxidasa (Yarlett, Martínez y col.; 2000). La ausencia de poliaminas aumenta la adherencia del parásito a las VECs y cuando se agrega putrescina exógena disminuye la adhesión a la célula blanco (García, Benchimol y col.; 2005). Por otra parte, las poliaminas regulan la expresión del gen *tvcp65*. En ausencia de poliaminas los niveles de expresión de *tvcp65* disminuyen, al igual que la actividad proteolítica y la cantidad de CP65 (Álvarez-Sánchez, Carvajal-Gamez y col.; 2008). El factor de inicio de la traducción 5A (TveIF-5A) está regulado de manera positiva por las poliaminas ya que la expresión del gen y la cantidad de la proteína disminuyen en ausencia de poliaminas, los cuales son recuperados cuando se adiciona putrescina exógena al medio (Carvajal-Gamez, Arroyo y col.; 2011).

El Zn²⁺ actúa como cofactor y como integrante de al menos 300 enzimas como aldolasas, deshidrogenasas, esterasas, peptidasas, fosfatasa alcalina, anhidrasa carbónica, superóxido-dismutasa así como DNA y RNA polimerasas. Por lo cual está involucrado en el metabolismo energético y de hidratos de carbono, en las reacciones de biosíntesis y degradación de proteínas, en procesos biosintéticos de ácidos nucleicos y compuestos hemo, así como en el transporte de dióxido de carbono (CO₂). Se sabe que el proteoma total de un organismo eucarionte contiene 8.8% proteínas que unen Zn²⁺, esto refleja la importancia y el requerimiento de este catión por los distintos organismos (Andreini, Banci y col.; 2006; Maret y Li 2009). El microambiente del tracto urogenital masculino de una persona normal tiene concentraciones de Zn²⁺ tricomonicidas. Estudios de proteómica sobre la interacción de *T. vaginalis* con células prostáticas DU-145 en presencia o ausencia de Zn²⁺ a una concentración de 1.6 mM identificaron mediante electroforesis en doble dimensión (2-DE) y espectrometría de masas (MS) 37 péptidos con expresión diferencial, cuya secuencia fue analizada por MASCOT presentando identidad para: fimbrinas (gi|123493533, gi|123397260), adenosil-homocisteinasa

(gi|123499896,gi|123488577), cistein-peptidasa parecida a asparaginil endopeptidasa (gi|123408789) y una metaloproteasa parecida a aminopeptidasa P (gi|123445672). El análisis del proteoma e inmunoproteoma, permitió la identificación de una metaloproteína de 50 kDa sobreexpresada e inmunodetectada diferencialmente por el suero de pacientes masculinos con tricomosis (Vázquez-Carrillo, Granados y col.; 2011).

1.4 PATOGENIA

La vagina tiene un pH normal de 3.8 a 4.4, este pH ácido se debe principalmente a la flora vaginal normal, en especial a la presencia del bacilo de Döderlein que utiliza los depósitos de glucógeno de la mucosa para producir ácido láctico (McGrory y Garber 1992). Los parásitos adheridos a las VECs o células HeLa inducen la síntesis de grandes cantidades de adhesinas y la síntesis de estas, está regulada positivamente por Fe^{+2} , el cual, controla la tasa de crecimiento y multiplicación de este parásito (Lehker, Arroyo y col.; 1991; Arroyo, Engbring y col.; 1992). Posteriormente, se da una señal de quimiotaxis para reclutar más parásitos al sitio inicial de la infección. La destrucción del epitelio vaginal provocada por la infección de *T. vaginalis* disminuye los depósitos de glucógeno lo que puede llevar a la eliminación del bacilo de Döderlein, ocasionando que el pH vaginal se vuelva alcalino, lo cual, favorece la proliferación del parásito y otros microorganismos (Yano, Yui y col.; 1983).

La capacidad de causar una infección por parte de *T. vaginalis* depende de su interacción con el huésped, teniendo en cuenta desde su conducta sexual en relación al uso de preservativos, número de parejas sexuales, etc.(Lichtenstein, Desmond y col.; 2008), así como su propio sistema inmunológico; factores hormonales, la coexistencia con la flora vaginal y la relativa concentración de los organismos presentes en la vagina (Kumar, Sharma y col.; 1990). Además, las características del microambiente en el tracto urogenital, que implican factores como pH, temperatura, presencia de cationes como el Zn^{2+} , el Fe^{2+} y las poliaminas

(Álvarez-Sánchez, Carvajal-Gamez y col.; 2008; Vázquez-Carrillo, Granados y col.; 2011). Otro factor importante son las propiedades de virulencia del parásito que son:

1.4.1 Adhesión

La citoadherencia es un paso fundamental para la patogenia de *T. vaginalis*. En este mecanismo están involucradas proteínas de adhesión y del citoesqueleto. Se han descrito cinco adhesinas de *T. vaginalis* nombradas de acuerdo a su peso molecular, AP65, AP51, AP33, AP23 y AP120 (65, 51, 33, 23 y 210 kDa respectivamente), la interacción de estas adhesinas con las VECs, a través de una vía ligando-receptor, induce un incremento de la síntesis de adhesinas así como un cambio en la morfología ameboide del parásito, esto le permite adaptarse a los ambientes adversos como la menstruación (Alderete y Garza 1988; Arroyo, Engbring y col.; 1992). La α -tubulina es una proteína involucrada en la morfología del parásito (Kucknoor, Mundodi y col.; 2005). La cisteín-proteasa de 30kDa (CP30) además de ser una molécula involucrada en la adhesión, es capaz de degradar diferentes proteínas como la hemoglobina, fibronectina y colágeno IV (Mendoza-López, Becerril-García y col.; 2000). Otra molécula es el lipofosfoglicano (TvLPG), proteína de superficie de *T. vaginalis*, rica en ramnosa que sirve de ligando para el receptor galectina-1 de la superficie de las VECs, por lo que esta interacción puede ser importante para la adhesión del parásito (Hirt, Noel y col.; 2007). *T. vaginalis* posee 8 familias con alrededor de 800 proteínas de superficie que incluyen aproximadamente 650 proteínas semejantes a BspA, caracterizadas por el repetido rico en leucina de *Treponema pallidum* (TpLLR). Estas proteínas se expresan en la superficie de ciertas bacterias patógenas mediando la adhesión celular y la agregación (Carlton, Hirt y col.; 2007). Otro grupo de proteínas involucradas en la adhesión son las parecidas a GP63, las cuales son homólogas a las leishmanolisinas y algunas presentan un dominio transmembranal lo que indica su participación en la cito-adhesión (Carlton, Hirt y col.; 2007).

Otras proteínas expresadas durante la citoadherencia a VECs son: la “GTPBP” proteína de unión a GTP involucrada en el endocitosis (Kucknoor, Mundodi y col.; 2005), la “PDI” proteína disulfuro isomerasa con función de chaperona siendo una proteína involucrada en el plegamiento de proteínas (Kucknoor, Mundodi y col.; 2005); la adhesina enolasa (TvENO-1) (Mundodi, Kucknoor y col.; 2008); la adhesina gliceraldehido-3-fosfato-deshidrogenasa (TvGAPDH) (Lama, Kucknoor y col.; 2009) y la adhesinas Pmp; TvSUB1; parecida a M17; y parecida a VSP (Hirt, Noel y col.; 2007)

1.4.2 Citotoxicidad

T. vaginalis posee una gran cantidad de proteinasas como metaloproteinasas, cisteín-proteinasas, serin-proteinasas, y aspartil-proteinasas, las cuales, son reguladas por factores como Zn^{2+} , Fe^{2+} , pH; agentes quelantes como el EDTA y EGTA e inhibidores como N-a-p-tosil-L-lisina cloro-metilcetona (TLCK) y leupeptina (Alderete y Garza 1985). Las cisteín-proteasas (CPs) son endopeptidass que contienen una cisteína en su sitio activo. *T. vaginalis* posee diferentes CPs involucradas en diversas propiedades de virulencia como adhesión, hemólisis, citotoxicidad y evasión de la respuesta inmune (Álvarez-Sánchez, Ávila-González y col.; 2000; Hernández-Gutierrez, Ortega-López y col.; 2003).

La cisteín-proteasa de 65 kDa (CP65) participa en la citotoxicidad hacia las células HeLa (Álvarez-Sánchez, Ávila-González y col.; 2000) y se encuentra localizada tanto en el citoplasma como en la superficie de la membrana; degrada componentes de la matriz extracelular como el colágeno IV y la fibronectina. Su actividad óptima se da a un pH de 5.5 y una temperatura de 37°C. A nivel de transcrito es regulado negativamente por Fe^{2+} y Zn^{2+} pero positivamente en presencia de poliaminas. A nivel postraducciona presenta hipusinación. Se han detectado anticuerpos contra esta proteína en suero de pacientes con tricomonosis, lo que sugiere su alta inmunogenicidad (Álvarez-Sánchez, Ávila-González y col.; 2000; Álvarez-Sánchez, Carvajal-Gamez y col.; 2008).

La cisteín-proteasa de 39 kDa (CP39) se encuentra en las secreciones de pacientes con tricomonosis, su función en la citotoxicidad es por su capacidad de degradar sustratos como el colágeno I, III, IV y V, la fibronectina, hemoglobina humana e inmunoglobulinas A y G. El pH óptimo en que realiza su actividad proteolítica es 7, pero trabaja en un amplio rango de pH 3.9 a 9. También se localiza tanto en citoplasma como en la superficie celular y es altamente inmunogénica, además en presencia de Zn^{2+} y Fe^{2+} se regulan negativamente a nivel postranscripcional y en presencia de poliaminas tiene regulación positiva. Esta proteína sufre modificaciones postraduccionales como son glicosilación e hipusinación (Hernández, Sariago y col.; 2004; Ramón-Luing, Rendón-Gandarilla y col.; 2010) (Datos no publicados de Carrillo y col.).

Otras proteínas involucradas en la citotoxicidad de *T. vaginalis* son: TvLPG (Paschinger, Hykollari y col.; 2012); TvCP30 (Alvarez-Sanchez, Avila-Gonzalez y col.; 2000; Álvarez-Sánchez, Ávila-González y col.; 2000); Fosfolipasa A2 (Lubick y Burgess 2004) y TvCP12 (León-Sicairos, León-Felix y col.; 2004).

1.4.3 Fagocitosis.

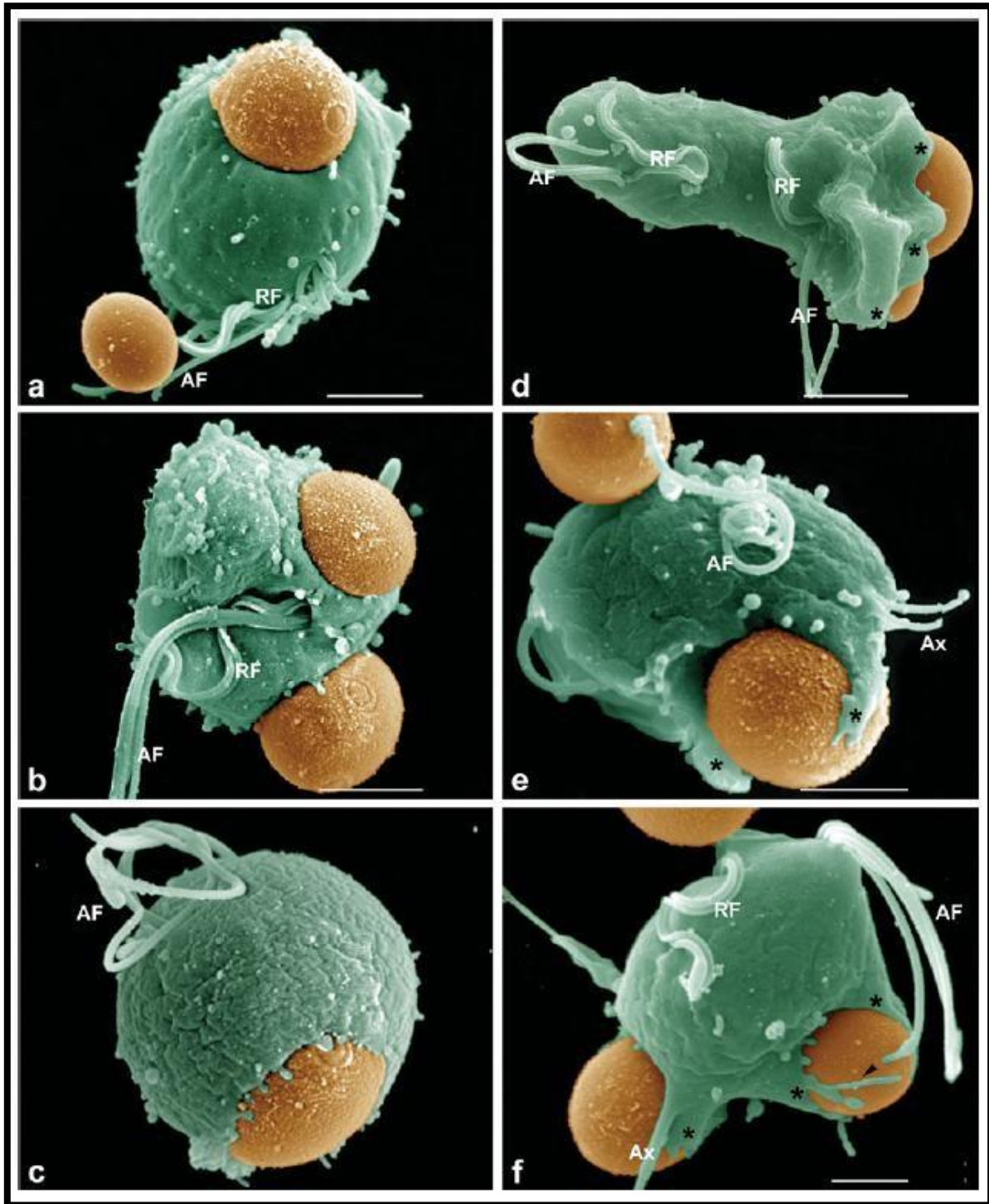
Se considera a *T. vaginalis* como una célula fagocítica debido a su capacidad de ingerir eficientemente partículas inertes y células por un mecanismo similar al que usan otras células fagocíticas (Benchimol, Batista y col.; 1990). Algunos estudios indican que la citofagocitosis constituye un mecanismo patogénico de *T. vaginalis*. Diferentes estudios por microscopía han demostrado que *T. vaginalis* puede fagocitar bacilos de Döderlein, levaduras, leucocitos, eritrocitos, fragmentos de VECs células prostáticas y espermatozoides (Benchimol, de Andrade Rosa y col.; 2008). Se han descrito dos formas en que se lleva a cabo la fagocitosis (**Fig. 4**); la primera es similar a la utilizada por los fagocitos profesionales en la cual, los flagelos y/o pseudopodias se extienden hacia la célula blanco para atraerla hacia el cuerpo del parásito y otra en la cual, células más pequeñas que el parásito como las levaduras se hunden en la membrana del parásito sin la formación de

pseudopodias o la participación de los flagelos. La interacción con los diferentes tipos celulares parece responder a receptores específicos como los de D-manosa y L-fucosa en la superficie membranal del parásito; una vez adentro, el contenido internalizado es degradado por lisosomas y los nutrientes utilizados por los parásitos (Pereira-Neves y Benchimol 2007). La capacidad fagocítica de *T. vaginalis* se ve disminuida en presencia de colchicina y citocalasinas B y D, lo cual indica un papel fundamental del citoesqueleto de actina durante la fagocitosis (Juliano, Monaco y col.; 1987; Pereira-Neves y Benchimol 2007).

1.4. 4 Evasión de la respuesta inmune.

Para que la infección por *T. vaginalis* no sea combatida, es necesario que el parásito evada la respuesta inmune del hospedero, ha sido demostrado que al interaccionar con las VECs, en presencia de Fe^{2+} la expresión de un inmunógeno de superficie, llamado P270 se regula negativamente. P270 es una proteína antigénica de la superficie membranal del parásito (Alderete 1999) y está

Figura 4.- Mecanismos de fagocitosis de *T. vaginalis*. Se muestran dos formas en que la ticomona (verde) fagocita células de levaduras (naranja). la primera forma involucra un proceso de hundimiento sin participación de extensiones de la membrana plasmática (a-c). La segunda forma es semejante a la fagocitosis clásica de los fagocitos profesionales, donde los pseudópodos (asteriscos) y filipodias (puntas de flecha) se extienden hacia la célula blanco (d-f), la cual es internalizada posteriormente (flechas blancas) (d). AF, flagelo anterior; Ax, axostilo; RF flagelo recurrente. La barra de escala es de 4 μ m. (Pereira-Neves y Benchimol 2007)



relacionada con la presencia de un virus de RNA de doble cadena (ds-RNA), cuya presencia ha sido asociado a los aislados más virulentos (Conrad, Gorman y col.;

2012). Otra proteína de superficie es P230 que se relaciona con cambios en la accesibilidad al epítipo del antígeno. Ambas proteínas son responsables de la variación fenotípica del parásito para evadir la respuesta inmune (Alderete 1999). Durante la menstruación el sistema del complemento de la mujer puede eliminar al parásito, no obstante, ante un ambiente rico en Fe^{2+} se sobreexpresan algunas cisteín-proteasas en la superficie que degradan al factor del complemento C3b. Del mismo modo, se expresan las cisteín-proteinasa de 60 y 39 kDa (CP60 y CP39) que degradan inmunoglobulinas IgG, IgA e IgM. Además, CP60 también degrada hemoglobina (Lehker y Alderete 2000). Otra de las formas en que *T. vaginalis* evade el sistema inmune es mimetizándose con las proteínas del hospedero. Algunas adhesinas de *T. vaginalis* comparten gran homología con enzimas metabólicas del hospedero, tales como AP65, AP51 y AP33 que tienen homología con la enzima málica, y las subunidades α y β de la succinil CoA respectivamente (Alderete, Millsap y col.; 2001). Es importante mencionar que *T. vaginalis* no sólo se mimetiza con proteínas codificadas en su genoma, sino también es capaz de tomar proteínas del hospedero y revestirse de ellas en su membrana, de manera que no pueda ser inmunorreconocido por el sistema inmune (Lehker y Alderete 2000; Alderete, Millsap y col.; 2001). Otro mecanismo de evasión consiste en liberar una gran cantidad de proteínas inmunogénicas solubles que pueden neutralizar los anticuerpos y a los linfocitos T citotóxicos. Además, tiene la capacidad de reclutar neutrófilos por la vía de la interleucina 8 (IL-8), pero también de inducirles apoptosis por la activación de la vía de la caspasa 3 dependiente de especies reactivas de oxígeno (ROS) (Song, Shin y col.; 2008).

1.4.5 Apoptosis

T. vaginalis induce apoptosis por dos formas, en la primera se requiere la adhesión a las VECs y neutrófilos lo que desencadena un incremento en la actividad de la caspasa-3 (proapoptótica) y disminución de Mcl-1 perteneciente a la familia de Bcl-2 (proteína antiapoptótica) en las células del hospedero. Se probó que la apoptosis está ligada a la adhesión, al tratar a los parásitos con un anticuerpo anti-

AP65, de esta forma, al inhibir la función de esta adhesina se redujo la apoptosis (Kang, Song y col.; 2006). También se ha demostrado que la infección por *T. vaginalis* causa desprendimiento de las VECs seguida por destrucción de las mismas como resultado de la apoptosis inducida por la secreción de CPs en el rango de los 30 kDa (CP2, CP3, CP4 y CPT) que tienen una función de porinas, estas proteínas se insertan en la bicapa lipídica de las células del hospedero formando canales transmembranales de un diámetro de 1.14 y 1.34 nm, los cuales conducen a la muerte por lisis osmótica (Dailey, Chang y col.; 1990; Fiori, Rappelli y col.; 1999; Kummer, Hayes y col.; 2008).

1.4.6 Hemólisis

Los eritrocitos son una fuente importante de nutrientes para *T. vaginalis* y su capacidad beta-hemolítica se correlaciona con su virulencia clínica. Se han identificado algunas moléculas involucradas en la hemólisis por *T. vaginalis* tales como CPs de superficie (TvCP65), proteínas formadoras de poro (TvCP4) y proteínas tipo fosfolipasa A (TvLIP) (Krieger, Poisson y col.; 1983; Dailey, Chang y col.; 1990; Carvalho, Freitas y col.; 2005).

II.- ESTADO DEL ARTE

2.1 EXPRESIÓN DE GENES

La regulación de la expresión génica es fundamental para todas las células, puesto que les permite llevar a cabo su desarrollo, homeostasis y adaptación al medio ambiente. En eucariontes, cada paso de la expresión génica, que incluye cambios en la estructura de la cromatina, transcripción, procesamiento del transcrito, su transporte al citoplasma, la traducción del RNAm a proteína, transporte, modificación y activación postraducciona; está sometido a una estricta regulación (Gómez, Esther Ramírez y col.; 2010).

La transcripción es el proceso mediante el cual, una molécula de RNA es sintetizada a partir de un DNA que sirve como templado. Este proceso se divide básicamente en 3 pasos: iniciación, elongación de la cadena y terminación, cada uno de estos pasos es susceptible de regulación (Gómez, Esther Ramírez y col.; 2010). Además de estos pasos básicos que tienen que ver con la síntesis de la molécula de RNA (pre-RNA_m) que va a dar origen al RNA_m, el transcrito necesita procesarse para poder cumplir su función biológica. Este procesamiento postranscripcional se da inmediatamente después de iniciar la transcripción, y consiste en tres modificaciones del transcrito primario. La primera modificación es el “capping”, el cual ocurre después de que se sintetizan los primeros 25-30 nucleótidos del pre-RNA_m y consiste en tres reacciones, en la primera, la enzima RNA trifosfatasa actúa sobre el nucleótido terminal del extremo 5' y remueve el fosfato- γ , después la enzima guanililtransferasa adiciona una guanina, la cual es posteriormente metilada por la acción de la enzima 7-metiltransferasa para finalizar la estructura llamada “cap”. En mamíferos las primeras dos reacciones se llevan a cabo por la misma enzima, mientras que en la levadura, estas reacciones son llevadas a cabo por enzimas diferentes. El “cap” ayuda a preservar al RNA_m de la acción de las exonucleasas 5'-3' presentes tanto en el núcleo como en el citoplasma (Shuman 2001; Hocine, Singer y col.; 2010). La segunda modificación es llamada “splicing” y consiste en la eliminación de intrones y la posibilidad de empalmar y combinar diferentes exones, esta reacción es catalizada por el “spliceosoma” que consiste en las pequeñas ribonucleoproteínas nucleares (snRNPs) U1, U2, U4, U5 y U6; en conjunto con alrededor de 125 proteínas diferentes. El “spliceosoma” reconoce motivos en la secuencia del preRNA_m que definen el 5' y 3' de los intrones. En los eucariontes superiores el sitio de “splicing” no sólo se define por estos motivos, sino que intervienen otros elementos reguladores en trans, incluyendo potenciadores y silenciadores de intrones y exones que reclutan o limitan el acceso a un determinado sitio de “splicing”. El resultado final de este proceso es eliminar secuencias no codificantes de los genes y dar formas alternativas de RNA_m que conducen a isoformas de proteínas o proteínas con diferentes funciones codificadas por un mismo gen (Hocine, Singer y col.; 2010). El paso final de la transcripción consiste

en el procesamiento del extremo 3'UTR del RNAm y se realiza en dos pasos importantes que son el corte de la cadena de RNA recién sintetizada y la adición de un polímero de adeninas de aproximadamente 200pb. Este proceso ha sido recientemente descrito en diferentes organismos, así como algunas de sus implicaciones biológicas, no obstante la regulación del proceso en sí mismo, aún es poco entendido, debido a la complejidad de las interacciones no sólo entre las moléculas que intervienen en el proceso, sino por su estrecha relación con moléculas que tienen que ver tanto en el inicio de la transcripción, el “splicing” y el inicio de la traducción (Proudfoot 2011). En un primer acercamiento para entender la regulación del procesamiento del extremo 3'UTR en diferentes organismos, el conocer los elementos que intervienen en dicho proceso es de primordial importancia.

2.2 GENERALIDADES DEL PROCESAMIENTO DEL EXTREMO 3'UTR DE LOS RNAm.

Mediante estudios de digestión con RNAasa pancreática que corta residuos de citosina y U y la RNAasa T1 que corta residuos de guanina se demostró la existencia de fragmentos de RNAm resistente a la acción de estas enzimas por lo que se dedujo que eran secuencias de A y que estos segmentos de RNAm no tenían origen en un templado de DNA, por lo cual se hipotetizó la existencia de una polimerasa que sintetizara dichos fragmentos *de novo*, a la cual se le llamó PAP (Polimerasa de Poli-A) (Lim, Canellakis y col.; 1970; Winters y Edmonds 1973; Winters y Edmonds 1973). En poco tiempo este tracto de poli-A se convirtió en el blanco para obtener RNAm de genes particulares mediante la tecnología del DNA recombinante. Se usaron oligos dT, los cuales complementan a los tractos de poli-A de los RNAm. El RNA fue extraído de muestras del tejido que expresa el gen de interés, de esta forma, el RNAm de la globina fue purificado de eritrocitos y el RNAm de la inmunoglobulina de las células B murinas. Posteriormente con el empleo de la transcriptasa reversa en *Escherichia coli* se crearon las primeras librerías de genes

y se hizo notorio el hecho de que para el mismo RNAm existen diferentes tamaños de colas de poli-A y de su secuencia 3'UTR (Proudfoot 2011).

Con el uso de los métodos de secuenciación, se pudo determinar la presencia de una secuencia constante en la mayoría de los RNAm que poseen un tracto de poli-A, esta secuencia fue llamada señal de poliadenilación (PAS) y consiste en los nucleótidos AAUAAA, los cuales se encuentran 20-30 nucleótidos previos al tracto de poli-A, mediante ensayos de restricción por exonucleasas se confirmó que esta secuencia es esencial para llevar a cabo la poliadenilación y de la misma forma se descubrió una secuencia rica en GU río abajo del sitio de poliadenilación llamadas Elemento Río Abajo (DSE) que también es indispensable para el proceso de corte y poliadenilación, además de otra secuencia rica en GU, río arriba (USE) de la PAS, la cual resultó ser un potenciador de dicho proceso(Proudfoot 2011).

Posteriormente en células HeLa se obtuvieron dos complejos proteínicos que se unen a estas secuencias, el primero fue el factor de especificidad de corte y poliadenilación (CPSF, cleavage and polyadenylation specificity factor) que se une a la PAS, además del factor estimulante del corte (CstF, cleavage stimulation factor) que se une al DSE. Otras proteínas como los factores de corte I y II, la polimerasa de poli-A fueron asociadas a esta maquinaria y fueron buscados sus elementos homólogos en *S. cerevisiae* donde se observó la pérdida de algunas de estas proteínas en extractos sometidos a temperatura restrictiva, lo cual dió un panorama nuevo de la importancia de este proceso y su funcionamiento (Proudfoot 2011).

El estudio del procesamiento del extremo 3'UTR del RNAm ha revelado su importancia en la diferenciación celular, desarrollo tisular y modulación de la expresión proteica. De esta forma se sabe que el tamaño del extremo 3'UTR incluyendo su cola de poli-A, determina la estabilidad y la vida media del RNAm, a mayor longitud encontramos mayor estabilidad y vida media, no obstante los RNAm con extremos 3'UTR cortos tienen una mayor traducción aunque su vida media es

más corta. Un ejemplo de que la longitud del extremo 3'UTR para el mismo RNAm juega un importante papel regulador, lo encontramos en las diferentes etapas del desarrollo, siendo que en las primeras etapas, un RNAm tiene un 3'UTR más corto mientras que en etapas más avanzadas, esta región del mismo RNAm es más larga (Ji, Lee y col.; 2009). Del mismo modo, se ha determinado esa diferencia entre tejidos diferenciados del mismo individuo (Sandberg, Neilson y col.; 2008). En células cancerígenas se ha observado que los oncogenes expresados tienen un extremo 3'UTR más corto que el de las células sanas homólogas, esto cobra relevancia a la luz de que es precisamente el extremo 3'UTR en donde se encuentran los blancos principales de los microRNAs (miRNAs). Del mismo modo, el segmento 3'UTR de los RNAm sirve como blanco para moléculas como los cationes Fe^{2+} y Zn^{2+} los cuales modulan la expresión génica (Lutz y Moreira 2011).

2.3 MECANISMOS ALTERNATIVOS DE POLIADENILACIÓN

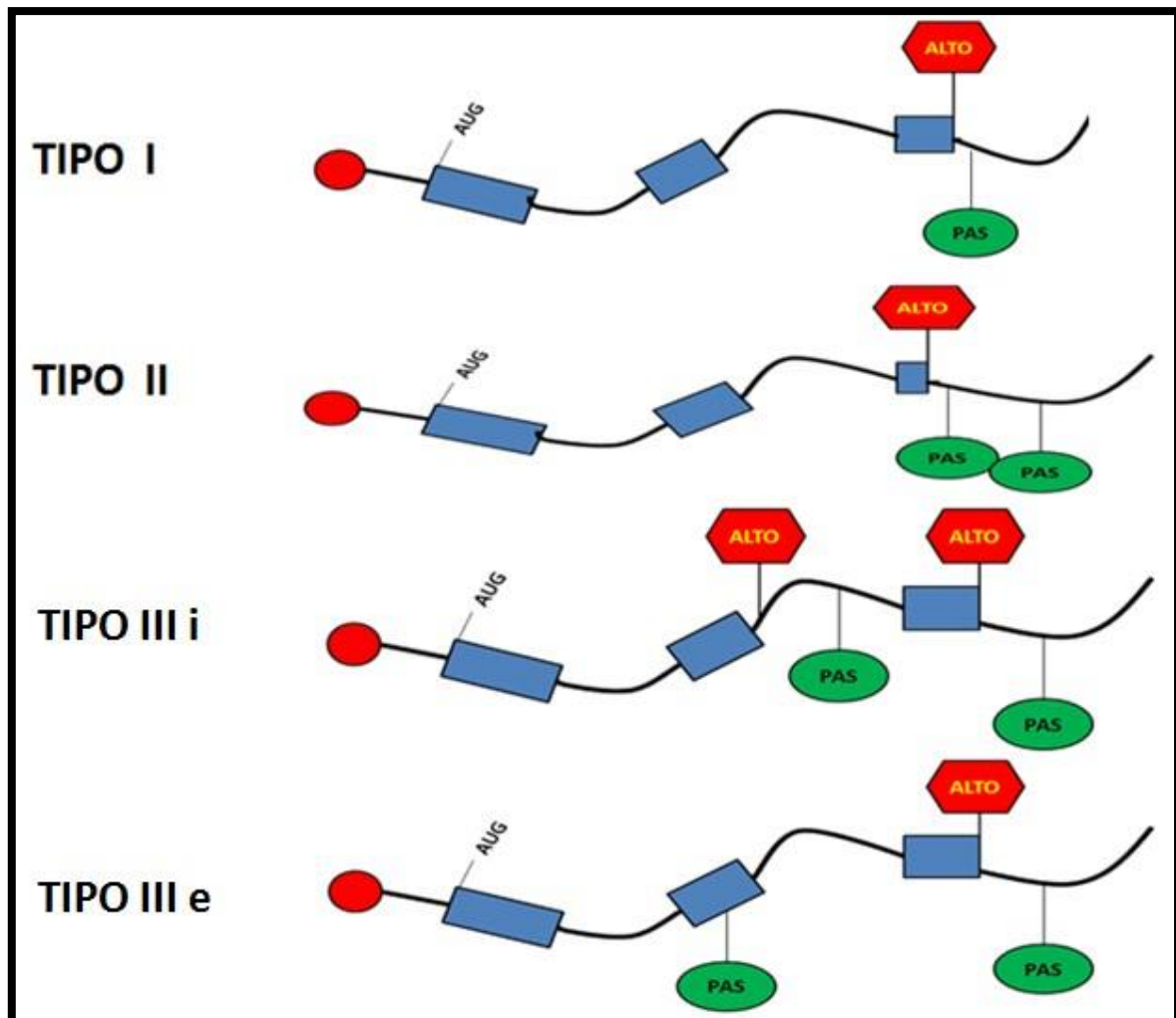
La presencia de más de una PAS en la secuencia de un gen, abre un abanico de posibilidades para llevar a cabo el procesamiento de su extremo 3'UTR. Esta posibilidad de un procesamiento alternativo del RNAm es un importante nivel regulatorio en la expresión génica, como se aprecia al considerar que más de la mitad de los RNAm de humano tienen poliadenilación alternativa, además, que este mecanismo de regulación está conservado evolutivamente en las plantas. Aunque aún no se entiende a detalle los mecanismos que controlan la poliadenilación alternativa; se pueden clasificar diferentes tipos de este procesamiento (**Fig. 5**) con diferentes implicaciones (Lutz y Moreira 2011):

El tipo de **poliadenilación I** corresponde al más común, se trata de un gen que presenta un codón de paro y en su 3'UTR presenta una sola señal de poliadenilación, nos da siempre una isoforma de RNAm. El tipo de **poliadenilación II** corresponde a un gen en cuya secuencia presenta un solo codón de paro, pero en su extremo 3'UTR tiene dos o más PAS, por lo que sus RNAm pueden tener un

extremo 3'UTR corto o largo pero no altera su marco de lectura, esto se ve con varios genes en metazoarios, los cuales se expresan en diferentes tejidos con diferentes longitudes de RNAm; el efecto final, son diferentes cantidades de la misma proteína. El tipo de **poliadenilación IIIi** corresponde a genes que por efecto de “splicing” tienen codones de paro alternativo y señales de poliadenilación ocultas en los intrones, el efecto de esto son distintas isoformas de proteínas. El tipo de **poliadenilación IIIe** se refiere a que por efecto de “splicing” puede quedar al descubierto una PAS dentro de un exón y esto provocar la poliadenilación temprana y producir proteínas truncas o con diferentes funciones (Lutz y Moreira 2011).

En estudios de activación de linfocitos T, los cuales, posterior a su activación, aumentan la poliadenilación de sus RNAm y la cantidad de proteínas expresadas; se determinó que las PAS proximales eran las más utilizadas y que esta variación en la poliadenilación alternativa no se detecta a las 6 hrs pero era evidente a las 48 h, sugiriendo que la selección de una PAS alternativa no es un evento inmediato y que probablemente sea necesario la transcripción de factores de corte y poliadenilación basales y/o auxiliares. Sin embargo, hasta el momento se desconoce como las señales de la superficie membranal se traducen a la maquinaria de corte y poliadenilación (Sandberg, Neilson y col.; 2008).

Figura 5.- Tipos de poliadenilación alternativa. En este esquema se representan diferentes RNAm con su cap (círculo rojo), exones (cuadros azules, intrones (línea negra). Los RNAm con poliadenilación tipo I sólo tienen una señal de poliadenilación (PAS) en su último exón. Los RNAm con poliadenilación tipo II tienen más de una señal de poliadenilación en su último exón. La poliadenilación tipo III está acoplada con el “splicing”, las señales tipo IIIi están en los intrones corriente arriba de la PAS canónica, mientras que las tipo IIIe tienen una o más PAS en exones corriente arriba.



2.4.- ELEMENTOS DEL PROCESAMIENTO DE LOS 3'UTR DE LOS RNAm DE HUMANOS.

2.4.1.-Elementos de secuencia

También llamados elementos *cis*, son secuencias específicas en casi todos lo pre-RNAm que son poliadenilados y las alteraciones de estos elementos afectan la eficacia del procesamiento del extremo 3'. En el humano los elementos de secuencia son cuatro, el primero define el sitio de poliadenilación y es llamado señal de poliadenilación (**PAS**), corresponde al hexámero AAUAAA que se encuentra en 80 a 90% de los RNAm secuenciados o una variante que es AUUAAA presente en el 10% de los casos, esta secuencia señal se encuentra de 10-35 bases de distancia corriente arriba del sitio de poliadenilación. El segundo elemento se encuentra alrededor de 40 bases corriente abajo del sitio de poliadenilación (**DSE**) este elemento se ha observado con dos secuencias, la primera es una secuencia rica en GU que contiene YGUGUUY (donde Y= pirimidina) y la segunda es una secuencia rica en U (UUUUU), el pre-RNAm puede tener alguna, ambas o ninguna de estas secuencias; este elemento se localiza aproximadamente a 30 nucleótidos del sitio de corte. El sitio de corte del pre-RNAm se sitúa entre PAS y DSE dentro de una región de 13 nucleótidos que no está estrictamente conservada, el corte se realiza en lado 3' de un residuo de adenosina que se encuentra seguido por un nucleótido que tiene como preferencia para hacer el corte A>U>C>>G. El nucleótido que precede a la adenina del sitio de corte es la citosina en el 59% de los casos, refiriéndose como el sitio de corte óptimo al dinucleótido **CA**. El elemento auxiliar río arriba (**USE**), se encuentra por arriba de PAS y frecuentemente consiste en una secuencia rica en U (UUUU) o similar (UGUA, UAUUA). La eficiencia del corte y poliadenilación aumenta con la presencia de este elemento, pueden existir otros elementos auxiliares río abajo del sitio de corte y poliadenilación, algunos se sabe que son ricos en G pero son secuencias no conservadas y que varían de distancia a CA (Millevoi y Vagner 2010) (Mandel y col., 2008; Danckwardt y col., 2008; Millevoi y col., 2010).

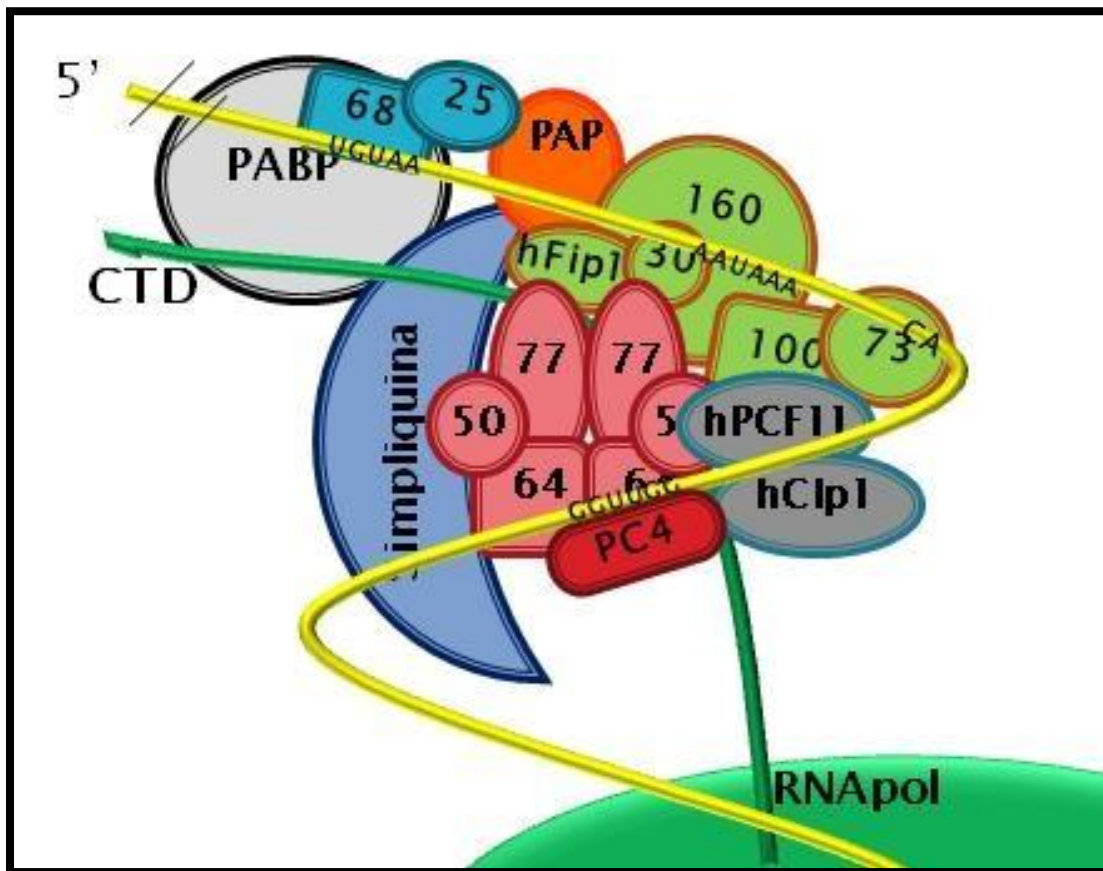
2.4.2 Maquinaria de procesamiento del extremo 3'UTR del RNAm del humano y sus complejos.

Alrededor de 14 proteínas conforman la maquinaria esencial de procesamiento del extremo 3'UTR, mediante microscopía electrónica se ha observado que el núcleo de la maquinaria de procesamiento ensamblada tiene una forma de riñón y una longitud de alrededor de 250 Å.(Shi, Di Giammartino y col.; 2009). Algunas de estas proteínas forman los complejos principales de este proceso (**Fig. 6**), que son: el factor de especificidad de corte y poliadenilación (CPSF, cleavage and polyadenylation factor), el factor estimulante del Corte (CstF, cleavage stimulation factor), el factor de corte I (CFIm), el factor de corte II (CFIIm), así como otras enzimas que no forman complejos como la polimerasa de poli-A (PAP, polyA polymerase) y la proteína de unión a la cola de poli-A (PABP, polyA binding protein), además de otras proteínas como la simpliquina y otros elementos como el extremo carboxilo terminal de la polimerasa II de RNA (CTDRNAPolII) que ayudan a anclar e integrar todos estos complejos y por lo tanto son indispensables para llevar a cabo el procesamiento del extremo 3'UTR de los RNAm (Mandel, Bai y col.; 2008).

