

**UNIVERSIDAD DE GRANADA**

**Instituto de Biotecnología**

**Programa Oficial de Doctorado en  
Biología Fundamental y de Sistemas**



**Estudio de la familia MASP (*Mucin Associated  
Surface Proteins*) de *Trypanosoma cruzi* y de su  
variabilidad interclonal.**

Tesis doctoral

**Víctor Seco Hidalgo**

2016

Editor: Universidad de Granada. Tesis Doctorales  
Autor: Víctor Seco Hidalgo  
ISBN: 978-84-9125-867-4  
URI: <http://hdl.handle.net/10481/43692>



**Estudio de la familia MASP (*Mucin Associated Surface  
Proteins*) de *Trypanosoma cruzi* y de su variabilidad  
interclonal.**

Memoria presentada

por


**Víctor Seco Hidalgo**

para optar al grado de Doctor en Biología

VºBº de los directores de Tesis

**Dr. Antonio Osuna**

Director

Handwritten signature of Dr. Antonio Osuna in blue ink, featuring a stylized 'A' and 'O'.

**Dr. Luis Miguel de Pablos**

Co-director

Handwritten signature of Dr. Luis Miguel de Pablos in blue ink, featuring a stylized 'L' and 'M'.

Granada, 2016



Parte de los resultados de esta tesis doctoral han sido publicados en:

- Seco-Hidalgo V, Osuna A, Pablos LM. 2015. To bet or not to bet: Deciphering cell to cell variation in protozoan infections. Trends Parasitol .
- Seco-Hidalgo V, De Pablos LM, Osuna A. 2015a. Transcriptional and phenotypical heterogeneity of trypanosoma cruzi cell populations. Open Biol 5(12):10.1098/rsob.150190.

Y presentados en los congresos:

- Curso Universidad Menéndez Pelayo, Santander Agosto 2015.
- XII Congreso centroamericano y del Caribe de Parasitología y Medicina tropical. Santo Domingo (República Dominicana), 2015.
- Congreso Internacional de Parasitología ICOPA XIII Ciudad Mexico, Agosto 2014.
- II Workshop de jóvenes biotecnólogos. Granada (España), 2014.
- I Workshop de jóvenes biotecnólogos. Granada (España), 2013.
- XVIII Congreso de la Sociedad Española de Parasitología. Canarias (España), 2013.
- BSPs (British Society for Parasitology) 50th Spring Meeting. Glasgow (Escocia), 2012.



El doctorando Víctor Seco Hidalgo y los directores de la tesis Antonio Osuna y Luis Miguel de Pablos garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Granada, 07 de enero de 2016

Director/es de la Tesis

Handwritten signature in blue ink, appearing to be 'Antonio Osuna'.

Fdo.: Antonio Osuna

Handwritten signature in blue ink, appearing to be 'Luis Miguel de Pablos'.

Fdo.: Luis Miguel de Pablos

Doctorando

Handwritten signature in blue ink, appearing to be 'Víctor Seco Hidalgo'.

Fdo.: Víctor Seco Hidalgo





*“Cualquiera que nunca haya cometido un error  
no ha intentado nunca nada nuevo”.*

*Albert Einstein*

*“Los muros están ahí por una razón...  
están ahí para darnos la posibilidad  
de demostrar cuanto deseamos algo.  
Porque los muros están ahí  
para detener a la gente  
que no quiere algo suficientemente...”*

*Randy Pausch*



# ÍNDICES



# ÍNDICE GENERAL

<b>ÍNDICES .....</b>	<b>I</b>
<b>ÍNDICE GENERAL .....</b>	<b>iii</b>
<b>ÍNDICE DE FIGURAS .....</b>	<b>vii</b>
<b>ÍNDICE DE TABLAS .....</b>	<b>x</b>
<b>GLOSARIO DE ACRÓNIMOS Y ABREVIATURAS .....</b>	<b>XI</b>
<b>RESUMEN.....</b>	<b>XIII</b>
<b>I. INTRODUCCIÓN GENERAL.....</b>	<b>1</b>
1. Enfermedad de Chagas .....	3
1.1 Aspectos epidemiológicos.....	3
1.2 Historia y descubrimiento .....	5
1.3 Modos de transmisión.....	6
1.4 Manifestaciones clínicas .....	8
1.5 Diagnóstico .....	9
1.6 Tratamiento.....	11
1.7 Programas de control y prevención de la enfermedad .....	13
2. <i>Trypanosoma cruzi</i> .....	14
2.1 Ciclo de vida .....	14
2.2 Biología y estructura celular .....	17
2.3 Genómica y control de la expresión génica .....	20
2.3.1 Genoma .....	20
2.3.1.a Genómica comparativa .....	23
2.3.2 Control de la expresión génica en tripanosomátidos .....	24
2.3.2.a Transcriptoma.....	27
2.3.3 Pseudogenes .....	27
2.3.3.a Aspectos generales.....	27
2.3.3.b Expresión y funcionalidad.....	30
2.3.3.c Pseudogenes en <i>T. cruzi</i> .....	32
2.3.4 Retrotransposones .....	34