2.4.2.1 Complejo CPSF

El factor de especificidad de corte y poliadenilación (CPSF) en el humano, está constituido por 5 subunidades, cuatro de ellas se diferencian por su peso molecular (160, 100, 73 y 30 KDa), y la quinta se denomina hFip1 debido a que primero fue descrita en *Saccharomyces cerevisiae*. La subunidad CPSF160 es la más conservada en los eucariontes y dirige la reacción de corte y poliadenilación al identificar a PAS, utilizando con mayor afinidad la secuencia AAUAAA, esta proteína contiene dos motivos de unión a RNA cerca de su extremo N-terminal (Murthy y Manley 1995). La subunidad CPSF-73 es importante para la viabilidad celular como lo demuestran estudios de disrupción de esta proteína (Chanfreau, Noble y col.; 1996). Contiene un dominio metalo- β -lactamasa en su región N-

Figura 6.- Maquinaria de poliadenilación del humano. Complejo CPSF de color verde con bordes cafés, con sus subunidades de 160, 100, 73 y 30 KDa y Fip1 en contacto con la señal de poliadenilación; complejo CstF de color rosa con bordes rojos y sus subunidades de 77, 64 y 50 kDa, la subunidad 64 en contacto con el elemento corriente abajo y la subunidad 77 interaccionando con CPSF; complejo CFI de azul con bordes grises con su heterodímero más frecuente (25/68), interaccionando con el elemento corriente arriba; complejo CFII de color gris Oxford con bordes azules con sus dos subunidades; las proteínas sin complejo, PAP1, PAPB, simpliquina y CP4 esta última en su posición entre CstF64 y el elemento corriente abajo. En amarillo RNA recién sintetizado con los elementos de secuencia.



terminal que coordina dos iones de Zn^{2+} , por lo cual, se ha postulado como la endonucleasa que realiza el corte del extremo 3' del pre-RNAm (Mandel, Bai y col.; 2008). La subunidad CPSF-100 contiene un dominio semejante al de la subunidad 73, pero sin la capacidad de coordinar iones de Zn^{2+} , su función no ha sido dilucidada hasta el momento (Dominski, Yang y col.; 2005). La subunidad CPSF-30 presenta cinco arreglos de dedos de Zn^{2+} que funcionan en las interacciones proteína-proteína coordinando la formación del complejo (Barabino, Hubner y col.; 1997). La función de hFip1 es básicamente mediar la interacción del complejo CPSF con PAP y CFIm. Posee un dominio de unión a secuencias ricas en U (Kaufmann, Martin y col.; 2004; Hu, Lutz y col.; 2005).

2.4.2.2 Complejo CstF

El factor estimulante de corte (CstF) es un heterohexámero constituido por dos subunidades CstF-64 que contienen un dominio de unión a RNA para la secuencia G/U del DSE, este dominio presenta gran flexibilidad por lo que puede interactuar con otras secuencias ricas en GU, en su dominio C-terminal interactúa con otros factores de transcripción regulando también ese proceso (Mandel, Bai y col.; 2008). La subunidad CstF-77 tiene una función de interacción con otras proteínas en especial con CPSF-160 justo en su dominio de reconocimiento de AAUAAA, lo que sugiere un papel regulador de su función teniendo en cuenta además de que su unión a CTD-RNAPolIII, aumenta cuando este último se encuentra fosforilado (Bai, Auperin y col.; 2007). La subunidad CstF-50 es requerida para la reacción de corte *in vitro*, además de interactuar con CTDRNAPolIII, presenta interacción con el coactivador de “splicing” SRm160 (Mandel, Bai y col.; 2008).

2.4.2.3 Complejo CFIm

El factor de corte de mamífero I (CFIm) parece ser único en eucariontes superiores, está formado por heterodímeros derivados de la interacción de la subunidad 25 y alguna de las otras subunidades, a saber, 59, 68 ó 72, siendo el heterodímero 25/68 el que, *in vitro*, restituye la actividad de corte, previamente suprimida al modificar este complejo. La subunidad 68 se une al RNA mediante dominios específicos en su extremo N-terminal, del mismo modo interacciona con elementos de “splicing” como U2AF. La subunidad CFIm 25 interactúa con PAP a través de su extremo C-terminal. La principal función de CFIm es el reconocimiento de secuencias que contengan UGUAA corriente arriba de la PAS, que sirven como su sustrato y define adecuadamente el sitio de poliadenilación (Ruegsegger, Blank y col.; 1998).

2.4.2.4 Complejo CFIm

El factor de corte II de mamíferos (CFIm) se compone de 2 subunidades que originalmente fueron descubiertas en *S. cerevisiae* (Ruegsegger, Blank y col.; 1998; de Vries, Bruin y col.; 2000). La primera es hPcf11 que contiene un dominio de unión a CTD-RNapolII y se une preferencialmente cuando este último se encuentra fosforilado. También se ha propuesto como un factor que libera la maquinaria de procesamiento del 3'UTR del CTD-RNapolII (Mandel, Bai y col.; 2008). hClp1 interactúa con CFIm y CPSF tiene una actividad RNA 5'cinasa que sirve para el "splicing" del tRNA y la activación de los RNA de interferencia (RNAi) y a pesar de presentar un dominio Walker A, no tiene una función de ATPasa (Weitzer y Martinez 2007).

2.4.4.5 Elementos proteicos que no forman parte de complejos.

La polimerasa de poli-A (**PAP**) es la enzima que agrega los residuos de adenosina al extremo 3' del preRNAm. Consta de 3 dominios, los dominios Nterminal, medio y C-terminal, el primer dominio coordina dos iones metal (Mg^{2+} ó Mn^{2+}) que son requeridos para la polimerización. El dominio C-terminal sirve para unirse a otras proteínas como hFip1. PAP reconoce los últimos 3 nucleótidos del pre-RNAm como sustrato para su unión y función (Balbo y Bohm 2007).

La proteína de unión a poli-A (**PABP**) se requiere para el correcto proceso de poliadenilación, se une directamente a trectos de 11-14 adenosinas y controla el tamaño de la cola de poli-A hasta que alcanza la longitud apropiada de alrededor de 200 a 300 bases e incrementa la eficiencia de la poliadenilación hasta 80 veces. Contiene cuatro repetidos de motivos que reconocen RNA (Keller, Kuhn y col.; 2000; Kerwitz, Kuhn y col.; 2003).

Otra proteína es la **Simpliquina**, la cual, fue descubierta en asociación con uniones estrechas y probablemente contribuye a formar el andamiaje de la maquinaria de procesamiento del extremo 3', al colocarse entre un gran número de proteínas y estabilizando la interacción entre los complejos CPSF y CstF (Keon, Schafer y col.; 1996; Takagaki y Manley 2000).

Un elemento muy importante para que se lleve a cabo el procesamiento del extremo 3'UTR de los RNAm eucariontes es el **CTD-RNAPolIII**, que como ya se ha señalado previamente, es el sitio de anclaje de la maquinaria. Su delección provoca un RNAm inestable que se degrada rápidamente. CTD-RNAPolIII se conforma de 52 heptapéptidos repetidos cuya secuencia consenso es YSPTSPS. Para cada uno de los repetidos, los residuos de serina son susceptibles de fosforilación y una vez fosforilados, sirven de puerto de anclaje para las maquinarias del "capping", "splicing", así como de procesamiento del extremo 3'UTR del pre-RNAm, con esta última, principalmente a través de las subunidades CPSF160 y CstF-50 (Danckwardt, Hentze y col.; 2008; Mandel, Bai y col.; 2008; Millevoy y Vagner 2010).

2.5. ELEMENTOS EN C/S DEL PROCESAMIENTO DEL EXTREMO 3'UTR DEL RNAm EN OTROS ORGANISMOS.

Los elementos de secuencia que sirven como reguladores de la poliadenilación del RNAm, son predecibles por medios bioinformáticos, debido a la posición conservada de sus secuencias cortas y la proporción de cada uno de los nucleótidos. Tomando en consideración estas características, se analizan un gran número de secuencias pertenecientes a un mismo tipo celular para encontrar esos elementos consensos. Varios grupos han reportado la identificación *in silico* de elementos conservados que podrían estar involucrados en la formación de la cola de poli-A en *Arabidopsis thaliana* (Loke, Stahlberg y col.; 2005), *Homo sapiens* (Hu, Lutz y col.; 2005), *Saccharomyces cerevisiae* (Graber, McAllister y col.; 2002) y

Entamoeba histolytica (Zamorano, López-Camarillo y col.; 2008) entre otros organismos.

La secuencia que predice el sitio de poliadenilación en humanos es CA, puede haber otro nucleótido que preceda a la A, pero C es por mucho, el antecesor más frecuente (59%), mientras que para *A. thaliana* y *S. cerevisiae*, se considera como secuencia consenso YA, donde Y es una pirimidina (C ó U) y el corte se da en el extremo 3' de A (Graber, McAllister y col.; 2002). En *E. histolytica* el sitio de poliadenilación se encuentra en un residuo de U, el cual, puede estar rodeado por motivos ricos en U (Zamorano, López-Camarillo y col.; 2008).

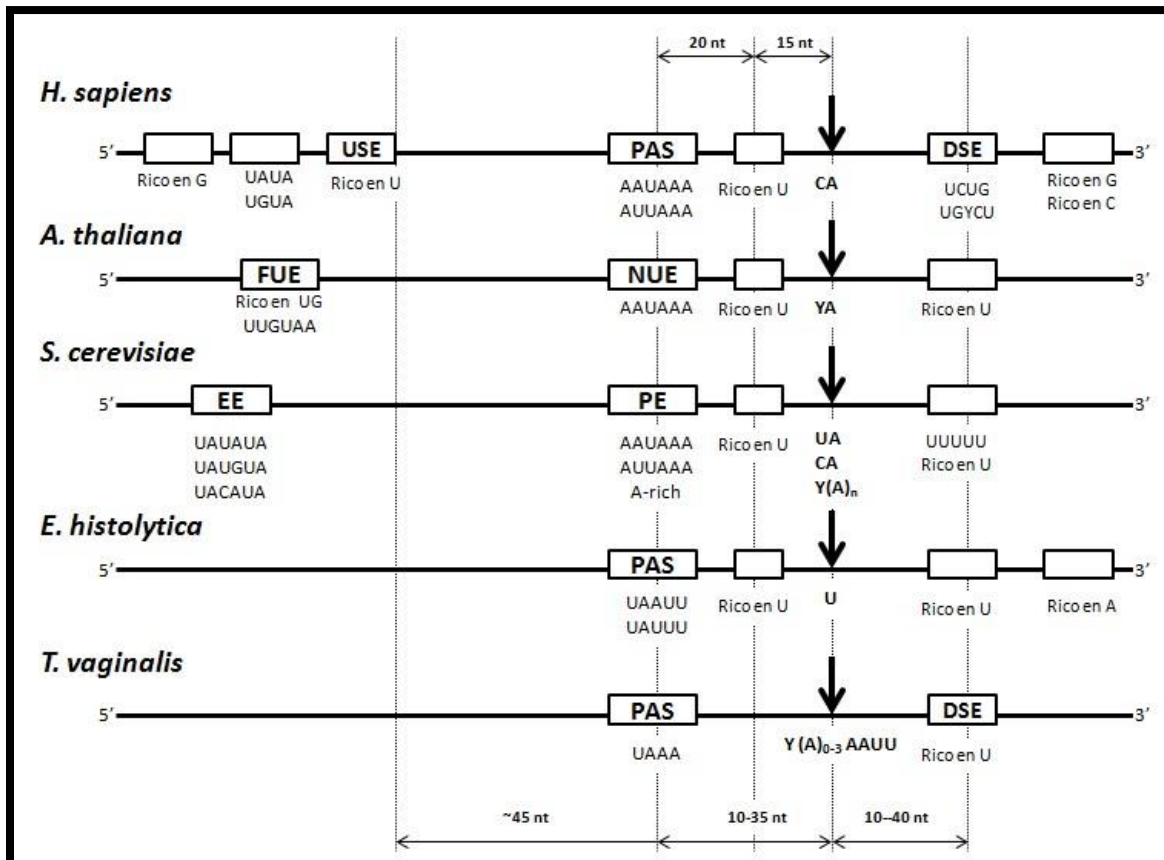
Se han detectado elementos reguladores corriente arriba y corriente abajo del sitio de corte y poliadenilación de los RNAm de estos organismos que tienen características muy semejantes entre sí, como una secuencia de 4 a 10 nucleótidos entre 10 y 35 bases corriente arriba del sitio de poliadenilación (**Fig. 7**), esta secuencia es llamada PAS en el humano y se caracteriza por el hexanucleótido AAUAAA o su variante más frecuente AUUAAA, mientras en *A. thaliana*, esta misma secuencia se denomina elemento cercano corriente arriba (NUE, near upstream element) y en *S. cerevisiae* se le conoce como elemento de posicionamiento (PE, positioning element), en este último organismo se observa con más frecuencia otros hexanucleótidos en la misma posición, pero que mantienen la característica de ser abundantes en A; mientras que en *E. histolytica* también se denomina PAS y está compuesta por los pentanucleótidos UAAUU y UAUUU, los cuales incluyen en algunos casos el codón de paro de los marcos de lectura. Otra secuencia corriente arriba muy conservada en estos organismos, está más allá de las 45 bases y se denomina USE en humano y tiene como característica tener un gran contenido de U. En *A. thaliana*, se le conoce como elemento lejano corriente arriba (FUE, far upstream element), el cual tiene como característica ser rico en UG, mientras que en *S. cerevisiae* se conoce este elemento de secuencia como elemento de eficiencia (EE, efficiency element) y es rico en UA. En *E. histolytica*, hasta la fecha, no se ha encontrado equivalentes de este elemento. Además de lo ya mencionado, en

humano existen otros dos elementos reguladores corriente arriba que sirven como elementos auxiliares, uno proximal a USE que se caracteriza por los tetranucleótidos UAUUA y UGUA, mientras que el otro se encuentra más distal y es rico en G (Graber, McAllister y col.; 2002; Hu, Lutz y col.; 2005; Loke, Stahlberg y col.; 2005; Zamorano, López-Camarillo y col.; 2008) (Graber, McAllister y col.; 2002; Hu, Lutz y col.; 2005; Loke, Stahlberg y col.; 2005; Zamorano, Lopez-Camarillo y col.; 2008).

De los elementos corriente abajo del sitio de poliadenilación el más conservado es el que se encuentra entre 10 a 40 nucleótidos de distancia y se caracteriza por ser rico en U; en el ser humano se le conoce como DSE cuya secuencia ya se mencionó previamente. En *A. thaliana*, *S. cerevisiae* y *E. histolytica* esta secuencia también se ha hallado pero no recibe una denominación especial (Zamorano, Lopez-Camarillo y col.; 2008). Mediante el programa computacional PROBE (Polyadenylation-Related Oligonucleotide Bidimensional Enrichment), en humano, se detectó otra secuencia rica en U que se sobrelapa con DSE, pero que se puede encontrar a partir del nucleótido 25 corriente abajo del sitio de poliadenilación.

Además, se encontró otro elemento a una distancia entre 50 y 100 nucleótidos de distancia del sitio de poliadenilación, cuya secuencia puede ser rica en G ó C (Hu, Lutz y col.; 2005). Por otro lado, en *E. histolytica*, también se detectó un elemento distal corriente abajo rico en A (López-Camarillo, Hernández de la Cruz y col.; 2010).

Figura 7.- Elementos *cis* para la poliadenilación de los RNAm en diferentes organismos. Esquema que muestra los elementos de secuencia reportados hasta el momento para el procesamiento del extremo 3' de los RNAm de *Homo sapiens*, *Saccharomyces cerevisiae*, *Arabidopsis thaliana*, *Entamoeba histolytica* y *Trichomonas vaginalis*. PAS, señal de poliadenilación (polyadenylation signal); USE, elemento corriente arriba (upstream element); DSE, elemento corriente abajo (downstream element); NUE, elemento cercano corriente arriba (near upstream element); FUE, elemento lejano corriente arriba (far upstream element); EE, elemento de eficiencia; PE, elemento de posicionamiento (positioning element). Las líneas delgadas indican posiciones similares, las flechas la distancia promedio entre los elementos señaladas como distancia en nucleótidos. Y, pirimidina (C ó T).



Del análisis de esos trabajos se puede observar que: 1) La organización de las señales de procesamiento del extremo 3'UTR de los RNAm, está fuertemente conservada a través de la escala evolutiva. 2) La gran incidencia de la señal de poliadenilación canónica, confirma su función primordial en la poliadenilación del pre-RNAm. 3) La presencia de múltiples señales de poliadenilación en varias secuencias, sugiere que el uso de distintas PAS puede ser un importante evento regulatorio de la expresión de genes en todos los organismos estudiados. 4) Las distancias en número de nucleótidos entre los elementos regulatorios de los

organismos, son muy semejantes. 5) Las PAS de algunos protozoarios incluyen el codón de paro de sus secuencias (Zamorano, López-Camarillo y col.; 2008).

2.6. MAQUINARIA DE PROCESAMIENTO DEL EXTREMO 3'UTR EN LOS RNAm DE OTROS ORGANISMOS.

2.6.1 *Saccharomyces cerevisiae*

Los factores proteicos que se necesitan para el procesamiento del extremo 3' de los RNAm, generalmente están conservados en todos los eucariontes; tal es el caso de mamíferos y levadura (**Tabla 3**). La maquinaria de levadura comprende el factor de corte y poliadenilación (CPF; cleavage and polyadenylation factor), el factor de corte IA (CFIA) y el factor de corte IB (CFIB). CPF contiene subunidades que son homólogas a las del CPSF en mamíferos, pero distribuidas en dos diferentes subcomplejos, CFII y PFI e incluyen muchos factores adicionales como son Pfs2, Ssu72, Mpe1, Glc7 y Ref2 (Mandel, Bai y col.; 2008). Pta1 es homólogo de la simplequina en levadura y está incluida en CFII, sirviendo como el andamiaje requerido tanto para el corte como para la poliadenilación (Zhao, Kessler y col.; 1999). CFIA contiene subunidades que son homólogos a las de CFII_m y CstF de mamíferos, excepto por la ausencia de CstF en levadura. Hrp1 es el único miembro de CFIB que no tiene un homólogo en mamíferos. A nivel funcional *in vitro*, la levadura requiere de CFIA, CFIB y CFII para realizar el corte del preRNAm, mientras que para la poliadenilación requiere CPF, CFIA, CFIB y Pap1.

La presencia del CTD-RNAPIII no es esencial para el procesamiento del extremo 3' en levadura, sin embargo, aumenta la eficiencia tanto del corte como de la poliadenilación. En la levadura, la subunidad Yhh1, homóloga de CPSF160 de mamífero, no se une a PE, sino a una región rica en A cerca del sitio de corte. Del mismo modo, CstF que se une al elemento rico en U/GU corriente abajo del sitio de poliadenilación, en levadura, su homólogo RnaP15 y su auxiliar RnaP14 son los que reconocen el PE (Millevoi y Vagner 2010).

2.6.2 *Arabidopsis thaliana*

Muchos de los genes que codifican para la maquinaria de procesamiento del extremo 3' UTR de los RNAm de mamífero y levadura, tienen homólogos en el genoma de la planta *Arabidopsis thaliana* (**Tabla 3**), a excepción del factor CFI de mamíferos y la subunidad Hrp1 (Hunt 2008). La maquinaria de corte y poliadenilación de la *A. thaliana* tiene un complejo CPSF (AtCPSF) que incluye AtCPSF30, AtCPSF73-I, AtCPSF73-II, AtCPSF100, AtCPSF160, AtFIPS5 y AtFY. La subunidad AtCPSF100 sirve como núcleo del complejo mientras que AtCPSF30 se une a AtCstF vía una interacción con AtFip, y además, medía la interacción entre CPSF y otros factores como AtCLPS3, AtSYM5 y AtCPSF4, debido a estas interacciones, la función de AtCPSF30 se considera distinta a la de sus homólogos en otros eucariontes (Rao, Dinkins y col.; 2009). Otra diferencia con mamíferos y levadura, es que algunos factores son codificados por pequeñas familias de genes en lugar de un solo gen, y varios de estos factores, no son esenciales sino que sólo afectan funciones biológicas específicas como el caso de CPSF73-II y FY que son proteínas específicas de planta involucradas en el desarrollo femenino de la planta y la floración (Millevoi y Vagner 2010).

2.6.3. *Entamoeba histolytica*

La secuenciación del genoma de este protozoo causante de la amibiasis, facilitó la el estudio de la expresión de genes y la determinación del repertorio completo de genes involucrados en el metabolismo del parásito. Mediante un análisis *in silico*, se buscó la maquinaria de procesamiento del extremo 3' del preRNAm y posteriormente se evaluó la expresión de algunos de esos factores por ensayos de RT-PCR. Se encontró que los factores involucrados en el procesamiento del extremo 3'UTR de los preRNAm de *E. histolytica* están bien conservados con respecto a los factores de mamíferos y levadura, lo que sugiere una gran conservación de esta maquinaria a través de la evolución (**Tabla 3**). El análisis arrojó como resultado, una lista de 16 genes que codifican para proteínas putativas que

intervienen en el procesamiento del extremo 3' del preRNA_m. Los nombres de los factores encontrados en *E. histolytica* son los mismos que sus homólogos en otros organismos, pero se les añade el prefijo Eh. Los factores putativos de corte y poliadenilación de *E. histolytica*, incluyen las cuatro subunidades de EhCPSF (160, 100, 73 y 30 kDa), con EhFIP1 y EhPAP, las tres subunidades de EhCstF (77,64 y 50 kDa), las dos subunidades de CFIm (CLP1 y PCF11) y la subunidad de 25 kDa de CFIm así como factores adicionales, tales como EhMpe1, EhPsf2, EhSsu72 y EhPC4. Se encontró que *E. histolytica* tiene algunas proteínas homólogas a factores presentes en levaduras pero no en mamíferos y viceversa, por ejemplo, un homólogo al factor EhPsf2 de levadura que no se encuentra en mamíferos; por estos hallazgos han considerado que la maquinaria de procesamiento del extremo 3' de los preRNA_m de *E. histolytica* está en una posición evolutiva intermedia levadura y mamíferos (López-Camarillo, Orozco y col.; 2005). Los factores de levadura Pta1p y Hrp1p, así como los factores de mamíferos Simpliquina, CFIm59 y CFIm68 no fueron encontrados. De manera interesante en este parásito se encontró la presencia de más de una secuencia que codifica para la proteína EhPAP. Una canónica con todas las características encontradas en otras PAPs de diferentes organismos y otras seis llamadas no canónicas que tienen homología con otros miembros de la familia de PAPs no convencionales de humano. Se ha visto que la PAP canónica se expresa en las fases del ciclo celular G1 y S, mientras que su expresión decae en las fases G2 y M (López-Camarillo, Hernández de la Cruz y col.; 2010). El estudio estructural del factor EhCFIm 25 demostró que pierde su capacidad hidrolítica, pero que presenta regiones que pueden servir de unión al RNA y a EhPAP. Se encontró un gen que codifica para la proteína PC4 de mamíferos o Sub1 de levadura (EhPC4/Sub1). PC4 se ha considerado como un anti-terminador de la transcripción, al interactuar con CstF64 impidiendo su unión al elemento *cis* río abajo del sitio de poliadenilación. El gen *Ehpc4* se encuentra sobre-expresado en trofozoitos virulentos de la clona A pero no se detecta en trofozoitos virulentos de la clona L6, sugiriendo un papel potencial en la virulencia del microorganismo (López-Camarillo, Hernández de la Cruz y col.; 2010).

Tabla 3.- Homología entre las maquinarias de procesamiento del extremo 3'UTR de los preRNAm de *Homo sapiens*, *Arabidopsis thaliana*, *Saccharomyces cerevisiae* y *Entamoeba histolytica*.

Complejo CPSF (*)				
Subunidad	<i>H. sapiens</i>	<i>A. thaliana</i>	<i>S. cerevisiae</i>	<i>E. histolytica</i>
160	CPSF160	AtCPSF160	Cft1p/Yhh1p	EhCPSF160
100	CPSF100	AtCPSF100	Cft2p/Ydh1p	EhCPSF100
73	CPSF73	AtCPSF73-I AtCPSF73-II	Brr5p/Ysh1p	EhCPSF73
30	CPSF30	AtCPSF30	Yth1p	EhCPSF30
Fip1	hFip1	AtFipS3 AtFIPS5	Fip1p	EhFIP1
Mpe1p	Rbbp6/PACT	----	Mpe1p	EhMpe1p
FY	WDR33	AtFY	Pfs2p	----
SYM5	----	AtSYM5	----	----
PCSF4	----	AtPCSF4	----	----
Swd2p	----	----	Swd2p	----
Ref2p	----	----	Ref2p	----

Complejo CstF (CF-I) *				
Subunidad	<i>H. sapiens</i>	<i>A. thaliana</i>	<i>S. cerevisiae</i>	<i>E. Histolytica</i>
77	CstF-77	AtCstF-77	Rna14p	EhCstF-77
64	CstF-64	AtCstF-64	Rna15p	EhCstF-64
50	CstF-50	----	----	EhCstF-50
Complejo CFI				
Subunidad	<i>H. sapiens</i>	<i>A. thaliana</i>	<i>S. cerevisiae</i>	<i>E. Histolytica</i>
25	CFIm 25	----	----	EhCFI 25
59	CFIm 59	----	----	----
68	CFIm 68	----	----	----

Complejo CFIIIm (CF-IA) *				
Subunidad	<i>H. sapiens</i>	<i>A. thaliana</i>	<i>S. cerevisiae</i>	<i>E. Histolytica</i>
Pcf11p	hPcf11	AtCLPS3	Pcf11p	EhPcf11
Clp1	hClp1	----	Clp1p	EhClp1
SYM5	----	AtSYM5	----	----

Proteínas que no forman complejos

Proteína	<i>H. sapiens</i>	<i>A. thaliana</i>	<i>S. cerevisae</i>	<i>E. Histolytica</i>
PAP	PAP	AtPAP	Pap1p	EhPAP
PABP	PABP	----	----	----
PC4	PC4	----	----	EhPC4
Simplekina	Simplekina	----	Pta1p ⁺	----
Hrp1p	----	----	Hrp1p	----
Ssu72p	----	----	Ssu72p	EhSsu72p
Glc7p	PNAS-120 PP1	----	Glc7p	----
Ref2p	----	----	Ref2p	----

* Nombre del complejo de mamífero y entre paréntesis el complejo correspondiente en levadura. Basado en (Mandel, Bai y col.; 2008; López-Camarillo, Hernández de la Cruz y col.; 2010; Millevoi y Vagner 2010).

2.7 PROCESAMIENTO DEL EXTREMO 3'UTR DE LOS RNAm DE

Trichomonas vaginalis.

Un análisis *in silico* de secuencias de cDNA de *T. vaginalis*, ha descrito los elementos de secuencia para el procesamiento del extremo 3'UTR del pre-RNAm (**Fig. 6**) destacando diferencias como la ausencia del elemento río arriba y que PAS está incluido en el codón de paro UAA de las secuencias, es decir, la señal de poliadenilación se constituye con la secuencia UAAA que incluye dicho codón (Espinosa, Hernández y col.; 2002). Esto fue demostrado mediante ensayos con un gen reportero, haciendo análisis mutacionales de la secuencia UAAA introducida en el gen y se comprobó que esta secuencia tiene exactamente la misma función que tiene el exanucleótido que sirve de PAS en los metazoarios, ya que al mover de posición esta secuencia, también cambió el sitio de poliadenilación. Los datos reportados sugieren que si bien esta señal dirige el proceso de corte y poliadenilación, no es por sí sola suficiente y sugiere la presencia de al menos dos elementos auxiliares ricos en U que pueden aumentar la especificidad del proceso. Además el punto de corte en lugar de ser un dinucleótido CA presenta una secuencia YA(0-3) AAUU (Y=pirimidina y el número de poliadeninas antes del corte varía hasta 3) (Fuentes, Barrera y col.; 2012).

III. JUSTIFICACION

Hasta el momento no se ha descrito la maquinaria de procesamiento del extremo 3'UTR de los RNAm de *Trichomonas vaginalis*, la cual, estrictamente coordinada, es importante para la expresión de los RNAm. Por lo que resulta importante e interesante dilucidar esta maquinaria.

IV. HIPÓTESIS

Trichomonas vaginalis posee una maquinaria de procesamiento del extremo 3'UTR de sus RNAm conservada evolutivamente en eucariontes.

V. OBJETIVOS

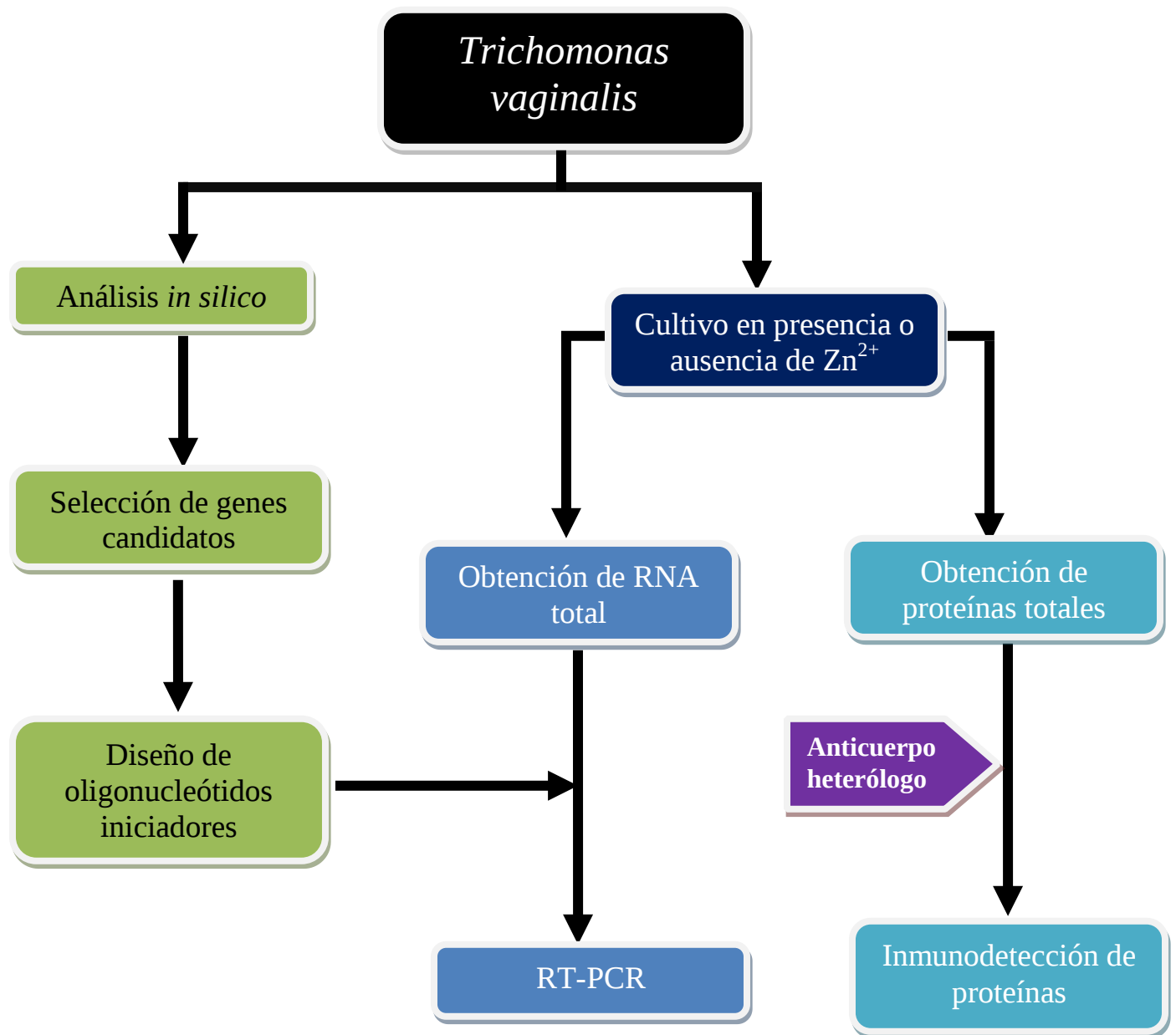
5. 1. OBJETIVO GENERAL:

Identificar la maquinaria del complejo de procesamiento del 3'UTR de los RNAm de *Trichomonas vaginalis*.

5. 2. OBJETIVOS PARTICULARES:

1. Identificar *in silico*, mediante genómica comparativa, las posibles proteínas que conforman la maquinaria de procesamiento del extremo 3'UTR de los RNAm de *T. vaginalis*.
2. Determinar la expresión de los RNAm de algunos de los genes que codifican para las proteínas involucradas en el procesamiento del 3'UTR de los RNAm de *T. vaginalis*.
3. Inmunodetectar la presencia de al menos una de estas proteínas en *T. vaginalis* mediante anticuerpos heterólogos.

VI. ESTRATEGIA EXPERIMENTAL



VII. MATERIAL Y MÉTODOS.

7.1 ANÁLISIS *IN SILICO*

Las maquinarias de procesamiento del extremo 3'UTR de los RNAm de las especies *Homo sapiens*, *Saccharomyces cerevisiae*, *Arabidopsis thaliana* y *Entamoeba histolytica* fueron elegidas debido a que son organismos eucariontes cuyas maquinarias han sido ampliamente estudiadas y al estar en diferentes reinos taxonómicos, existe mayor diversidad de proteínas en número y función para comparar. Los números de acceso de cada proteína de estas maquinarias se obtuvieron de las referencias Mandel y col. 2008; Lopez-Camarillo y col. 2010; y Millevoi y Vagner 2010. Con estos números de acceso se obtuvieron las secuencias de aminoácidos de cada una de ellas en las bases de datos NCBI (<http://www.ncbi.nlm.nih.gov/>), UniProt de ExPASy (<http://expasy.org/>), y Amoeba DB (<http://amoebadb.org/amoeba/>), (**Anexo2**). Posteriormente, cada secuencia de aminoácidos fue alineada en la base de datos de secuencias curadas de Tricho DB (<http://trichdb.org/trichdb/>) y se obtuvieron las secuencias con un valor de E más óptimo (≤ 0.001) (El **valor de e** se refiere a la menor probabilidad de que otra secuencia se parezca a esta), con mayor porcentaje de identidad (%I) y porcentaje de similitud (%S).

Se realizó modelaje de algunas proteínas seleccionadas con el programa Phyre 2 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) y se buscaron los dominios correspondientes con el programa Pfam (<http://pfam.sanger.ac.uk/>).

7.2 CULTIVO DE *T.vaginalis*

El cultivo de parásitos se hizo a partir del aislado, HGMN01 que corresponde a un paciente masculino diagnosticado por cultivo (proporcionado por la Dra. Rossana Arroyo). Fue sembrado en medio TYM (Extracto de maltosa-tripticase de levadura), al cual, se le adiciona 10% de suero de caballo que previamente es inactivado por calor durante 30 min a 56°C. El cultivo se mantiene a 37°C en cajas de 25cm². Se realizan pases diarios considerando su tiempo de replicación de cada 4 hrs en

promedio. El aislado HGMN01 es trabajado con una semana de pases. En estas condiciones la viabilidad de los parásitos evaluada por la metodología de Azul de Tripano es >98%.

7.3 CONDICIÓN DE CULTIVO EN Zn^{2+}

Doce horas antes de realizar el ensayo, se prepararon en dos cajas de cultivo de 25cm² con 27mL de medio TYM y se agregaron 1.72mL de ZnCl₂ para obtener un medio de cultivo con una concentración de 1.6mM de Zn²⁺. Como control, se utilizó una caja de 25cm² con medio TYM sin ZnCl₂. Al día siguiente se sembraron 1x10⁶ millones de parásitos para las cajas control y 2x10⁶ de parásitos para las cajas que contienen medio con Zn²⁺. Se dejaron crecer durante 24 hrs para posteriormente ser cosechados. Estas condiciones se utilizaron en todos los experimentos.

7. 4 EXTRACCIÓN DE DNA

Un total de 20 x 10⁶ parásitos fueron contados y centrifugados a 1800 rpm/ 5 min/ 4°C; posteriormente se agregó 300µl de buffer de lisis (Tris-HCl 10mM pH 8.0, SDS 0.2%, NaCl 100mM, EDTA mM) y 300 µl de fenol, se agitó por inversión y se centrifugó a 13,000rpm/ 5 min/ 4°C. La fase acuosa se transfirió a un tubo nuevo de 1.5 ml, inmediatamente después se adicionó fenol, se mezcló suavemente, se centrifugó 13,000 rpm/ 5 min/ 4°C, y el sobrenadante se transfirió a un tubo nuevo de 1.5 ml. Luego, se agregó fenol-cloroformo, se centrifugó a 13,000 rpm/ 5 min/ 4°C y la fase acuosa se transfirió a un tubo nuevo para ser precipitada con etanol absoluto (2¹/₂ volumen) y acetato de sodio 3M (¹/₁₀ del volumen). Seguido, se incubó a -20°C por 2 hrs y se centrifugó a 13,000rpm / 15 min/ 4°C. Posteriormente, se decantó el sobrenadante y se lavó la pastilla con etanol frío al 70% se centrifugó a 13,000rpm/ 15 min/ 4°C, se decantó el sobrenadante y dejó secar la pastilla para finalmente resuspenderla en libre de nucleasas. Finalizada la extracción, se cuantificó por espectrofotometría a 260nm y se corrió en un gel de garosa al 1%

teñido con Red Gel 1X y se visualizó el gel en un fotodocumentador Universal Hood II de BIO-RAD para verificar su calidad e integridad.

7.5 EXTRACCIÓN DE RNA.

Para extraer RNA, se usaron 20×10^6 parásitos y se centrifugaron a 1800 rpm/5min/4°C, posteriormente, se agregó 1 mL de TRIzol frío, se mezcló por inversión e incubó a temperatura ambiente (TA) durante 10 min, posteriormente 30 min a 4°C. Después se agregaron 200 µl de cloroformo y se mezcló fuertemente a TA, se centrifugó a 10,000 rpm/ 15 min/4°C, se recuperó la fase acuosa a la cual se le añadió 500 µL de isopropanol, se mezcló e incubó 15 min a TA para posteriormente, centrifugar a 10,000 rpm/ 15 min/ 4°C. Se decantó el sobrenadante y lavó la pastilla dos veces con etanol al 75% en agua DEPC, en cada lavado, los tubos se centrifugaron a 10,000 rpm/ 15 min/ 4°C. Una vez finalizados los lavados, se decantó el sobrenadante y se dejó secar la pastilla a TA, y se resuspendió en agua DEPC. La cuantificación se realizó por espectrofotometría 260 nm y se corrió en un gel de agarosa al 1% preparado con TBE 1X, usando como buffer de electroforesis TBE al 0.5%, posteriormente fue teñido con Red para verificar la calidad e integridad del RNA en fotodocumentador Universal II BIO-RAD.

7.8 DISEÑO DE OLIGOS

En vista de poder realizar en otro tiempo una PCR multiplex en el equipo GenExpression Beckman Coulther, los oligonucleótidos se ajustaron a las especificaciones para el protocolo del mismo, es decir, con un tamaño de amplicón entre 100 y 400 pb; que su Tm estuviera cercana a los 52°C; que no formaran homodímeros ni heterodúplex, tanto con su par como con los otros oligos diseñados. Lo anterior se logró con diferentes herramientas informáticas que se encuentran en la página http://mybio.wikia.com/wiki/Primer_design_tools.

Finalmente se les agregó un adaptador universal al extremo 5' de los Forward (5'AGGTGACACTATAGAATA-3') y de los Reverse (5'-GTACGACTCACTATAGGG A-3'). Una vez diseñados, cada iniciador fue posteriormente sometido a otro blastn en la misma base de datos, para cerciorarse de que fuera específico. Solicitamos su síntesis a Integrated DNA Technologies (IDT).

7.7 RT-PCR

Una vez extraído el RNA total de las condiciones de cultivo en ausencia y presencia de Zn^{2+} , se trató con DNAasa Promega^{MR} 1U/ μ g de RNA durante 2 hrs. Se cuantificó el RNA tratado y para 1 μ g de RNA total se adicionó 1 μ L de oligo DT y agua DEPC cbp 12 μ L y se incubó 15 min a 70°C. Después se colocó 5 min en hielo y se le agregaron 4 μ L de buffer para RT 5X, 2 μ L de dNTPs y 2 μ L de DTT. Esta mezcla se incubó 2 min a 42°C, inmediatamente se agregó 1 μ L de retrotranscriptasa reversa (Superscrip II, Invitrogen) y se incubó 90 min a 42°C. Luego se puso 10 min a 70°C para inactivar la enzima RT y finalmente se cuantificó por espectrofotometría.

Como control de carga se estandarizó (α -tubulina) y posteriormente se realizaron las PCRs correspondientes al cDNA de cada condición en estudio (con y sin Zn^{2+}), con el kit Kappa Biosystem. Siguiendo las indicaciones del proveedor, el volumen final de la reacción fue de 12 μ L, con los siguientes componentes: 6 μ L de 2X KAPA Taq ReadyMix + dyeDNA Polymerase (KAPA BYOSYSTEMS); 0.4 μ L de iniciador Forward, 0.4 μ L de iniciador Reverse, ambos a una concentración de 10mM/ μ L; 1 μ g de cDNA y agua inyectable cbp 12 μ L.

Programa usado para PCR:

Desnaturalización inicial:	95°C	2min	
Desnaturalización:	95°C	30 s	} ciclos
Alineamiento	52°C	30 s 35	
Extensión	72°C	1min	
Extensión final	72°C	2min	

Los productos de PCR se corrieron en un gel de agarosa al 3%, teñido con Red Gel y con un marcador de 100 pb. Los geles se analizaron por densitometría con el programa Quantity One.

7.8 EXTRACCIÓN DE PROTEÍNAS TOTALES DE *T. vaginalis*.

A partir de una pastilla de 20×10^6 parásitos, los cuales se centrifugaron a 1800 rpm/5 min/4°C, se le hizo dos lavados resuspendiendo en 1 ml de PBS-1X y volviendo a centrifugar a 1800 rpm/ 5 min/ 4°C y tirando el sobrenadante. Después del último lavado, se decantó el sobrenadante y secamos perfectamente la pastilla, posteriormente se resuspendió la pastilla en 900 μ L de PBS-1x y se agregaron 100 μ L de ácido tricloroacético (TCA) para que quedara a una concentración de 10% y luego se incubó 16 hrs a 4°C. Al término del tiempo de incubación, se realizaron dos lavados más con PBS 1X, secando las paredes del tubo y la pastilla final se resuspendió en 400 μ L de buffer de carga 1X. Una vez obtenida la muestra, se corrieron 10 μ L en geles de poliacrilamida al 16% en 1 dimensión para ver la integridad de la muestra mediante tinción con Azul de Coomassie, los extractos fueron guardados a -70°C hasta su uso.

7.9 ELECTROFORESIS EN DOBLE DIMENSIÓN

1×10^8 parásitos se utilizaron para este ensayo, fueron centrifugados y lavados con PBS pH 7.0. Posteriormente, se lisaron en buffer de hidratación (7 M urea, 4% CHAPS, 70 mM DTT, 2% IPG buffer pH 4–7, trace bromophenol blue) y se agregaron inhibidores de proteasas. Se cuantificó por Bradford y se hidrató de manera pasiva por 16h. Al término de ese tiempo, se realizó el isoelectroenfoque (IEF), con la siguiente programación: Método de 4 pasos: 1) 500 V por 1 hr; 2) 1000V por 1 hr; 3) 6000V por 2:50hrs; 4) hasta 20,000 V/hr. Posteriormente se equilibraron las tiras con DTT (10mg/ml) 15 min en agitación y posteriormente IAA (10 mg/ml) durante 15 min y se cargó la tira en un gel de poliacrilamida al 12% para correr la electroforesis en segunda dimensión. Los geles se tiñeron con Sypro Ruby.

7.10 WESTERN BLOT

Se prepararon geles en espejo de extractos de proteínas en 1D y en 2D de parásitos crecidos bajo las condiciones de ausencia y presencia de Zn^{2+} , uno de los geles en 1D fue teñido con Azul de Coomassie y uno de los geles en 2D con Sypro Ruby; los geles espejo se transfirieron a una membrana de nitrocelulosa y se bloquearon con leche al 5% O.N.. Al día siguiente se lavaron 3 veces con PBS1X durante 15 min cada lavado, se incubó con el anticuerpo primario policlonal heterólogo de *E. histolytica* anti-PC4 (donado por el Dr. César López Camarillo) a una dilución 1:500 durante toda la noche y al día siguiente se lavó con PBS-Tween 0.1% incubó con el anticuerpo secundario anti-conejo conjugado con HRP a una dilución 1:3000 durante/2 hrs, se realizaron 4 lavados con PBSTween al 0.1%, posteriormente se reveló por quimioluminiscencia. Para el caso de los geles en 2D se escanearon las imágenes del gel teñido con Sypro Ruby y el WB-2D, con PD Quest se analizaron los resultados para identificar la mancha candidata.

VIII. RESULTADOS

8.1 Detección *in silico* de las posibles proteína que integran la maquinaria de procesamiento del extremo 3'UTR de los RNAm de *T. vaginalis*.

De acuerdo a la búsqueda bibliográfica de los componentes de la maquinaria de procesamiento del extremo 3'UTR descritas en otros organismos como *Homo sapiens*, *Saccharomyces cerevisiae*, *Arabidopsis thaliana* y *Entamoeba histolytica*, se obtuvieron los números de acceso de las diferentes bases de datos para obtener la secuencia de aminoácidos de cada elemento de estas maquinarias, una vez obtenidas las secuencias, estas se utilizaron como modelos para hacer el análisis comparativo de la posible maquinaria de procesamiento del extremo 3'UTR de los RNAm de *Trichomonas vaginalis*, mediante BLASTp en la base de datos Trich DB.

Del análisis bioinformático de cada una de las proteínas que participan en la maquinaria de procesamiento de los diferentes organismos, se encontraron desde 3 hasta 50 posibles candidatos dentro del genoma de *Trichomonas vaginalis*; los parámetros considerados para obtener a las posibles proteínas que participan en la maquinaria de procesamiento de los RNAm en *T. vaginalis* fueron el mayor valor de e mas optimo (≤ 0.001), mayor porcentaje de identidad (%I) y porcentaje de similitud (%S).

De las 19 proteínas de *H. sapiens* mas estudiadas que s encuentran en la maquinaria de procesamiento de los RNAm, el BLASTp dió como resultados 19 secuencias conservadas en el genoma de *T. vaginalis*. Para el caso de la proteína CPSF30 en *T. vaginalis*, se encontraron dos proteínas con valores predictivos muy semejantes (**Tabla 4**). Por otra parte, las secuencias de *T. vaginalis* que corresponden a las proteínas CFIm59 y 68, tuvieron un valor predictivo > 0.001 , y una identidad de < 30 , por lo que ya no fueron consideradas para los siguientes procesos.

4.- **Tabla** Elementos de la maquinaria de corte y poliadenilación de los RNAm en *H. sapiens* y secuencias relacionadas en *T. vaginalis*.

SUBUNIDAD	No. DE ACCESO	No. de SR en Tv	SECUENCIA MÁS PARECIDA	Valor de e	I%	S%
CPSF160	Q10570	50	TVAG_077030	2.7 e-25	23	42
CPSF100	Q9P2I0	8	TVAG_363680	4.7 e-60	32	52
CPSF73	Q9UKF6	22	TVAG_437970	8.5 e-137	40	61
CPSF30	O95639	37	TVAG_047060	1.9 e-29	50	66
			TVAG_094550	3.9e-29	50	66
FIP1	Q6UN15	3	TVAG_463950	1.9e-06	42	59
PAP-b	Q9NRJ5	17	TVAG_388620	5.5e-78	38	59
Simplexina	Q92797	15	TVAG_494090	0.00036	21	40
CstF77	Q12996	21	TVAG_343030	8.9e-29	31	49
CstF64	P33240	50	TVAG_021720	1.0e-13	45	60
CstF50	Q05048	50	TVAG_374220	2.3e-19	28	42
CFIm25	043809	5	TVAG_475950	1.6e-25	29	56
CFIm59	Q8N684	22	TVAG_097790	0.0013	29	50
CFIm68	Q16630	50	TVAG_258530	0.0018	36	58
CLP1	Q92989	9	TVAG_162880	6.2e-15	23	43
PCF11	094913	50	TVAG_100880	8.0e-05	36	47
Rrbp6	Q7Z6E9	50	TVAG_480160	2.0e-24	20	41
PNAS-120	Q9NP77	50	TVAG_229140	1.6e-25	32	54
PC4	P53999	8	TVAG_255210	1.3e-07	43	66
PABP	P11940	50	TVAG_385330	1.7e-53	33	53

Subunidad= Nombre de la proteína en la maquinaria del organismo de origen; **No. de acceso**= clave con la que se localiza la secuencia en la base de datos Swiss-Prot; **No. SR en Tv**= Número de secuencias obtenidas por el blastp en la base de datos de Trich DB; **Secuencia más parecida**= aquella secuencia obtuvo los mejores valores predictivos en el blastp; **I%**=porcentaje de identidad; **%S**= porcentaje de similitud

De las 26 proteínas que se encontraron para la maquinaria de *S. cerevisiae*, se obtuvieron candidatas para todas las proteínas descritas (**Tabla 5**), sin embargo, debido a que los valores predictivos de las proteínas candidatas en *T. vaginalis* a ser homólogas de Pcf11p, Ref2, Pta1, Pt1p y Sic1p; no cumplían con los valores predictivos antes mencionados, estas cinco secuencias no se tomaron en consideración por lo que se eliminaron. Además, se encontraron dos secuencias con valores cercanos para la proteína Glc7p y para la proteína Rna15p por lo que se tomaron en cuenta.

Tabla 5.- Elementos de la maquinaria de corte y poliadenilación de los RNAm en *S. cerevisiae* y secuencias relacionadas en *T. vaginalis*.

SUBUNIDAD	No. DE ACCESO	No. de SR en TV	SECUENCIA MÁS PARECIDA	Valor de e	I%	S%
Cft1p/Yhh1p	Q06632	9	TVAG_077030	0.00012	21	37
Cft2p/Ydh1p	Q12102	50	TVAG_363680	4.1e-12	25	47
Brf5/Ysh1p	Q06224	43	TVAG_437970	4.8e-96	42	61
Yth1p	Q06102	16	TVAG_094550	9.5e-19	33	56
Pap1	P29468	17	TVAG_388620	1.8e-69	34	56
Fip1	P45976	18	TVAG_463950	6.4e-06	56	75
RNA15	P25299	50	TVAG_267990	1.1e-09	32	65
			TVAG_023279	2.7e-09	28	62
RNA14	P25298	50	TVAG_343030	1.1e-14	23	42
Pcf11	P39081	50	TVAG_166780	0.022	29	50
Clp1	Q08685	15	TVAG_317830	2.0e-06	25	43
Pta1p	Q01329	2	TVAG_264230	0.70	30	51
Hrp1p	Q99383	50	TVAG_245890	6.1e-15	30	50
Pfs2p	P42841	50	TVAG_121040	3.9e-45	29	50
Ssu72p	P53538	46	TVAG_229140	3.1e-29	35	50
Mpe1	P35728	19	TVAG_265310	2.7e-05	32	53
Sub1	P54000	13	TVAG_255210	3.6e-05	40	56
Pti1p	P39927	33	TVAG_245890	0.0099	24	48
Swd2p	P36104	50	TVAG_016940	2.0e-11	24	43
Syc1p	Q08553	5	TVAG_128630	0.99	27	51
Glc7p	P32598	50	TVAG_101240	1.6e-96	56	75
			TVAG_379450	2.0e-96	57	74
Ref2p	P42073	46	TVAG_211330	0.11	28	47
Pab1	P04147	50	TVAG_385330	6.9e-55	37	56

Subunidad= Nombre de la proteína en la maquinaria del organismo de origen; **No. de acceso**= clave con la que se localiza la secuencia en la base de datos Swiss-Prot; **No. SR en Tv**= Número de secuencias obtenidas por el blastp en la base de datos de Trich DB; **Secuencia más parecida**= aquella secuencia obtuvo los mejores valores predictivos en el blastp; **I%**=porcentaje de identidad; **%S**= porcentaje de similitud

Para la maquinaria de *Arabidopsis thaliana*, se obtuvieron 15 proteínas que se han descrito como parte de la maquinaria de procesamiento del extremo 3'UTR de los RNAm de este organismo (**Tabla 6**) y para cada una de ellas se obtuvieron proteínas relacionadas en *T. vaginalis*, excepto para la proteína SYM5. Mientras que

para la proteína de *T. vaginalis* que corresponde a la proteína PCFS4, los valores predictivos no fueron los esperados para seguir el análisis correspondiente.

Tabla 6.- Elementos de la maquinaria de corte y poluadenilación de los RNA, en *A. thaliana* y secuencias relacionadas en *T. vaginalis*.

SUBUNIDAD	No. DE ACCESO	No. de SR en TV	SECUENCIA MÁS PARECIDA	Valor de e	I%	S%
AtCPSF160	Q9FGR0	4	TVAG_077030	1.5e-22	23	39
AtCPSF100	Q9LKF9	25	TVAG_363680	6.6e-80	33	54
AtCPSF73-I	Q9C952	22	TVAG_437970	1.4e-127	41	60
AtCPSF73-II	K8GUU3	18	TVAG_183200	2.6e-71	29	50
AtCPSF30	AEE31221	3	TVAG_047060	4.0e-13	36	57
AtFipS5	AED96990	50	TVAG_210560	6.7e-16	20	42
AtPAP	F4KE92	50	TVAG_107950	5.7e-06	23	42
AtCstF77	Q8GUP1	12	TVAG_343030	1.1e-51	29	50
AtCstF64	Q9M033	42	TVAG_254160	0.0010	16	30
AtCLPS3	Q9SR06	13	TVAG_162880	6.0e-22	27	47
AtPCFS4	Q0WPF2	7	TVAG_388070	0.14	28	46
FY	F4K3Y7	50	TVAG_121040	1.4e-63	35	52
PABP5	Q05196	50	TVAG_385330	6.2e-54	36	57

Subunidad= Nombre de la proteína en la maquinaria del organismo de origen; **No. de acceso**= clave con la que se localiza la secuencia en la base de datos Swiss-Prot; **No. SR en Tv**= Número de secuencias obtenidas por el blastp en la base de datos de Trich DB; **Secuencia más parecida**= aquella secuencia obtuvo los mejores valores predictivos en el blastp; **I%**= porcentaje de identidad; **%S**= porcentaje de similitud.

Para *E. histolytica* se obtuvieron 20 proteínas de la maquinaria de procesamiento del extremo 3'UTR de los RNAm, todas tuvieron secuencias relacionadas en la base de datos Trich DB; sin embargo se eliminaron 3 secuencias por su bajo valor predictivo (**Tabla 7**).

Un total de 70 secuencias candidatas para las proteínas que participan en la maquinaria de procesamiento del extremo 3'UTR de *T. vaginalis* fueron analizadas. Estas secuencias se procesaron de acuerdo a su valor predictivo; las secuencias repetidas se eliminaron dejando sólo el más alto valor predictivo; las secuencias con un valor de e >0.001 y con la misma factorización permanecieron en el análisis. Al

final, se identificaron 31 proteínas que probablemente estén involucradas en el procesamiento del extremo 3'UTR de los RNAm de *T. vaginalis* (**Tabla 8**).

Tabla 8.- Elementos de la maquinaria de corte y poliadenilación de los RNAm en *E. histolytica* y secuencias relacionadas en *T. vaginalis*.

SUBUNIDAD	No. DE ACCESO	No. de SR en TV	SECUENCIA PARECIDA e	Valor de MÁS	I%	S%
EhCPSF160	EHI_106110	50	TVAG_077030	2.0e-14	24	45
EhCPSF100	EHI_033130	9	TVAG_186280	0.23	27	44
EhCPSF73	EIH_136170	9	TVAG_437970	6.5e-91	39	59
EhCPSF30	EIH_067580	9	TVAG_047060	1.3e-14	34	47
			TVAG_094550	5.6e-14	30	46
EhFIP1	EIH_083400	10	TVAG_463950	6.8e-09	41	60
EhPAP	EIH_012040	47	TVAG_388620	3.0e-45	28	52
EhCstF77	EIH_098370	16	TVAG_343030	3.1e-09	31	51
EhCstF64	EIH_151990	50	TVAG_211580	4.3e-07	32	52
EhCstF50	EIH_152770	50	TVAG_493960	3.6e-14	24	42
EhCFIm25	EIH_077110	29	TVAG_495750	1.8e-22	31	56
EhClp1	EIH_008100	13	TVAG_162880	5.2e-08	22	40
EhPcfl1	EIH_045130	15	TVAG_322760	0.16	43	56
EhPfs2	EIH_170080	50	TVAG_121040	2.7e-53	31	49
EhMpe1	EIH_014000	5	TVAG_450960	0.997	50	61
EhSsu72	EIH_027340	29	TVAG_229140	6.5e-20	29	49
EhPC4/Sub1	EIH_192520	25	TVAG_255210	3.9e-06	40	72
EhPABP	EHI_198750	50	TVAG_385330	1.1e-22	24	45

Subunidad= Nombre de la proteína en la maquinaria del organismo de origen; **No. de acceso**= clave con la que se localiza la secuencia en la base de datos Swiss-Prot; **No. SR en Tv**= Número de secuencias obtenidas por el blastp en la base de datos de Trich DB; **Secuencia más parecida**= aquella secuencia obtuvo los mejores valores predictivos en el blastp; **I%**=porcentaje de identidad; **%S**= porcentaje de similitud.

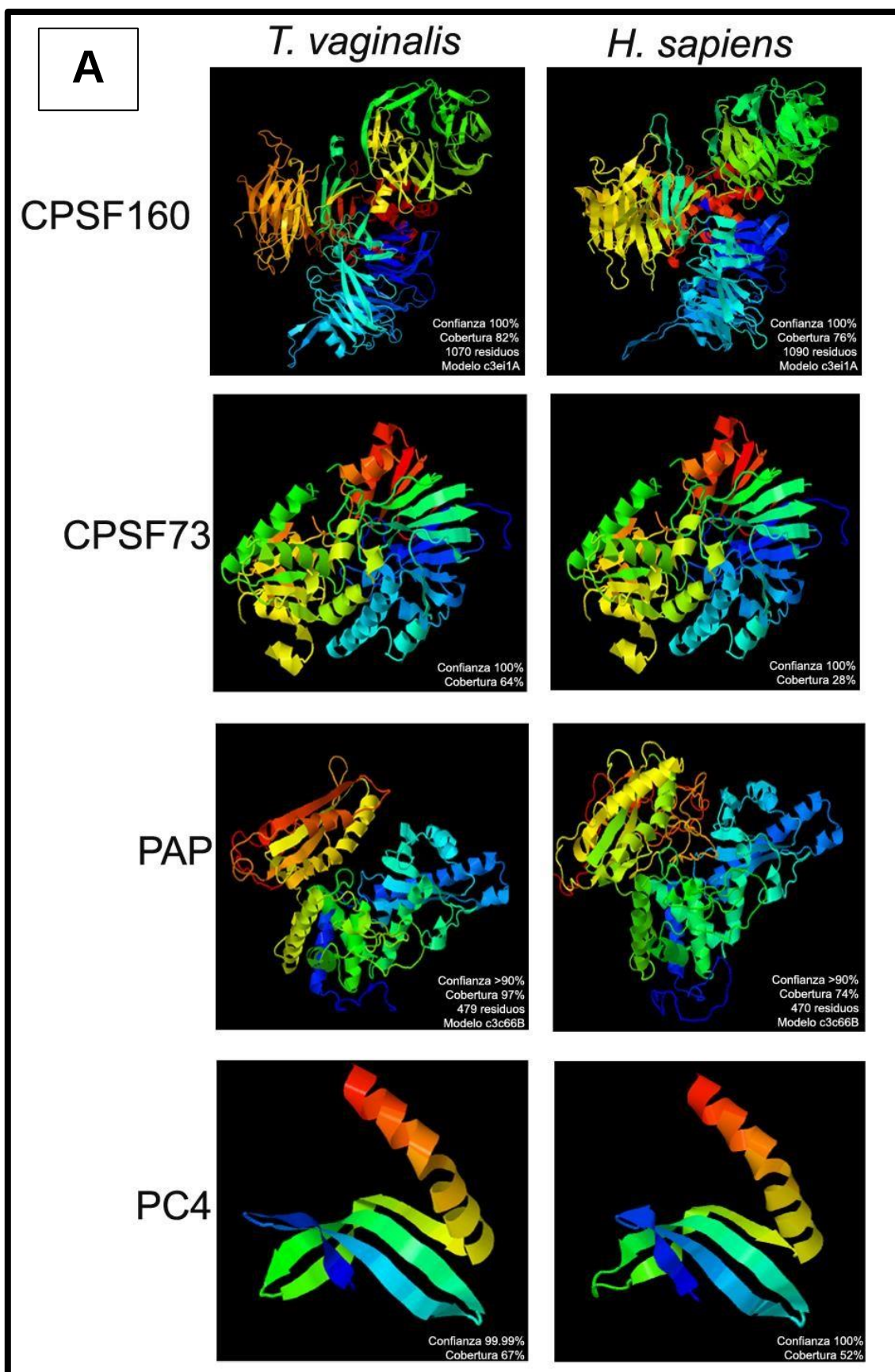
De las 31 proteínas elegidas, se decidió trabajar con nueve candidatas las cuales son: TvPSF2p, TvCPSF160, TvFIP1, TvCF1m25, TvPAP1, TvPC4, TvPAB1. TvCPSF73-II y TvCPCF73, entre la cuales se encuentra nproteínas con homologia a proteínas de planta. Con la secuencia de aminoácidos se predijo la estructura terciaria de cada una de ellas utilizando el programa PHYRE, comparándolas con su homóloga más cercana (**Fig 8**), encontrando una gran similitud entre ellas, esto se tomó en cuenta para elegir a la proteína candidata para ser inmunodetectada dentro del proteoma de *T. vaginalis*.

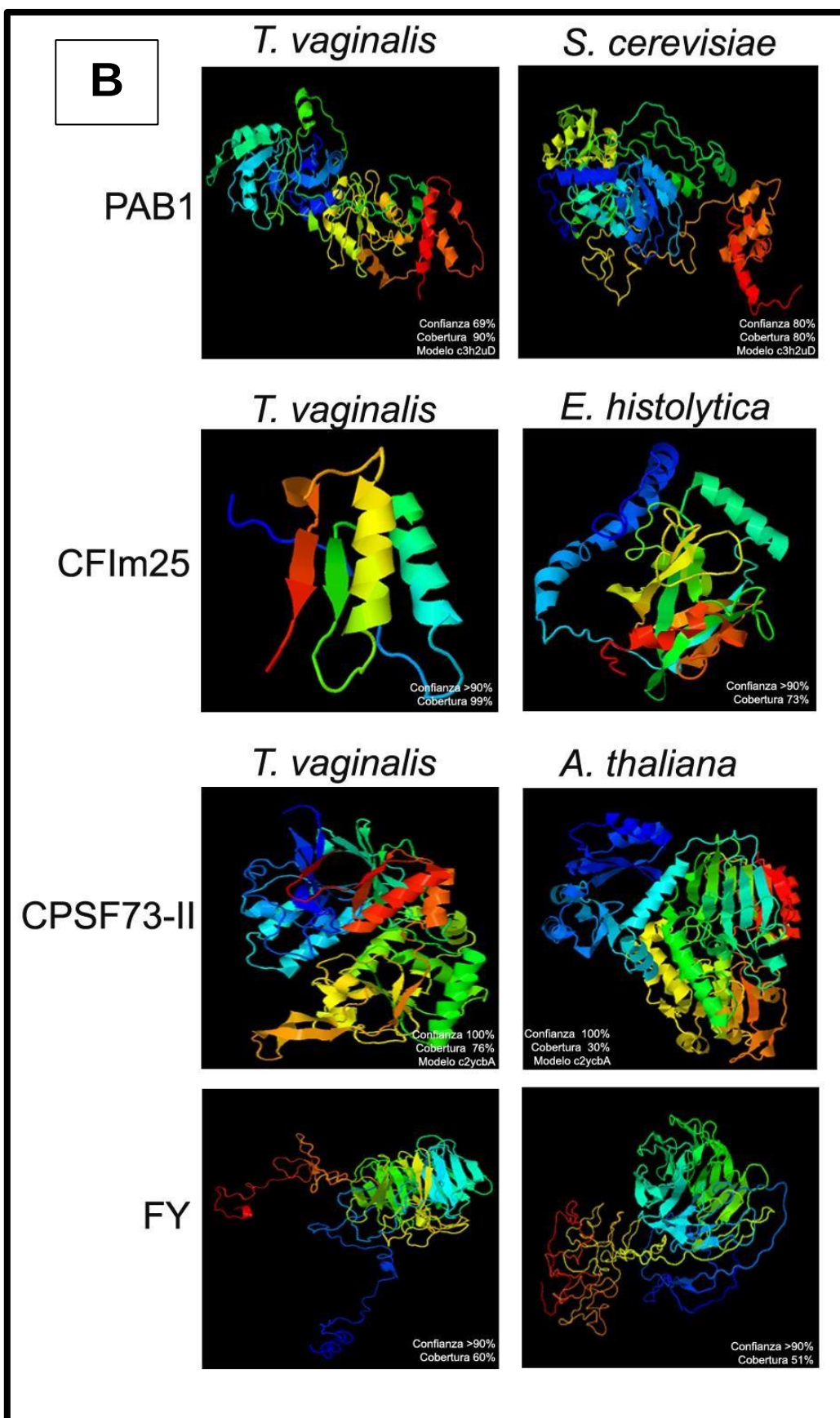
Tabla 9. Proteínas detectadas *in silico* como candidatas a formar parte de la maquinaria de procesamiento del extremo 3'UTR de los RNAm de *Trichomonas vaginalis*. La tabla muestra el número de acceso de la base de datos TrichDB, el organismo con que se obtuvo la mejor comparación, %I= % identidad; %S= % similitud, e= valor de e. A las proteínas candidatas se les colocó el prefijo Tv.

Proteína	No. de acceso	Organismo comparado	%I	%S	e
<u>Complejo CPSF</u>					
TvCPSF160	TVAG_077030	<i>H. sapiens</i>	23	42	2.7 e-25
TvCPSF100	TVAG_363680 33	<i>A. thaliana</i>		54	6.6 e-80
TvCPSF73-I	TVAG_437970	<i>H. sapiens</i>	40	61	8.5 e-137
TvCPSF-II	TVAG_183200	<i>A. thaliana</i>	29	50	2.6 e-71
TvCPSF-I	TVAG_047060	<i>H. sapiens</i>	50	66	1.9 e-29
TvCPSF-II	TVAG_094550	<i>H. sapiens</i>	50	66	3-9 e-29
TvFIP1	TVAG_463950	<i>H. sapiens</i>	42	59	1.9 e-06
TvFIPS5	TVAG_210560	<i>A. thaliana</i>	20	42	6.7 e-16
<u>Complejo CstF</u>					
TvCstF77	TVAG_343030	<i>A. thaliana</i>	29	50	1.1 e-51
TvCstF64-I	TVAG_021720	<i>H. sapiens</i>	45	60	1.0 e-13
TvCstF64-II	TVAG_211580	<i>E. histolytica</i>	32	52	4.3 e-07
TvCstF50-I	TVAG_374220	<i>H. sapiens</i>	28	42	2.3 e-19
TvCstF50-II	TVAG_493960	<i>E. histolytica</i>	24	40	3.6 e-14
TvRna15p-I	TVAG_267990	<i>S. cerevisiae</i>	32	65	1.1 e-09
TvRna15p-II	TVAG_023270	<i>S. cerevisiae</i>	28	62	2.7 e-09
<u>Complejo CFI</u>					
TvCFI25	TVAG_495750	<i>E. histolytica</i>	31	56	1.8 e-22
<u>Complejo CFII</u>					
TvCLPS3	TVAG_162880	<i>A. thaliana</i>	27	47	6.0 e-22
TvCLP1	TVAG_317830	<i>S. cerevisiae</i>	25	43	2.0 e-06
TvPCF11	TVAG_100880	<i>H. sapiens</i>	36	47	8.0 e-05

<u>No asociadas a complejos</u>					
TvPAP-I	TVAG_388620	<i>H. sapiens</i>	38	59	5.5 e-78
TvPAP-II	TVAG_107950	<i>A. thaliana</i>	23	42	5.7 e-06
TvPAB1	TVAG_385330	<i>S. cerevisiae</i>	37	56	6.9 e-55
TvPC4	TVAG_255210	<i>H. sapiens</i>	43	66	1.3 e-07
TvSimplequina	TVAG_494090	<i>H. sapiens</i>	21	40	0.00036
		<i>A. thaliana</i>			
TvFY	TVAG_121040		35	52	1.4 e-63
TvSsu72	TVAG_229140	<i>S. cerevisiae</i>	35	50	3.1 e-29
TvHrp1	TVAG_245890	<i>S. cerevisiae</i>	30	50	6.1 e-15
TvMpe1	TVAG_265310	<i>S. cerevisiae</i>	32	53	2.7 e-05
TvSw2p	TVAG_016940	<i>S. cerevisiae</i>	24	43	2.0 e-11
TvGlc7p-I	TVAG_101240	<i>S. cerevisiae</i>	56	75	1.6 e-96
TvGlc7p-II	TVAG_379450	<i>S. cerevisiae</i>	57	74	2.0 e-96
TvhRrbp6				41	2.0 e-24
	TVAG_480160	<i>H. sapiens</i>	20		

Figura 8: Panel A. Estudio comparativo de las estructuras tridimensionales de CPSF160, CPSF73, PAP y PC4 de *T. vaginalis* contra sus homólogos en *H. sapiens*. Las estructuras terciarias se obtuvieron mediante modelación utilizando el programa Phyre 2. En cada caso se indica el porcentaje de cobertura que se obtuvo de la secuencia primaria comparada con el cristal utilizado para generar el modelo. Los valores de confianza obtenidos para los modelos son superiores al 90%. **Panel B.** Estudio comparativo de la estructura tridimensional de PAB1, CFIm25, CPSF73-II y FY de *T. vaginalis* comparadas con sus homólogos.





8.2 Expresión del RNAm de algunos de los genes que codifican para las proteínas candidatas de *T. vaginalis* involucradas en el procesamiento del 3'UTR de los RNAm.

Para las proteínas TvPSF2p, TvCPSF160, TvFIP1, TvCF1m25, TvPAP1, TvPC4, TvPAB1. TvCPSF73-II y TvCPCF73 de la probable maquinaria de procesamiento del extremo 3'UTR de *T. vaginalis* se obtuvo la secuencia del RNAm que las codifica de la base de datos Trich DB (<http://trichdb.org/trichdb/>). A cada secuencia de RNAm, se le realizó un BLASTn en la misma base de datos para obtener aquellas secuencias de RNAm que tuvieran similitud con los genes blanco, estas secuencias fueron alineadas con el blanco correspondiente para determinar las zonas menos conservadas y diseñar los iniciadores correspondientes para amplificar los genes mediante RT-PCR. El ensayo de RTPCR de los genes *tvpsf2p*, *tvcpsf160*, *tvfip1*, *tvcf1m25*, *tvpap1*, *tvpc4*, *tvpab1*, *tvcpsf73-II*, *tvpcf73*, se realizó, obteniendo el RNA total del aislado HGMN01 de *T. vaginalis* crecidos en presencia y ausencia de Zn²⁺. En la **Fig. 9** se muestran los amplicones de los genes que codifican para elementos de la posible maquinaria de procesamiento del 3'UTR en *T. vaginalis*; donde se observan amplicones de 186, 399, 229, 262, 177, 209 y 168 pb que corresponden al tamaño esperado para los fragmentos de los transcritos de *tvpsf2p*, *tvcpsf160*, *tvfip1*, *tvpap1*, *tvpc4*, *tvcf1m25* y *tvpab1*. El gen de α -tubulina (α -tub) se utilizó como control de carga (159 pb). Los resultados muestran que al menos los 9 genes analizados se expresan en *T. vaginalis*. De manera interesante los amplicones de 338 y 291 pb, que corresponden a *tvcpsf73* y *tvcpsf73-II* se observaron en condiciones normales, sin embargo cuando los parásitos son crecidos en presencia de Zn²⁺ la expresión de estos genes no se observa.

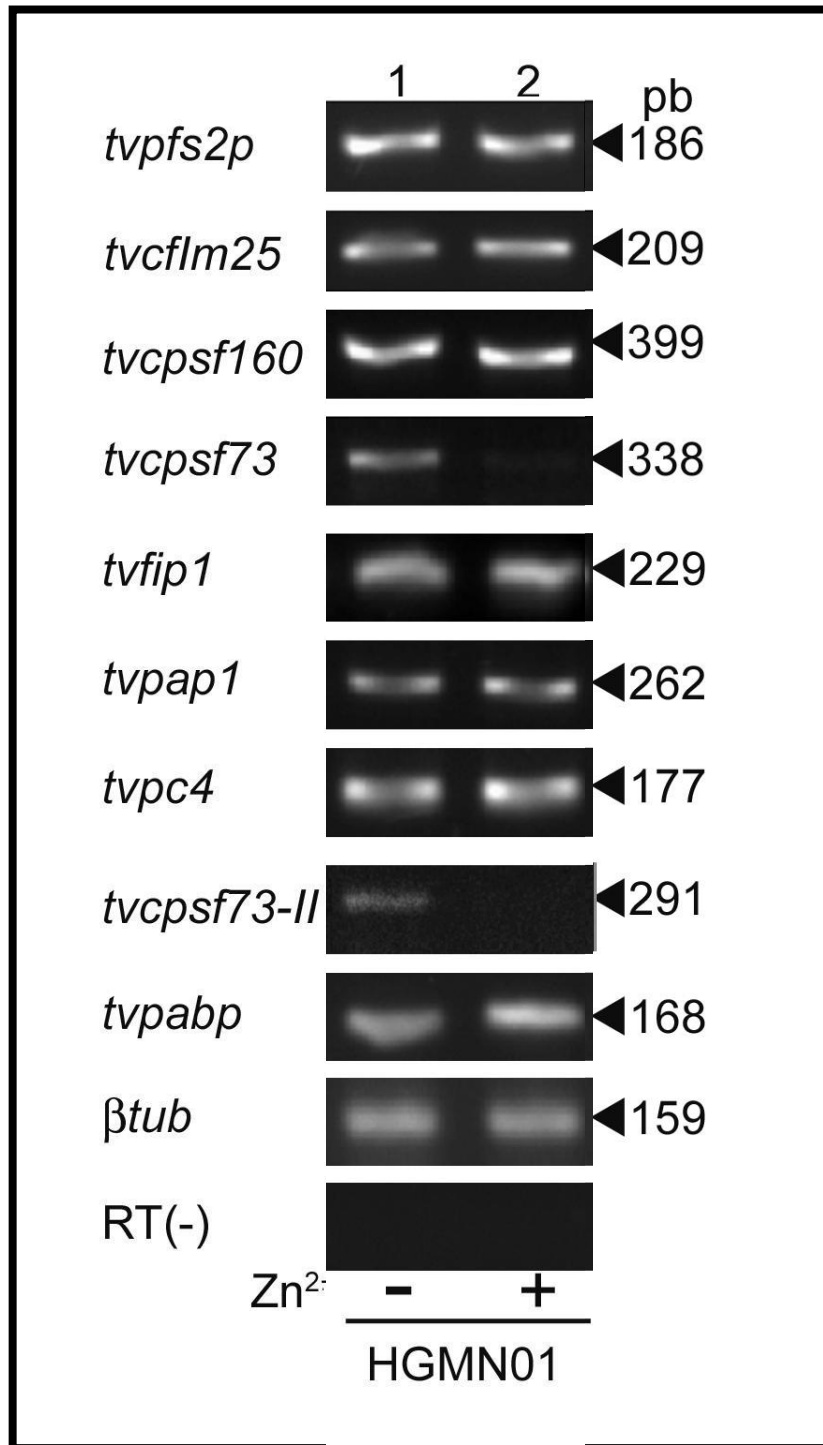
Tabla 10.- Características de los oligos iniciadores utilizados para la RTPCR.
Se señala el nombre propuesto del gen en *T. vaginalis*, su código de acceso en la base de datos Trich DB, el tipo de oligo, su secuencia, el tamaño del oligo y el tamaño del amplicón. En rojo se señala la secuencia del adaptador universal.

AMP*, el tamaño esperado del amplicón en pb.

GEN	ACCESO	TIPO DE OLIGO	SECUENCIA DEL OLIGO	Pb	AMP*
<i>tvcpf160</i>	TVAG_077030	Forward	5'-AGG TGA CAC TAT AGA ATA CAA TCG AGG ATT GCT TGG3'	36	399
		Reverse	5'-GTA CGA CTC ACT ATA GGG ACT CAG TCC GCA AAG AGA T-3'	37	
<i>tvcpf73</i>	TVAG_437970	Forward	5'-AGG TGA CAC TAT AGA ATA GAC GCA ATT GAT CCT GC 3'	35	338
		Reverse	5'-GTA CGA CTC ACT ATA GGG ACA GGG TAG CAT GTC ATC T-3'	37	
<i>tvfip1</i>	TVAG_463950	Forward	5'-AGG TGA CAC TAT AGA ATA AGC CAA AGG ATG AGA ATC C-3'	37	229
		Reverse	5'-GTA CGA CTC ACT ATA GGG AGT TGA ATT AAT GCG CGT GT-3'	38	
<i>tvclm25</i>	TVAG_495750	Forward	5'-AGG TGA CAC TAT AGA ATA CGT ATT CCC TCG ATC GG-3'	35	209
		Reverse	5'-GTA CGA CTC ACT ATA GGG ATA CCC ATG TCC CAA AGC3'	36	
<i>tvpap1</i>	TVAG_388620	Forward	5'-AGG TGA CAC TAT AGA ATA CCG AAT TGC AGG CTA GTA3'	36	262
		Reverse	5'-GTA CGA CTC ACT ATA GGG ATT ACG ACA CCT AAA CGG TAA G-3'	40	
<i>tvpc4</i>	TVAG_255210	Forward	5'-AGG TGA CAC TAT AGA ATA GAC ATC CTA CGT CTT TCC TT-3'	38	177
		Reverse	5'-GTA CGA CTC ACT ATA GGG ATT CCT TTC TTG CCA GGA A-3'	37	
<i>tvpsf2p</i>	TVAG_121040	Forward	5'-AGG TGA CAC TAT AGA ATA GTT CCA CAT TCC AGC TTC3'	36	186
		Reverse	5'-GTA CGA CTC ACT ATA GGG AGC TTC TCA ACG TTG TTG TAG-3'	39	
<i>tvcpf73II</i>	TVAG_183200	Forward	5'-AGG TGA CAC TAT AGA ATA CTT TGT CCA AGT TCG TGG3'	36	291
		Reverse	5'-GTA CGA CTC ACT ATA GGG AAT ATT TCT TGG AGA CCG ACC-3'	39	
<i>tvpa1</i>	TVAG_385330	Forward	5'-AGG TGA CAC TAT AGA ATA CAG TTC AGG AAC CCA GA-3'	35	168
		Reverse	5'-GTA CGA CTC ACT ATA GGG AAA GTG GTT TGA TGT AGA CGT -3'	39	
<i>tv□-tubulina</i>	TVAG_148390	Forward	5'-AGG TGA CAC TAT AGA ATA CAT TGA TAA CGA AGC TCC TTT ACG AT-3'	44	159
		Reverse	5'-GTA CGA CTC ACT ATA GGG AGC ATG TTG TGC CGG ACA TAA CCA T-3'	43	

Figura 9.- Amplicones de los transcritos de los genes candidatos del aislado HGMN01 que participan en la maquinaria de procesamiento de *T. vaginalis*.

RT-PCR a partir de parásitos crecidos en ausencia (carriles 1 y 3) o presencia (carriles 2 y 4) de Zn^{2+} . Los transcritos amplificados son *tvpsf2p* de 186 pb; *tvclm25*, 209 pb; *tvcpf160*, 399 pb; *tvfip1*, 229 pb; *tvap1*, 262 pb; *tvpc4*, 177 pb; *tvabp*, 168 pb; y β -*tub* 159 pb. de las condiciones evaluadas. Las flechas indican el tamaño molecular de los transcritos. Control negativo (RT(-)) de α -tubulina.



8.3 INMUNODETECCIÓN DE PC4 EN EXTRACTOS DE *T. vaginalis*.

La proteína TvPC4 se inmunolocalizó utilizando un anticuerpo heterólogo policlonal de *E. histolytica* que reconoce a la proteína PC4. Extractos proteicos totales de *T. vaginalis* del aislado HGMN01 se separaron mediante SDS-PAGE en un gel al 15% y un gel espejo se transfirió a una membrana de nitrocelulosa para realizar un WB.

De manera interesante, se inmunodetectó una banda de aproximadamente 11 kDa (**Fig 10**) sugiriendo que *T. vaginalis* aislado HGMN01 expresa a la proteína TvPC4, como control del experimento se utilizó el suero pre-inmune con el cual no se inmunodetectó ninguna banda, también se utilizó la proteína recombinante EhPC4 de *E. histolytica* inmunodetectando una banda de aproximadamente 22 kDa que corresponde al tamaño molecular de la proteína EhPC4 previamente descrita, esta diferencia en tamaño es debido a que EhPC4 presenta un dominio rico en lisina, como control se utilizó el suero pre-inmune el cual no detectó alguna banda.

Para obtener una mejor resolución de la proteína y distinguirla entre otras proteínas que pudieran tener el mismo peso molecular, se decidió realizar un 2DWB, para lo cual se obtuvo extractos proteicos totales de *T. vaginalis* del aislado HGMN01, los cuales, se separaron mediante una electroforesis de doble dimensión resolviendo un total de 470 manchas, concordando con previos estudios reportados por el grupo de investigación. Un duplicado de este gel fue transferido a una membrana de nitrocelulosa para realizar un 2D-WB utilizando el anticuerpo α -EhCP4 (1:100) (**Fig 11**) De manera interesante, dos manchas fueron inmunoreconocidas por el anticuerpo heterólogo α -PC4 en extractos de parásitos crecidos en ausencia y presencia de Zn^{2+} .

Ambos proteomas de los parásitos crecidos en ausencia y presencia de Zn^{2+} se analizaron mediante el programa PD-Quest, determinamos que las manchas inmunodetectadas por el anticuerpo heterólogo tienen un pI similar de 5.1, sin embargo varían en tamaño molecular, la mancha indicada con el número 1, tienen un peso molecular aproximadamente a 12.3 kDa, mientras que la mancha número 2 es de 11.7 kDa.

Figura 10.- Inmunodetección de la proteína TvPC4 en *T. vaginalis* HGMN01 en 1-D.

Western blot de A) Extractos proteicos totales del aislado HGMN01 separados mediante SDS-PAGE en un gel al 15% (carriles 2 y 3) usando el anticuerpo antiPC4 heterólogo de *E. histolytica* (carril 5) y el suero preinmune (carril 4). La cabeza de flecha indica la banda con peso molecular de 11 kDa inmunoreconocida por el anticuerpo anti-PC4. Marcadores de tamaño molecular (carril 1). B) Proteína recombinante EhPC4 de *E. histolytica* que se inmunodetectó con el anticuerpo anti-EhPC4 como control npositivo. Cabeza de flecha que indica la banda inmunodetectada por el anticuerpo anti-EhPC4 de *E. histolytica* como control negativo.