2.4 Filogenia y diversidad genética .....	37
2.5 Bases moleculares y mecanismos de entrada en la célula hospedadora .....	41
2.5.1 Aspectos generales.....	41
2.5.2 Modos de internalización: Movilización de Calcio y lisosomas y despolimerización de la actina .....	42
2.6 Proteínas de superficie e invasión celular .....	47
2.6.1 Aspectos generales.....	47
2.6.2 “Mucin-associated surface proteins” (MASPs) .....	49
2.7 Poblaciones celulares: mosaicos de variaciones entre células .....	51
2.7.1 Variabilidad clonal en poblaciones de células cancerígenas y células madre .....	52
2.7.2 Heterogeneidad celular en poblaciones de protozoos parasitarios.....	54
<b>II. OBJETIVOS/OBJECTIVES.....</b>	<b>57</b>
<b>III. MATERIALES Y MÉTODOS.....</b>	<b>63</b>
1. Cultivo y manejo de <i>T. cruzi</i> .....	65
1.1 Mantenimiento del cultivo de <i>T. cruzi</i> .....	65
1.2 Criopreservación de <i>T. cruzi</i> .....	65
1.3 Obtención <i>in vitro</i> de formas tripomastigotes metacíclicas .....	66
1.4 Curvas de crecimiento y metaciclogénesis espontánea.....	67
2. Cultivo y manejo de células Vero .....	67
2.1 Mantenimiento del cultivo de células Vero .....	68
2.2 Criopreservación de células Vero.....	69
3. Infección de células Vero con tripomastigotes metacíclicos y obtención de tripomastigotes sanguíneos.....	69
4. Clonación de la cepa Pan4 de <i>T. cruzi</i> .....	70
5. Cultivo y manejo de <i>Escherichia coli</i> .....	71
5.1 Cultivo de <i>E. coli</i> .....	71
5.2 Preparación de células competentes .....	72
5.3 Criopreservación de <i>E. coli</i> .....	72
6. Técnicas de Biología Molecular .....	73
6.1 Extracción de ADN genómico de <i>T. cruzi</i> .....	73
6.2 Extracción de ARN total de <i>T. cruzi</i> .....	73

6.3 Extracción de ARNm de <i>T. cruzi</i> .....	74
6.4 Cuantificación de ADN y ARN .....	75
6.5 Retrotranscripción del ARNm .....	75
6.6 Análisis del tiempo de vida del ARNm en <i>T. cruzi</i> .....	76
6.7 Electroforesis de ADN en geles de agarosa .....	76
6.8 Purificación productos de PCR desde geles de agarosa.....	77
6.9 Elaboración de librería de expresión MASP ( <i>3-step MASP library</i> ) .....	77
6.9.1 Amplificación de (pseudo)genes MASP (3- step nested PCR) .....	78
6.9.2 Purificación de los productos de PCR.....	80
6.9.3 Clonación en pGEM®-T Easy .....	80
6.9.3.a Reacción de ligación.....	80
6.9.3.b Transformación de bacterias competentes.....	80
6.9.3.c Rastreo de clones transformantes .....	81
6.9.3.d Purificación del ADN plasmídico.....	82
6.9.3.e Secuenciación de ADN.....	83
6.9.4 Anotación de genes MASP .....	83
6.10 Análisis de los niveles de expresión MASP.....	83
6.10.1 RT-qPCR.....	83
6.10.2 Electroforesis capilar.....	85
6.10.3 Northern blot .....	85
6.10.3.a Obtención de las sondas.....	85
6.10.3.b Electroforesis, hibridación y cuantificación de la señal.....	86
7. Técnicas de análisis proteico e inmunológico.....	88
7.1 Técnicas de análisis proteico.....	88
7.1.1 Electroforesis de proteínas. SDS-PAGE .....	88
7.1.1.a Obtención de homogenados proteicos .....	88
7.1.1.b Cuantificación de la concentración de proteínas .....	89
7.1.1.c Elaboración de geles y condiciones de electroforesis.....	89
7.1.1.d Tinción de los geles.....	90
7.2 Técnicas inmunológicas.....	91
7.2.1 Diseño de péptidos inmunogénicos.....	91
7.2.2 Acoplamiento del péptido antigénico al carrier.....	91
7.2.3 Inmunización y obtención de sueros policlonales .....	92
7.2.4 Titulación de los sueros policlonales (ELISA).....	92
7.2.4 Western blots.....	93