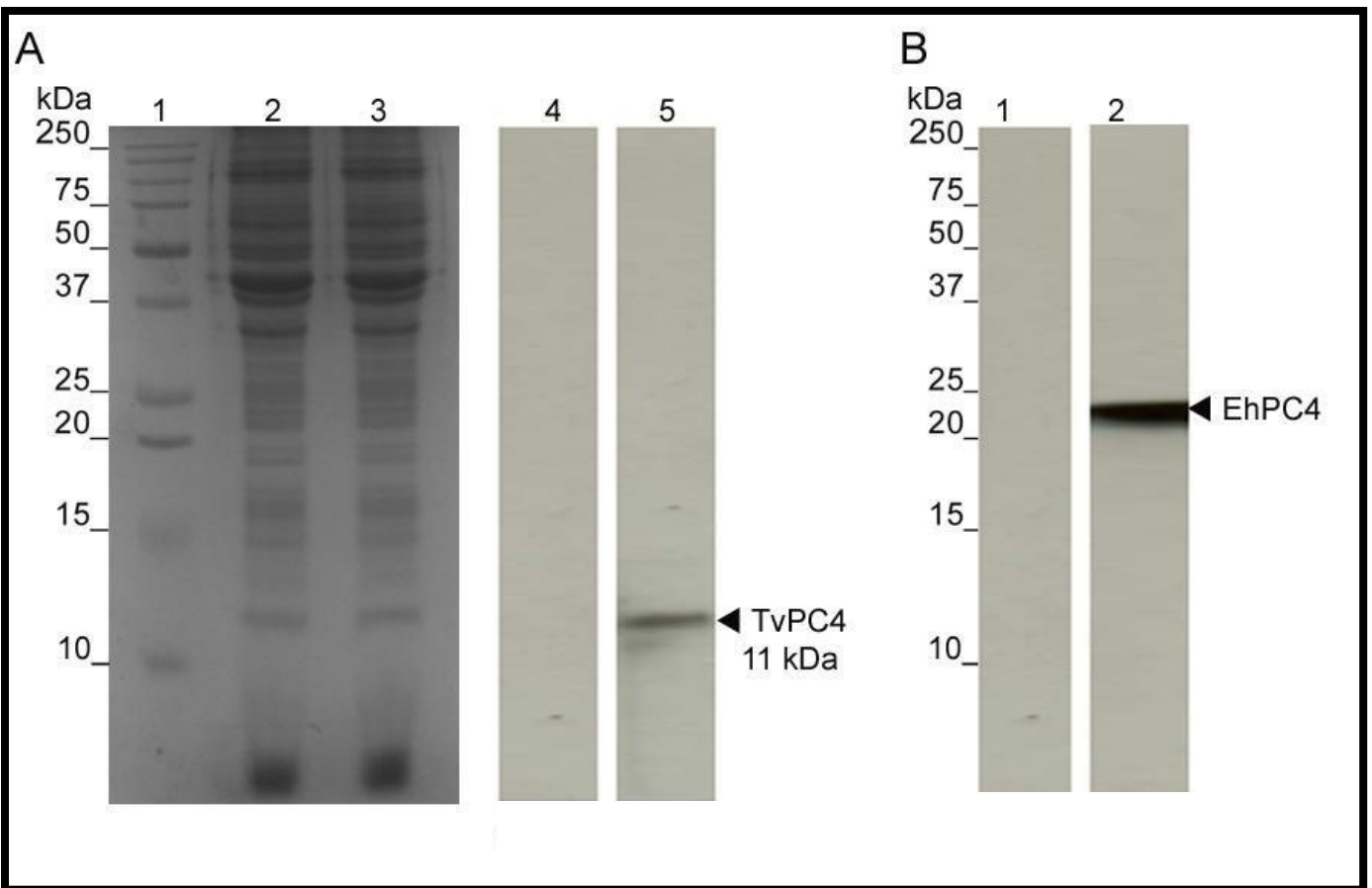
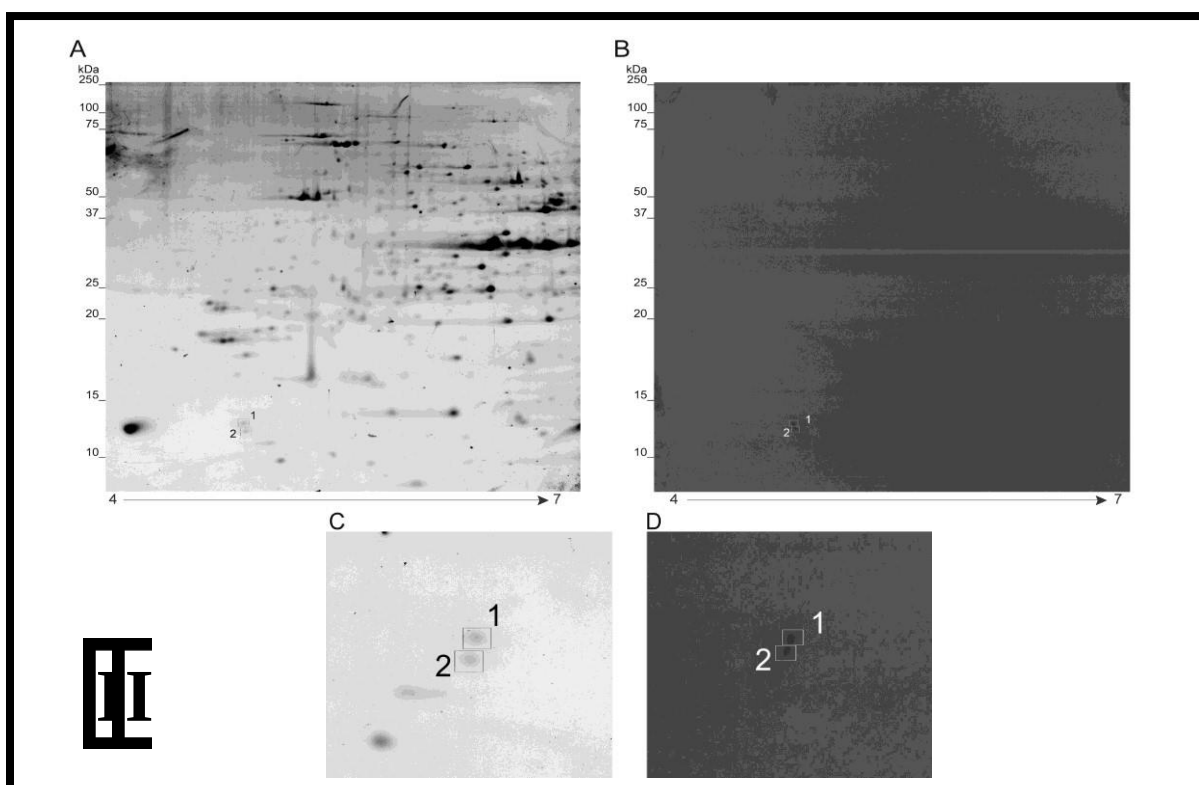
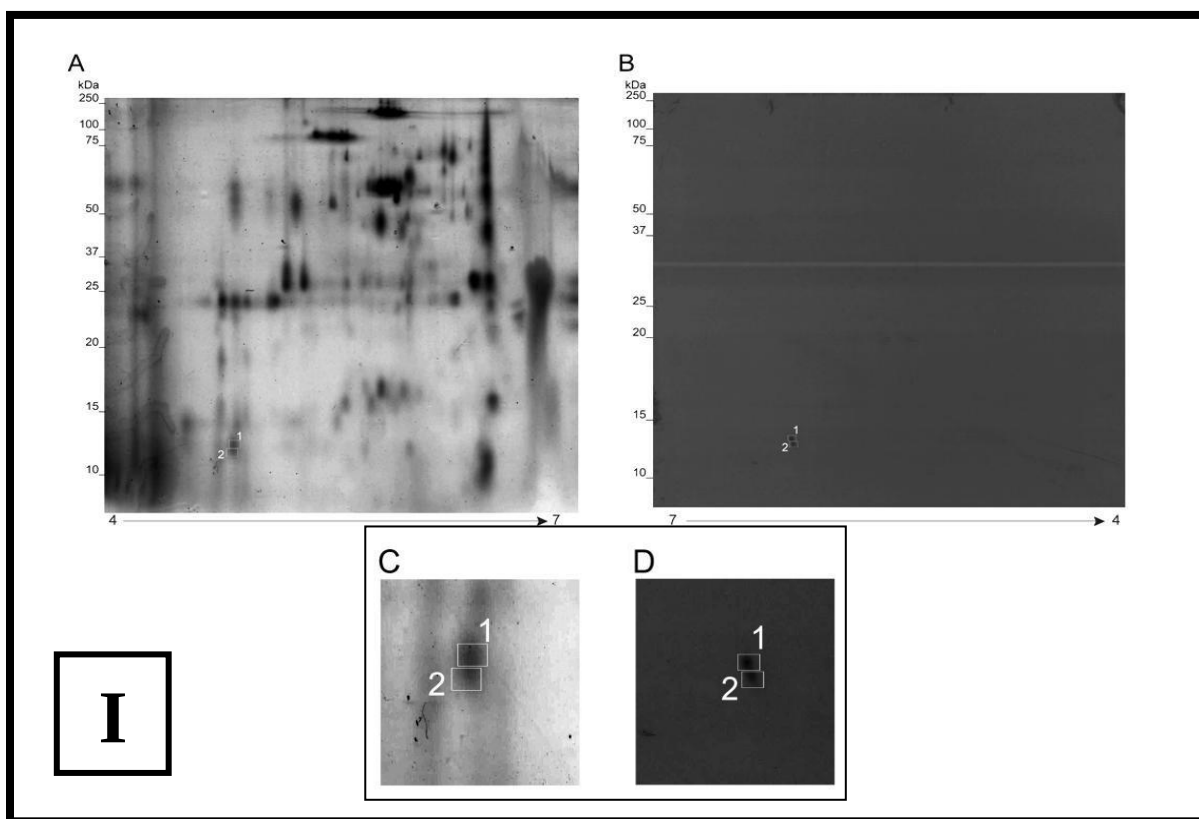


Figura 11.- Inmunoreconocimiento de PC4 en el aislado HGMN01 de *T. vaginalis* en 2-D. PANEL I. A) Extractos proteicos totales en condiciones normales. B) WB del proteoma de *T. vaginalis* HGMN01, utilizando el anticuerpo heterólogo α -PC4 (1-100). C) y D) Acercamiento de las manchas reconocidas por el anticuerpo. **PANEL II,** A) Extractos proteicos totales en condiciones de cultivo con Zn^{2+} B) WB del proteoma de HGMN01 con Zn^{2+} , utilizando el anticuerpo heterólogo α -PC4 (1-100). C) y D) Acercamiento de las manchas reconocidas por el anticuerpo.



Adicionalmente, para determinar la influencia del Zn^{2+} sobre la expresión de la proteína TvPC4 en *T. vaginalis*. Se realizó una electroforesis de doble dimensión de los extractos de parásitos del aislado HGMN01 crecidos en presencia de Zn^{2+} , obteniendo un total de 417 manchas, un gel espejo fue transferido a una membrana de nitrocelulosa para realizar un 2DE-WB (**Fig.11** Panel II,B) Las figuras 12,II, C yD, muestran un acercamiento de la región del proteoma y del 2DE-WB, respectivamente, donde fueron localizados las manchas inmunoreconocidas.

De manera interesante, el anticuerpo heterólogo α -PC4 inmunoreconoció dos manchas de punto isoeléctrico y tamaños moleculares similares a los encontrados en ausencia del catión, para determinar si existían diferencias en la expresión de la proteína TvPC4 se realizó un análisis densitométrico con el programa Melany Viewer 7, obteniendo como resultado que no existe una diferencia significativa de la proteína en condiciones analizadas (**Fig.12**).

Debido al resultado obtenido de dos proteínas inmunodetectadas por el anticuerpo anti-PC4, con igual punto isoeléctrico y diferente peso molecular analizamos la presencia de posibles modificaciones postraduccionales. A partir de la secuencia de aminoácidos, utilizando como herramienta, el programa YinOYan, del servidor Expasy. El análisis de la secuencia de aminoácidos de TvPC4 nos mostró la presencia de dos posibles sitios de glicosilación en el aminoácido 17 (T) y en el aminoácido 95 (S), de manera interesante, el aminoácido 17 también es susceptible de ser fosforilado (**Fig 13**).

Figura 12.-. Análisis por Melanie de la expresión de TvPC4 de *T. vaginalis* en presencia y ausencia de Zn^{2+} .

Mancha 1 y 2 del aislado HGMN01. En ambos casos analizados mediante el programa Melanie Viewer versión 7.0.

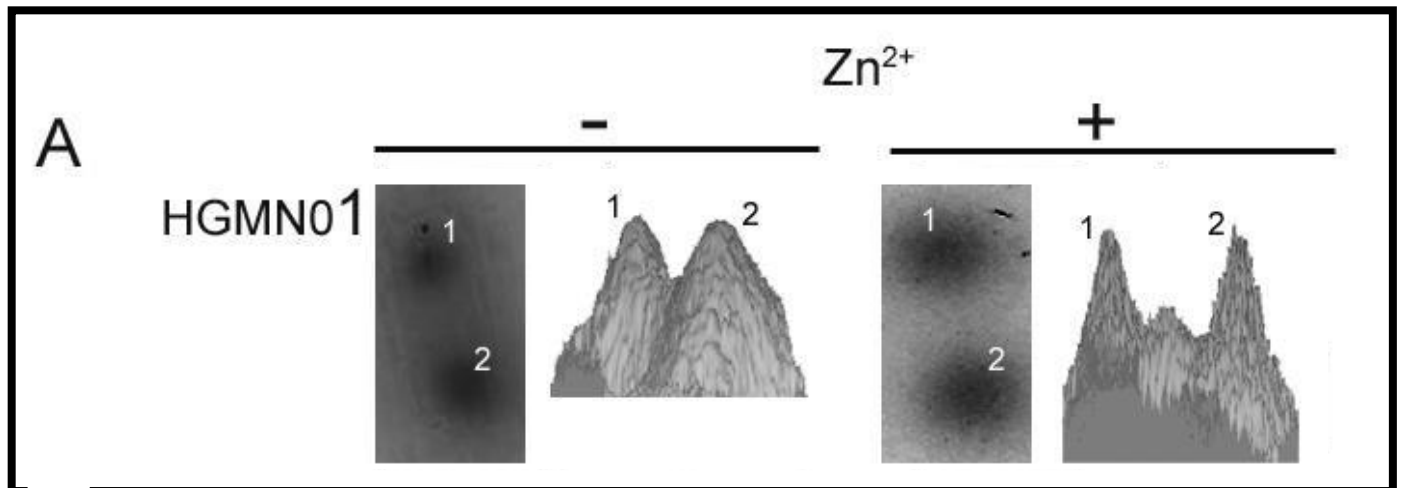
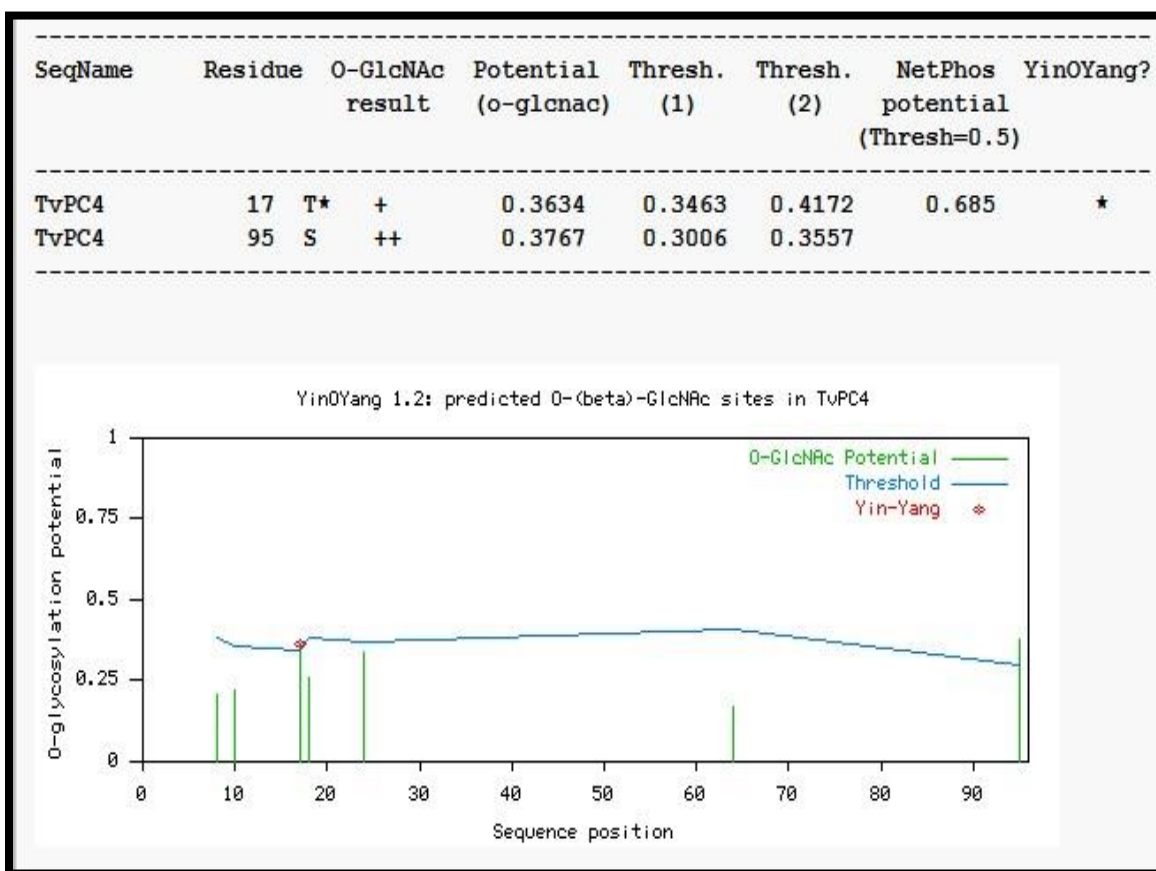


Figura 13.- Predicción de modificaciones postraduccionales de TvPC4 de *T. vaginalis*. Estructura primaria de TvPC4 analizada mediante el programa YingOYang en donde se muestra los posibles sitios de glicosilación de TvPC4 en los residuos T17 y S95. Las líneas verticales verdes indican todos los posibles sitios de O-glicosilación. La línea azul horizontal indica el punto de corte de confianza y el punto rojo indica los puntos susceptibles de fosforilación.



IX. DISCUSION

El complejo CPSF en humanos, responsable del reconocimiento de PAS, está compuesto por varias subunidades, de manera interesante, analizando el genoma de *T. vaginalis* determinamos posibles homólogos a estas subunidades. Adicionalmente identificamos, en este parásito de transmisión sexual, más de una secuencia homóloga para CPSF30, asimismo reportamos que *T. vaginalis* presenta dos proteínas distintas con la misma función, que son hFip1 y AtFipS5 y una secuencia homóloga reportada en *Arabidopsis thaliana* denominada CPSF73II, la cual tiene una alta homología con la proteína de mamíferos y su función es vital para la planta, en especial durante el control de la floración (Hunt 2008), por este motivo, el hecho de identificar a un homólogo de esta proteína en *T. vaginalis* nos llevó a especular la posibilidad de que *T. vaginalis* posea dos proteínas alternativas para la misma función.

Sin embargo, esta no fue la única proteína propia de plantas que identificamos mediante análisis *in silico*, cabe mencionar que las proteínas CPSF73-II y TY, cuyos homólogos están presentes en el genoma de *T. vaginalis*, son consideradas proteínas muy específicas de plantas (Millevoi y Vagner 2010). Previamente se ha reportado que *T. vaginalis* posee algunos genes que son característicos de plantas, tal es el caso de los genes que codifican para las proteínas TvLEGU-1 y TvLEGU-1 (León-Félix, Ortega-López y col.; 2004) una proteinasa tipo legumaina.

Por lo anterior, nos pareció interesante estudiar la expresión de estos peculiares genes; no obstante no se detectó RNAm para *tvcpf73-II* bajo las condiciones aquí descritas. Aunque amplificando a *tvcpf73-II* utilizando DNA genómico como molde, se logró obtener el amplicón de tamaño esperado. Adicionalmente, realizamos en análisis de secuencias de expresión (ESTs) para *tvcpf73-II* en la base de datos TrichDB, sin embargo, no encontramos reportes para *tvcpf73* ó para *tvcpf73-II*. Cabe mencionar que los ESTs almacenados en la base de datos consultada

corresponden a parásitos provenientes de cultivos sincronizados, en estadio pseudoquiste, en su forma ameboide o en la fase G1/S del ciclo celular, lo cual nos indicaría que los RNAm de *tvcpstf73* ó para *tvcpstf73-II* no se expresan bajo las condiciones reportadas en la base de datos TrichDB. Los datos obtenidos en el laboratorio sugieren que estos RNAm pudieran expresarse en alguna otra condición no evaluada en este estudio.

Como previamente mencionamos, el complejo CstF está conformado por varias subunidades. De manera particular la subunidad 77 identificada en el genoma de *T. vaginalis*, posee mayor homología con la subunidad correspondiente de *A. thaliana*, Por otro lado identificamos en el parásito, dos secuencias homólogas para la subunidad 64 y dos para la subunidad 50, dichas secuencias muestran homología con distintas especies,

Además, *T. vaginalis* posee dos secuencias homólogas a Rnap15 de *Saccharomyces cerevisiae*, que es homóloga a la subunidad 64 en mamíferos y que participa en el reconocimiento de los elementos auxiliares río abajo y ricos en A en mamíferos y en levadura, respectivamente (Mandel, Bai y col.; 2008; Millevoi y Vagner 2010). Nuestros hallazgos sugieren que el parásito posee proteínas alternativas para conformar los complejos involucrados en la maquinaria de procesamiento del RNAm, ya sea que sean proteínas alternativas que se expresan bajo algún estímulo particular o bien de proteínas que reconozcan diferentes elementos auxiliares aún no descritos. Cabe señalar que dentro de los elementos auxiliares descritos para *T. vaginalis* sólo se ha reconocido el elemento auxiliar río abajo pero se ha sugerido la existencia de al menos otros dos elementos que ayuden a llevar el proceso de corte y poliadenilación en forma eficiente y específica (Fuentes, Barrera y col.; 2012), por lo cual, la identificación de proteínas de unión a elementos auxiliares puede hablarnos de un mecanismo de control o de regulación que activaría a *tvcsstf77* y *tvcsstf64* bajo ciertas condiciones y a *tvnap14* y *tvnap15* en otra circunstancia.

El complejo llamado Factor de Corte I de mamíferos, consta de 3 subunidades en el ser humano (Mandel, Bai y col.; 2008), sin embargo en el genoma de *T. vaginalis* únicamente identificamos a la subunidad 25. De manera interesante, en el parásito protozoario *Entamoeba histolytica* también se identificó a esta subunidad (López-Camarillo, Orozco y col.; 2005), la cual comparte una alta homología con la identificada para trichomonas. Esto nos puede indicar que esta proteína tiene una función semejante en ambos parásitos y que no puede ser la misma que lleva a cabo en los mamíferos, pues en estos últimos, su función es reconocer el elemento auxiliar río arriba para estimular el corte del RNA recién sintetizado.

Actualmente no se ha descrito ningún elemento río arriba en protozoarios, por lo que sugerimos que la función de CFI25 en los protozoarios podría ser únicamente la de unir PAP con la maquinaria de corte y poliadenilación, sin embargo, faltaría evidencia experimental que soporte esta teoría.

En levadura el Factor de Corte I se subdivide en un pequeño complejo llamado IA que está formado por Rnap14, Rnap15, Clp1 y Pcf11; que reconoce elementos de respuesta río arriba del punto de corte y el factor llamado FCIB compuesto por una sola subunidad llamada Hrp1p, que reconoce un elemento río arriba muy lejano al punto de corte. Esta última subunidad tiene una homóloga en *T. vaginalis* (Millevoi y Vagner 2010). Cabe mencionar que Rnap14 y Rna15 que son homólogos de CstF 77 y 64 respectivamente tienen una función similar que es reconocer elementos auxiliares pero en diferente posición.

El Factor de Corte II en humanos fue descrito después de que sus dos subunidades fueran descritas en *S. cerevisiae*, ambas proteínas tienen secuencias similares en *T. vaginalis*; PCF11 cuya secuencia tiene más similitud con la del humano y Clp1p que tiene dos secuencias, una se asemeja más con la descrita en levadura y otra a su homóloga CLPS3 de *A. thaliana*. Estas proteínas tienen la función de unir a los factores de corte con el complejo CPSF y son necesarias para la terminación adecuada del extremo 3'UTR de los RNAm.

Existen proteínas que no están involucradas en la formación de un complejo, pero que tienen una función importante dentro de la maquinaria de procesamiento del extremo 3'UTR de los RNAm de *T. vaginalis*. Un ejemplo de ello es la proteína PAP, *T. vaginalis* presenta dos secuencias correspondientes para esta proteína, la de mayor valor predictivo es la que muestra mayor similitud con la de humano, mientras que otra secuencia muestra más semejanza a la de *A. thaliana*. En este sentido en *E. histolytica* se ha reportado la presencia de una proteína PAP canónica y de seis no-canónicas (López-Camarillo, Hernández de la Cruz y col.; 2010) por lo cual no resultó extraño la identificación de dos proteínas PAP en el parásito causante de la tricomonosis.

La simplequina es una proteína que sirve de andamiaje para el ensamble de toda la maquinaria, tiene su homóloga en *S. cerevisiae* denominada Pta1 y en *A. thaliana* conocida como SYM5 (Mandel, Bai y col.; 2008). La simplequina identificada en *T. vaginalis* tiene más semejanza con la de humano.

Posteriormente, seleccionamos a un grupo de genes que probablemente codifiquen para elementos *trans* de la maquinaria de procesamiento del extremo 3'UTR de *T. vaginalis* y se determinó la expresión del transcrito. De acuerdo a los antecedentes reportados para *E. histolytica*, PC4 una proteína involucrada en la maquinaria de procesamiento en este protozooario, presenta una expresión diferencial; adicionalmente diferentes condiciones de cultivo de *T. vaginalis* modifican la expresión de varias proteínas (Santi-Rocca, Weber y col.; 2008; Vázquez-Carrillo, Granados y col.; 2011). Por lo anterior, evaluamos la expresión de los transcritos seleccionados utilizando diferentes aislados, además de determinar el efecto del Zn^{2+} sobre la expresión de dichos RNAm. Nueve genes fueron analizados: *tvcpstf73*, *tvcpstf73-II*, *tvpsf2p*, *tvclm25*, *tvcpstf160*, *tvfip1*, *tvpc4*, *tvpap1* y *tvpabp*. De manera interesante, a nivel de RNAm estos genes no presentan diferencias significativas en su expresión.

Adicionalmente, la proteína homóloga a PC4 en *T. vaginalis* tiene más parecido con la de humano, cuya función es evitar la interacción de CstF64 con el elemento río abajo, inhibiendo el corte del RNA. De manera interesante, PC4 presenta una expresión diferencial en aislados de *E. histolytica* provenientes de un absceso hepático en comparación con otros aislados no patógenos (Santi-Rocca, Weber y col.; 2008). Por este motivo se decidió determinar si existe diferencia en la expresión de PC4, tanto a nivel de RNAm y proteína, en el aislado HGMN01 proveniente de cultivos en presencia o ausencia de Zn^{2+} . En este caso, no encontramos diferencias aún en presencia de Zn^{2+} ni en RNA ni a nivel de proteína por lo que este elemento de la maquinaria parece ser un elemento de expresión basal en *T. vaginalis*.

X. CONCLUSIONES

- El análisis *in silico* predice la presencia de 31 elementos proteicos para la maquinaria de corte y poliadenilación en *Trichomonas vaginalis*.
- En el genoma de *Trichomonas vaginalis* se encuentran presentes los ORF que codifican para proteínas involucradas en la maquinaria de procesamiento de poliadenilación de RNAm como *tvpsf2p*, *tvcpf160*, *tvcpf73*, *tvcpf 73-II*, *tvfip1*, *tvcl25*, *tvap1*, *tvpc4* y *tvabp*.
- Los transcritos de *tvpsf2p*, *tvcpf160*, *tvfip1*, *tvap1*, *tvpc4* y *tvabp* cuyos productos proteicos posiblemente se encuentren relacionados con la maquinaria de procesamiento de poliadenilación del RNAm, son expresados en *T. vaginalis* bajo las condiciones de cultivo evaluadas.
- Los transcritos de *tvcpf73* y *tvcpf73-II* no son detectados bajo en presencia de Zn^{2+} .
- La proteína TvPC4 se encuentra presente en los aislado HGMN01 de *T. vaginalis*. De manera interesante, TvCP4 fue inmunolocalizada en los proteomas de ambas condiciones, encontrando dos isoformas de esta proteína de masa molecular diferente (11.7 kDa y 12.3 kDa) pero similar pI (5.1).

- La diferencia de tamaño molecular posiblemente se atribuya a que TvPC4 puede presentar glicosilaciones en los residuos T17 y S95.

Teniendo como base los resultados obtenidos podemos proponer el siguiente modelo de la maquinaria de procesamiento del extremo 3'UTR de los RNAm de *T. vaginalis* en el que participan el complejo CPSF, conformado por las subunidades 30, 100, 160, 73 y 73-II, la proteína Fip1 que se mantiene en contacto con la señal de poliadenilación; complejo CstF de 77, 64 y 50 kDa, la subunidad 64 en contacto con el elemento corriente abajo y la subunidad 77 interaccionando con CPSF; complejo CFI formado por un heterodímero más frecuente (25/68), interaccionando con el elemento corriente arriba; las proteínas sin complejo, PAP1, PAPB, simplequina y CP4 esta última en su posición entre CstF64 y el elemento corriente abajo. Las proteínas Rrp6, SSu72p, Sw2p, Hrp1, Mpe1 interaccionan con CPSF y PAPB. En negro RNA recién sintetizado con los elementos de secuencia (**Fig 14**).

Figura 14.- Modelo de la maquinaria de procesamiento del extremo 3'UTR de los RNAm de *T. vaginalis*. En rojo se observan los elementos *cis*, teniendo en cuenta que existen al menos dos elementos extras que aún no están descritos. En gris oscuro se muestran los elementos proteícos cuyo RNAm se ha detectado en este parásito. En gris claro tenemos a TvCPSF73-II que no presenta transcrito por lo cual se puede eliminar. Y en blanco se observan el resto de los elementos que faltan por estudiar.



XII.- REFERENCIAS

- Abdolrasouli, A., A. Amin y col.; (2007). "Persistent urethritis and prostatitis due to *Trichomonas vaginalis*: A case report." Can J Infect Dis Med Microbiol **18**(5): 308310.
- Alderete, J. F. (1999). "Iron modulates phenotypic variation and phosphorylation of P270 in double-stranded RNA virus-infected *Trichomonas vaginalis*." Infect Immun **67**(8): 4298-4302.
- Alderete, J. F., M. Benchimol y col.; (2002). "The complex fibronectin--*Trichomonas vaginalis* interactions and Trichomonosis." Parasitol Int **51**(3): 285-292.
- Alderete, J. F. y G. E. Garza (1985). "Specific nature of *Trichomonas vaginalis* parasitism of host cell surfaces." Infect Immun **50**(3): 701-708.
- Alderete, J. F. y G. E. Garza (1988). "Identification and properties of *Trichomonas vaginalis* proteins involved in cytoadherence." Infect Immun **56**(1): 28-33.
- Alderete, J. F., K. W. Millsap y col.; (2001). "Enzymes on microbial pathogens and *Trichomonas vaginalis*: molecular mimicry and functional diversity." Cell Microbiol **3**(6): 359-370.
- Alderete, J. F., J. Nguyen y col.; (2004). "Heme-iron increases levels of AP65-mediated adherence by *Trichomonas vaginalis*." Microb Pathog **36**(5): 263-271.
- Alvarez-Sanchez, M. E., L. Avila-Gonzalez y col.; (2000). "A novel cysteine proteinase (CP65) of *Trichomonas vaginalis* involved in cytotoxicity." Microb Pathog **28**(4): 193-202.
- Álvarez-Sánchez, M. E., L. Ávila-González y col.; (2000). "A novel cysteine proteinase (CP65) of *Trichomonas vaginalis* involved in cytotoxicity." Microb Pathog **28**(4): 193-202.
- Álvarez-Sánchez, M. E., B. I. Carvajal-Gamez y col.; (2008). "Polyamine depletion downregulates expression of the *Trichomonas vaginalis* cytotoxic CP65, a 65-kDa cysteine proteinase involved in cellular damage." Int J Biochem Cell Biol **40**(11): 2442-2451.
- Álvarez-Sánchez, M. E., E. Solano-González y col.; (2007). "Negative iron regulation of the CP65 cysteine proteinase cytotoxicity in *Trichomonas vaginalis*." Microbes Infect **9**(14-15): 1597-1605.
- Andreini, C., L. Banci y col.; (2006). "Counting the zinc-proteins encoded in the human genome." J Proteome Res **5**(1): 196-201.
- Arroyo, R., J. Engbring y col.; (1992). "Molecular basis of host epithelial cell recognition by *Trichomonas vaginalis*." Mol Microbiol **6**(7): 853-862.
- Arroyo, R., A. González-Robles y col.; (1993). "Signalling of *Trichomonas vaginalis* for amoeboid transformation and adhesion synthesis follows cytoadherence." Mol Microbiol **7**(2): 299-309.
- Bai, Y., T. C. Auperin y col.; (2007). "Crystal structure of murine CstF-77: dimeric association and implications for polyadenylation of mRNA precursors." Mol Cell **25**(6): 863-875.
- Balbo, P. B. y A. Bohm (2007). "Mechanism of poly(A) polymerase: structure of the enzyme-MgATP-RNA ternary complex and kinetic analysis." Structure **15**(9): 1117-1131.

- Barabino, S. M., W. Hubner y col.; (1997). "The 30-kD subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are RNA-binding zinc finger proteins." Genes Dev **11**(13): 1703-1716.
- Beach, D. H., G. G. Holz, Jr. y col.; (1990). "Fatty acid and sterol metabolism of cultured *Trichomonas vaginalis* and *Tritrichomonas foetus*." Mol Biochem Parasitol **38**(2): 175-190.
- Beach, D. H., G. G. Holz, Jr. y col.; (1991). "Phospholipid metabolism of cultured *Trichomonas vaginalis* and *Tritrichomonas foetus*." Mol Biochem Parasitol **44**(1): 97-108.
- Benchimol, M., C. Batista y col.; (1990). "Fibronectin- and laminin-mediated endocytic activity in the parasitic protozoa *Trichomonas vaginalis* and *Tritrichomonas foetus*." J Submicrosc Cytol Pathol **22**(1): 39-45.
- Benchimol, M., I. de Andrade Rosa y col.; (2008). "Trichomonas adhere and phagocytose sperm cells: adhesion seems to be a prominent stage during interaction." Parasitol Res **102**(4): 597-604.
- Borchardt, K. A., S. al-Haraci y col.; (1995). "Prevalence of *Trichomonas vaginalis* in a male sexually transmitted disease clinic population by interview, wet mount microscopy, and the InPouch TV test." Genitourin Med **71**(6): 405-406.
- Campbell, L., V. Woods y col.; (2008). "Evaluation of the OSOM Trichomonas rapid test versus wet preparation examination for detection of *Trichomonas vaginalis* vaginitis in specimens from women with a low prevalence of infection." J Clin Microbiol **46**(10): 3467-3469.
- Carlton, J. M., R. P. Hirt y col.; (2007). "Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*." Science **315**(5809): 207-212.
- Carter, J. E. y K. C. Whithaus (2008). "Neonatal respiratory tract involvement by *Trichomonas vaginalis*: a case report and review of the literature." Am J Trop Med Hyg **78**(1): 17-19.
- Carvajal-Gamez, B. I., R. Arroyo y col.; (2011). "Putrescine is required for the expression of eif-5a in *Trichomonas vaginalis*." Mol Biochem Parasitol **180**(1): 8-16.
- Carvalho, S. E., D. Freitas y col.; (2005). "*Trichomonas vaginalis*: identification of a triacylglycerol acylhydrolase." Exp Parasitol **111**(4): 260-263.
- Conrad, M. D., A. W. Gorman y col.; (2012). "Extensive genetic diversity, unique population structure and evidence of genetic exchange in the sexually transmitted parasite *Trichomonas vaginalis*." PLoS Negl Trop Dis **6**(3): e1573.
- Crouch, M. V. y J. F. Alderete (2001). "*Trichomonas vaginalis* has two fibronectin-like iron-regulated genes." Arch Med Res **32**(2): 102-107.
- Chanfreau, G., S. M. Noble y col.; (1996). "Essential yeast protein with unexpected similarity to subunits of mammalian cleavage and polyadenylation specificity factor (CPSF)." Science **274**(5292): 1511-1514.
- Dailey, D. C., T. H. Chang y col.; (1990). "Characterization of *Trichomonas vaginalis* haemolysis." Parasitology **101 Pt 2**: 171-175.
- Danckwardt, S., M. W. Hentze y col.; (2008). "3' end mRNA processing: molecular mechanisms and implications for health and disease." EMBO J **27**(3): 482-498.
- De Jesús, J. B., P. Cuervo y col.; (2007). "A further proteomic study on the effect of iron in the human pathogen *Trichomonas vaginalis*." Proteomics **7**(12): 1961-1972.
- de Vries, D. S.,

- M. C. Bruin y col.; (2000). "[Congenital amegakaryocytic thrombocytopenia: indication for allogeneic stem cell transplantation]." Ned Tijdschr Geneeskd **144**(33): 1596-1598.
- Dominski, Z., X. C. Yang y col.; (2005). "A CPSF-73 homologue is required for cell cycle progression but not cell growth and interacts with a protein having features of CPSF-100." Mol Cell Biol **25**(4): 1489-1500.
- Espinosa, N., R. Hernández y col.; (2002). "Separable putative polyadenylation and cleavage motifs in *Trichomonas vaginalis* mRNAs." Gene **289**(1-2): 81-86.
- Fast, N. M. y W. F. Doolittle (1999). "*Trichomonas vaginalis* possesses a gene encoding the essential spliceosomal component, PRP8." Mol Biochem Parasitol **99**(2): 275-278.
- Fichorova, R. N. (2009). "Impact of *T. vaginalis* infection on innate immune responses and reproductive outcome." J Reprod Immunol **83**(1-2): 185-189.
- Fiori, P. L., P. Rappelli y col.; (1999). "The flagellated parasite *Trichomonas vaginalis*: new insights into cytopathogenicity mechanisms." Microbes Infect **1**(2): 149-156.
- Fuentes, V., G. Barrera y col.; (2012). "Functional analysis of sequence motifs involved in the polyadenylation of *Trichomonas vaginalis* mRNAs." Eukaryot Cell **11**(6): 725-734.
- García, A. F., M. Benchimol y col.; (2005). "*Trichomonas vaginalis* polyamine metabolism is linked to host cell adherence and cytotoxicity." Infect Immun **73**(5): 2602-2610.
- Gómez-Conde, E., R. Mena-López y col.; (2000). "*Trichomonas vaginalis*: chromatin and mitotic spindle during mitosis." Exp Parasitol **96**(3): 130-138.
- Gómez, C., M. Esther Ramírez y col.; (2010). "Regulation of gene expression in protozoa parasites." J Biomed Biotechnol **2010**: 726045.
- Graber, J. H., G. D. McAllister y col.; (2002). "Probabilistic prediction of *Saccharomyces cerevisiae* mRNA 3'-processing sites." Nucleic Acids Res **30**(8): 1851-1858.
- Guenther, P. C., W. E. Secor y col.; (2005). "*Trichomonas vaginalis*-induced epithelial monolayer disruption and human immunodeficiency virus type 1 (HIV-1) replication: implications for the sexual transmission of HIV-1." Infect Immun **73**(7): 4155-4160.
- Gunderson, J., G. Hinkle y col.; (1995). "Phylogeny of trichomonads inferred from small subunit rRNA sequences." J Eukaryot Microbiol **42**(4): 411-415.
- Hernández-Gutierrez, R., J. Ortega-López y col.; (2003). "A 39-kDa cysteine proteinase CP39 from *Trichomonas vaginalis*, which is negatively affected by iron may be involved in trichomonal cytotoxicity." J Eukaryot Microbiol **50 Suppl**: 696-698.
- Hernández, H., I. Sariego y col.; (2004). "Monoclonal antibodies against a 62 kDa proteinase of *Trichomonas vaginalis* decrease parasite cytoadherence to epithelial cells and confer protection in mice." Parasite Immunol **26**(3): 119-125.
- Heyworth, P. G., W. E. Gutteridge y col.; (1982). "Purine metabolism in *Trichomonas vaginalis*." FEBS Lett **141**(1): 106-110.
- Hirt, R. P., C. J. Noel y col.; (2007). "*Trichomonas vaginalis* surface proteins: a view from the genome." Trends Parasitol **23**(11): 540-547.
- Hobbs, M. M., D. M. Lapple y col.; (2006). "Methods for detection of *Trichomonas vaginalis* in the male partners of infected women: implications for control of trichomoniasis." J Clin Microbiol **44**(11): 3994-3999.
- Hocine, S., R. H. Singer y col.; (2010). "RNA processing and export." Cold Spring Harb Perspect Biol **2**(12): a000752.
- Honigberg, B. M., P. K. Gupta y col.; (1984). "Pathogenicity of *Trichomonas vaginalis*:

- cytopathologic and histopathologic changes of the cervical epithelium." Obstet Gynecol **64**(2): 179-184.
- Hu, J., C. S. Lutz y col.; (2005). "Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation." RNA **11**(10): 1485-1493.
- Hunt, A. G. (2008). "Messenger RNA 3' end formation in plants." Curr Top Microbiol Immunol **326**: 151-177.
- Ji, Z., J. Y. Lee y col.; (2009). "Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development." Proc Natl Acad Sci U S A **106**(17): 7028-7033.
- Juliano, C., G. Monaco y col.; (1987). "Action of anticytoskeletal compounds on in vitro cytopathic effect, phagocytosis, and adhesiveness of *Trichomonas vaginalis*." Genitourin Med **63**(4): 256-263.
- Kang, J. H., H. O. Song y col.; (2006). "*Trichomonas vaginalis* promotes apoptosis of human neutrophils by activating caspase-3 and reducing Mcl-1 expression." Parasite Immunol **28**(9): 439-446.
- Kaufmann, I., G. Martin y col.; (2004). "Human Fip1 is a subunit of CPSF that binds to Urach RNA elements and stimulates poly(A) polymerase." EMBO J **23**(3): 616-626.
- Keller, R. W., U. Kuhn y col.; (2000). "The nuclear poly(A) binding protein, PABP2, forms an oligomeric particle covering the length of the poly(A) tail." J Mol Biol **297**(3): 569-583.
- Keon, B. H., S. Schafer y col.; (1996). "Symplekin, a novel type of tight junction plaque protein." J Cell Biol **134**(4): 1003-1018.
- Kerwitz, Y., U. Kuhn y col.; (2003). "Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA." EMBO J **22**(14): 3705-3714.
- Krieger, J. N., M. A. Poisson y col.; (1983). "Beta-hemolytic activity of *Trichomonas vaginalis* correlates with virulence." Infect Immun **41**(3): 1291-1295.
- Kucknoor, A., V. Mundodi y col.; (2005). "*Trichomonas vaginalis* adherence mediates differential gene expression in human vaginal epithelial cells." Cell Microbiol **7**(6): 887-897.
- Kucknoor, A. S., V. Mundodi y col.; (2005). "Adherence to human vaginal epithelial cells signals for increased expression of *Trichomonas vaginalis* genes." Infect Immun **73**(10): 6472-6478.
- Kumar, P., N. K. Sharma y col.; (1990). "Trichomoniasis and candidiasis in consorts of females with vaginal discharge." Indian J Sex Transm Dis **11**(2): 54-56.
- Kummer, S., G. R. Hayes y col.; (2008). "Induction of human host cell apoptosis by *Trichomonas vaginalis* cysteine proteases is modulated by parasite exposure to iron." Microb Pathog **44**(3): 197-203.
- Lama, A., A. Kucknoor y col.; (2009). "Glyceraldehyde-3-phosphate dehydrogenase is a surface-associated, fibronectin-binding protein of *Trichomonas vaginalis*." Infect Immun **77**(7): 2703-2711.
- Langley, J. G., J. M. Goldsmid y col.; (1987). "Venereal trichomoniasis: role of men." Genitourin Med **63**(4): 264-267.
- Lehker, M. W. y J. F. Alderete (2000). "Biology of trichomonosis." Curr Opin Infect Dis **13**(1): 37-45.

- Lehker, M. W., R. Arroyo y col.; (1991). "The regulation by iron of the synthesis of adhesins and cytoadherence levels in the protozoan *Trichomonas vaginalis*." J Exp Med **174**(2): 311-318.
- Leitsch, D., D. Kolarich y col.; (2009). "*Trichomonas vaginalis*: metronidazole and other nitroimidazole drugs are reduced by the flavin enzyme thioredoxin reductase and disrupt the cellular redox system. Implications for nitroimidazole toxicity and resistance." Mol Microbiol **72**(2): 518-536.
- León-Félix, J., J. Ortega-López y col.; (2004). "Two novel asparaginyl endopeptidase-like cysteine proteinases from the protist *Trichomonas vaginalis*: their evolutionary relationship within the clan CD cysteine proteinases." Gene **335**: 25-35.
- León-Sicairos, C. R., J. León-Felix y col.; (2004). "*tvcp12*: a novel *Trichomonas vaginalis* cathepsin L-like cysteine proteinase-encoding gene." Microbiology **150**(Pt 5): 1131-1138.
- Lichtenstein, B., R. A. Desmond y col.; (2008). "Partnership concurrency status and condom use among women diagnosed with *Trichomonas vaginalis*." Womens Health Issues **18**(5): 369-374.
- Lim, L., Z. N. Canellakis y col.; (1970). "Metabolism of naturally occurring homopolymers. II. Characterization of adenine-rich polynucleotides of mouse liver ribosomes." Biochim Biophys Acta **209**(1): 128-138.
- Liston, D. R., J. C. Carrero y col.; (1999). "Upstream regulatory sequences required for expression of the *Trichomonas vaginalis* alpha-succinyl CoA synthetase gene." Mol Biochem Parasitol **104**(2): 323-329.
- Liston, D. R. y P. J. Johnson (1998). "Gene Transcription in *Trichomonas vaginalis*." Parasitol Today **14**(7): 261-265.
- Liston, D. R. y P. J. Johnson (1999). "Analysis of a ubiquitous promoter element in a primitive eukaryote: early evolution of the initiator element." Mol Cell Biol **19**(3): 2380-2388.
- Loke, J. C., E. A. Stahlberg y col.; (2005). "Compilation of mRNA polyadenylation signals in *Arabidopsis* revealed a new signal element and potential secondary structures." Plant Physiol **138**(3): 1457-1468.
- Lopez-Camarillo, C., O. N. Hernandez de la Cruz y col.; (2010). "Recent insights in premRNA 3'-end processing signals and proteins in the protozoan parasite *Entamoeba histolytica*." Infect Disord Drug Targets **10**(4): 258-265.
- López-Camarillo, C., O. N. Hernández de la Cruz y col.; (2010). "Recent insights in premRNA 3'-end processing signals and proteins in the protozoan parasite *Entamoeba histolytica*." Infect Disord Drug Targets **10**(4): 258-265.
- López-Camarillo, C., E. Orozco y col.; (2005). "*Entamoeba histolytica*: comparative genomics of the pre-mRNA 3' end processing machinery." Exp Parasitol **110**(3): 184-190.
- López, L. B., M. B. Braga y col.; (2000). "Strategies by which some pathogenic trichomonads integrate diverse signals in the decision-making process." An Acad Bras Cienc **72**(2): 173-186.
- Lubick, K. J. y D. E. Burgess (2004). "Purification and analysis of a phospholipase A2-like lytic factor of *Trichomonas vaginalis*." Infect Immun **72**(3): 1284-1290.
- Lutz, C. S. y A. Moreira (2011). "Alternative mRNA polyadenylation in eukaryotes: an effective regulator of gene expression." Wiley Interdiscip Rev RNA **2**(1): 22-31.

- Lutz, C. S. y A. Moreira (2011). "Alternative mRNA polyadenylation in eukaryotes: an effective regulator of gene expression." Wiley Interdiscip Rev RNA **2**(1): 23-31.
- Mabey, D., J. Ackers y col.; (2006). "Trichomonas vaginalis infection." Sex Transm Infect **82 Suppl 4**: iv26-27.
- Mandel, C. R., Y. Bai y col.; (2008). "Protein factors in pre-mRNA 3'-end processing." Cell Mol Life Sci **65**(7-8): 1099-1122.
- Maret, W. y Y. Li (2009). "Coordination dynamics of zinc in proteins." Chem Rev **109**(10): 4682-4707.
- Martincova, E., L. Voleman y col.; (2012). "Live imaging of mitosomes and hydrogenosomes by HaloTag technology." PLoS One **7**(4): e36314.
- McGrory, T. y G. E. Garber (1992). "Mouse intravaginal infection with *Trichomonas vaginalis* and role of *Lactobacillus acidophilus* in sustaining infection." Infect Immun **60**(6): 2375-2379.
- Mendoza-López, M. R., C. Becerril-García y col.; (2000). "CP30, a cysteine proteinase involved in *Trichomonas vaginalis* cytoadherence." Infect Immun **68**(9): 4907-4912.
- Millevoi, S. y S. Vagner (2010). "Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation." Nucleic Acids Res **38**(9): 2757-2774.
- Moodley, P., D. Wilkinson y col.; (2002). "*Trichomonas vaginalis* is associated with pelvic inflammatory disease in women infected with human immunodeficiency virus." Clin Infect Dis **34**(4): 519-522.
- Moreno-Brito, V., C. Yáñez-Gómez y col.; (2005). "A *Trichomonas vaginalis* 120 kDa protein with identity to hydrogenosome pyruvate:ferredoxin oxidoreductase is a surface adhesin induced by iron." Cell Microbiol **7**(2): 245-258.
- Mundodi, V., A. S. Kucknoor y col.; (2008). "Immunogenic and plasminogen-binding surface-associated alpha-enolase of *Trichomonas vaginalis*." Infect Immun **76**(2): 523-531.
- Murthy, K. G. y J. L. Manley (1995). "The 160-kD subunit of human cleavagepolyadenylation specificity factor coordinates pre-mRNA 3'-end formation." Genes Dev **9**(21): 2672-2683.
- Paschinger, K., A. Hykollari y col.; (2012). "The N-glycans of *Trichomonas vaginalis* contain variable core and antennal modifications." Glycobiology **22**(2): 300-313.
- Pattullo, L., S. Griffith y col.; (2009). "Stepwise diagnosis of *Trichomonas vaginalis* infection in adolescent women." J Clin Microbiol **47**(1): 59-63.
- Pereira-Neves, A. y M. Benchimol (2007). "Phagocytosis by *Trichomonas vaginalis*: new insights." Biol Cell **99**(2): 87-101.
- Petrin, D., K. Delgaty y col.; (1998). "Clinical and microbiological aspects of *Trichomonas vaginalis*." Clin Microbiol Rev **11**(2): 300-317.
- Proudfoot, N. J. (2011). "Ending the message: poly(A) signals then and now." Genes Dev **25**(17): 1770-1782.
- Ramón-Luing, L. A., F. J. Rendón-Gandarilla y col.; (2010). "Immunoproteomics of the active degradome to identify biomarkers for *Trichomonas vaginalis*." Proteomics **10**(3): 435-444.
- Rao, S., R. D. Dinkins y col.; (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 **10**: 51.

- Rasmussen, S. E., M. H. Nielsen y col.; (1986). "Morphological studies of the cytotoxicity of *Trichomonas vaginalis* to normal human vaginal epithelial cells in vitro." Genitourin Med **62**(4): 240-246.
- Rueggsegger, U., D. Blank y col.; (1998). "Human pre-mRNA cleavage factor Im is related to spliceosomal SR proteins and can be reconstituted in vitro from recombinant subunits." Mol Cell **1**(2): 243-253.
- Sandberg, R., J. R. Neilson y col.; (2008). "Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites." Science **320**(5883): 1643-1647.
- Santi-Rocca, J., C. Weber y col.; (2008). "The lysine- and glutamic acid-rich protein KERP1 plays a role in *Entamoeba histolytica* liver abscess pathogenesis." Cell Microbiol **10**(1): 202-217.
- Schwebke, J. R. y D. Burgess (2004). "Trichomoniasis." Clin Microbiol Rev **17**(4): 794803, table of contents.
- Schwebke, J. R., R. A. Desmond y col.; (2004). "Predictors of bacterial vaginosis in adolescent women who douche." Sex Transm Dis **31**(7): 433-436.
- Shafir, S. C., F. J. Sorvillo y col.; (2009). "Current issues and considerations regarding trichomoniasis and human immunodeficiency virus in African-Americans." Clin Microbiol Rev **22**(1): 37-45, Table of Contents.
- Shi, Y., D. C. Di Giammartino y col.; (2009). "Molecular architecture of the human premRNA 3' processing complex." Mol Cell **33**(3): 365-376.
- Shuman, S. (2001). "Structure, mechanism, and evolution of the mRNA capping apparatus." Prog Nucleic Acid Res Mol Biol **66**: 1-40.
- Smith, L. M., M. Wang y col.; (2002). "*Trichomonas vaginalis* infection in a premature newborn." J Perinatol **22**(6): 502-503.
- Solano-González, E., E. Burrola-Barraza y col.; (2007). "The trichomonad cysteine proteinase TVCP4 transcript contains an iron-responsive element." FEBS Lett **581**(16): 2919-2928.
- Song, H. O., M. H. Shin y col.; (2008). "*Trichomonas vaginalis*: reactive oxygen species mediates caspase-3 dependent apoptosis of human neutrophils." Exp Parasitol **118**(1): 59-65.
- Sutcliffe, S., J. F. Alderete y col.; (2009). "Trichomonosis and subsequent risk of prostate cancer in the Prostate Cancer Prevention Trial." Int J Cancer **124**(9): 2082-2087.
- Sutcliffe, S., E. Giovannucci y col.; (2006). "Plasma antibodies against *Trichomonas vaginalis* and subsequent risk of prostate cancer." Cancer Epidemiol Biomarkers Prev **15**(5): 939-945.
- Swygard, H., A. C. Sena y col.; (2004). "Trichomoniasis: clinical manifestations, diagnosis and management." Sex Transm Infect **80**(2): 91-95.
- Takagaki, Y. y J. L. Manley (2000). "Complex protein interactions within the human polyadenylation machinery identify a novel component." Mol Cell Biol **20**(5): 1515-1525.
- Tasca, T., C. D. Bonan y col.; (2005). "Heterogeneity in extracellular nucleotide hydrolysis among clinical isolates of *Trichomonas vaginalis*." Parasitology **131**(Pt 1): 71-78.

- Thurman, A. R. y G. F. Doncel (2011). "Innate immunity and inflammatory response to *Trichomonas vaginalis* and bacterial vaginosis: relationship to HIV acquisition." Am J Reprod Immunol **65**(2): 89-98.
- Torres-Machorro, A. L., R. Hernandez y col.; (2006). "The 5S ribosomal RNA gene from the early diverging protozoa *Trichomonas vaginalis*." Mol Biochem Parasitol **145**(2): 269-273.
- Vanacova, S., W. Yan y col.; (2005). "Spliceosomal introns in the deep-branching eukaryote *Trichomonas vaginalis*." Proc Natl Acad Sci U S A **102**(12): 4430-4435.
- Vázquez-Carrillo, L. I., L. I. Q. Granados y col.; (2011). The effect of Zn²⁺ on prostatic cell cytotoxicity caused by *Trichomonas vaginalis*.
- Weitzer, S. y J. Martinez (2007). "The human RNA kinase hClp1 is active on 3' transfer RNA exons and short interfering RNAs." Nature **447**(7141): 222-226.
- Winters, M. A. y M. Edmonds (1973). "A poly(A) polymerase from calf thymus. Characterization of the reaction product and the primer requirement." J Biol Chem **248**(13): 4763-4768.
- Winters, M. A. y M. Edmonds (1973). "A poly(A) polymerase from calf thymus. Purification and properties of the enzyme." J Biol Chem **248**(13): 4756-4762.
- Yano, A., K. Yui y col.; (1983). "Immune response to *Trichomonas vaginalis*. IV. Immunochemical and immunobiological analyses of *T. vaginalis* antigen." Int Arch Allergy Appl Immunol **72**(2): 150-157.
- Yarlett, N., M. P. Martínez y col.; (2000). "Dependence of *Trichomonas vaginalis* upon polyamine backconversion." Microbiology **146** (Pt 10): 2715-2722.
- Zamorano, A., C. Lopez-Camarillo y col.; (2008). "In silico analysis of EST and genomic sequences allowed the prediction of cis-regulatory elements for *Entamoeba histolytica* mRNA polyadenylation." Comput Biol Chem **32**(4): 256-263.
- Zamorano, A., C. López-Camarillo y col.; (2008). "In silico analysis of EST and genomic sequences allowed the prediction of cis-regulatory elements for *Entamoeba histolytica* mRNA polyadenylation." Comput Biol Chem **32**(4): 256-263.
- Zhao, J., M. Kessler y col.; (1999). "Pta1, a component of yeast CF II, is required for both cleavage and poly(A) addition of mRNA precursor." Mol Cell Biol **19**(11): 7733-7740.

XII. APÉNDICE.

Acrilamida al 30% bisacrilamida al 0.8%

Acrilamida 29.2g

Bisacrilamida 0.8g

Disolver en 40ml de agua destilada. Una vez disuelto aforar hasta 100ml.

Filtrar a travez de papel filtro "Whatman" No.1

Almacenar a 4°C hasta un mes, previamente protegido de la luz.

Tris-HCl/SDS 4x, pH8.8 (1.5M Tris-Cl,0.4%SDS)

Trizma-base (Sigma) 9.08g

SDS (Sigma) 0.2g

Disolver en 20 ml de agua destilada y ajustar a pH 8.8 con HCl

Aforar a un volumen final de 50 ml con agua destilada. Filtrar la solución almacenar a 4°C.

Tris-HCl/SDS 4x, pH6.8 (1.5M Tris-Cl,0.4%SDS)

Trizma-base (Sigma) 3.025g

SDS (Sigma) 0.2g

Disolver en 20 ml de agua destilada y ajustar a pH 6.8 con HCl

Aforar a un volumen final de 50 ml con agua destilada. Filtrar la solución almacenar a 4°C.

Persulfato de amonio al 10% (Bio-Rad)

10 mg (0.01 g) en un ml de agua destilada.

TEMED (Bio-Rad)

0.01 ml (10µl)

SDS-amortiguador 2X

Trizma – base 1.52g

SDS 2g

Disolver en 50 ml de agua destilada y ajustar el pH a 6.8. Posteriormente adicionar:

Glicerol 20 ml

Agua destilada (aforar a 100 ml)

Amortiguador de muestra 1X (BM 1X)

SDS amortiguador 2X 500µl

β-mercaptoetanol 5% 50µl

Azul de bromofenol 50µl

(stock 10 mg/ml)

Agua destilada 400µl

Amortiguador de electroforesis 5X (stock)

Glicina (Sigma) 72g

Trizma-base (Sigma) 15.1g

SDS (Sigma) 5.0g Aforar
con agua destilada a 1litro.

Amortiguador de electroforesis 1X

Amortiguador de electroforesis 5X 100 ml

Agua destilada 400ml

Solución Colorante de proteínas (Azul brillante de Coomassie) 0.05%

Metanol 500 ml

Ácido acético glacial 100 ml

Agua destilada 400 ml

Azul de Coomassie 0.5g

Disolver toda la noche. Filtrar con papel "Whatman" No.1. Almacenar a temperatura ambiente.

Solución decolorante de geles de poliacrilamida

Metanol	165 ml
Ácido acético glacial	50 ml Agua
destilada	785 ml

Almacenar a temperatura ambiente.

Preparación de Gelatina al 9%

0.0324 g en 10 ml de agua destilada incubar a 37°C para solubilizar.

Gel separador

Poliacrilamida /Bispoliacrilamida al 30%	4.50 ml
4XTris-Cl/SDS pH 8.8	3.75 ml
Gelatina al 9% o Agua destilada	6.75 ml
Persulfato de amonio al 10%	0.05 ml
TEMED	0.01 ml

Gel concentrador

Poliacrilamida /Bispoliacrilamida al 30%	0.65 ml
4XTris-Cl/SDS pH 6.8	1.25 ml
Agua destilada	3.05 ml
Persulfato de amonio al 10%	0.025 ml
TEMED	0.005 ml

Buffer de activación de proteasas pH 4.5 3M

Acetato de sodio	13.61 g
DTT	0.154g
Ácido acético	20 ml
Agua destilada	980 ml

Se disuelve en 500 ml de acetato de sodio, DTT posteriormente se adiciona el ácido acético y se ajusta el pH y se afora a 1litro.

Tritón 2.5%

Tritón X-100	25 ml
Agua destilada	1000 ml

TDSET

Tris-HCl pH 7.8 1M	1 ml
DOC 10%	2 ml
SDS	0.5 ml
EDTA pH 8.0	2.0 ml
Tritón X-100	1.0 ml

Se mezclan los componentes en 50 ml y se afora a 100 ml.

PBS-EDTA 0.2%

PBS 1X pH 7.0	500 ml
EDTA	1g

Se disuelve el EDTA y se esteriliza por autoclave 121°C/20min

Solución Amortiguadora de Rehidratación

Urea 7M, CHAPS (3-[(3-colamidopropil)dimetilamonio]-1-propanosulfonato) 4% w/v, DTT 70mM, azul de bromofenol al 0.001%, anfolinas (Bio-Lyte pH 3-10, BioRad) 0.1%.

Rehidratar el CHAPS, la urea y el azul de bromofenol con la muestra de proteína, agregar el DTT y H₂O_d cbp el volumen final deseado de la mezcla. Para tiras gradiente inmobilizado de pH (IPG) de 7cm, mezclar 120µL de la solución amortiguadora de rehidratación con las anfolinas. Las tiras IPG se rehidratan de 10 a 16h a temperatura ambiente.

Solución Amortiguadora de Equilibrio

Urea 6M, Glicerol 30% v/v, Tris pH8.8 50mM, SDS 2% w/v, Azul de bromofenol 0.002%. Las tiras IPG se equilibran 15min en solución amortiguadora de equilibrio con DTT (15 mg/mL) y posteriormente con IAA (37.5mg/mL).

Solución de agarosa para sellar

Agarosa 0.5%w/v, Buffer de corrida 5X 10mL, Azul de bromofenol 0.002%w/v, H₂O_d 40mL. Fundir la agarosa en el H₂O_d y almacenar a 4°C.

Tinción con plata

	Solución	Tiempo
1. Fijación	EtOH 40%, CH ₃ COOH 5%	4h o toda la noche
2. Lavado	EtOH 5%	2x20min
3. Pre-tratamiento	H ₂ O	3x5min
	Na ₂ S ₂ O ₃ ·5H ₂ O 0.2g/L	2min
4. Lavado	H ₂ O	3x20s
5. Impregnación	AgNO ₃ 2g/L, formaldehído 37% 0.4 mL/L	25min
6. Lavado	H ₂ O	2x20s
7. Revelado	NaCO ₃ 30g/L, Na ₂ S ₂ O ₃ ·5H ₂ O 4mg/L, formaldehído 37% 0.4mL/L	10min o según la aparición de los puntos
8. Lavado	H ₂ O	2x2min
9. Paro	CH ₃ COOH 10%	10min

Las soluciones de pre-tratamiento, impregnación y revelado deben prepararse al instante. No es recomendable rehusar más de tres veces las soluciones de fijación y paro. Los geles pueden almacenarse en la solución de paro a 4°C hasta por una semana. Antes de secar los geles entre dos hojas de papel celofán, se recomienda hacer un lavado con solución de glicerol al 10% durante 10min.

ANEXO 2

Secuencias de proteínas de *Homo sapiens*

CPSF160

>gi|23503048|sp|Q10570.2|CPSF1_HUMAN RecName: Full=Cleavage and polyadenylation specificity factor subunit 1; AltName: Full=Cleavage and polyadenylation specificity factor 160 kDa subunit; Short=CPSF 160 kDa subunit

MYAVYKQAHPPPTGLEFSMYCNFFNNSERNLVVAGTSQLYVYRLNRDAEALTKNDRSTEGKAHREKLELAA
SFSFFGNVMSMASVQLAGAKRDALLLSFKDAKLSVVEYDPTGTHDLKTLSTLHYFEEPELRDGFVQNVHTPR
VRVDPDGRCAAMLVYGTRLVVLPRRESLAEHEGLVGEGQRSSFLPSYIIDVRALDEKLLNIIDLQFLH
GYEPTLLILFEPNQTPGRVAVRQDTCISIVAISLNTQKVHPVIWSLTSLPFDCTQALAVPKPIGGVVV
FAVNSLLYLNQSVPPYGVALNSLTGTGTAFLRTQEGVRITLDCAQATFISYDKMVISLKGGEIYVLTLI
TDGMRSVRAFHFDDKAAASVLTSMVTMEPGYLFLGSRGLGNSLLKYTEKLQEPASAVREAADKEEPPSK
KKRVDATAGWSAAGKSVPQDEVDEIEVYGSEAQSGTQLATYSFEVCDASILNIGPCANAAVGEPAFLSEEF
QNSPEPDLEIVVCSGHGKNGALSVLQKSIRPVVTTTFELPGCYDMWTVIAPVRKEEEDNPKGEGTEQEPS
TTPEADDDGRRHGFLILSREDSTMILQTGQEIEMELDTSGFATQGPTVFAGNIGDNRYIVQVSPLGIRLLE
GVNQHLHFIPVDLGAPIVQCAVADPYVIMSAEGHVTMFLKSDSYGGRHRLALHKPPLHHQSKVITLCL
YRDLSGMFTTESRLGGARDELGGRSGPEAEGLGSETSPVTDDEEEMLYGDSGSLFSPSKEEARSSQPPA
DRDPAPFRAEPHTWCLLVRENGTMEIYQLPDWRLVFLVKNFPVGQVRVLDSSFGQPTTQGEARREEATRO
GELPLVKEVLLVALGSRQSRPYLLVHVDQELLIYEAFPHDSQLGQGNLKVRFKKVPHNINFREKKPKPSK
KKAEGGGAEEGAGARGRVARFRYFEDIYGYSGVFICGSPHWLLVTGRGALRLHPMAIDGPVDSFAPFHN
VNCPRGFLYFNRQELRISVLPAYLSYDAPWPVRKIPLRCTAHYVAYHVESKVYAVATSTNTPCARIPRM
TGEEKEFETIERDERYIHPQQEAFSIQLISPVSWEAIPNARIELQEWEHVTCMKTVSLRSEETVSGLKGY
VAAGTCLMQGEEVTCRGRILIMDVIEVVPEPGQPLTKNKFVLYEKEQKGPVTALCHCNGHLVSAIGQKI
FLWSLRASELTGMAFIDTQLYIHQMISVKNFILAADVMKSISSLRYQEESKTSLVSRDAKPLEVYSVDF
MVDNAQLGFLVSDRDRNLMVYMLPEAKESFGGMRLRRADFHVGAHVNTFWRTPCRGATEGLSKKSVVW
ENKHITWFATLDGGIGLLLPMQEKTYRLLMLQNALTTMLPHHAGLNPRAFRMLHVDRTLQNAVRNVLD
GELLNRYLYLSTMERSELAKKIGTTPDIILDDLLETDRVTAHF

CPSF100

>gi|51338827|sp|Q9P2I0.2|CPSF2_HUMAN RecName: Full=Cleavage and polyadenylation specificity factor subunit 2; AltName: Full=Cleavage and polyadenylation specificity factor 100 kDa subunit; Short=CPSF 100 kDa subunit

MTSIIKLTTLSGVQEEESALCYLLQVDEFRLLDCGWDEHFSMDIIDSLRKHVHQIDAVLLSHPDPLHLGA
LPYAVGKLGLENCAIYATIPVYKMGQMFYDLYQSRHNTEDFTLFTLDDVDAAFDKIQQLKFSQIVNLKKG
GHGLSITPLPAGHMIGGTIWKIVKDGEIEIVYAVDFNHKREIHLNGCSLEMLSRPSLLITDSFNATYVQP
RRKQRDEQLLTNVLETLRGDNVLIADVDTAGRVLELAQLLDQIWRTKDAGLGVYSLALLNNVSYNVVEFS
KSQVEWMSDKLMRCFEDKRNPFQFRHLSLCHGLSDLARVSPKVVLASQPDLECGFSRDLFIQWCQDPK
NSIILTYRTTPGTARFLIDNPSEKITEIELRKRVKLEGKELEEYLEKEKLKKEAAKKLEQSKEADIDSS
DESDIEEDIDQPSAHKTKHDLMMKGEGSRKGSFFKQAKKSYPMFPAPPEERIKWDEYGEIIPKPEDFLVPEL
QATEEEKSKLESGLTNGDEPMDQDLSDVPTKCISTTESIEIKARVTYIDYEGRS DGDSIKKIINQMKPRQ
LIIVHGPPEASQDLAECCRAFGGKDIAKVMYKPKLHETVDATSETHIYQVRLKDSLVSLLQFCKAKDAELAW
IDGVLDMRVSKVDTGVILEEGELKDDGEDSEMQVEAPSDSSVIAQQKAMKSLFGDDEKETGEESEIIPTL
EPLPPHEVPGHQSVFMNEPRLSDFKQVLLREGIQAEEFVGGVLVCNNQVAVRRTETGRIGLEGCLCQDFYR
IRDLLYEQYAIIV

CPSF73

>gi|18203503|sp|Q9UKF6.1|CPSF3_HUMAN RecName: Full=Cleavage and polyadenylation specificity factor subunit 3; AltName: Full=Cleavage and polyadenylation specificity factor 73 kDa subunit; Short=CPSF 73 kDa subunit; AltName: Full=mRNA 3'-end-processing endonuclease CPSF-73
MSAIPAEESDQLLIRPLGAGQEVGRSCIILEFKGRKIMLDCGIHPGLEGMDALPYIDLIDPAEIDLLLLIS
HFHLDHCGALPWFLQKTSFKGRTFMTHATKAIYRWLLSDYVKVSNISADDMLYTETDLEESMDKIETINF
HEVKEVAGIKFWCYHAGHVLGAAMFMIEIAGVKLLYTGDFSRQEDRHLMAAEIPNIKPDILIIESTYGT
IHEKREEREARFCNTVHDIIVNRGGRGLIPVFALGRAQELLLILDEYWQNHPELHDIPIYYASSLAKKCMA
VYQTYVNAMNDKIRKQININNPVFVKHISNLKSMDFDDIGPSVVMASPGMMQSGLSRELFESWCTDKRN
GVIIAGYCVEGTLAKHIMSEPEEITTMGQKLPLKMSVDYISFSAHTDYQQTSEFIRALKPPHVILVHGE
QNEMARLKAALIREYEDNDEVHIEVHNPRNTEAVTLNFRGEKLAQVMGFLADKKPEQGQVRVSGILVKRNF
NYHILSPCDLSNYTDLAMSTVKQTQAIPYTGPFNLLCYQLQKLTGDVEELEIQEKPAKLVFKNITVIQEP
GMVVLEWLANPSNDMYADTVTTVILEVQSNPKIRKGAVQKVSCKLEMHVYSKRLEIMLQDIFGEDCVSVK
DDSILSVTVDGKTANLNLETRTVECEESEDESREMVELAAQRLYEALTPVH

CPSF30

>gi|37999487|sp|O95639.1|CPSF4_HUMAN RecName: Full=Cleavage and polyadenylation specificity factor subunit 4; AltName: Full=Cleavage and polyadenylation specificity factor 30 kDa subunit; Short=CPSF 30 kDa subunit; AltName: Full=NS1 effector domain-binding protein 1; Short=Neb-1; AltName: Full=No arches homolog
MQEIIASVDHIKFDLEIAVEQQLGAQPLPFPMDKSGAAVCEFFLKAACGKGGMCPFRHISGEKTVVCKH
WLRGLCKKGDQCEFLHEYDMTKMPECYFYSKFGECSNKECPFLHIDPESKIKDCPWYDRGFCKHGPLCRH
RHTRRVICVNYLVGFCPEGPSCFKFMHPRFELPMGTTEQPPLPQQTQPPAKQSNNPPLQRSSSLIQLTSQN
SSPNQQRTPQVIGVMQSQNSSAGNRGPRPLEQVTCYKCGEKGHYANRCKTGHLAFLSGQ

FIP1

>gi|74749365|sp|Q6UN15.1|FIP1_HUMAN RecName: Full=Pre-mRNA 3'-endprocessing factor FIP1; Short=hFip1; AltName: Full=FIP1-like 1 protein; AltName: Full=Factor interacting with PAP; AltName: Full=Rearranged in hypereosinophilia
MSAGEVERLVSELSSGGTGGDEEEEWLYGGPVDVHVHSDLAKDLDENEVERPEEENASANPPSGIEDETAE
NGVPKPKVTETEDDSDDSDSDDEDDVHVTIGDIKTGAPQYGSYGTAPVNLNIKTGGRVYGTGTGKVKGVD
LDAPGSINGVPLLEVDLDSFEDKPWRKPGADLSDFNYGFNEDTWKAYCEKQKRIRMGLEVIPVTSTTNK
ITAEDCTMEVTPGAIEIQDGRFNLFKVQQGRTGNSEKETALPSTKAEFTSPPSLFKTGLPPSRNSTSSQSQ
TSTASRKANSSVGKQWDYRGRAESPDRLRLPGAIDVIGQTITISRVEGRRRANENSNIQVLSERSATEVD
NNFSKPPPPFPFGAPPTLPPPPFLPPPTVSTAPPLIPPPGFPPPPGAPPPSLIPTIESGHSSGYDSRS
ARAFPYGNVAFPHLPGSAPSWPSLVDTSKQWDYYARREKDRDRERDRDRERDRDRERERTREERERERD
HSPTPSVFNSDEERYRYREYAERGYERHRASREKEERHRERRHREKEETRHKSRSNSRRRHESEEGDSH
RRHKHKKSKRSKEGKEAGSEPAPEQESTETPAE

PAP-b

>gi|18203318|sp|Q9NRJ5.1|PAPOB_HUMAN RecName: Full=Poly(A) polymerase beta; Short=PAP-beta; AltName: Full=Polynucleotide adenyltransferase beta; AltName: Full=Testis-specific poly(A) polymerase
MPFPVTTQGPPQPAPPPNRYGVSSPISLAVPKETDCLLTQRLIETLRPFGVFEEEEELQRRILVLEKLNN
LVKEWIREISESKSLPQSVIENVGGKIFTFGSYRLGVHTKGADIDALCVAPSHVDRSDFFTSFYAKLKLQ
EEVKDLRAVEEAFVPVIKLCFDGIEIDILFARLALQTIPEDLDRDDSLKKNLDIRCIRSLNGCRVTDEI
LHLVFNIDNFRLLTALRAIKLWAKCHNIYSNIGLGGVSWAMLVARTCQLYPNAVASTLVRKFFLVFSEWE
WPNPVLKKEPEERNLNLFPVWDPRVNPSPDRYHLMPIITPAYPQQNSTYNVSISTRMVMIEEFKQGLAITHE
ILLSKAWSKLFEPSPFFQKYKHYIVLLASASTEKQHLEWVGLVESKIRILVGSLEKNEFITLAHVNPQS
FPAPKENPDMEEFRTMWVIGLGLKKPDNSEILSIDLTYDIQSFTDTVYRQAVNSKMFEMGMKITAMHLRR
KELHQLLPHHVLQDKKAHSTEGRRLLTDLNDSSFDLSAGCENSMSVPSSTSTMKTGPLISSSQGRNSPALA
VMTASVANIQATEFSLQQVNTNESSGVALNESIPHAVSQPAISPSPKAMVARVVSSTCLISHPDLQETQQ
QTYLIL

Simplexina

>gi|71153180|sp|Q92797.2|SYMPK_HUMAN RecName: Full=Symplekin
MASGSGDSVTRRSVASQFFTQEEGPGIDGMTTSERVVDLLNQAALITNDSKITVLKQVQELIINKDPTLL
DNFLDEIIAFQADKSIEVRKFVIGFIEEACKRDIELLLKLIANLNMLLRDENNVNVKKAILTMTQLYKVA
LQWMVKSRISELQEACWDMVSAMAGDIILLDSDNDGIRTHAIKFVEGLIVTLSPRMADSEIPRQEH
ISLDRIPRDHPYIQYNVLWEEGKAALQLLKFMVHPAISSINLTALGLSANIARQRPFMFSEVIQAYET
LHANLPPTLAKSQVSSVRKNLKLHLLSVLKH PASLEFQAQITTLVLDLGTQAEIARNMPSSKDTRKRPR
DDSDSTLKKMKLEPNLGEDDEDKDLPEGPSGTSKASQISGQSDTDITAEFLQPLLPDNLVANLVLISMV
YLPEAMPASFQAIYTPVESAGTEAQIKHLARLMATQMTAAGLGPGVEQTKQCKEKPKEKVVKTESVLIK
RRLSAQGGQAI SVVGLS SSMSPLEEEAPQAKRRPEPIIPVTQPRLAGAGGRKKIFRLSDVLKPLTDAQVEA
MKLGAVKRILRAEKAVACSGAAQVRIKILASLVTQFN SGLKAEVLSFILEDVRLRLDLAFALYQYENAY
LAAGASGLDKYEDCLIRLLSGLQEKPDQKDGIFTKVVLEAPLITESALEVVRKYCEDESRTYLG MSTLR
DLIFKRPSRQFQYLHVLLDLSSHEKDKVRSQALLFIK RMYEKEQLREYVEKFALNYLQLLVHPNPPSVLF
GADKDEVAAPWTEETVKQCLYLYLALLPQN HKLIHELAAVYTEAIADIKRTVLRVIEQPIRGMGMNSPE
LLLLVENC PKGAETLVTRCLHSLTDKVPSPPELVKVRVDLYHKRLPDVRFLIPVLNGLEKKEVIQALPKL
IKLNPIVVKVFNRLLGTQHGE GNSALSPLNPGELLIALHNIDSVKCDMKSI IKATNLCFAERNVYTSEV
LAVVMQQLMEQSP LPMMLMRTVIQSLTMYPRLG GFVMNILSRLIMKQVWKYPKVWEGFIKCCQRTK PQSF
QVILQLPQQQLGAVFDKCP ELREPLLAHVRSFTPHQQAHPNSIMTILEASGKQEP EAKEAPAGPLEEDD
LEPLTLAPAPAPRPPQDLIGLRLAQEKALKRQLEEEQKLKPGGVGAPSSSSPSPSPSARPGPPPSEEAMD
FREEGPECETPGFISMDDDSGLTEAALLDSSLEGPLPKETAAGGLTLKEERSPQTLAPVGEDAMKTPSP
AAEDAREPEAKGNS

CstF77

>gi|71153231|sp|Q12996.1|CSTF3_HUMAN RecName: Full=Cleavage stimulation factor subunit 3; AltName: Full=CF-1 77 kDa subunit; AltName: Full=Cleavage stimulation factor 77 kDa subunit; Short=CSTF 77 kDa subunit; Short=CstF-77
MSGDGATEQAAEYVPEKVKKA EKKLEENPYDLDAWSILIREAQNPIDKARKTYERLVAQFPSSGRFWKL
YIEAEIKAKNYDKVEKLFQRCLMKVLHIDLWKCYLSYVRETGKGLPSYKEKMAQAYDFALDKIGMEIMSY
QIWVDYINFLKGVEAVGSYAENQRITAVRRVYQRGCVNPMINIEQLWRDYNKYEEGINIHLAKKMIEDRS
RDYMNARRVAKEYETVMKGLDRNAPSVPQNT PQEAQQVDMWKYIQWEKSNPLRTEDQTLITKRVMFAY
EQCLLVLGHPDIWYEEAQYLEQSSKLLAEKGMNNAKLFSDEAANIYERAISTLLKKNMLLYFAYADYE

ESRMKYEKVHSIYNRLAIEDIDPTLVYIQYMKFARRAEGIKSGRMIFKKAREDTTRTRHHVYVTAALMEY
YCSKDKSVAFKIFELGLKKYGDIPYVLAYIDYLSHLNEDNNTRVLFERVLTSGSLPPEKSGEIWARFLA
FESNIGDLASILKVEKRRFTAFKEEYEGKETALLVDYKFMMDLYPCSASELKALGYKDVSRAKLAAIIPD
PVVAPSIVPVLKDEVDRKPEYPKPDTQQMIPFQPRHLAPPGLHPVPGGVFPVPPAAVVLMLKLLPPPICFQ
GPFVQVDELMEIFRRCKIPNTVEEAVRIITGGAPELAVEGNGPVESNAVLTKAVKRPNEDSDEDEEKGA
VPPVHDIYRARQQKRIR

CstF64

>gi|461847|sp|P33240.1|CSTF2_HUMAN RecName: Full=Cleavage stimulation
factor subunit 2; AltName: Full=CF-1 64 kDa subunit; AltName:
Full=Cleavage stimulation factor 64 kDa subunit; Short=CSTF 64 kDa
subunit; Short=CstF-64
MAGLTVRDPVAVDRSLRSVVFVGNIPYEATEEQDKDIFSEVGPVVSFRLVYDRETGKPKGYGFCEYQDQETA
LSAMRNLNREFSGRALRVDNAASEKNKEELKSLGTGAPVIESPYGETISPEDAPESISKAVASLPPEQM
FELMKQMKLCVQNSPQEARNMLLQNPQLAYALLQAQVVMRIVDPEIALKILHRQTNIPTLIAGNPQPVHG
AGPGSGSNVSMNQNPQAPQAQSLGGMHVNGAPPLMQASMGGVPAPGQMPAAVTGPGPGSLAPGGGMQA
QVGMPGSGPVSMERGQVPMQDPRAAMQRGSLPANVPTPRGLLGDAPNDPRGGTLLSVTGEVEPRGYLGPP
HQGPPMHVHPGHESRGPPPHELRGGPLPEPRPLMAEPRGPMLDQRGPPLDGRGGRDPRGIDARGMEARAM
EARGLDARGLEARAMEARAMEARAMEARAMEARAMEVRGMEARGMDTRGPVPGPRGPIPSGMQGPSPINM
GAVVPQGSRQVPVMQGTGMQGASIQGGSQPGGFSPGQNVTPQDHEKAALIMQVLQLTADQIAMLPPEQR
QSILILKEQIQKSTGAP

CstF50

>gi|461848|sp|Q05048.1|CSTF1_HUMAN RecName: Full=Cleavage stimulation
factor subunit 1; AltName: Full=CF-1 50 kDa subunit; AltName:
Full=Cleavage stimulation factor 50 kDa subunit; Short=CSTF 50 kDa
subunit; Short=CstF-50
MYRTKVGLKDRQQLYKLIISQLLYDGYISIANGLINEIKPQSVCAPSEQLLHLIKLGMENDDTAVQYAIG
RSDTVAPGTGIDLEFDADVQTMSPSEASEYETCYVTSHKGPCRVATYSRDGQLIATGSADASI KILDTERM
LAKSAMP IEVMMNETAQNMENHPVIRTLYDHVDEVTCLAFHPTEQILASGSRDYTLKLFDSKPSAKRA
FKYIQEAEMLRSISFHPSGDFILVGTQHPTLRLYDINTFQCFVSCNPQDQHTDAICSVNYNSSANMYVTG
SKDGCIKLWDGVSNRCTTFEKAHDGAEVCSAIFSKNSKYILSSGKDSVAKLWEISTGRTLVRVTGAGLS
GRQVHRTQAVFNHTEDYVLLPDERTISLCCWDSRTAERRNLLSLGHNNIVRCIVHSPTNPGFMTCSDDFR
ARFWYRRSTTD

CFIm25

>gi|74735411|sp|O43809.1|CPSF5_HUMAN RecName: Full=Cleavage and
polyadenylation specificity factor subunit 5; AltName: Full=Cleavage and
polyadenylation specificity factor 25 kDa subunit; Short=CFIm25;
Short=CPSF 25 kDa subunit; AltName: Full=Nucleoside diphosphate-linked
moiety X motif 21; Short=Nudix motif 21; AltName: Full=Pre-mRNA cleavage
factor Im 25 kDa subunit
MSVVPNRSQTGWPRGVTQFGNKYIQQTKPLTLERTINLYPLTNYTFGTKEPLYEKDSSVAARFQRMREE
FDKIGMRRTVEGLVIVHEHRLPHVLLQLGTTFFKLPGGELNPGEDEVEGLKRLMTEILGRQDGVLDQDWV
IDDCIGNWWRPNFEPPQYPYIPAHITKPKHEKKLFLVQLQEKALFAVPKNYKLVAAPLFELYDNAPGYGP
IISSLPQLLSRFNFIYN