8. <i>Hierarchical Clustering</i> y estimación de las distancias entre la librería MASP y diversas cepas de <i>T. cruzi</i> .....	94
9. Búsqueda de sitios de inserción TcTREZO en la librería MASP.....	95
10. Resonancia magnética nuclear de protones (H NMR) .....	96
11. Análisis estadístico .....	97
<b>IV. RESULTS.....</b>	<b>99</b>
1. MASP expression library .....	101
2. Search of TcTREZO specific insertion sites in the 3-step MASP library .....	112
3. 3-step MASP library hierarchical cluster analysis.....	114
4. Analysis of the MASP clonally variant expression .....	122
4.1 RT-qPCR .....	123
4.2 Capillary electrophoresis .....	123
4.3 Northern blot .....	129
4.4. MASP RNA turnover assay .....	130
4.5 Western blot.....	132
5. Analysis of the phenotypic variations of clonally isogenic populations of cells.....	134
5.1 Growth curve analysis .....	135
5.2 Spontaneous metacyclogenesis analysis.....	136
5.3 Analysis of the phenotypical clonal variation by H NMR.....	137
<b>V. DISCUSIÓN.....</b>	<b>141</b>
<b>VI. CONCLUSIONES/CONCLUSIONS.....</b>	<b>169</b>
<b>VII. BIBLIOGRAFIA.....</b>	<b>175</b>

## ÍNDICE DE FIGURAS

Figura 1.1. Índices de prevalencia de la enfermedad de Chagas en América Latina .....	3
Figura 1.2. Flujo migratorio desde países endémicos para la enfermedad de Chagas. ....	4
Figura 1.3. Representación esquemática del ciclo de vida de <i>T. cruzi</i> .....	16
Figura 1.4. Diferentes vistas del kinetoplasto y de la organización general de las formas de <i>T. cruzi</i> .....	19
Figura 1.5. Diagrama representativo de las principales estructuras y orgánulos encontrados en la forma epimastigote de <i>T. cruzi</i> .....	20
Figura 1.6. Distribución de genes ortólogos (COGs) entre los tres kinetoplastidos .....	21
Figura 1.7. Representación esquemática del proceso de maduración del ARNm por trans-splicing y poliadenilación típico de los tripanosomátidos .....	26
Figura 1.8. Mecanismos para la formación de pseudogenes.....	30
Figura 1.9. Representación esquemática y localización de un retrotransposon TcTREZO. ....	36
Figura 1.10. Comparación de los modelos de (A) Doble Hibridación y (B) Árbol Ancestro para explicar los mecanismos de intercambio genético durante la evolución clonal de <i>T. cruzi</i> .....	39
Figura 1.11. Distribución geográfica aproximada de las diferentes DTUs de <i>T. cruzi</i> , con los dos ciclos de transmisión, selvático y doméstico .....	40
Figura 1.12. Modelo integrado de la invasión de <i>T. cruzi</i> en células no fagocíticas.....	44
Figura 1.13. Evolución de un proceso cancerígeno desde la singularidad de una única célula.....	53
Figura 3.1. Representación de un transcrito MASP estándar indicando la posición de los cebadores que se han empleado para la amplificación de MASP mediante 3-step nested PCR .....	79
Figure 4.1. Construction of the MASP expression library from trypomastigotes of the Pan4 strain by 3-step nested PCR: agarose electrophoresis gels (1%) for RNA, cDNA (quality controls) and for the 3-step nested PCR for MASP amplification .....	101
Figure 4.2. Agarose electrophoresis gels (1%) with the different MASP members of the 3-step MASP library .....	106
Figure 4.3. SNPs analysis by multiple alignment among different MASP sequences from the 3-step MASP library with lengths of 432 bp (A), 638 bp and 639 bp (B), 648 bp (C) and 651 bp (D).....	111

Figure 4.4. Multiple alignment for the TcTREZO specific insertion sites respect with the 3-step MASP library.....	113
Figure 4.5. Hierarchical cluster analyses based on the 3-step MASP library according to the pairwise distances with the ortholog sequences in CL Brener, Dm28c, Sylvio X10/1 and Marinkellei genomes. ....	115
Figure 4.6. SNP's analysis by Clustal multiple alignment of MASP expressed sequences in different libraries and strains of <i>T. cruzi</i> .....	119
Figure 4.7. Boxplots representing the pairwise distances between the 3-step MASP library and the ortholog sequences of CL Brener, Dm28c, Sylvio X10/1 and Marinkellei. ....	121
Figure 4.8. Representation of the cloning process with three passages of the trypomastigotes through VERO cells and further cloning procedure to obtain the final clonal independent cell lines from the parental line. ....	122
Figure 4.9. Quantification of the MASP clonal expression by qRT-PCR in the epimastigotes forms of all the clones and the parental strain .....	123
Figure 4.10. 3-step nested PCR for clonal MASP expression analysis in the epimastigote form. Agarose electrophoresis gels (1%) for RNA, cDNA (quality controls) and for the 3-step nested PCR for MASP amplification .....	125
Figure 4.11. Analysis of MASP clonal expression resolved by capillary electrophoresis.....	128
Figure 4.12 Northern blot analysis of total RNA extracts from the epimastigotes forms of all the clones and the parental strain hybridized with radiolabelled <i>masp19</i> conserved probe (A) and <i>masp20</i> specific probe (B). ....	130
Figure 4.13. Northern blot analysis of total RNA extracted from epimastigote and trypomastigote forms of Pan4. ....	131
Figure 4.14. Clonally variant MASP expression at protein level .....	133
Figure 4.15. Protein extracts from the isogenic parasite populations electrophoresed and stained with Coomassie brilliant blue.....	134
Figure 4.16. Analysis of the phenotypical clonal variation in the growth curves of epimastigote cell lines. ....	135
Figure 4.17. Analysis of phenotypical clonal variations in spontaneous metacyclogenesis yields at day 9, 10 and 11 of the culture.....	136
Figure 4.18. Full view of the <sup>1</sup> H NMR spectrums for the different clonal lines of epimastigotes.....	137

Figure 4.19. Full area under the peak obtained from the H NMR spectrums for acetic acid produced and excreted during the epimastigote cultures of the isogenic lines.....	138
Figure 4.20. Full area under the peak obtained from the H NMR spectrums for pyruvic acid produced and excreted during the epimastigote cultures of the isogenic lines.....	139
Figure 4.21. Full area under the peak obtained from the H NMR spectrums for L-alanine produced and excreted during the epimastigote cultures of the isogenic lines.....	140
Figura 5.1. Representación esquemática del polimorfismo en el porcentaje del genoma ocupado por las familias multigénicas TS, Mucina, DGF-1 y MASP de <i>T. cruzi</i> al comparar cepas con diferentes orígenes filogenéticos.....	144
Figura 5.2. Representación esquemática del proceso de clonación en el que se visualiza la manera en que AP o BH actuarían sobre la variabilidad clonal de <i>T. cruzi</i> .....	161
Figura 5.3. Representación esquemática de los estímulos ambientales que pueden incidir sobre la variabilidad clonal de una población de epimastigotes de <i>T. cruzi</i> durante su trayecto por el aparato digestivo del triatomino. ....	163

## ÍNDICE DE TABLAS

Tabla 1.1. Familias multigénicas de la cepa CL Brener de <i>T. cruzi</i> .....	22
Tabla 1.2. Porcentaje de pseudogenes en diferentes organismos, incluido <i>T. cruzi</i> .....	33
Tabla 1.3. Nomenclatura actual para las divisiones de <i>T. cruzi</i> .....	37
Tabla 3.1. Combinación de cebadores empleados para la realización de la 3-step <i>nested PCR</i> .....	78
Tabla 3.2. Composición de las soluciones de gel separador y gel concentrador para SDS-PAGE.....	90
Tabla 4.1. Sequencing results of the 3-step Pan4 library .....	102
Tabla 4.2. List of MASP genes from the 3-step MASP library and their respective orthologs in the different strains.....	103
Tabla 4.3. List of MASP pseudogenes from the 3-step MASP library and their respective orthologs in the different strains.....	105
Tabla 4.4. Distances obtained with MEGA6.06 among the members of the 3-step MASP library and the ortholog sequences identified in CL Brener, Dm28c, Sylvio X10/1 and Marinkellei strains of <i>T. cruzi</i> .....	120
Tabla 4.5. Percentages of identity of <i>masp19</i> probe with 23 transcripts of the 3-step MASP library .....	129