CFIm59

>gi|74759932|sp|Q8N684.1|CPSF7_HUMAN RecName: Full=Cleavage and polyadenylation specificity factor subunit 7; AltName: Full=Cleavage and polyadenylation specificity factor 59 kDa subunit; Short=CFIm59; Short=CPSF 59 kDa subunit; AltName: Full=Pre-mRNA cleavage factor Im 59 kDa subunit
MSEGVDLIDIYADEEFNQDPEFNNTDQIDLYDDVLTATSQPSDDRSSSTEPPPPVRQEPSPKPNKTPAI
LYTYSGLRNRRRAAVYVGSFSWTTDQQLIQVIRSIGVYDVVELKFAENRANGQSKGYAEVVVASENSVHK
LLELLPGKVLNGEKVDVRPATRQNLQFEAQARKRECVRVPRGGIPRAHSRDSADSADGRATPSENLP
SSARVDKPPSVLPYFNRPPSALPLMGLPPPIPPPPPLSSSFVPPPPPGIHYQHLMPPPPRLPPHLAVP
PPGAIPPALHLNPAFFPPPNATVGPPPDYMKASAPYNHHGSRDSGPPPPSTVSEAEFEDIMKRNRRAISS
AISKAVSGASAGDYSDAIETLLTAIAVIKQSRVANDERCRLISSLKDCLHGIEAKSYSVGASGSSSRKR
HRSRERSPSRSRESSRRHRDLLHNEDRHDDYFQERNREHERHRDRERDRHH

CFIm68

>gi|88909266|sp|Q16630.2|CPSF6_HUMAN RecName: Full=Cleavage and polyadenylation specificity factor subunit 6; AltName: Full=Cleavage and polyadenylation specificity factor 68 kDa subunit; Short=CFIm68; Short=CPSF 68 kDa subunit; AltName: Full=Pre-mRNA cleavage factor Im 68 kDa subunit; AltName: Full=Protein HPBRII-4/7
MADGVDHIDIYADVGEFQEAEEYGGHDQIDLYDDVISPSANNGDAPEDRDYMDTLPPTVGDDVGKGAAP
NVVYTYTGKRIALYIGNLTWTTDEDLIEAVHSLGVNDILEIKFFENRANGQSKGFALVGVGSEASSKKL
MDLLPKRELHGQNPVVTCPNKQFLSQFEMQSRKTTQSGQMSGEGKAGPPGGSSRAAFPQGGGRGRFRPGA
VPGGDRFPGPAGPGGPPPPFPAGQTPPRPPLGPPGPPGPPGPPPPGQVLPPLAGPPNRGDRPPPPVLF
GQPFQPPPLGLPPLPPGPPPPVPGYGGPPGPPPPQGGPPPPPGFPFRPPGPLGPPLTLAPPPHLPGPPGA
PPPAPHVNPAAFFPPPTNSGMPTSDSRGPPPTDPYGRPPPYDRGDYGGPGREMDTARTPLSEAEFEEIMNR
NRAISSAISRAVSDASAGDYGSAIETLVTAISLIKQSKVSADDRCKVLISLQDCLHGIESKSYGSGSR
RERSRERDHSRSREKSRRHKSRSRDRHDDYYRERSRERERHRDRDRDRDRERDREREYRHR

CLP1

>gi|13431366|sp|Q92989.1|CLP1_HUMAN RecName: Full=Polyribonucleotide 5'-hydroxyl-kinase Clp1; AltName: Full=Polynucleotide kinase Clp1; AltName: Full=Pre-mRNA cleavage complex II protein Clp1
MGEEANDDKKPTTKFELERETELRFEVEASQSVQLELLTGMAEIFGTELTRNKKFTFDAGAKVAVFTWHG
CSVQLSGRTEVAYVSKDTPMLLYLNTHTALEQMRRQAEKEEERGPRVMVVGPTDVGKSTVCRLLLNYAVR
LGRRPTYVELDVGQGSVSIPTMGALYIERPADVEEGFSIQAPLVYHFGSTTPGTNIKLYNKITSRLADV
FNQRCVENRRASVSGCVINTCGWVKSGYQALVHAASAFEVDVVVLDQERLYNELKRDLPHFVRTVLLP
KSGGVVERSDFRRECRDERIREYFYGFRGCFYPHAFNVKFSVDVKIYKVGAPTIPDSCLPLGMSQEDNQL
KLVPVTPGRDMVHLLSVSTAEGTEENLSETSVAGFIVVTSVDLEHQVFTVLSPAPRPLPKNFLLIMDIR
FMDLK

PCF11

>gi|160370000|sp|O94913.3|PCF11_HUMAN RecName: Full=Pre-mRNA cleavage complex 2 protein Pcf11; AltName: Full=Pre-mRNA cleavage complex II protein Pcf11
MSEQTPAEAGAAGAREDACRDYQSSLEDLTFNSKPHINMLTILAEENLPFAKEIVSLIEAQTAKAPSSSEK
LPVMYLMDSIVKNVGREYLTAFATKLVATFICVFEEKVDENTRKSFLKLRSTWDEIFPLKKLYALDVRVNS
LDPAWPIKPLPPNVNTSSIHVNPKFLNKSPEEPSTPGTVVSSPSISTPPIVPDIQKNLTQEQLIRQQLLA
KQKQLELELQKKLELELEQAKAQLAVSLSVQQETSNLGPGSAPSKLHVSQIPPMVAKAPHQVPVQSEKSR
PGPSLQIQDLKGTNRDPRNLNRISQHSHGKDQSHRKEFLMNTLNQSDTKTSKTIPSEKLNSSKQEKSKSGE
KITKKELDQLDSKSKSKSKSPSPLKNKLSHTKDLKNQESSEMRSDMNKRDPRLKKHLQDKTDGKDDDDVK
EKRKTAEEKDKDEHMKSSSEHRLAGSRNKIINGIVQKQDTITEESEKQGTKPGRSSTRKRSRSPKSRSP
IIHSPKRRDRRSPKRRQRSMSPTSTPKAGKIRQSGAKQSHMEEFTPPSREDRNAKRSTKQDIRDPRMKK
TEERPQETTNQHSKSGTEPKENVENWQSSSKSAKRWKSGWEENKSLQQVDEHSKPPHLRHRESWSSTKG
ILSPRAPKQQHRLSVDANLQIPKELTLASKRELLQKTSERLASGEITQDDFLVVVHQIRQLFQYQEGVR
EEQSRSPFNDRFPLKRPYEDSDKPFVDSASRFAGLDTNQRLTALAEDRPLFDGSPSRPSVARDGPTKMIF
EGNPSLSPRIDGPPTPASLRFDGSPGQMGGGGGLRFEGPQGGQLGGGCPLRFEGPPGPVGTPLRFEGPIGQ
AGGGGFRFEGSPGLRFEGSPGGLRFEGPGGQPVGGLRFEGHRGQPVGGLRFEGPHGQPVGGLRFDNPRGQ
PVGGLRFEGGHGPGSAAIRFDGPHGQPGGGIRFEGPQLQGGVGMRFEGPHGQSVAGLRFEGQHNQLGGNL
RFEGPHGQPGVGIRFEGPLVQGGGMRFEGSPVGGGLRIEGLGQGGPRFEGCHALRFDGQPGQPSLLP
RFDGLHGQPGPRFERTPGQPGQRFDGPPGQVQPRFDGVPQRFDGPHQQAASRFDIPLGLQGTFRDNHP
SQRLSVSVFNQTGYPNDPPGNAFNAPSQGLQFQRHEQIFDSPQGNFNGPHGPGNQSFNPLNRASGHYF
DEKNLQSSQFGNFGNIPAPMTVGNIIQASQVLSGVAQPVAFGQGGQFLPVHPQNPGFVQNPSGALPKAYP
DNHLSQVDVNEFLSKLLKTGILKLSQTDSATQVSEVTAQPPPEEEEDQNEQDQVDPDLTNFTVEELKQRY
DSVINRLYTGICYSKCMRFTTSQTDVYADHLWDHYRQNRTEKDVSRKVTHRRWYYSLTDWIEFEEIADL
EERAKSQFFEKVEEVVLKTQEAAKEKEFQSVAPAGAVESCEICQEQQFEQYWDEEEEWHLKNAIRVD
GKIYHPSCYEDYQNTSSFDCTPSPSKTPVENPLNIMLNIVKNELQEPDCSPKVKEERIDTPPACTEESIA
TPSEIKTENDTVESV

Rbbp6

>gi|74762440|sp|Q7Z6E9.1|RBBP6_HUMAN RecName: Full=E3 ubiquitin-protein
ligase RBBP6; AltName: Full=Proliferation potential-related protein;
AltName: Full=Protein P2P-R; AltName: Full=Retinoblastoma-binding Q
protein 1; Short=RBQ-1; AltName: Full=Retinoblastoma-binding protein 6;
AltName: Full=p53-associated cellular protein of testis
MSCVHYKFSSKLNVDYTFDGLHISLCDLKKQIMGREKLKAADCQLITNAQTKEEYTDNALIPKNSSV
IVRRIPIGGVKSTSKTYVISRTEPAMATTKAIDSSASISLAQLTKTANLAEANASEEDKIKAMMSQSGH
EYDPIINYMKKPLGPPPPSYTCFRCGKPGHYIKNCPTNGDKNFESGPRIKKSTGIPRSFMMEVKDPNMKGA
MLTNTGKYAIPTIDAEAYAIGKKEKPPFLPEEPSSSSSEDDPIPELCLICKDIMTDAVVI PCCGNSYC
DECIRTALLESDEHTCPTCHQNDVSPDALIANKFLRQAVNNFKNETGYTKRLRKQLPPPPPIPPRPLI
QRNLQPLMRSPISRQDPLMIPVTSSSTHPAPSISLTSNQSSLAPPVSGNPSSAPAPVPDITATVSISV
HSEKSDGPFRRSDNKILPAAALASEHSKGTSSIAITALMEEKGYQVPVLGTPSLLGQSLHGLIPTTGP
VRINTARPGGGRPGWEHSNKLGYLVSPQQIRRGERSCYRSINRGRHHSERSQRTQGPSLPATPVFVPVP
PPPLYPPPPHTLPLPPGVPPPQFSPQFPFPPAGYSVPPPGFPPAPANLSTPWVSSGVQTAHSNTIPT
TQAPPLSREEFYRQRLKEEEKKSKLDEFTNDFAKELMEYKIKQKERRRSFSRSKSPYSGSSYSRSSY
TYSKSRSGSTRSRYSRFSRSHRSYSRSPYPYPRGRGKSRNYSRSHGYHRSRSPYPYRRYHSRS
RSPQAFRGQSPNKRNVQGETEREYFNRYREVPPYDMKAYYGRSVDFRDPFEKERYREWERYEWYK
YYKGYAAGAQPRPSANRENFSERFLPLNIRNSPFTGRREDYVGGQSHRSRNIGSNYPEKLSARDGHNQ
KDNTKSKEKESENAPGDGKGNKHKHRRKRRKGESEGFNLPELLETSRKSREPTGVEENKTDLSFLVPSR
DDATPVRDEPMDAESITFKSVSEKDKRERDKPAKAGDKTKRNDGSVSKKENIVKPAKGPQEKVDGERE
RSPRSEPPIKKAKEETPKTDNTKSSSSSQKDEKITGTPRKAHKSASKEHQETKPVKEEKVKKDYSKDVKS

EKLTTKEEKAKKPNEKNKPLDNKGEKRKRKTEEGVVDKDFESSMKISKLEVTEIVKPSPKRKMEPDTEK
MDRTPEKDKISLSAPAKKIKLNRETGKKIGSTENISNTKEPSEKLESTSSKVKQEKVKGVRRKVTGTEG
SSSTLVDTSTSSSTGGSPVRKSEKTDTKRTVIKTMEEYNNDNTAPAEDVIMIQVPQSKWDKDDFESEE
EDVKSTQPISSVGKPAVIVKVNSTKPSNIVKYPEKESEPESEKIQKFTKDVSHI IQHEVKSSKNSASSEK
GKTKDRDYSVLEKENPEKRKNSTQPEKESNLDRLNEQGNFKSLSQSSKEARTSDKHDSTRASSNKDFTPN
RDKKTDYDTREYSSSKRRDEKNELTRRKDPSRNKDSASGQKNKPREERDLPKKGTGDSKKSNSSPSRDR
KPHDHKATYDTKRPNEETKSVDKNPCKDREKHVLEARNNKESGKNLLYILNPPETQVEKEQITGQIDKS
TVKPKPQLSHSSRLSSDLTRETDEAAFEFDYNESEDSESNVSVKEEESSGNISKDLKDKIVEKAKESLDTA
AVVQVGISRNSHSSPSVSPSRSHSPSGSQTRSHSSASSASSQDSKKKKKKKEKKKKKKKKKKKKKKH
AGTEVELEKSQKHKKKKKKSKKNKDKEKEKEKDDQKVKSVTV

PNAS-120

>gi|74752877|sp|Q9NP77.1|SSU72_HUMAN RecName: Full=RNA polymerase II
subunit A C-terminal domain phosphatase SSU72; Short=CTD phosphatase
SSU72
MPSSPLRVAVVCSSNQNRSM EAHNILSKRGFSVRSFGTGTHVKLPGPAPDKPNVYDFKTTYDQMYNDLLR
KDKELYTQNGILHMLDRNKRIKPRPERFQNCCKDLFDLILTCEERVYDQVVEDLNSREQETCQPVHVNVND
IQDNHEEATLGAFLICELCQCIQHTEDMENEIDELLQEFEEKSGRTFLHTVCFY

PC4

>gi|1709514|sp|P53999.3|TCP4_HUMAN RecName: Full=Activated RNA polymerase
II transcriptional coactivator p15; AltName: Full=Positive cofactor 4;
Short=PC4; AltName: Full=SUB1 homolog; AltName: Full=p14
MPKSKELVSSSSSGSDSDSEVDKCLKRKKQVAPEKPVKKQKTGETSRALSSSKQSSSRDDNMFIQIKMR
YVSVRDFKGVKVLIDIREYWMDPEGEMKPGRKGISLNPEQWSQLKEQISDIDDAVRKL

PABP

>sp|P11940|PABP1_HUMAN Polyadenylate-binding protein 1 OS=Homo sapiens
GN=PABPC1 PE=1 SV=2
MNPSPAPSYPMASLYVGDLPDVTEAMLYEKFSPAGPILSIRVCRDMITRRSLGYAYVNFQ
QPADAERALDTMNFVVIKGPVRIMWSQRDPSLRKSGVGNIFIKNLDKSIDNKALYDTFS
AFGNILSCKVVCDENGSKGYGFVHFETQEAAERAIEKMNGMLLNDRKVFVGRFKSRKERE
AELGARAKEFTNVYIKNFGEDMDDERLKD LFGKFGPALSVKVMTDESGKSKGFGFVSFER
HEDAQKAVDEMNGKELNGKQIYVGRAQKKVERQTELKRKFQMKQDRITRYQGVNLYVKN
LDDGIDDERLRKEFSPFGTITSKVMMEGGRSKGFGFVCFSSPEEATKAVTEMNGRIVAT
KPLYVALAQRKEERQAHLTNQYMQRMASVRAPNPVINPYQPAPPSGYFMAAIPQTQNR
AYYPPSQIAQLRPSRWTAQGARGPHFQNMPGAIRPAAPRPPFSTMRPASSQVPRVMSTQ
RVANTSTQTMGPRPAAAAAATPAVRTVPQYKYAAGVRNPQHLNAQPQVTMQQPAVHVQ
GQEPLTASMLASAPPQEQKQMLGERLFLIQAMHPTLAGKITGM LLEIDNSELLHMLESP
ESLRSKVDEAVAVLQAHQAKEAAQKAVNSATGVPTV

Secuencias de proteínas de *Arabidopsis thaliana*

AtCPSF160

>gi|290457637|sp|Q9FGR0.2|CPSF1_ARATH RecName: Full=Cleavage and polyadenylation specificity factor subunit 1; AltName: Full=Cleavage and polyadenylation specificity factor 160 kDa subunit; Short=AtCPSF160; Short=CPSF 160 kDa subunit
MSFAAYKMMHWPTGVENCASGYITHSLSDSTLQIPIVSVHDDIEAEWPNPKRGIGPLPNVVITAANILEV
YIVRAQEEGNTQELRNPKLAKRGGVMDGVYGVLSLELVCHYRLHGNVESIAVLPMGGGNSSKGRDSIILTF
RDAKISVLEFDDSIHSLRMTSMHCFEGPDWLHLKRGRESFPRGPLVKVDPQGRCGGVLVYGLQMIILKTS
QVGSGLVGDDDAFSSGGTVSARVESSYIINLRDLEMKHVKDFVFLHGYIEPVIVILQEEEHTWAGRVSWK
HHTCVLSALSINSTLKQHPVIWSAINLPHDAYKLLAVPSPIGGVLVLCANTIHYHSQSASCALALNNYAS
SADSSQELPASNFSVELDAAHGTWISNDVALLSTKSGELLLLTLIYDGRAVQRLDLSKSKASVLASDITS
VGNSLFFLGSRGLGDSLLVQFSCRSGPAASLPGLRDEDEDIEGEGHQAKRLRMTSDTFQDTIGNEELSLFG
STPNNSDSAQKSFSFAVRDSLNVNVPVKDFAYGLRINADANATGVSKQSNYELVCCSGHGKNGALCVLRQ
SIRPEMITEVELPGCKGIWTVYHKSSRGHNADSSKMAADEDEYHAYLIISLEARTMVLETADLLTEVTES
VDYYVQGRITIAAGNLFGRRRVIQVFEHGARILDGSFMNQELSFGASNSESNSGSESSTVSSVSIADPYVL
LRMTDDSIIRLLVGPSTCTVSISSPSVLEGSKRKISACTLYHDKGPEPWLKASTDAWLSSGVGEAVDSV
DGGPQDQGDIIYCVVCYESGALEIFDVPSFNCVFSVDKFASGRRHLSDMPIHELEYELNKNSEDNTSSKEI
KNTRVVELAMQRWSGHHTRPFLFAVLADGTILCYHAYLFDGVDSTKAENSLSENPAALNSSGSSKLRNL
KFLRIPLDTSTREGTSDGVASQRITMFKNISGHQGGFLLSGSRPGWCMLFRERLRFHSQLCDGSIAAFTVL
HNVCNHNHGFYIVTAQGVLIKICQLPSASIYDNYWPVQKIPLKATPHQVTYYAEKNLYPLIVSYPVSKPLNQ
VLSSLVQDEAGQQLDNHNMSSDDLQRTYTVEEFEIQILEPERSGGPWETKAKIPMQTSEHALTVRVVTTLL
NASTGENETLLAVGTAYVQGEDVAARGRVLLFSFGKNGDNSQNVVTEVYSRELKGAISAVASIQGHLLIS
SGPKIILHKWNGTELNGVAFFDAPPLYVVMNVVKSFILLGDVHKSIFYLSWKEQGSQSLSLAKDFESLD
CFATEFLIDGSTLSLAVSDEQKNIQVFYIAPKMIESWKGLKLLSRAEFHVGAVHVSFRLRLQMVSSGADKI
NRFALLFGTLDGSFGCIAPLDEVTFRRLQSLQKKLVDAVPHVAGLNPLAFRQFRSSGKARRSGPDSIVDC
ELLCHYEMLPLEEQLELAHQIGTTRYISILKDLVDLSVGTSFL

AtCPSF100

>gi|18203240|sp|Q9LKF9.2|CPSF2_ARATH RecName: Full=Cleavage and polyadenylation specificity factor subunit 2; AltName: Full=Cleavage and polyadenylation specificity factor 100 kDa subunit; Short=AtCPSF100; Short=CPSF 100 kDa subunit; AltName: Full=Protein EMBRYO DEFECTIVE 1265; AltName: Full=Protein ENHANCED SILENCING PHENOTYPE 5
MGTSVQVTPLCGVYNENPLSYLVSIDGFNFLIDCGWNDLFDTSLEPLSRVASTIDAVLLSHPDTHIGA
LPYAMKQLGLSAPVYATEPVHRLGLLTMYDQFLSRKQVSDFDLFTLDDIDSAFQNVIRLTYSQNYHLSGK
GEGIVIAPHVAGHMLGGSIWRIKDGEDVIYAVDYNHRKERHLNGTVLQSFVRPAVLITDAYHALYTNQT
ARQQRDKFEFLDTISKHLEVGGNVLLPVDTAGRVLELLLILEQHWSQRGFSFPIYFLTYVSSSTIDYVKSF
LEWMSDSISKSFETSRDNAFLLRHVTTLLINKTDLNAPPGPKVVLASMASLEAGFAREIFVEWANDPRNL
VLFTETGQFGTLARMLQSAPPPKFVKVTMSKRVPPLAGEELIAYEEEEQNRLKREEALRASLVKEEETKASH
GSDDNSSEPMIIDTKTTHDVIGSHGPAYKDILIDGFVPPSSSVAPMFPYYDNTSEWDDFGEIINPDDYVI
KDEDMDRGAMHNGGDVDGRLDEATASLMLDTRPSKVMSNELIVTVSCSLVKMDYEGRS DGRS IKSMAIAHV
SPLKLVLVHAI AEATEHLKQHCLNNICPHVYAPQIEETVDVTS DLCA YKVQLSEKLMSNVIFKKLGDSEV
AWVDSEVGKTERDMRSLLPMPGAASPHKPVVLVGDLKIADFKQFLSSKGVQVEFAGGGALRCGEYVTLRKV
GPTGQKGGASGPQQILIEGPLCEDYYKIRDYLYSQFYLL

AtCPSF73-I

>gi|75262219|sp|Q9C952.1|CPSF3_ARATH RecName: Full=Cleavage and polyadenylation specificity factor subunit 3-I; AltName: Full=Cleavage and polyadenylation specificity factor 73 kDa subunit I; Short=AtCPSF73-I; Short=CPSF 73 kDa subunit I
MASSSTSLKRREQPISRDGDQLIVTPLGAGSEVGRSCVYMSFRGKNILFDCGIHPAYSGMAALPYFDEID
PSSIDVLLITHFHIDHAASLPYFLEKTTFNQGRVFMTHATKAIYKLLLTIDYVKVSKVSVEDMLFDEQDINK
SMDKIEVIDFHTQVEVNGIKFWCYTAGHVLGAAMFMVDIAGVRILYTGDISREEDRHLRAAELPQFSPDI
CIIESTSGVQLHQSRHIREKRFTDVIHSTVAQGGRLIPAFALGRAQELLLILDEYWANHPDLHNIPIYY
ASPLAKKCMNAVYQTYILSMNDRIRNQFANSNPFVFKHISPLNSIDDFNDVGPSVVMATPGGLQSGLSRQL
FDSWCSDDKKNACIIPGYMVEGTLAKTIINEPKVETLMNGLTAPLNMQVHYISFSAHADYAQTSTFLKELM
PPNIILVHGEANEMMRKQKLLTEFPDGNTKIMTPKNCSVEMYFNSEKLAKTIGRLAEKTPDVGDVTSG
ILVKKGFYQIMAPDELHVFSQLSTATVTQRITIPFVGAFGVIKHRLEKIFESVEFSTDEESGLPALKVH
ERVTVKQESEKHISLQWSSDPISDMVSDSIVALILNISREVPKIVMEEEDAVKSEEENGKKVEKVIYALL
VSLFGDVKLGENGKLVIRVDGNVAQLDKESGEVESEHSGLKERVVAFERIQSAVKPIPLSAS

AtCPSF73-II

>gi|332278175|sp|Q8GUU3.2|CPS3B_ARATH RecName: Full=Cleavage and polyadenylation specificity factor subunit 3-II; AltName: Full=Cleavage and polyadenylation specificity factor 73 kDa subunit II; Short=AtCPSF73-II; Short=CPSF 73 kDa subunit II; AltName: Full=Protein EMBRYO SAC DEVELOPMENT ARREST 26
MAIDCLVLGAGQEIGKSCVVVTINGKKIMFDCGMHMGCDHNRYPNFSLISKSGDFDNAISCIIITHFHM
DHVGALPYFTEVCGYNGPIYMSYPTKALSPLMLEDYRRVMVDRRGEEELFTTTHIANCMKKVIAIDLKQT
IQVDEDLQIRAYYAGHVLGAVMVYAKMGDAIIVYTGDMNTTDRHLGAAKIDRLQLDLLISESTYATTIR
GSKYPREREFLQAVHKCVAGGGKALIPSFALGRAQELCMLLDDYWERMNIKVPIYFSSGLTIQANMYKYM
LISWTSQNVEKEHNTHNPFDFKNVKDFDRSLIHAPGPCVLFPATPGMLCAGFSLEVFKHWPSPNLVALP
GYSVAGTVGHKLMAGKPTTVDLYNGTKVDVRCKVHVQVAFSPHTDAKGIMDLTKFLSPKNVVLVHGEKPSM
MILKEKITSELDIPCFVPANGETVSFASTTYIKANASDMFLKSCSNPNFKFSNSTQLRVTDHRTADGVLV
IEKSKKAKIVHQDEISEVLHEKNHVSLAHCCPVKVKGESEDDDDVDLIKQLSAKILKTVSGAQIHESENC
LQVASFKGSLCLKDKCMHRSSSSSEAVFLCCNWSIADLELGWEIINAIKLNH

AtCPSF30

>gi|332193100|gb|AEE31221.1| cleavage and polyadenylation specificity factor CPSF30 [Arabidopsis thaliana]
MEDADGLSFDGEGGLDSGPVQNTASVPVAPPENSSSAAVNVAPTYDHSSATVAGAGRGRSFRQTVCRHWL
RGLCMKGDACGFLHQFDKARMPICRFFRLYGEQDCVYKHTNEDIKECNMYKLGFCPNGPDCRYRHAK
LPGPPPPVEEVLQKIQQLTITYNYGTNRLYQARNVAPQLQDRPQQGVPMQGGPQESGNLQQQQQQQPQQSQ
HQVSQTLIPNPADQTNRTSHPLPQGVNRYFVVKSNRENFEFSVQQGVWATQRSNEAKLNEAFDSVENVI
LIFSUNRTRHFQGCAMTSRIGGYIGGNWKHEHGTAYGRNFSVKWLKLCESFHKTRNLRNPYNENLP
VKISRDCQELEPSVGEQLASLLYLEPDSELMAISIAAEAKREEEKAKGVNPESRAENPDIVPFEDNEEEEE
EEDESEEEEEESMAGGPQGRGRGRGIMWPPQMPLGRGIRPMPGMGGFPLGVMGPGDAFPYGPGGYNGMPD
PFGMGPRPFPYGPFRGGDFRGPVPGMMFPGRPPQQFPHGGYGMMGGGRGPHMGGMGNAPRGGRPMPYPP
ATSSARPGPSNRKTPERSDERGVSGDQONQDASHDMEQFEVGNLSLRNEESESEDEDEAPRRSRHGEGKKR
R

AtFipS5

>gi|332009607|gb|AED96990.1| FIP1 [V]-like protein [Arabidopsis thaliana]
MEEDDEFGLYSDVLQPFQPPVVLPPPPPLPHRSIDLNLRSQDQDVSEPNSAPISRVSDNDVAVKLSTQDA
TRQAIVDGGGDDKDMSEFDIEPDADSTPTIPGLFVTGALPGLATDRGVSQVTTTRIEQQVGGGGDGGYGGQ
GEGDDWDSSEDDLQIVLNDSSRNVMIIGADRRSRMGDNEDDDDEDEDPLVIVADTDPNQPMEEQMWGE
DGLQGIEGDGKDGGEAGKSGPGGATGPPKAGYSSHGYHFFHSQFKYVRPGAAPIPGGAASVGGPSSGQV
RPPANLGPAGRGRGDWRPLGMRNASAAQKGFHPWGSNTAGRLDFTLPSHKTI FEVDIDSFEKPPWRY
PGVEMTDYFNFGLNEESWKDYCKQLDQHRIQTTMQSRIRVYESGRTDQGYDPLPELAAATGAQGV PVD
SSNLVKPDSVQGDSAKVPANVRPTLPPGRPIPVETGSGERLPSIDTRAPMRDLDAIIEIVCQDSHEDEP
SGENGTDQADSSLPGENVPVETSYVNNKRPDTESAEHSQAQDEPHKNLLKKQDDEISRSTDSGQSRSSS
PVGDRGTRSSSVDRDVGGEAGKDAEMGEELKMSFTSPQSAVQEDDGGESKTERSSSESSKARSGSHRDFQ
QEEDVIQDKHSSRPANNRKQYDNNAPHQSRKNQDRGKEMERTRAASKGGRENSNPHMELDSTYIYSIASR
EDFDKRKERDVGAVWRRKEDDPYSRRGGDEGSRKRDREDDPGFRQRGKMRENEIRSKDDQVPSRKHMDD
AGMRNIYEPDDHINKRRKDEEYLRRSRPEKNEISYQRESMSRVKREDDRLHQQKRDVQHKIRDDFDDH
GSLRQRDDIYMRDGNERLRERDVLDKLKLPHEDGISARGRERQVAVRGHRGSEDRSSRMKDEYKASDKE
HVTKDTLRHAKQTKRRDYPGEESSSHHRGHEDFSARTDNIVNNEKKPRQERTGAKIDKFIDTLDGQRLQD
RKHKDSRRKIKEQREGTESLSKQGEQNGSSVVTGSKGTNDARNCRSEIPHQPNTAKRHKENASSGDEIHD
SKRGRTKLERWASHKEREDAVSAKSSSISSKLEEKENNTNGRLSEPVHGSIGKSRDVTEEKIGHDLADTK
DGSEKGPGRHLDTVEKLLKRRSERFKLPMPTEKDDTGVKKMESETLPSAKIEGPVDSEGEYVWDERSCVR
IGREYA

AtPAP

>tr|F4KE92|F4KE92_ARATH Polynucleotide adenylyltransferase family protein
OS=Arabidopsis thaliana GN=At5g23690 PE=3 SV=1
MAISSVGFACRSFSPVRTLQSHCLWKIRFNTVAAAIETMDESDFATDSDRQDKVSDEPR
DREWKQLNSKDLGLSSSMIAKSTRKVLNGLKSKGHDVYLVGGCVRDLILKRTPKDFDILT
SAELREVVRTFPRCEIVGRRFPICHVHIGDDLIEVSSFSTSAQNSSRNTRTECKESSGSD
GDEDCIRLNNCLQRDFTINGLMFDPYAKVVYDYLGGMEDIRKAKVRTVIHAGTSFHQDCA
RILRAIRIAARLGFMRSKETAHFINKLSLLVQRLDKGRILMEMNYMLAYGSAEASRLLLW
KFGILEILLPIQAAYLARSGFRRRDKRTNMLLSLFANLDKLLAPDRPCHSSLWIAILAFH
KALADKPRSPIVVAAFSLAVHNCGDILEAVEITKKITRPHDKSFFELVEPEENLDFQTLL
DEVMDLDASIEDALNQMTDAYFISKAMSAYPQAPYSDLVFIPLQLYLRAGRIFDCVKNEE
TRIGFEAKQGSKIEYGLNSGYFPEIRHV FARVVDTVFPLNLSQEL

AtCstF77

>tr|Q8GUP1|Q8GUP1_ARATH Putative uncharacterized protein At1g17760
OS=Arabidopsis thaliana GN=CSTF77 PE=2 SV=1
MADKYIVEEAEALAKRALHSPIAQATPIYEQLLSLYPTSARFWKQYVEAQMAVNNDATK
QIFSRCLLTCLQVPLWQCYIRFIRKVYDKKGAEGQEETTKAFEFMLNYIGTDIASGPIWT
EYIAFLKSLPALNLDLHRKTALRKVYHRAILTPTHHVEQLWKDYENFENTVNRQLAKG
LVNEYQPKFNSARAVYRERKKYIEEIDWNMLAVPPTGTSKEETQWVAWKKFLSFEKGNPQ
RIDTASSTKRIIYAYEQCLMCLYHPDVWYDYAEWHVKS GSTDAAIKVFQ RALKAIPDSE
MLKYAFAEMEE SRGAIQS AKKLYENILGASTNSLAHIQYLRLRRAEGVEAARKYFLDAR
KSPSCTYHVYIAFATMAFCIDKEPKVAHNIFEEGLKLYMSEPVIILKYADFLTRLNDDRN
IRALFERALSTLPVEDSAEVWKRFIQFEQTYGDLASILKVEQRMKEALSGKEEGSSPPE
SSLQDVVSRYSYM DLWPCTSNLDHLARQELLVKNLNKKAGKTNLPHVPAAIGSVASSSK
VVYPDTSQM VVQDP TKKSEFASSANPVAASANTFPSTVTATATHGSASTFDEIPKTTTP
ALVAFLANLPIDVGTPNVDVVLISICLQSDFTGTQTVKQSFAAKGNPPSQNDPSGPTRGV

SQRLPRDRRATKRKDSRQEEDDTATVQSQPLPTDVFRLRQMRKARGIATSSQTPTGSTS
YGSAFSGELSGSTG

AtCstF64

>tr|Q9M033|Q9M033_ARATH Putative uncharacterized protein T1008_110
OS=Arabidopsis thaliana GN=T1008_110 PE=4 SV=1
MASYSRARKDLANSKSA TELP PKLQRLRYMRRDLQKDDSVFPTELLPHLFDLLSDQFG
AVRK FVAEILGEIGLKYVELIPEIVPLLIKSLEDETPAVARQVIACGADLFRSTLERVAV
QGLHSSELNDLLESSWTWLIKFKDEICSVAFKQGN SGVKLCAMKFVEALILLYTPHEGIE
ADFNISILRGGHPVLKIGDLSIEASQKLGLLLDQLRH PAAKSLNSSTIIIVLINSLSVAK
KRPAYCGRILPVLLSLDPLSFLKGVYAAATNLALKTVFLSCLKCTHPAAAPDRLTSALKE
IEGGGQAAKADL FYKTNGSIQDKDSVEDTKVSVEENPLCASSDVAESNLSRKRS GSEYN
IDLNGDASDGKRARITPSVSEESTDGLNGNDGVSLPRVASTSTGPSDSRGVSDSGPAQQL
VGLFGTLVSQGEKAIGSLEILISSISADLLTDVVMANMHNIPNCSSYADGTDELVMNMC
IVGSDAQIKYPPSFVAGVLSLSTAFPPIAALINPHNEDEEVY SVHVDQQMFP AEDARTPP
GLLATCDTSFPENEESENTVSPQNVHYIGNRESGIPGLESSAQHDGSGALVTNVLSSTNVE
AASKNQNASFSGKLLVDVIPSMSVDKLEEFSPKAVGTVASASQFVLPKISAPVVDLSDEE
KDSLQKL VFLRIVEAYKQISMSGGSQLRFSLLAHLGVEFPSELDPWKILQEHVLS DYLNH
EGHELTVRVLYRLYGEAEAEQDFFSSTTAASAYESFLLTVAEALRDSFPPSDKSLSKLLG
DSPHLPKSVLMLLESFCCPGSGEVEKDLQHGD RVTQGLSAVWSLILMRPGIRNDCLNIAL
QSAVHHLEEIRMKAIRLVANKLYLSL FITEQIEEFKADRLFSVVSDDCDKMDL DLKSPPN
KPQHSISGMSMETPSEATSSSTS VTEAQRCLSLYFALCTKVLRIFTILRLMTNLVFNIYK
NASDPVKQAIHLQIPILVRTMGSSSELLKIIADPPSGSDNLLIQVLQTLTEGPTPSSELI
LTIRKLFDTRIKDVEILFPILPFLPRDDVLRIFPHMVNLPMEKFQVALSRVLQGSSQSGP
VLSPSEALIAIHSIDPARDGIPLKQVTDACNTCF AQRQTFTQQVLAGVLNQLVQQIPLPM
LFMRTVLQAIGA FPA LSDFILEILSRLVSKQIWKYPKLWVGFLKCTQTTQPQSYKVLLQL
PPLQLGNALT KIPALRAPLTAHASQPEIQSSLP RSTLAVLGLVPDSQGTQTSQVQANETQ
TSQEQQQQQASEPQQTSQSQQVSVPLSHSQVDHQEPSQVVASQSQSSPIGTVQSAMSQSQ
NSPIDTGRSEMSQSQNSPIDTGRSEMSQSQNSPIDTGRSEMSQSQNSPIDTGRSEMSESQ
SSPIGQSQSSPIGTGQSDMSQTPQVSDSSAPEPTSHTRTSDPQASSQTLRDDDEKIDDTA
TSENEVTEIEKSKESSEEEEEEEEEEEEE

AtCLPS3

>tr|Q9SR06|Q9SR06_ARATH CLP-like protein 3 OS=Arabidopsis thaliana
GN=F7O18.15 PE=2 SV=1
MAYGGPSMNPPALSGAVPGSANLKQVKLERESELRIEVSEEPLRLRVVNGTAEIFGSELP
PEIWRTFPPRMKFVAVFTWYGATIEMDGVTETDYTADETPMVSYINVHAILDARRRFAKAS
TSNDPESSQGPRVIVVGPTDSGKSTLT KMLLSWAAKQGWRPTFVDLDVGQGSITIPGSIA
AAPIEMPLDPVEGFPLDMALVYYYGHASPNMNVELYKALVKELAQVLEKQFVGNPESRAA
GMVINTMGWIEGIGYELLLHAIDTFNASVVLVLGQEKLFSRLKDVLRSKSNVDVVKLHKS
GGVVARVKEVRKRSRNFKIQEYFYGLSKELSPYANTSSFS DLQVFRIGGGPQAPKSALPA
GSTSVSNPLRVTPVNIDDRDLLHSVLAVSYAE EPDQIISSNVSGFVYVTEVNVQKKKITY
LAPSPGTLPSKLLVAGSLAWLESV

AtPCFS4

>tr|Q0WPF2|Q0WPF2_ARATH PCF11P-similar protein 4 OS=Arabidopsis thaliana
GN=PCFS4 PE=2 SV=1
MDSEKILNPR LVSINSTSRKGMSVELPQKPPPPPSLLDRFKALLNQREDEF GGGGEEVLPP

SMDEIVQLYEVLGELTFNSKPIITDLTIIAGEQREHGEIANAICTRILEAPVEQKLPS
LYLLDSIVKNIGRDYGRYFSSRLPEVFCLAYRQAHPSLHPSMRHLFGTWSSVFPPVLRK
IDMQLQLSSAANQSSVGASEPSQPTRGIHVNP KYLRRLPSAAENNLRGINSSARVYGQN
SLGGYND FEDQLES PSSLSTPDGFTRRSNDGANPSNQAFNYGMGRATSRDDEHMEWRRK
ENLGQGN DHERPRALIDAYGVDTSKHVTINKPIRDMNGMHSMVTPWQNT EEEEFDWEDM
SPTLDRSRAGEFLRSSVPALGSVRARPRVGNTSDFHLDS DIKNGVSHQLRENWSLSQNY P
HTSNRVDTRAGKDLKVLASSVGLVSSNSEFGAPPFDSIQDVNSRFGRALPDGTWPHLSAR
GPNSLPVPSAHLHHLANPGNAMS NRLQGKPLYRPENQVSQSHLNDMTQQNQMLVNYLPSS
SAMAPRPMQSL LTHVSHGYPPHGSTIRPSLSIQGGEAMHPLSSGVL SQIGASNQPPGGAF
SGLIGSLMAQGLISLNNQAGQG PLGLEFDADMLKIRNES AISALYGDLP RQCTTCGLRF
KCQEEHSHKMDWHVTKNRMSKNHKQNPSRKW FVSASMWLSGAEALGAEAVPGFLPTEPTT
EKKDDEDM AVPADEDQTSCALCGEPFEDFYSD ETEEW MYKGAVYMN APEESTTDMDKSQL
GPIVHAKCRPESNGGDMEEGSQRKKMRS

FY

>tr|F4K3Y7|F4K3Y7_ARATH Flowering time control protein FY OS=Arabidopsis
thaliana GN=FY PE=4 SV=1
MMRQSSASSTNINPDYHHPSGPFDPNVDSFGAKMRKHTQRRAVDYTSTVVR YIQFNELI
LYMELKVESVYSIYVFVQARTWQRDSRDRTTLQPTPAAAVDMLPTVAYS DNPSTSFAAKF
VHASLNKNRCSINRVLWTPSGRRLITGSQSGEFTLWNGQSFNFEMILQAHDQPIRSMVWS
HNENYMVSGDDGGTLKYWQNNMNNVKANKTAHKESIRDLSFCKTDLKFCSCSDDTTVKVW
DFTKCVDES SLTGHGWDVKSVDWHPTKSLLVSGGKDQLVKLWDTRSGRELCSLHG HKNIV
LSVKWNQNGNWLLTASKDQIIKLYDIRTMKELQSFRGHTKDVTSLAWHPCHEEYFVSGSS
DGSICHWIVGHENPQIEIPNAHDNSVWDLAWHPIGYLLCSGSNDHTTKFWCRNR PADNPR
DVL MQNQGYNEQGFGRQPDNFQPSEASPIPGA FVPGLTRNEG TIPGIGIAMPFDASSQGD
HKQPLPGSMALGAPPLPPGPHPSLLGSGQQQGYQQQQHQGHPPQMLPMPNMPHHQLPPS
SHMPLPHPHLPRPMQMPPHGHMPPPSMPMSHQMPGSMGMQGGMN PQMSQSHFMGAPSGVF
QGQPNSSGGPQMPYQGRGGFNRPQMIPGYNNPFQQQQQPPLPPGPPNNNQHQ

PABP5

>sp|Q05196|PABP5_ARATH Polyadenylate-binding protein 5 OS=Arabidopsis
thaliana GN=PAB5 PE=1 SV=3
MAAAVASGIAPTAMVDQVIPNQPTVAAAAPPPFPAVSQVA AVAAAAAAAEALQTHPNSS
LYVGDLDPVNESHL LDFNQVAPVHNLRVCRDLTHRSLGYAYVNFANPEDASRAMESLN
YAPIRDRPIRIMLSNRDPSTRLSGKGNVFIKNLDASIDNKALYETFSSFGTILSCKVAMD
VVGRSKGYGFVQFEKEETAQAAIDKLNGMLLNDKQVFVGHFVRRQDRARSESGAVPSFTN
VYVKNLPEITDDELKKTFGKYGDISSAVVMKDQSGNSRSFGFVN FVSPEAAAVAVEKMN
GISLGEDVLYVGRAQKKS DREEELRRKFEQERISRFEKLQGSNLYLKNLDDSVNDEKLKE
MFSEYGNVTSCKVMNNSQGLSRGFGFVAYS NPEEALLAMKEMNGKMIGRKPLYVALAQRK
EERQAHQLSLFTQIRSPGTMSPVPSPMSGFHHHPGGPMSGPHHPMFI GHNGQGLVPSQP
MGYGYQVQFMPGMRPGAGPPNFMMPFPLQRQTQPGPRVGFRRGANNMQQQFQQQQMLQQN
ASRFMGAGNRRNGMEASAPQGI IPLPLNASANSHNAPQRSHKPTPLTISKLASDLALAS
PDKHPRMLGDHLYPLVEQQEPANA AKVTGM LLEMDQAEILHLL ESPEALKAKVSEALDVL
RRSADPAAVSSVDDQFALSSE

Secuencias de proteínas de *Saccharomyces cerevisiae*

Cft1p

```
>sp|Q06632|CFT1_YEAST Protein CFT1 OS=Saccharomyces cerevisiae (strain
ATCC 204508 / S288c) GN=CFT1 PE=1 SV=1
MNVYDDVLDATVVVSHSLATHFTTSDYEELLVVRTNILSVYRPTRDGKLYLTDEFKFKHGLI
TDIGLIPQKDSPLSCLLLCTGVAKISILKFNTLTNSIDTSLHYYEGKFKGKSLVELAKI
STLRMDPGSSCALLFNNDIIAFLPFHVKNKNDDEEEDEDEDENIDDSelihSMNQKSQGTN
TFNKRKRRTKLGDKFTAPSVVLVASELYEGAKNIIDIQFLKNFTKPTIALLYQPKLVWAGN
TTISKLPTQYVILTNIQPAESATKIESTTIAFVKELPWLHTIVPVSNGAIIIVGTNELA
FLDNTGVLQSTVLLNSFADKELQKTKIINNSSLIMFREKNNTTSIWIPSSKSKNGGSNND
ETLLMLDKSNIYYIQMEAEGRLLIKFDIFKLPIVNDLLKENSNPKCITRLNATNSNKNM
DLFIGFGSGNALVLRNLNLKSTIETREAHNPSSGTSNLMDDINDDDEEMDDLYADEAPEN
GLTTNDSKGTVEVTVQPFDIELLSSLRNVGPITSLTVGKVSSIDDVVKGLPNPNKNEYSLV
ATSGNGSGSHLTVIQTSVQPEIELALKFISITQIWNLIKIKGRDRYLITTDSTKSRSDIYE
SDNNFKLHKGGRLRRDATTVYISMFGEEKRIIQVTTNHLVLYDTHFRRLTTIKFDYEVH
VSVMDPYILVTVSRGDIKIFELEEKNRKLLKVDLPEILNEMVITSGILKSNMCNEFLI
GLSKSQEEQLLFTFVTADNQIIFFTKDHNDRIQNLNGVDQLNESLYISTYQLGDEIVPDP
SIKQVMINKLGHNDKKEEYLTILTFGGEIYQYRKLPQRRSRFYRNVTRNDLAITGAPDNAY
AKGVSSIERIMHYFPDYGYSVIFVTGSVPYILIKEDDSTPKIFKFGNIPLVSVTPWSE
SVMCVDNIKARVYTLTNDNMYGNKLPLKQIKISNVLDDYKTLQKLVYHERAQLFLVSY
CKRVPYEALGEDGEKVIGYDENVPHAEGFQSGILLINPKSWKVIDKIDFPKNSVVMERS
SMIQINSKTKRKREYIIAGVANATTEDTPPTGAFHIYDVIEVPEPGKPDNYKLKEIFQ
EEVSGTVSTVCEVSGRFMISQSQKVLVRDIQEDNSVIPVAFLDIPVFVTDSSKSGNLLII
GDAMQGFQFIGFDAEPYRMISLGRSMKFKQTMSEFLVNGGDMYFAATDADRNVHVLKYA
PDEPNLSLQRLVHCSSFTLHSTNSCMMLLPNEEFGSPQVPSFQNVGGQVDGVSFQKIVP
LSEEKYRRLVYIQQQIIDRELQGLNPRMERLANDFYQMGHSMRPMLDFNVIRRFCGLA
IDRRKSIAQKAGRAHFAEAWRDIINIEFSMRSLCQGK
```

Cft2p

```
>sp|Q12102|CFT2_YEAST Cleavage factor two protein 2 OS=Saccharomyces
cerevisiae (strain ATCC 204508 / S288c) GN=CFT2 PE=1 SV=1
MTYKYNCCDDGSGTTVGSVVRFDNVTLIDPGWNPSKVSYEQCIKYWEKVIPEIDVILS
QPTIECLGAHSLLYNFTSHFISRIQVYATLPVINLGRVSTIDSASAGVIGPYDTNKL
LEDIEISFDHIVPLKYSQLVDLRSRYDGLTLLAYNAGVCPGGSIWICISTYSEKLVYAKRW
NHTRDNILNAASILDATGKPLSTLMRPSAIITTLDRFGSSQPFKKRSKIFKDTLKKGLSS
DGSVIIPVDMGKFLDLFTQVHELLFESTKINAHTQVPVLILSYARGRTLTYAKSMLEWL
SPSLLKTWENRNNTSPFEIGSRIKIIAPNELSKYPGSKICFVSEVGALINEVVIKVGNSE
KTTLILTTPSFECASSLDKILEIVEQDERNWKTFPEDGKSFLCDNYISIDTIKEEPLSKE
ETEAFKVQLKEKKRDRNKKILLVKRESKKLANGNAIIDDNGERAMRNQDILVENVNGVP
PIDHIMGGDEDDDEEEENDNLLNLLKDNSEKSAAKNTEVPVDIIIQPSAASKHKMFN
PAKIKKDDYGTVDFTMFLPDSDNVNQNSRKRPLKDGAKTTSVPNEEDNKNEEDGYNM
SDPISKRSKHRASRYSGSGTGEAENFDNLDKIDKTLKRTISTVNVQLKCSVILNL
QSLVDQRSASIIWPSLKSRIKIVLSAPKQIQNEEITAKLIKKNIEVNMPLNKIVEFSTTI
KTLDISIDSNLDNLLKWQRISDSYTVATVVGRLVKESLPVNNHQKTASRSKLVLKPLHG
SSRSHTGALSIGDVRLAQLKKLLTEKNYIAEFKGEGTLVINEKVAVRKINDAETIIDGT
PSELFDTVKKLVTDMLAKI
```

Brr5

>sp|Q06224|YSH1_YEAST Endoribonuclease YSH1 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=YSH1 PE=1 SV=1
MERTNTTTFKFFSLGGSNEVGRSCHILQYKGKTVMLDAGIHPAYQGLASLPFYDEFDLISK
VDILLISHFHLDAASLPYVMQRTNFQGRVFMTHPTKAIYRWLLRDFVRVTSIGSSSSSM
GTKDEGLFSDEDLVDSFDKIETVDYHSTVDVNGIKFTAFAHAGHVLGAAMFQIEIAGLRVL
FTGDYSREVDRLNSAEVPPSSNVLIVESTFGTATHEPRLNRERKLTQLIHSTVMRGGR
VLLPVFALGRAQEIMLILDEYWSQHADELGGGQVPIFYASNLAKKCMSVFQTYVNMND
IRKKFRDSQTNPFIFKNISYLRNLEDFQDFGPSVMLASPGMLQSGLSRDLLERWCPEDKN
LVLITGYSIEGTMAKFIMLEPDTIPSINNPEITIPRRCQVEEISFAAHVDFQENLEFIEK
ISAPNIILVHGAEANPMGRKLSALLSNFASLKGTDNEVHVFNPRNCVEVDLEFQGVKVAKA
VGNIVNEIYKEENVEIKEEIAAKIEPIKEENEDNLDSQAEGKLVDEEEHKDIIVSGILVS
DDKNFELDFLSLSDLREHHPDLSTTILRERQSVRVNCKKELIYWHILQMFGEAEVLQDDD
RVTNQEPKVKEESKDNLTNTGKLILQIMGDIKLTIVNTLAVVEWTQDLMNDTVADSI
LMNVDSAPASVKLSSHSCDDHDHNNVQSNAQGKIDEVERVKQISRLFKEQFGDCFTLFLN
KDEYASNKEETITGVVTIGKSTAKIDFNNMKILECNSNPLKGRVESLLNIGGNLVTPLC

Yth1p

>sp|Q06102|YTH1_YEAST mRNA 3'-end-processing protein YTH1 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=YTH1 PE=1 SV=1
MSLIHPDTAKYPFKFEPFLRQEYSFSLDPDRPICEFYNSREGPKSCPRGPLCPKKHVLPI
FQNKIVCRHWLRGLCKKNDQCEYLHEYNLKRMPECVFFSKNGYCTQSPDCQYLHIDPASK
IPKCENYEMGFCPLGSSCPRRHIKKVFCQRYMTGFCPLGKDECDMEHPQFIIPDEGSKLR
IKRDEINTRKMDEEKERRLNAIINGEV

Pap1

>sp|P29468|PAP_YEAST Poly(A) polymerase OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=PAP1 PE=1 SV=1
MSSQKVFGITGPVSTVGATAAENKLNDSLILQELKKEGSFETEQETANRVQVLKILQELAQ
RFVYEVSKKKNMSDGMARDAGGKIFTYGSYRLGVHGPSDIDTLVVVPKHVTREDDFTVF
DSLLRERKELDEIAPVPDAFVPIIKIKFSGISIDLICARLDQPVPLSLTSLDKNLLRNL
DEKDLRALNGTRVTDEILELVPKPNVFRIALRAIKLWAQRRAVYANIFGFPGGVAWAMLV
ARICQLYPNACSAVILNRFFIILSEWNWPQPVILKPIEDGQLQVRVWNPKIYAQDRSHRM
PVITPAYPSMCATHNITESTKKVILQEFVRGVQITNDIFSNKSWANLFKNDFFFRYKF
YLEITAYTRGSDEQHLKWSGLVESKVRLLVMKLEVLGKIAHPFTKPFESSYCCPTEDD
YEMIQDKYGSHKTETALNALKLVTDENKEEESIKDAPKAYLSTMYIGLDFNIENKKEKVD
IHIPCTEFVNLCRSFNEDYGDHKVFNLALRFVKGYDLPDEVFDENEKRPSKSKSRKNLDA
RHETVKRSKSDAASGDNINGTTAAVDVN

Fip1

>sp|P45976|FIP1_YEAST Pre-mRNA polyadenylation factor FIP1 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=FIP1 PE=1 SV=1
MSSSEDEDDKFLYIGSDSELALPSSKRSRDDEADAGASSNPDIVKRQKFDSPVEETPATAR
DDRSDEDIYSDDSDSDSLEVIISLGPDPTRLDAKLDSYSTAATSSSKDVISVATDV
SNTITKTSDERLITEGEANQGVTTATTVKATESDGNVPKAMTGSIDLKKEGIFDSVGITTI

DPEVLKEKPWRQPGANLSDYFNYGFEFTWMEYLHRQEKLQQDYNPRRILMGLLSLQQQG
KLNSANDTDSNLGNIIDNNNNVNNANMSNLNSNMGNSMSGTPNPPAPPMHPSFPPLPMFG
SFPPFPMPGMMPPMNQQPNQNNQNSK

Pta1p

>sp|Q01329|PTA1_YEAST Pre-tRNA-processing protein PTA1 OS=Saccharomyces
cerevisiae (strain ATCC 204508 / S288c) GN=PTA1 PE=1 SV=1
MSSAEMEQLLQAKTLAMHNNPTEMLPKVLETTASMYHNGNLSKLLPLAKFFTQLVLDVV
SMDSPIANTERPFIAAQYLP LLAMAQSTADVLVYKNIVLIMCASYPLVLDLVAKTSNQE
MFDQLCMLKKFVLSHWRTAYPLRATVDDDETVEQWLAQIDQNIGVKLATIKFISEVVLSQ
TKSPSGNEINSSTIPDNHPVLNKPALSEAKRLLDMLLNYLIEEQYMVSSVFIGIINSLS
FVIKRRPQTTIRILSGLLRFNVDKFPLEGKSDLNYKLSKRFFVERAYKNFVQFGLKNQII
TKSLSSSGSGSSIYSKLTKISQTLHVICEETKSKGILNFDPSKGNSKKTLSRQDKLKYISL
WKRQLSALLSTLGVSTKTPTPVSA PATGSSTENMLDQLKILQKYTLNKASHQGNTFFNNS
PKPISNTYSSVYSLMNSSNSNQDVTQLPNDILIKLSTEAILQMDSTKLITGLSIVASRYT
DLMNTYINSVPSSSSSKRKSDDDDDGNDNEEVGNDGPTANSKKIKMETEPLAEEPEEPED
DDRMQKMLQEEESAQEIISGDANKSTSAIKEIAPPFEPDSLQDEKLKYLKLTCKLFELS
GRQDTRAKSSSSSILLDDDDSSWLHVLIRLVTRGIEAQEASDLIREELLGFFIQDFE
QRVSLII EWLNEEWFFQTS LHQDPSNYKKWSLRVLES LGPFLENKHRRFFIRLMSELP SL
QSDHLEALKPICLDPARSSSLGFQTLKFLIMFRPPVQD TVRDLLHQLKQEDEGLHKQCD SL
LDRLK

Rna15p

>sp|P25299|RNA15_YEAST mRNA 3'-end-processing protein RNA15
OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=RNA15 PE=1
SV=1
MNRQSGVNAGVQNNPPSRVVYLGSIPYDQTEEQILDLC SNVGPVINLKMMDPQTGRSKG
YAFIEFRDLESSASAVRN LNGYQLGSRFLKCGYSSNSDISGVSQQQQQQYNNINGNNNNN
GNNNNNSNGPDFQNSGNANFLSQKFPELPSGIDVNINMTTPAMMISSELAKKPKEVQLKF
LQKFQEWTRAHPEDAVSLLELCPQLSFVTAELLLTNGICKVDDLIPLASRPQEEASATNN
NSVNEVVDP AVLNKQKELLKQVLQLNDSQISILPDDERMAI WDLKQKALRGEFGAF

Rna14p

>sp|P25298|RNA14_YEAST mRNA 3'-end-processing protein RNA14
OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=RNA14 PE=1
SV=2
MSSSTTPDLLYPSADKVAEPSDNIHGDELRLRERIKDNPTNLSYFQLIQYLETQESYAK
VREVYEQFHNTFFFYSPAWTLQLKGELARDEFETVEKILAQCLSGKLENNDLSLWSTYLD
YIRKNNLITGGQEARAVIVKAFQLVMQKCAIFEPKSSSFWNEYLNLFLEQWKPFNKWEEQ
QRIDMLREFYKMKLCVPFDNLEKMWNRYTQWEQEINSLTARKFIGELSAEYMKARSLYQE
WLNVTNGLKRASPINLRTANKKNIPQPGTSDSNIQQLQIWLNIKW ERENKMLMLSEDMLS
QRISYVYKQGIQYMI FSAEMWYDYSMYISENSDRQNILYTALLANPDSPSLTFKLSECYE
LDNDSESVSNCFDKCTQTLLSQYKKIASDVNSGEDNNT EYEQELLYKQREKLT FVFCVYM
NTMKRISGLSAARTVFGKCRKLKRILTHDVYVENAYLEFQNNQNDYKTA FKVLELGLKYFQ
NDGVYINKYLD FLIFLNKDSQIKTLFETSVEKVQDLTQLKEIYKKMISYESKFGNLNNVY
SLEKRFFERFPQENLIEVFTSRYQIQNSNLIKKLELT YMYNEEEDSYFSSGNGDGHGGSY
NMSSSDRKRLMEETGNNGNFSNKKFKRDS ELPTEVLDLLSVIPKRQYFNTNLLDAQKLVN
FLNDQVEIPTVESTKSG

Pcf11

>sp|P39081|PCF11_YEAST Protein PCF11 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=PCF11 PE=1 SV=2
MDHDETEVIVKDFNSILEELTFNSRPIITTLTKLAEENISCAQYFVDAIESRIEKCMPKQK
LYAFYALDSICKNVGSPYTIYFSRNLFNLYKRTYLLVDNTRTKLINMFKLWLNPNDTGL
PLFEGSALEKIEQFLIKASALHQKNLQAMLPTPTVPLLLRDIDKLTCLTSERLKNQPNDE
KLKMKLLVLSQLKQELKREKLTNLALKQVQMLRQVFSQDQQVLQERMRYHELQOQOQOQOQ
QOQOQOQOQOQOQYHETKDMVGSYTQNSNSAIPLFGNNSDTTNQONSLSLSSSLFGNISGVE
SFQEIEKKKSLNKINNLYASLKAEGLIYTPPKESIVTLYKKLNHGSNYSLDSHEKQLMKN
LPKIPLNDILSDCKAYFATVNIDVLNPNPSLQQLSEQTLLQENPIVQNNLIHLLYRSKPNK
CSVCGKRFGNSESEKLLQNEHLDWHFRIINTRIKGSQNTANTGISNSNLNTTTTRKNIQSR
NWYLSDSQWAAFKDDEITSTKHKNDYTDPHANKNIDKSALNIHADENDEGSVDNTLGS DR
SNELEIRGKYVVVPETSQDMAFKCPICKETVTGVYDEESGEWVWKNTIEVNGKYFHSTCY
HETSQNSSKSN SGKVLDDLKLVTK

Clp1

>sp|Q08685|CLP1_YEAST Protein CLP1 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=CLP1 PE=1 SV=1
MASLPGIDEHTTSEELITGDNEWHKLVI PKGSDWQIDLKAEGLIVKVNSGIVEIFGTEL
AVDDEYTFQNWKFPIYAVEETELLWKCPDLTTNTITVKPNHTMKYIYNLHFMLEKIRMSN
FEGPRVVIVGGSQTGKTSLSRTLCSYALKFNAYQPLYINLDPQQPIFTVPGCISATPISD
ILDAQLPWTGQSLTSGATLLHNKQPMVKNFGLERINENKDLYLECISQLGQVVGQRLHLD
PQVRRSGCIVDTPSISQLDENLAELHHIIEKLNVNIMLVLCSETDPLWEKVKKTFGPELG
NNNIFFIPKLDGVS AVDDVYKRS LQRTSIREYFYGSLDTALSPYAIGVDYEDLTIWKPSN
VFDNEVGRVELFPVTITPSNLQHAIITFAERRADQATVIKSPILGFALITEVNEKRRK
LRVLLPVPGRRLPSKAMILTSYRYLE

Hrp1p

>sp|Q99383|HRP1_YEAST Nuclear polyadenylated RNA-binding protein 4
OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=HRP1 PE=1
SV=1
MSSDEEDFNDIYGDDKPTTTEEVKKEEEQNKAGSGTSQLDQLAALQALSSSLNKLNNPNS
NNSSSNNSNQDTSSSKQDGTANDKEGSNEDTKNEKKQESATSANANANASSAGPSGLPWE
QLQQTMSQFQQPSSQSPPPQQVQTKEERSKADLSKESCKMFIGGLNWDTTEDNLREYFG
KYGTVTDLKIMKDPATGRSRGFGFLSFEKPSSVDEVVKTQHILDGKVIDPKRAIPRDEQD
KTGKIFVGGIGPDVRPKFEFEFFSQWGTIIDAQLMLDKDTGQSRGFGFVTYDSADAVDRV
CQNKFIDFKDRKIEIKRAEPRHMQQKSSNNGGNNGNMNRGNGNFGNQGDFNQMYQNPM
MGGINPMMNPQAMTDYYQKMQEYYQQMQKQTGM DYTQMYQQMQQMAMMMPGFAMPPNAM
TLNQPQQDSNATQGS PAPS DSDNNKSN DVQTIGNTSNTDSGSPPLNLPNGPKGPSQYND
HNSGYGYNRDRGRDRDRDRDYNHRSNGNHRNRGRGRGGYNRRNNGYHPYNR

Pfs2p

>sp|P42841|PFS2_YEAST Polyadenylation factor subunit 2 OS=Saccharomyces
cerevisiae (strain ATCC 204508 / S288c) GN=PFS2 PE=1 SV=1
MDGHNQNQYQNQNQIQSQQPPLKKYVTQRRSVDVSSPYINLYNRRHGLPNLVVEPETS
YTIDIMPPNAYRGRDRVINLPSKFTHLSSNKVKHVIPAIQWTPEGRRLVVATYSGEFSLW

NASSFTFETLMQAHDSAVTTMKYSHSDSDWMISGDADGMIKIWQPNFSMVKEIDAAHTESI
RDMAFSSNDSKFVTCSDDNILKIWNFSNGKQERVLSGHHWDVKSCDWHPEMGLIASASKD
NLVKLWDPRSGNCISSILKFKHTVLKTRFQPTKGNLLMAISKDKSCRVFDIRYSMKELMC
VRDETDYMTLEWHPINESMFTLACYDGSCLKHFDLLQNLNEPILTIPYAHDKCITSLSYNP
VGHIFATAAKDRTIRFWTRARPIDPNAYDDPTYNNKKINGWFFGINNDINAVREKSEFGA
APPPPATLEPHALPNMNGFINKKPRQEIPGIDSNIKSSTLPGLSI

Ssu72

>sp|P53538|SSU72_YEAST RNA polymerase II subunit A C-terminal domain
phosphatase SSU72 OS=Saccharomyces cerevisiae (strain ATCC 204508 /
S288c) GN=SSU72 PE=1 SV=1
MPSHRNSNLKFCSTVCASNRRSMESHKVLQEAGYNVSSYGTGSAVRLPGLSIDKPNVYSF
GTPYNDIYNDLLSQSADRYKSNGLLQMLDRNRRLKKAPEKWQESTKVFDFVFTCEERCDFD
AVCEDLMNRGGKLNKIVHVINVDIKDDDENAKIGSKAILELADMLNDKIEQCEKDDIPFE
DCIMDILTEWQSSHSQPLPSLYAPSY

Mpe1

>sp|P35728|MPE1_YEAST Protein MPE1 OS=Saccharomyces cerevisiae (strain
ATCC 204508 / S288c) GN=MPE1 PE=1 SV=1
MSSTIFYRFKSRNTSRILFDGTGLTVFDLKREIIQENKLGDTDFQLKIYNPDTEEEYD
DDAFVPIRSTSVIVKRSPAIFSVHSRLKGNVGAAALGNATRYVTGRPRVLQKRQHTAT
TTANVSGTTEEEERIASMFATQENQWEQTQEEMSAATPVFFKSQTNKNSAQENEGPPPGY
MCYRCGGRDHKWKNCPTNSDPNFEGKRIRRTTGIPKKFLKSIEIDPETMTPEEMAQRKIM
ITDEGKFVVQVEDKQSWEDYQKRENQIDGDETIWRKGHFKDLPDDLKCPLTGGLLRQP
VKTSKCCNIDFSKEALENALVESDFVCPNCETRDILLDSLVPDQDKEKEVETFLKKQEEL
HGSSKDGNPETKKMKLMDPTGTAGLNNNTSLPTSVNNGGTPVPPVPLPFGIPPFPMFPM
PFMPPTATITNPHQADASPKK

Sub1

>sp|P54000|SUB1_YEAST RNA polymerase II transcriptional coactivator SUB1
OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=SUB1 PE=1
SV=1
MSYYNRYRNKRSDNNGGNLSNSNNNNGMPSGLSASDAIFDLGKNKRVTVRQFRNINLI
DIREYYLDSSTGEMKPGKKGISLTEDLYDELLKHRLNIDEALRRLGSKRPKTKMVRLLSD
DEYEDDNNNDSTNNDKDKNGKDKNSPKKRREDKSKASNESHDLPRSKKKKPPAPTLLPH
EENIQNAEREANATLIIPGQAGRKQQEERKQKEKEEAEAKAKAVAEQKEAKEKEKIAE
PEPEPVPTLQAKKEDIVSNINESKDANSSDEEFAQSLEAEMNKAEDDISEEE

Pti1p

>sp|P39927|PTI1_YEAST Protein PTI1 OS=Saccharomyces cerevisiae (strain
ATCC 204508 / S288c) GN=PTI1 PE=1 SV=2
MTDPRRRTGRHFLTPENLSSTLQITNLPPEWNQDIITSVVAGSGPVIDIKAKNDPRTGKL
TGVLFDYLTSKDCKRAWEILNRIENFPVKIEQIIPPNYKDHLRETANKNSQKQVLQLNRD
SYPFEAGLELFPFEMVTEVPIPRRPPPPQAANNTNSVSNNTNIQFPDILSKASKHLPSFQD
GSIIAPDKISQNLSKIPPLQLIEIISNLKILSNQENIQKSQLESFLDTNSDITISVTQAL

LEMGFIDYSVVTKVLKSQVGEAPSLSSNNTSNTSPVSVIRNNTPLHVPSNEVSNNPNN
MPLNVAMPMPMSTPPFIPLPLQQQPFQFAPPGPFMPPAQGPSMGQPVLANQLGQVQQQNI
SSTEGPSNANKANDSGTINMAKLQLLPENQQDMIKQVLTLTTPAQIQSLPSDQQLMVENFR
KEYII

Swd2p

>sp|P36104|SWD2_YEAST COMPASS component SWD2 OS=Saccharomyces cerevisiae
(strain ATCC 204508 / S288c) GN=SWD2 PE=1 SV=1
MTTVSINKPNLLKFHVKSQFQPEKDCGPVTSLNFDNQGFLLTSSSNDTMQLYSATNCK
FLDTIASKKYGCHSAIFTHAQNECIYSSTMKNFDIKYLNLETNQYLRYSFGHGALVNDLK
MNPVNDTFLSSSYDESRLWDLKISKPVIIIPSLVPNCIAYDPGLVLFALGNPENFEIGL
YNLKKIQEGPFLLIKINDATFSQWNKLEFSNNGKYLLVGSSIGKHLIFDAFTGQQLFELI
GTRAFPMREFLDGSGACFTPDGEFVLGTDYDGRIAIWNHSDSISNKVLRPQGFIPCVSHE
TCPRSIAFNPKYSMFVTADETVDFYVYDE

Syc1p

>sp|Q08553|SYC1_YEAST Protein SYC1 OS=Saccharomyces cerevisiae (strain
ATCC 204508 / S288c) GN=SYC1 PE=1 SV=1
MDLPKDKSDRTHQRINLNNSGTDRTNDLYLHIVQTFGCIETTATENATKLLMLGDVEVEI
SASSVSIEWTQKSMISQTIADSIVIMIIGLCASDKNVLSESELKERNHNWVKIQELQNLF
REQFGDSFSIDEGIGKKENVKNGSVTIGKSKATIDFSTMKLIDCNSNPLKGRVESILSIG
QKLTTPLC

Glc7

>sp|P32598|PP12_YEAST Serine/threonine-protein phosphatase PP1-2
OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=GLC7 PE=1
SV=1
MDSQPVVDVNIIDRLLEVRSKPGQQVDLEENEIRYLCSKARSIFIKQPILLELEAPIKI
CGDIHGQYYDLLRLFEYGGFPPESNYLFLGDYVDRGKQSLETICLLLAYKIKYPENFFIL
RGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIIDEKIFCMHGGLSPDLN
SMEQIRRVMRPTDIPDVGLLCDLLWSDPDKDIVGWSENDRGVSFTFGPDVVNRFLQKQDM
ELICRAHQVVEDGYEFFSKRQLVTLFSAPNYCGEFDNAGAMMSVDESLLCSFQILKPAQK
SLPRQAGGRKKK

Ref2p

>sp|P42073|REF2_YEAST RNA end formation protein 2 OS=Saccharomyces
cerevisiae (strain ATCC 204508 / S288c) GN=REF2 PE=1 SV=2
MSAPVPQLVNISHALQASTIQQIRLDMVDFNKDCKLSSIQLARIDKYIDSLQAALNQFTK
DNLHIERKEKNVTEADIQLYSGLSMYLDYLNQLIKLKHEKQHHSTPPIANDVSLDFFVN
QLPKFSPEERKNYIDNLILNKNSHNRLSKMDGLVDAVINLCVLDTSVAENVRSYMKLLDT
LGFQKGSNSTGTKANLKKKLASSKAKIKDSEKEKEKEKDKSKVKMKTCLKPSPLLNNDDK
NSSPSPTASTSSMKKLKSGLFNKNEAKSTESLPTSSKKKLSFSKYLNKDDADMTKLGTKR
SIDVDVKVNPEASTVASNIISSTSGSSTTTVATPASSEEPLKKKTKISVQDSNVQSILR
NGKPKKARISSIKFLDDSQLIKVYGDDLPNQGLQVSPTQLKKILKPFKEGEPKEIILFED
MSIKLKPLDLMFLKNTNSDDYMDISETKGGPIHCETRTPLIYRKNFNHFNPDNLNKRPPRE
PIEFDLNGNTNSTPTIAKAFGKNSLLLRKDRGGLPYKHVPIVKNKYPPRPVH

PAB1

>sp|P04147|PABP_YEAST Polyadenylate-binding protein, cytoplasmic and nuclear OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=PAB1 PE=1 SV=4

MADITDKTAEQLENLNIQDDQQAATGSESQSVENSSASLYVGDLEPSVSEAHLYDIFSP
IGSVSSIRVCRDAITKTS LGYAYVNFNDHEAGRKAIEQLNYTPIKGRLCRIMWSQRDPSL
RKKGSGNIFIKNLHPDIDNKALYDTFSVFGDILSSKIATDENGKSKGFGFVHFEEEGA
EAIDALNGMLLNQGEIYVAPHLRKRERDSQLEETKAHYTNLYVKNINSETTDEQFQELFA
KFGPIVSASLEKDADGKLKGFVNYEKHEDAVKAVEALNDSELNGEKLYVGRAQKKNER
MHVLKKQYEAYRLEKMAKYQGVNLFVKNLDDSVDEKLEEEFAPYGTITSAKVMRTENGK
SKGFGFVCFSTPEEATKAITEKNQQIVAGKPLYVAIAQRKDVRRSQLAQQIQARNQMRYQ
QATAAAAAAAGMPGQFMPPMFYGVMPFPRGVFPNGPNPQQMNPMMGMPKNGMPPQFRNGP
VYGVPPQGGFPRNANDNNQFYQQKQKQALGEQLYKKVSAKTSNEEAAGKITGMILDLPPQ
EVFPLLESDELFEQHYKEASAAYESFKKEQEQQTEQA

Secuencias de protein de *Entamoeba Histolytica*

EhCPSF160

>EhCpsf160|EHI_106110

MTWLLHETVVNPQGVESHIIAKFTGKYEQLITSIRNQFVCYDITPQGLQVRCKEVTEANI
IQIGNVEVNHVMMLVLLFKEAKVSVLRYDETNNKFVIHSLHCFELPLKRMQEGLTPTTYT
DPRLLIDKRGRCISLICDRLMWVIPLGLDKTSYSINLEKFGINRIIDCIVLDGYDLPSV
AFLHMKIPTWEGRIVNTGETTNEIIILSLEPDVIHERQDIVATISYQFSYVPYNALQIVD
CYPTNGLLILTINSIIYLSTTSFESFILPFGKFFVIPKNINGPLSSFQILQMOTKIMNSV
KSIFKLTNHLIIIFSMNGESYYVHLLSIANRICDVIITNSPYKYHPTTFTISSNHLFIGS
TVHDSYIYNYEIVEYGGKQHEISQHINQEIIRSKNLLFRHNEMEEEEYPFDEPQPVSIPOE
LTVEFPQLDYIVSICNVHEYTLLEENNMISIVLCCGIQPFISKIVKLSRSIQPENIMGVK
ESECIFYKNDNIEVNKNGIFLNSINICSTEILGSIKRVKKEDDYLIIEKSDAKWCVFDINK
KEIIIQEEGINELSFLPSSNEMEEEGEGNNFIKLKNGKVYKGEMVLEGIGERISQWVKEE
TKKGKEVVDIECVSIIIEGKIDLIITKEGILCYRSMNDKGKFERLKVEAPQTFNCSIKP
HINRIEINNKKHYLINGNCSYFIEYNYYPPIVIPFSIKVIKEVIVLNNFTCYLDTDDGKL
IVKLNKQNIIDKYPLPFTRVDIQGTPHHLVSYKNKLFVTVSFTTVLDTDCPPQLPHSSNN
QQEPPVLEAQPLNSDVLLEERYRIINLTNDYEFKKGQFICTAKLLKFKVGKQLKNYLV
VGVNKQTTEDNPVGKGTIIFNIENQIQLINKIGDGKKS VHAVNEIGGFLAVASGNELELI
ERVDETRWIKKCFSDISILINSIEYLPLKVMERGNEKECYLILLSDFYRSVLLLLFKPYD
YTVIPLGKDARNIHCIDSTFIITKDYFSVLEFDSEQNLSLLNYSSAATEQLSIFEIDATF
NLGMNLLKFTRLWNGKGYIYMYVTVEGSVGYISVVEEKIYQVLRQINIKMNREPWHFAGT
NAEYRFEKGYGMGFGTRKHVFLDGDMLKQFRLLNEEQQKRVCLRNTSINDVFKLLSTGL
QYNTFLQYNN

EhCPSF100

>EhCPSF100|EHI_033130

MKFNRPMVLLTDIGGLGKCQEMLNVMMEIRKIVLRKGCCLLPVECGGRIMEYMEMVYIS
CDVDINRVIKDAFYCISVADQIKEMNKTIMEWVRDGTITNFKNISFQDNLEKLKRD
DARKPCVVFATLASFSVGSISRDLVKICRDKRNGIIFLEPPQKGTVASEIEEFASPMNQ
KTSVKVDVWHTEDYTEELKEIEALKEQKAKEENENKEHQIEEEKVQEKEIEEIEEIEE
SEEDKKTMMQIKTDLYEDDLEDVGGLFPPFFHNKVETTVYGDISTFKIEMNVEEPLLGNKD

VDEVNEIEDYPRKYVKEEEEELIISCTVASYDVSAEMETTKIRSIIARLIPRNLVVFVSALE
PNGIEWFKQNLPHSNIYGFNNTTEIMTTICPVTTPSETFTIDDSLLTVMKLNHFKEFNLGPI
DAKVEDGTLLPITRQQRKRHESVYIGELPIKVLTKAIETENIDVKIVNGTIVCANGTITV
SKEPSDVPKFIVRGRMNKAFFKVRKIVAEQFCIL

EhCPSF73

>EhCPSF73|EHI_136170
MSTTVVLDNEIQSTLINGNYLEIRPLGAGREVGRSCFILKYMGHNIMLDCGVHPAKPHGE
AALPLFEHADIDSIELLCVTHYHVDHCASLPYLILERQFKGKVLMTPTKEIFGELFKEF
HQMSSSTIQPPKSVNPKEVIDRIDTIKFHELQYNGMKIWCNAGHILGAAMFCIEINGVK
ILYTGDFSGETDRHLQAAEVPPFQIDVMMCESTYGIIEQESRIDRENAFIRQIIEILKRG
GKCLIPVFSLGRAQEFELILEEYWQNHKDLWSVSIFFFSSIAKKCTTYFEKFTSFMNQEL
RKKTKQAFDFKFIREGSSSVDDGAIDYKPCVVMASPGMLQDGISRKIFERWCTDKKNGVI
IPGYCVEGTAKDLILDSTKPFINSEGDQVVPKCSVTEISFCAHSDFAHTRKFIGNVKPK
HLVLIHGEGKSMEQLHNALKKEYPELNIYMPCNTQPIKIPKHEVRLLGEIASNILKM
EEEEKKEKEIIFNAVYLQNTKDVDTVVVKAKDVERFSSFSHSNEIYQHLYFPCLIPKID
ILKGFYPETIFNEENNTAQIKGIQIKYENNDLCISWQNDIITDILIDHIIIVLLEMGISE
KKDRPRKIPQLEDVSLDLVLDLLSNKFESAMILKNKDDVQKAYDNINQINKDLISKNETN
DEQQKDEIKDEKSKDENKNEVKDETKQESVKEVVEQKIEEEFCIKENVPYILIHENSNSV
LVEWPSLTPFGFDKDLLKRLNSVLPKILLTMKNTF

EhCPSF30

>EhCPSF30|EHI_067580
METLLELNNVMKMDISRVFREIKTERDPRTIKSNEKTIVCQHWLRGMCRKGANCDFLHRL
DEERTPACHHFVKYKCKECPFKHEDPQKAIPCEWYKRGFCCKHGKCKKHGHPVKLMCP
LFYLGFCPYGKKCKFVHPTISVPRETIEKTQPRGERFPRDGDEKNANFKHEFTKPKRD

EhFIP1

>EhFip|EHI_083400
MQQNTQPKPQTPSDESYRFYQLLLFNFSHFVADEDFNRYWEVFTGEKGFETLAPEDKPWE
TEGANLEEFNYGFNEKTWSAYAKKLRELYKNSLAEKEITTVDTTVKTKQAIMDLEPKNK
IKNEPKIEESKEEPKKEERRERGNDKHEERKTRQSEKRESRDKESKRESRDRERRDKESR
DRESRRERDEHKRREHESRTEHSRRSRRRERSPRKR

EhPAP

>EhPAP|EHI_012040
MACELKEFLVKNDMYPMGDQVEKKRKAISKMTHEYIQQWGKKIYIESAGVADDDVKAATI
YVYGSYRLNVYGNNSDIDACIVSNSTITRDDFYDGLYAELLNPNPDVKELKQIPSKRSPHL
SMIYLNIEFDLNFRTAYTSLPDNLDILNENILKNMDELDTRAINGVRNTDIIDAFVPNH
SEEAFRVMVRTIKLWTKKRGYGYVYCFNLGISIEILVAQVISENYQLDNVRLLEKFFQV
YSSWDWLRTVPLGTNDFFNDKKKEGVIQILTPASPSENAFAFSITKFSLEMIKRELKRGQ
EIVHEFSSEGVDWAKLFKPRHLFCGYIIFIEFIVTSSSLEGLTTTIGKFESGLVNLKMG
LSEIEEIIIEANVIPNGFLDEENEKYFYVGMNVKRDPCVDISTPLNSFLSIVNSGKDLIV
DASIKKRSEIPTKFAHSVKRSITKSQEIKETSSQVPSSAITETFDIPTKPSIEQQLKAKE
ENSIPNEEKKEQLKKEMKQEANTIVKNSSTDDDFMKRFTRKN

EhCstF77

>EhCstF77|EHI_098370

MQNTPNEEAWESKFSALPFFSDEEQIKQGRELYKRYLGEYPTAVNRWCEYIDLEMKYGHN
DSEVEEIEFRKCLVQVPDVEIAKRYIKYINTCYDDTEREDIDDIELARFKKIIEGAYSYAI
KIVGLDLNASTIYREFMEFLSKSRSNEVTKKILMHNLTTRIPMNERQEMFLEYEEKLDEEE
KKIEHQVFEATKSSQMRLTSYYNRIEQQKSDHINYCTKSIRSLSLLTYIETINYES
MIEGQPRYDDGKYTTNGTDKQYENCPSIQMKARQERIYYAMNKMLLTfYRYPQAWIICA
QYFERRGNIELAINFLERGRAAVNCP LLKLYQSFCFMIQGNTEKALSLLDNNSELEQVFK
LKIAARAYRINEFRKLFTTIQTNLKPYAFIAATEIEYFIFGRKTASLKLfQEALKRYPNN
LPIIKSYRFLNLITPSRFIPNALKTYGDYDQYMKNAILLKFKHLPKPKD TTQFLMSE
DLMHSIGEFDDHQISGDSSLSQFKWKPIEREVNKDIIREFIYILPPPNIYRGRKIDVVQ
LMDILNNLVN

EhCstF64

>EhCstF64|EHI_151990

MSSSEEKSI ESEEVKEEVKEEVEKKDKDKDKDKKEKKEKKEKDKDKKEKKEKHKDKKEK
KEKHHHHSKKEEEEKKEEEEKKEESKMEEHSKEEKEDESHSEEEKEENQSEEEKEKK
EEKKDNETKEFIPVEDVKLVSVFVKNIPIYGWATNDMYQYFASSGPVKTRVITNKETGKSR
GFGYLD FVDLEAANKFVEEHQGEEDVGRPLFLDLADGKKGGDKDNNGGKFGAFGSSNKTG
KQHTYKKGKQGGDNVCFNCGKPGHMSRECPEPRKDNRKGGFKRSFNENKTETKNDPKHIRF
D

EhCstF50

>EhCstF50|EHI_152770

MTEVP EEHSNDFIPSLPPSFQTNFGNERFVDLLKDILKEYTLLYEDNKALIAQLEGLKVQ
PNVRNVKFYQSDVVDTKPKPKQKSTYSHYVSKKTTEHWKFLHNYILHRDCILD LACSPWD
MSIFATASSDRSARISTAESGKCLQFLVHPRPVSSLMFHPTEHLLATGCGDGIVRLF KIG
VTIDEDHKEQLPTEIKIGDIVSGVCLGND SIYTATWGGVVGVDLHGEHISSISSIGI
NDHITSLASDVYSPILSCTATDGSLRLIDIRTPNHARIECINAHS DSCISAQFCEQGKTI
VSTGFDKIVKIWDIKMMKAPKAMYKISYPSKINISQNDKFLVQAEDKSWIIIDSNGLRHG
KLKIGADVLQREHLRLQICSSFSYDESVLYTAGLDRHVIAWGVQTN

EhCFIm25

>EhCFIm25|EHI_077110

MEEQEKRTTEPMQEEEEIKKEEAIEKRTTESMQEEEIKKKEEELKKKERNEHRNYPPLLK
IYPIENYQIDKKEKLDKLKHQTFGYQMDQLKISVEKNHVPRTSVYGVILVHKNNFP HLLV
LQSNLSMDLKDEIHLVGGRLKIGEDDPVEGLKRKL RKKMSMEYITHYEIGELLGTfYRIE
YDKNLYPYIPVHVSQVKEIINIYMIHLVEKCDFKIFD TDKLSSIPLFALHNNFEKYNITL
CSIPTLVSRYLMIYG

EhClp1

```
>EhClp1|EHI_008100
MTTSTNDIKTVHVNKGKELRIITGENFTFKILKGQMEWNGCELSNGTAQSIQQNTWLSFY
AIEESDVEIVNSKSCYIGIKSVVEEYYSIFNNLNDNRSHFAEDEPAQIILVTGPARGKT
TFALHQLNYQVQYGYNPLFVDLSLSTNIIITLPGCIGCTVVSNTFSPTTQLDYNFNQPLVY
CCGTTKIEDKNPLYHYHIELLAANCLTKMNSDQSVKKS GMIIRCPFISMEVIKKVNSFK
VDLVYVMDSEQLLNEIESSQFDHCVTSSFVNKPAGIPLCIEYNSFIRRLQHRVNNYFYG
CSPESPLCPRSSVLLSSSAKVVTIKSLNNENAGFLPIGQPASQKPYVEEKDRQDMKKNDV
YALLNCTTVDEVLNDDTTNVIGFVTVENIGISAEDQNILSLQIEILSPHPIEELPSTVLVA
TGMKIEHDI
```

EhPcf11

```
>EhPcf11|EHI_008100
MNSVNSIPIPNDELTLQCFKKVIESIKNDQGIITSLAQMAVGNIHADLYVEIIYREYMN
STNSEKRLCLLYVIDMIVKIDINRIYVNLFIHPHITSIFIETYNANETIRLALWQLYLTW
TNVFPAQTLNQIRTSMTISQGYRNSPAFQVRSTQPPQSQPTIIPVSYDSQGIIPISVNT
QVEPLQNASFTPVRRYKQPKQYNNSPLTTPRRTPRKQSKPSSPINVSIDMCDNSSLPSP
TDPAYKETLHPTTSPSVSTSPVQIMSPAPTSIPVDVPIQMSPMYSQTNVITEPVSTIH
QTTTPTIHQPQHVLGECIPICGGRFVDMERHVQYHQVLSACEGLOSSKPEESWNEQSFIO
QEDNYQQPSCETNSYTSQPLLCSICNKPLVVVQSQNGELLLKDGAIIDNLGQFVHQCLSN
SSTTAQEL
```

EhPfs2

```
>EhPfs2|EHI_170080
MAEEEQCNQFKKYQDYFEKQSDLMKKYDETLFDRWLHGEKGVIDQPPQHQGERRCIDFNC
CHSEMDSTYSMYCQTSRPYCHQHYSYTRYLQPPILCTSNPGINITTKFVHTSINKNKTP
VYCMRWTPDGRRIVTGNEKGEFTFWNGLQFNFACLSGAHDTAVRCMTWSHSGGHMVSGEQ
GSGNGSSSIYWNPSLAPVVTLKTAHSELIREVSFSPNDLRFASCSDDKTIGIWD FNKQVC
EIRFDESENAVYSVDWHPTESLLLSSSKGKVRIWDPRLEKEKVGMFSPHKTEINKVRWNKN
GKWFLTC SKDFKIILHDIRMFNKPLMIFEKHMKDVTIVNWHPIQEDFFVSGGANGVIYFW
DTEH SKPIGEIPVAHDGAIWDLQWHPVGHLLASCAHDQTTKFWSRDRCGDSLEIKKYQGT
NRDYDPIIPMYDDDLQPYLAPVFIPNQINPSRINGFHDI
```

EhMpe1

```
>EhMpe1|EHI_014000
MSVSFSYTARFFSRQRELVISKPNALFVDAKQKIIEDMKNCNISIEPSQLVLTVG NREVN
ELDQIIPNTVIRVSVKEIEEKKEVHKEIATTKGVYIDNVERPTTQLINTRPKKQQRQLPP
GYICNRCHKPGHYIKDCPENNNPAFDKKEAAGLPMEMYEDAPDDDP TAKTLKDGRVRRLR
DMNGQGFKPTISK NEMSEEFDIPKLLKCKKCGLLTQACFMRCCKNTLCEECLNEVRQKE
TCTMCGKQIDISKDIIIDKKTREEVQKYFDSINEKIGKHLTATNAFQMKNKNLSTKNTET
KTNPTTEENELAHQIKIIDISELFQGNRMEQRGRYHGKALSELHIDCNFNGKLIGVGED
PDEIIQKS FHPNEEEHKRSRSHSRERNQHEKKRS
```

EhSsu72

```
>EhSsu72|EHI_027340
MDIFQPKTNHFSESIIVICRSNQNRISIAHCRLKQLGYNVSSFSGTGVTILPGATRNFGE
QFGIPYKDIKNSLPKDEASVEYYRKNKIYQMLDRNANLKSAPEKFQDCKKEFNIIITLDY
SVFMDVLEYFDSRESSTGSLCYVFNLTVMDQLDTAEKGAIEVSEFLKLMESDSDWMNNID
HVIKSFYSKTKINVLHSMQMY
```

EhPC4/Sub1

```
>EhPC4/Sub1|EHI_027340
MSDKKPKHKSQDEEKEVEKTDKKKAKKEETEESEEKKVIHKDKKGIKKKAKKEQEDSSS
EESEEEKVSKKKAKKEKKEELKLPFDGDKYVQLGERKYVRLNQFRGTKYIDVREFYERDG
ELKPGQKGISLKDYEFEELVNNIDKIKKWK
```

EhPABP

```
>tr|C4LWS1|C4LWS1_ENTHI Polyadenylate-binding protein, putative
OS=Entamoeba histolytica GN=EHI_198750 PE=4 SV=1
MATQTVPQSVAFAYENENKTIFVNNLPTSVTETVLFAEVFKDFPNQVSNVHINRDSRFDT
TIAYINMNSHEAALQAIKKLNGTLIEGKKVNMFWSLRDFKQRTQNETNLYVKYIKKSVTQ
QKLQEVFEKYGHTLSVKLSSNEKNESNGFGYVYCTVEECQAAVDARDEIAKEIGDEGFL
IAKFCKQPKNLKTNLYIGNIDKTVTEEQFLKYFETFGPIKCGADGKKFYRFLYDEKYEKN
KGFVNYENTEDAQKAVTAPRHCVLGDSDILVTYYKTKKERKQEFKQHANEIKMNATGKYK
EYNFHVKTDKAGITEGEIRDALQDCGGIYSIKIKYVNSEPTDVVYICFDNAQSAHQFIQK
AQVLGWTANPFKTKQERSNANNYYNGTGIPGFPVNFLQYFQLFPSVFNTSMNRQPKPGNK
QPRNVASQNAKKKSKVDNNIKKQPKSIEKAPTVTTEEMKNNLGDQLYDFILGLEDYDEEI
TGRITGVLLESIEYSELNNMLRNKIEELKNVIKEVYENLKKINN
```

SECUENCIAS DE LAS PROTEINAS DE LA MAQUINARIA DE PROESAMIENTO DEL EXTREMO 3'UTR DE LOS RNAm DE *T. vaginalis*

TvCPSF160

>TvCPSF160|TVAG_077030

MNVTHRTICDSPAPESAIRCHLPFSKADCIVTANGNKLQVYSTKEDDLRLVWEKKFWGEI
FGVYRHKSGGEYDSIIIVGCDTSKVIVLQVIDNDLKETEYHEFNRPGPPEPDPPKPERPFD
ISTRNRNTIMDADPTGTCLALLLAQNILYVLPLANKIKIESTEKAGDEYHSSWKVIKDA
FAYDVHTDFKSPLYRIRDMVFLDGYKNPTLAIHELIPWVSVRLPLQKSTVAVSIVSPPL
KKKETVLISASIDKVTMWTSRALPHNSFGLVHVPDPIGGFLVLSKNAIYMDHTNIVALA
LNKLAYLDDEVPVDITANGPGCHELYSKVGTAKSHILLTVDQHYLSILTLYHNGVKVT
NLSLNVNLNLEFHPSCFLSLNYTNNRSLVFMGSTTHDSTLSEIIIEIEEEELASFLVDHV
MTETQNDLYQKFFKSLPRPKTSLKVKSFSVNNISYIRQLGTVANATPCLKLSQIDENEEQ
IAMALACGFKKYGLHFVRSIDPNLVSEFNIRGVTGIYSSNFYILISTDEGTSVLYA
DSQNLPEINDEFTSQVSKEKTLFAGDFNNFFIQTITVVRRAIKYSQEDGFSHTDLYSGSS
PAVLDAISSNALVVLFDNYTICAYDRALEPMTPLTNKYFSRIAIFGKYL FVYCTNGVLR
MFSVETMTEVATFDNFAFPDFIAPLGKDEIPLQPSIGIVDISFVVFESVLLVLTPRD
GPAFFYQWVESKLFRRRIKMHRFAVTSRKNIDIIPFNGISDVFDGCFVTDPSNPFFILSE
NGYPRIIPLSESFNCKPRAFAPVMRGPYKNHFLIADEDQIRLCNLENIKPEHNFFIIDGC
IVERIPVGMTVRRIAYCQNPNCVAFIASHPEPFTTENEEKIDVEVYENLQVHYQEPPSPA
KVYPDEDYETIPKWNEERYSLFLYTKDGLQQMVDYANHEIVNTVQFVHTTTPMPEDGITLL
NTYLAVGSGFLSQPEKMMRGVLYIYQIRYMQNDEGFNEITLRPLYNETNKIYKNPIIEIT
DNSGYMAIFCGNLLYLMRFFNENTVKIEAFLVGRFFASSIVSLKNYLLYADSYEGFEVAR
WRKYGKKLISMARDTMTKLPLSAAFLQYEDCLGGVVFDDDGNAHIFDVDEYAIPADAVVR
KSIFYIGGRAISSGQFPKAVTQATQQNPNEEIDEELLQLQTKIGGHIAWYVTTTHGKIGA
FTPIDENDRHKLGVQSAYEKSLCGLSHLEYRSGKFKNMIEQDIFNQSPKNVIDCDMLID
LIEDMPDHLKFATKGLRTQDFLSELRKIYNNGVNVFQ*

TvCPSF100

>TvCPSF100|TVAG_363680

MESDTTSSISFQPLSGAQSTTPFAYLLHVDEFTFLLDGWTEDFRLEDIQTQIEICSHVNA
VLLSHASIEHIGALPYLCSHGLSAPIFATMPIPALGSLLIYDSYLNIRDEEEFKEFNAND
IDQAFQKINRMTYQQSEQLDGKNITITPYNAGNTLGGTVWRIVKGQNEVIYSVSVGDHSK
YLSSFSLESGLHPTLWILDARGPESHDRGKEDEFWRQIFGKLNGGKTIIFPTDGVSGSLE
VISRLKEQWKKNWKKIYFLSHSSPAVLKNAQSLSNYLSLDIQEKINSGEYPPEFNDPD
LSYFSCVTSIKDIDFSQGCVVISSTDTLGRGFSRKLFLDKANSNLIIFTQREPPYSLAE
ALRTNNAHRTFRFIIKHREPLTGEELVKFMEKQSALQEKANEIEGDIDDESDEVSQENIE
NSSQIAQSLKKHFFQFKRKETSDLSYGANIVVENYLKGANPMAPSKMDTSKMIDSSLTQ
QNFIQELVYKPSKFMITQYDYNFVGTAFFWNLERTSDYSTIAYNVTSFNPTDIIIGAKK
ENCEELMKILKGKSPQNTRIYIPAIGEKVSLQRDLTTRKISLSRALLSGIDFVNCGVNDI
AYIEATLKADEHQFVQARPVESSAGHQATFVGTIDMSQLSSKLDLSGINNDFKAGGVLE
CGRRRVKVRLVNEKSITVEGMICPDYIKVRNAIQDLLKMI*

TvCPSF73-I

>TvCPSF73-I|TVAG_437970
MQEVQVGEDKLTVMPLGAGQEVGRSCIILKYHRKRVMLDCGIHPAYENFGGLPFIDAIDP
AKIDVLLITHFHIDHITAVPWFLTQTNFSGPCFMTHTTKTISKTLVDYVGVSGRGSEEP
NLFTRADVANVQNMITAVNYHQTVTHQGIKMTCPAGHVLGACMWLVEIDGVKVLTYGDF
SLENERHLQGAIEPKSLSGEIIIRPDVLIMESTHGLARIESRVDREYRFIDNVTKIIKRGG
RCLIPIFALGRAQELLIILDEYWESHPEYNGVPIYYGSNLAKQAIAAYNAFYQDHNSRVV
TAKGKFEFSYVKYIRDYDFDDSLPCVVLCS PAMLQNGMSRKIFEAWCSNSVNGLIIPGYI
VDGTLQPQVLKMPAEITTLSGKIIIPRKISIDYVSFSGHADFNQTSRFITELKPKRIVLIH
GVCGLMMQLKEKLLQM FVEDGLEVYTPGLCEKATMWFQSNPSAIITGGLSDNKNISGII
VRKDGQNMIMSASELSTHTTLKTLNATMSQEVPI NKPLSEYTKLLQNEFGKVQVVS KKL
LIMERLMLIYKDKETAILKWNADPTLDYLANAVTSLIMCYNDSGEKISLSDTTPEDL FVQ
KLQEALRERYSKDIEFDPDSQMFMFQISEKSVIIICIDSESPSGVKVECDKQDERLELELN
ELATTLHEFCQPLTLHQVR*

TvCPSF73-II

>TvCPSF73-II|TVAG_183200
MHSGTKKKKEVHCQLICFGAGGEVGRSSILVEIGSKKVLLDCGVNFTATDEKDRLPAYQD
PFPKVDLVLIISHIHTDHLAAPYLT EVLKCAQAPVYMTRASQMMMPIMLDDFLKV TENPPY
KAEDLTNCKPKIKVVEFYSRFEAAPGIFVQAFPA GHILGAACFFVQVRGLSFIYTGDFSA
IADHHLSGHAVPRLFPDLLITESTYGNQVRDSIAKRERSFVQMVHQVVGEGGKVLIPVFA
VGRLQEICLMLEDYWNRMGYTEPIYYTTNLGENCMKVYKQCVNWMNPTVQTNLFDNGSTA
FKFTYSRNFNPKKSKIDESRGLVMLATSGMLNPGTPAFNFFVNEKWYDDPRNMVIFPGYC
GPNTFGRAVLTRDLTTNRVQFTSRPAMTVDIIKCKVERISFSAHADQFEIISLCDRVR
PSKVVTIHGDKPSVDALATRITQTTGIPAEAPRNNAKVTTPTKNPNVISISRDCVN FALG
QPSACEGVILKDRDRRYKMISLNDYKKNYSINVYKLNKRRVKSAVSKEKVAEMLKKLSI
EYVINETESVCKVIAGEITISFLEEYLEVSFPVGQKTLANKLCCLIAQ*

TvCPSF30-I

>TvCPSF30|TVAG_047060
MWYFHNLQNQQQSL EFKFESVLKEVDQNEENDLKDYINKLSKKDRNSFHQVVCDKWVNT
CVHGDRCSLHEYDIDRMKKCFWEKYHECSNKFECIFRHELTDRIGTDCQYYIRGYCKH
GDKCKRKHTPRDAVCLNYLAGFCPDGPKCLFAHPKWIDDSGIDKNV*

TvCPSF30-II

>TvCPSF30-II|TVAG_094550
MWYFHNLQNQQQNLEFKFEPVLREVDENEERDLKNYISNLSRNEKNNFHQVVCDKWINT
CVHGDRCSLHEYDIDRMKKCFWEKFHECSNKFECIFRHELTDRIGTECKYYNSGFCKH
GDKCNRKHTPRDAICLNYLAGFCPDGPRCLFAHPKWVDDIRIEKEN*

TvFip1

```
>TvFip1|TVAG_463950
MSSEQEAEKLTDEEKPKDENPIDENSKEEGNDEYDSDEYGSSEEESESSESDDIELDQDL
NVQAQMSAIAYSTRALIQPGPLNNYEAWITQSDIVTQGSKNPWEIINPSGVHPRNNKKS I
FYDLSALPNDVESKPWTAPGADITDWFNYGFTEETWEEYRKTMLNRIKNIEYEQEIKSL
TGDTRQQVKRN*
```

TvFipS5

```
>TvFipS5|TVAG_210560
MMELFEKTKDSTKSLGDICLDITDAFSGFPMNQWFDLETKTDEYGSFKIAEIEIAENDE
EEEEYVEEDTTEFDTLGRGSI SHDSSSDSINEIRNSRLESGDSFP I IENPNNQSI VTN I
KIADNHEDEYEQYYYVEEDVTEESS I KIRIRQTDSL I ANTSYDSRQEFKITRTRRS D
SDKNKEQSKSYTTIYRMKRDKSHDVVSKIPLSVSKFDPKFPVDT S IFASYKIHRDERQE
ENLTDQDLKKMFEQIYINDLQPTTINIEDYNPDEDSQNKTLIVRTSIEKEEQEKIERERK
EKEEKERKQKEENEAKEKERKEREAKEKQKEEKERIERERKEKEEREKVEKEKKEKER
ERKQKEEKKEKEERERKEKEEKERKQKEEKKEKEERERKQKEEKKEKKEKERERKE
KEEKEKKEREKEKTEKEKKEREKERIERERKEKERKEKEEKREKEERERKEREEME
RKLKEEKEKAKEKKEREEQERKEKERIEKERREKEQKDKEEKERKEKEEREAKEKAKE
QKERERLERAKEKREKEEKIEKERERKEKEEREAREKAKEKREEREKAERERKEKEQK
EKEEREKAKEQRIEREQKEKEAREAKERAKEERERKEKEQKEKERIERERKEKEAREAK
EKEEKEKAEREIKEKEERERKQKEEKERLEREKEREKEKIELEARKKAEREQKEREKE
EKRELEEKQKEKEERERIEREEKEKAQQRIERERKEKERIEQELKEKERQEKEKQE
ERE I IEKFMKEGREKAKEKQSLKSVRRKEQERIEVVKVSKEKKSKTSRNIQVETKEEL
QNSAENHHEMLQFQINTDSISEIEMENQKESEKVARILMEIEEIESSTESSAAMKEHRM
SLKDISMIPQSKQTSSSSSDLKSKEVKIAKNVSSSSTTKEKKQKSVSIRLSSRTKSTINV
SKQVHHTSTTTISSHTAPSSRNLSNELTEEKENS PKKETNSEKKESPVKKEDTGKSESDK
IEAIARYRYILQRNADLIKQVEQSKTPNPILRLRMKASIEEIDSIMEELSK*
```

TvCstF77

```
>TvCstF77|TVAG_343030
MDPERFSYLKEVYEQIQRLPGPITDENRKLYDTFLEAFPGAFNIWIKYISEEVQISTDN
FDRVEKLFYRCLPNVPSVDLFTFYLYIRSKPNLSLEDIMAAYDYALDKVGLDIKACPIY
SEYIAFLNSHRSQIEGDLDKCRKVYHRAFFVPM DGLPILNKEYHDFENERAQQVAPSLFK
EWDGHFKTTKQAYAEKNII TQRLKTTDFETGPGLLSVLFYWRFFILKESQNEVHVKPDQ
LHAFVEYAYRKALGPLRYQWIIWHEFAQHYLSYNDAAAIRTYAEAVSILPHNLMLS FAY
AELLESRKRVSEALPVYRNIIKNSLNI SDATLSQIQLLKFLQRT EGP DAMRREFVMAFEK
GNFTYHLLIAVAEIE N AVNLDRESAKNILHAGLQIYKTD SAYIEEYVSLAIKLN DIEGLI
DAYDNVCQLHKDKRMELSWRLLDFFRLANDKSGGLSSSLSRVVKNLSTLLSDYKGQ RSLQ
TNEAPRKTLDLIKEYIPTLASYEKNINLVFEFALSTHGDDFENAVKSELGDDIFSGDLVR
RTIESLVNSHSAKDLYQAI SLFGI IDKAIVI IRNIMEKIDEI IHRIENDILAQDKTISRE
KLKLRAFFLPDGYN*
```

TvCstF64-I

```
>TvCstF64-I|TVAG_021720
MSGDEGSNHESDHQSGAEEQQKPEQKEEQAEQPQQDDKEKKPETITVHVGNL SFETTEES
LKIKFSEFGNVISSRIPVRAQSGKSKGFGFIEYATQEEADKAI AEMNKKEFEGRTIVVDI
SRGRSERRSSERRGRRDDYDRRSDDRRDDRRRRDYDDYDRRSRRRDYDDYDDRRSRRRDY
DDYDDYRGRGRDYDDRRRDSRY*
```

TvCstF64-II

```
>TvCstF64-II|TVAG_211580
MSEEGENPSPQQTKTEEPAAQKEEAPKGISVHIGNLPFEYTEAKLKEIFGKYGEVLDSKI
PTNPAGRSRGYGFVSFALKEDAQKAVDELNNKEMEGRKVEVAFAGDKKSPRSDRRDRKP
RDDRRDYDRRRADDDRYRNDRRDRDRYSRRYEDDYDRRRYDDYDRRSDDRRDYRRDDY
DRRDRRYDDYDRRYDERRYDDDRRKDDDLRRDYGRDERY*
```

TvCstF50-I

```
>TvCstF50-I|TVAG_374220
MSQLKIQKLNLRFPPLTVVLNDNGETKTENIELPPLSVLTDLHVLAEKIVSTYEFFNT
KHTKKIEAALNQILDMQTIIDYSDFQLQHMHQMHAMPLSDCSFNKMGDMFVTSSYDHTACV
WNTETGEQIAKFVGHTLPVYSCCFNNPYGTLVATGSFDKTA AIWSVAENKLVYRLGGQSS
TDGHTKEVVTVAFDPDSQYVATGSMSDKAKIWDVQTGQLLSLEEHTGEIVSVQFHPSEP
LLLTSFSDKTARLWDIRTGDCISALRGHKRETCAAYFNSAGTNIVTGS LDSTVRVWDVRQ
ALAIHVLKGHTSEVVAVAYSLDGSKVASSIDKTARVWSTTTGECIAICEGHTDEVGKVT
FNPQGTQVLTASDDFTCRIWDVNTGRELECLSGHQDYVTTCAFNYS GDRIITASKDNTVR
EWRAMKTEH*
```

TvCstF50-II

```
>TvCstF50-II|TVAG_493960
MSWMPKSPRGSREVLKAIHVTFYPPKLTLEIEQNGEIANHELDLQSLDSTTNLET LADQI
IAEESSLKPKHRAKLIQTLESMVKMQGIEYNTYVLARI IKAHEMPLTDCCFNKMGGFLFAT
ASNDKTCRIWDTLSGKELHCLMGHKQTVNVCCFNNPIGDL LGTGSADKTSKIWKVGTGEC
LYTFTGHTDQVINIKFENETKTFASSSSDHTVRLYDLTTGQFHTELLGHTDIVPHIEFSN
DGEIILTGSFDNTVKLWDIRSGSEISSLNGHTDDIFAAHFD FPCNKVLSASQDGTALIWD
LRTNQAI AIMDGHGGGCTDACWSADGNYIATGAGDAVARIWDIDGKKLFDCRGHTGEINR
VMFSPQSTRLLTASVDNSCKLWDTSTGLCVDTLKSHTNEI VCAAFNYS GKYIITASQDNT
VIQWEASITDHRC*
```

TvRna15p-I

```
>TvRna15p-I||TVAG_267990
MHNPNQNPPIAQPRARAAFRGPRPQSESFQRTPASGRRPRSQVITPSTNVFINYIPRFT
EQDLRLNCSQYGEIISSKIMINLETGQSKCFGVVKFRELSQAHA AIQAIDGMSIGNKRLL
AKYAESQEKHEKLSTMLYIKHLPLSIDVNGVYQIFSRFGEITQVSPHSVDSTDPQYWRCF
VRYQTQEAATEAMSMNNQIIIVEGSRPIHVKYADEARLSGSFVLPPTGSTPSAPTEDEQRK
LLPSFFFN*
```

TvRna15p-II

```
>TvRna15p-II|TVAG_023270
MSEDRNLAAAPRRRINRWERPKEAFVKDRIQTKKTRSQTIEPSTNVFINYIPPNYNQDD
LKNLCSQYGEIVCCKIMINLETGQSRCFGFVRFKELSQAAAAIRGIDGMTIGNKKLLAKY
AESHERQEIIASSMLYIKKLPSVSIEDVRRMFGVYGRIIQCVQHSVTSLDNESWRCFIQY
YSQEEATNAMHFMMNKIVVEGTKPIHVRYADETRLSGNYILPPPVMTEDLPQIDQHQLLP
SFFFN*
```

TvCFIm25

```
>TvCFIm25|TVAG_495750
MNRIPSIGLGTWKSPPRSDDLVAIVYAVKEAGYVHIDCAAVYGNEDIVGQAFKKLWDMGI
KREQLWITSKL*
```

TvCLPS3

```
>TvCLPS3|TVAG_162880
MSIVDIHPNYDLRFRLSLGDVVKLKVVEGVVWLFGWQIPILETETLFRDESFTLSSFPVPSK
VMLIESAVPYYSAMSQIPKIIERQLQKEVIIIGGSKSSGKTTLARWILNTRFIQGVKTFY
VNADTYQAPFATNGCIGLAYIDRPIDNNGFQLIDPILLYFGHSDIKQSRKTLYEGQLLQI
RDNILHKLEALENPQDVCIVIDYPTIREPFMGESLQEVMMMFGATHFVTLGDDALGELAP
KTSDIIVQKLPVLPALYINETIKSRLVNKAIREYFYGTPECPLKPITKTIQRSEFEAYS
LGTNLNISKIMLPSGADSPDPKVLNKIGITQMLNKRILAVIPQVPRPDVWKMNIIVGFYCL
ISVDEMKEEITILVPSEVPPPESYLVAGTIEFNEV*
```

TvClp1p

```
>Clp1p|TVAG_317830
MQTTTTTTTDPPIRKRFDSDSEARFLLTPGQSITLSVDSGSVFYNGWQLPETGQVVFRDKDF
TLMSYKDSYTLKFMGKLPRFYAAKRIPQESVPAYDLISQKSVVAVCGPPLSGKTTACQTI
INSALHGVFPQHPKEKPIPIYVNLDPAQAPFCPPGSIGALPITKPIDNTGFKIENPLTYFF
GHTEIDEAHRDRYTDLVHELATYVKQRRQAIQSDDGGVVIDLPSPTNEHVLPGLAQAIKD
FEVTHIVCIGDDRLVATFQRSFPWISVWGMPALGAAHDDEISIRTAQRS�DTKRYFDGDD
SADLVSMYTRFTNRDDFKLYKVDKREPEEVPISTSLMASVVAIVQRPVSNELWKQNVLG
YLTITKVQNEDNIEVLKPKQPLPNNVQFIIGSIKYFIK*
```

TvPcf11

```
>TvPcf11|TVAG_100880
MSYNDRSYNGRNDGPGGYQDYNQNGNKYDYNQGNRYNDGHNGNFNGPNGNQGNYNQGGGR
FNGPNGPNRNYNQGGRYPGDYNQGGRGYNGPRGGPGNYGPNGGYNDNNYNGPRGGPGYG
GPNGPNSGYNDNGPRGGPGNYSPPNGGYNDNNYNGPRGGPNGGFNGPNGGGYNGPNGPRND
PNDKGRRNDNFNGNGNFNDRGGYGNDRNGYGNNDNRNGYNNDRNNHGNQNDRRNYSDDRQ
NFDGRGPRGGNQQIHDEPVDAPTTEELDEV RDILKGSNVTDTHTSTVSRNVQDEPW*
```

TvPAP1

```
>TvPAP1|TVAG_388620
MLPQYEPQKPKQTEEGPTGAELQASIELQRFIDKYDAPATEEQKTKRQEVLSKIKILINE
FVRIIAKKTGNTAEGLENVESQLVPYGSYRLGVVSEGSIDDCIVVAPSFVKRSDFNTFY
EMMINHPSVKELVKVENAYSPIMSFHYDIEIDLSFASLAMTTVAEDTDFADDHILDNVD
EESVRSNLNRRVNDMLLQLVPNQDIFRVVLRFLRIWAKKRCVYGNVFGYLGGVNLAILAA
FICQRYPKKSPAYIIVMFFNDLATWTWPEPIYINTPTTGNRSNWDPTRQQDHKDCMPIIT
PAYPTINSLHSATKSSRNRMTEVEFKRGFKVASACLKSEASWERIIDSPKFFVRHRKFVEV
LVVATEEKVYTEFMKFVQSRVFLAINIETQLRHVSHAYIWPVEFPANGYRGNEFCSVFY
IGLSYHVPKDENQSRTLDTSPAKKWVDDIRQKAPPNADVILDVVDTKDLPSFLFPNGER
PEVKNVKKVEQPAK*
```

TvPAP2

```
>TvPAP2|TVAG_107950
MIDENIRKLESGEAFSKLDSEVIDIIKRTVTEILPKNVIVRIAGGWVRDKILGKENDDID
IAISGATCTEFSQTMVKVLGPNTKVVILEANHEQSKHLNNTARVCLFRDFWLDICGLRSES
YSEGETSEGTAQSDARHRDFSMNALFFNLQTMKIEDYVKGVS IKDKVIRTPISAIEDFT
DDPNRIVRCFRFVAKYDFQIDQEILDAIPKIIPLFLHNITKDRIATELTKMLHAKSAFFA
MQHLISSGMLSPIFDESLSLWLDMEKTLQVRVRVHQIQPDLTGNDSMITYFAAVLLDLEE
KKAVPDPEKKNKMPAITYQIVRSMRLQNALADNVKKILHGKLMVEKIFKDGMRNVNVGR
FVMSTGPLFYLVRLQLQTDLLDFYDNKLKPFIEEQNLSKSYEMKQILDGVTLASVHGK
PGPGLKDLIDQ MIDWQFANPNGTSDDYKKYIAEKKGH*
```

TvPC4

```
>TvPC4|TVAG_255210
MAEKGLDSYSVQEKKETS YVFPLSDDEKAKKRIQVHKFKGKILFDIRELYCKDDEWLPGK
KGISLRVEDFKKLKELMPLVEEAIVALGGNLDDDDSE*
```

TvSimplequina

```
>TvSimplequina|TVAG_494090
MTQDTNLGYDVKSLMAKFEKEASSAVNWKNTATILTNLGKIVTTLPEKAIDESNFLQMF
AKFSEFLESHAEEKITLQTKILIVKILQNSAIIMTKKQFITQADVAISIISDQFENECKKE
TPNVGYLITLLDAFFHFIRFQIGDNKEYVETINNLNNVVPKLSSSKKTSRLRYALQLDITK
NDSILAWEKDLRKLKLTSLIEIAITTTTRNTNAIPRTPRDNETRLKVQNTLLTYQPDIVLAL
VIEGFKKLRLSPNFKLSETVNTTIQEIMSIQEDVDYDMHEMNIEITSEYIVEQGLEKMAQ
RMKFTSQKVQSHFLAMICTTDQSEGEKYLNCTGRCTETAIRCASTDISSLLPFLINWVC
YEYVKS PRERFVRITEKVAVAICGKLSEKDAPQESSILVFLRSLPAIDPSLFSVLADQVN
RSPNTSRIIIFAMANVVQNTILIQKDSLNSIFELCISPDQLIRKNAIEIMINSFYAGRLF
VTEIEDYSKDKLKQGSLEEDQNSAERYLHIFFEVLKQNVNLLLNLLEMYDKMNEKVQSSV
RTHLLTQMPQIGFN YDIISGLFAAQNAENVKLIHFILDNLSKINTIPSQITVLIKSEFEK
TNDGRFLIPIIQTLTETEFFNYIPIFLT LRKSGRNRAFEIYIMQKLKNQAKEPRVKLISE
LLKPIEAKPYKSQIAALNYCMSKKDVIDYTTIAAAVNKTMTSSKKNLDLICEPLLFIVDE
YSENSKHVEKICQDLIKMNVHFNEFAWPKMITLFEKMMNMFKIYNLPLEKIEEVAKKCPRL
LTFMKNDSNKKKVPKKVLQLLETL*
```

TvFY

>TvFY|TVAG_121040
MSENFEEKNLIEMINYGGDEEPQERPQNNMRQPDQNYDGRRFRKAI SRRWCDHRQAFELY
TQARPYNEYDVFPQNVSIGYRNLLPPFCYDTSGISICAKFIAHKNNKFRASKALAWQPD
SLRVVVGDDQGNITMFNGLAFFYEFHIPASNTGIHDLIWITHSGDYLLTVDAMKQQQSPT
IKIWQANYNNVEKRSINCTKDIHQITISPTDRKFALC SSDYTVKVWDLATLTEEMVFTGH
GEAVRTVDWHPSQSLICSGGKDNCLKFFDPREDKPINTLTVHRNVVTRVHWNPNGNLLAS
CGRDSMVYITDVRTMSP IAECKGHENDVFAVRWHPIREDLLVSAGYKGDILWTTGSNEP
IYSRMKGHNKAVFQLKFNPMTILASTGTGEGVVKFWVRNMC GTDLLGSYSEANKKRDVEK
KRDVEKSDIPGLQRYNPLHGLPDDVEEDEESVVP TTSRNYVTPSQFHQVQSQPNYDDDED
EDDYAQGGVADDDNND*
*

TvSsu72p

>TvSsu72p|TVAG_229140
MTSNNQLAFACVCASNVRSMEGHRLKEAGFNVSSYGTNDYVKMPGASNVPLQFDFGAT
YKEILSKMESENNQYYEEQGLIKMLRDDANTKEKPERWSSTFN PATLKYFDVIFTYDNNV
MDRVLNEFRENGNRTFQISHVVNIQTPDSRDNAILSRQYTLRLAQMLSAATDLTENLESI
VDQFN MENGNPKDSFVASFHIVSY*
*

TvHrp1p

>TvHrp1p|TVAG_245890
MAIRFHDQNKDCTLFLCDISEHVTEEILTELFMQVGPVVFVNI PRDRITNRMNGYAFVEF
RTEQDCMYALSVMQGVKLFVPLKLSANSTPSTGDEL DVGAKLYIGNLSQDVNDGNLLQT
FRQFGNVLHARVVVDPATGKSLGHGFVAYDSFDAADKAKKAMNGEYFGGPITVSYAYKS
GTKSGEKHGDRSERSVAPNAQIATQLKKIKRAVNAIPDVG*
*

TvMpe1

>TvMpe1|TVAG_265310
MRSRIEYCFCSKKKSSYVTINQSTVALKQLINQIKEQEQLADQDNVTLVVKNRQTGEVYK
NEYEFINRNTSVIVYRVPLTSKIYLET KPVAQPAEQAAPQDQTAE EQAAKLAAEAQAAA
EEVPDENSLNKLEARKQEDEFYKQNI PRSYILRPLQGKIILPYNEVAYFQRPEDPPFA
RNRWTLLEDLQRWMCPI CNRIMTHATLTPCCKVAYCDSCITEALLNEDQRCPNKDCRRTGI
VPSMLTPATQLRNEI KEWQKFDVNKLKTTEEFNMSEQNRDEQHERSVHHEEVQEFKIPD
KPATGEQRHRDDTPRTRDDRPRSRDDRERSRDDPVEPRNRDERVTRDDRVRTRDDERV
RTRDDERVTRRDD EMVTRTS DERVTRDEDERVTRDDERVTRDEDRVRSRDEDDKPRS
RNDDRDRDRDRDRDRDRDRDRDRDRERRRDDRRERRRDDDDDRDRDRDRDRDRGRDR
HRDDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRGRDRDRDRGRNRDRRK*
*

TvSw2p

```
>TvSw2p|TVAG_16940
MSAPSSRDRKSNPSREEAALRSYFPAMKLLPDKQFHSDVAFRQTGEHFAVISENNMISLF
DVATCNLVKSIIPNRKYQC�KLLFHPDSVRIILNSFNDNDHKARLLNVETDGFDLYYTGPN
GQITSMATTDACLFVSEDERLHVFDVKTAEVITKIDCPRESQVATHPNGKCMVLANTSE
LSLYDIRNMDQSVCSIHFS PGCKVAPHFGEIGNLLITGSKYAKLHSLRDL SVIRTFAQL
EDEFFTTGACFTADEKYALVGSSDNSILAYDVPHSRKDTPLYEKT VLTGHISPVKGIAVS
NRYFNMVVSFGEDCLFWSVDIDSFNSLGYKE*
```

TvGlc7p1

```
>TvGlc7p1|TVAG_101240
MDLDQIEQKLFVTINKPLGTPSGITVEEINYL CQTVKEIYLSQPILLELKAPLTIVGDVH
GQFHD LIRIFESKAYPPNANFLFLGDYVDRGKQSIETVCLLF IYKIKYPDHFFMLRGNHE
CTYINRLYGFYDECISNFDIDVYTTTFSEVFNCLPIAAI IEDKIFCVHGGISPENSLDQI
RDVQRPFEIPEEGLIADLLWSDPSADATDWENNERGVSWCFGPQQVEDFLTKFGFDLVCR
AHQAIMDGFEFPPFESQSI VTVFSVPNYCNEYENKGAIMHVEENLFCNFTVLEPKKYDDD
YEFEIRAGSPPRGSTCERNYNQENVVVS DV*
```

TvGlc7p2

```
>TvGlc7p2|TVAG_379450
MLDEIEQKLLALKNQPMGHIAGLTETEIQFLCKNVSDIFMEQPC LLELKAPLTIVGDIHG
QYQDLLRVFDSTGYPPNSNYLFLGDYVDRGQNSIENVCLLFCYKIKYPNTFYMLRGNHEC
SYINRLYGFYDECIQNYSTDIWRLFCDFVNRLPIAAI IEEKIFCVHGGLSPSLNSLDEIK
NIERPIEVPEEGLLCDLLWSDPNMDVDDWEDNERGTSYCFGAQQVEDFLTKFGFDLVCR
HQAIMEGYEFPFPDRQSI VTVFSAPNYCNEYMNKGAILSVDQLFCSFTILNPMKYESDF
NLQPRPGTTPPRATTGEQESEFTVATNE*
```

TvRrbp6

```
>TvRrbp6|TVAG_480160
MFLLLFVALRATEIRANSKDPKPIDALNK PAPSTGNGKEIQAPLTKSNTTKVGNTNATLP
TNTNPNPLPTKNATVNKNVNI PGKNLTQKLNKTAAIPSEGVAKVNTTAPAVAKPINSSKV
PQINQTVKPKVVVPPN PALPAMKEVVNTTTK IANTTAKPQLKTGNQTAKVNSTNASTTVP
SGEKPKTSPISRNTVEKKDNETKKKDNETKKKDNETKKKDNETEKKDNETEKKDNETKKK
DNETEKKDNETEKKDNETEKKDNETEKKDNETEKKDNETKKKDNETEKKDNETEKKDNET
EKKDNETEKKDNETEKKDNETKKKDNETEKKDNETEKKDNETEKKDNETEKKDNETEKKD
NETKKKDNETKKKDNETEKKDNETKKKDNETEKKDNETEKKDNETEKKDNETEKKDNETE
KKDNETEKKDNETKKKDNETKKKDNETEKKDNETEKKDNETEKKDNETEKKDNETKKKDN
ETEKKDNETEKKDNETEKKDNETKKKDNETEKKDNETEKKDNETEKKDNETEKKDNETEK
KDNETKKKDNETKKKDNETEKKDNETEKKDNETEKKDNETEKKDNETEKKDNETEKKDNE
TEKKDNETKKKDNETKKKDNETKKKDNETEKKSEEKDGEKKGETSNTPIKANQT TKTN
TTAKPVLKPGLNQTTKVNTS IPEKKEENKEEEEKEIQSSEVSSESESELDSESPIGSESSC
SSEMSEESAISAESIPEPSISASSESSSSSSSSSSSESSEMSEVEEDKDTEPKPTPSPE
KKDEKEDDDKKEKKKEEEEKKEEKKEEDDKKDEKKDKDDDEKEEEEKKKEKEDKKDDDEKD
DDEKIEEKKEEKKEEKKEDKPKKEEKKEEKKEDKPKKEKKEDKPKKEEKKEEKKEDKPKK
EKKEDKPKKEEKKEEKNEEKKEEKKEDKPKKEEKKEKKDDDEDEEDAKQQNNPANKGKNE
IVTESYISSKGFFFLVSIIVGVVAVAGAAYYFIRMRRRPDEQQEEPLLITDEFY*
```

TvPABP

>TvPABP|TVAG_385330

MATLTTQIFVGDLPKIDIDESFLKQFFAESGEIIPNSGIVLKKHKTLDSSFAFVTFATHEQ
AAKAMKNLNYTKLDGRPIRILWCDTETKAIKSGRGS�FINGLDESIEVSQ LHDAFSNFG
EIIISCKIPLTDGKSRGYGYIQFRNPENAEKAKLELADASINGKPIKIDNYNRKSRNPDE
DFTNVYIKPLPVDQVKTDDEDLRKIFEPFGEIQNPSLKKDENGNSKGFGFCNFKLHEDAVK
AVEGLNGQEMFGVTLQVNRLMSKREKELYNIKKHTEKVEKFAEETNGRNLYVKSFSKDVT
DKEFQEYFSKFGEIELFKIERVPETKESKGFGFVLYKSKEDAQNAIEMAMLERLHGDLPY
VGFFQTKAAHERVKVKNTHPRPSAVPAIAGNLMENSMTTIARQYQPVEDNSPKGQLKQLL
QEKGVLTNLKRLSSVSAEQAVNLVKEQSKFEIWYSHNAN*

El jurado designado por la Universidad Autónoma de la Ciudad de México aprobó esta tesis el día 31 de Julio del 2012, en la Ciudad de México D.F para obtener el Grado de Maestro en Ciencias Genómicas, al M.C. Miguel Ángel Del Moral Stevenel.

Dra . María Elizbeth Álvarez Sánchez_____

Dra. Mavil López Casamichana_____

Dr. César López Camarillo_____

Dr. Máximo Berto Benítez Martínez_____