## Glosario de acrónimos y abreviaturas

---

<b>(p/v)</b>	Relación peso/volumen
<b>(v/v)</b>	Relación volumen/volumen
~	Aproximado
°C	Grados Celcius
µg	Micro gramos
µl	Micro litros
µm	Micro metro
<b>18S</b>	Molécula de ARNr de la subunidad del ribosoma de eucariotas
<b>Aa</b>	Aminoácido
<b>ADN</b>	Acido desoxirribonucleico
<b>ADNk</b>	ADN kinetoplástico
<b>AP</b>	Plasticidad adaptativa. Del inglés, <i>adaptive plasticity</i>
<b>APS</b>	Persulfato amónico
<b>ARN</b>	Acido ribonucleico
<b>ARNr</b>	ARN ribosómico
<b>ARNm</b>	ARN mensajero
<b>ARN pol II</b>	ARN polimerasa II
<b>BH</b>	Apuesta diversificada. Del inglés, <i>bet-hedging</i>
<b>CNV</b>	Variaciones en el número de copias. Del inglés, <i>copy number variation</i> .
<b>CCNV</b>	Variaciones en el número de copias cromosómicas. Del inglés, <i>chromosomal copy number variation</i>
<b>CVC</b>	Complejo vacuolar contráctil
<b>Da</b>	Daltons
<b>DNAsa</b>	Desoxirribonucleasa
<b>EDTA</b>	Ácido Etilendiaminotetraacético
<b>g</b>	Gramos
<b>FS</b>	Estrategias fijas . Del inglés, <i>fixed strategies</i> .
<b>GPI</b>	Glicosilfosfatidilinositol
<b>GFP</b>	Green Fluorescent Protein



<b>H<sub>2</sub>O<sub>2</sub></b>	Peróxido de hidrógeno
<b>HCl</b>	Ácido clorhídrico
<b>K<sub>2</sub>MnO<sub>4</sub></b>	Manganato de potasio
<b>kb</b>	Kilo bases
<b>kDa</b>	Kilo Daltons
<b>L</b>	Litro
<b>M</b>	Concentración molar
<b>mA</b>	Mili amperios
<b>mg</b>	Mili gramos
<b>ml</b>	Mili litros
<b>mm</b>	Mili metros
<b>mM</b>	Mili Molar
<b>PBS</b>	Tampón Fosfato salino
<b>PFR</b>	Bastón paraflagelar (del inglés, <i>paraflagellar rod</i> )
<b>pH</b>	Potencial de Hidrógeno
<b>PI3K</b>	Fosfatidilinositol 3- kinase
<b>ARNasa</b>	Ribonucleasa
<b>rpm</b>	Revoluciones por minuto
<b>SDS</b>	Dodecil sulfato sódico
<b>SDS-PAGE</b>	Electroforesis en gel de poliacrilamida con Dodecilsulfato sódico
<b>SNP</b>	Polimorfismo de nucleótido simple, del inglés <i>Single nucleotide polymorphisms</i>
<b>TS</b>	Familia multigénica de las Transialidasas
<b>TriTryp</b>	Parasitos kinetoplastidos: <i>T. cruzi</i> , <i>T. brucei</i> y <i>Leishmania</i> .
<b>TEMED</b>	N, N, N', N'-tetrametiletlenodiamina
<b>V</b>	Voltios
<b>Vh</b>	Voltaje de hall
<b>VSG</b>	Del inglés, <i>Variant Surface Glycoprotein</i>
<b>x g</b>	Fuerza relativa de centrifugación, del inglés <i>Relative Centrifugal Force (RCF)</i>





## **RESUMEN**



*Trypanosma cruzi*, es un importante patógeno responsable de una de las enfermedades parasitarias más importantes en humanos, la enfermedad de Chagas. Esta parasitosis, que tradicionalmente ha sido afincada en el continente Latinoamericano, es exportada a países no-endémicos a través de movimientos migratorios, alcanzando una cifra de infectados superior a los 8 millones (Organización Mundial de la Salud ; Rassi, Rassi, Marin-Neto 2010; World Health Organization 2015).

Este kinetoplástido presenta un ciclo de vida complejo, necesitando de dos hospedadores (insecto vector y hospedador mamífero) y de al menos, cuatro estados morfológicos y funcionales (uno de ellos obligatoriamente intracelular).

La heterogeneidad y diversidad genética de las poblaciones naturales de *T. cruzi* junto con la variabilidad propia del hospedador humano, hacen que la enfermedad de Chagas presente un cuadro clínico variable. Las principales patologías que se derivan de esta enfermedad son la cardiopatía y las manifestaciones mega del aparato digestivo como megacolon , megaesófago mega duodeno o mega coledoco.

En la actualidad existen dos fármacos (Nifurtimox y Benznidazol) que permiten el tratamiento de esta enfermedad, ambos ligados a una inherente toxicidad y a una eficacia limitada a los primeros compases de la infección.

Para poder avanzar en el conocimiento aplicado como por ejemplo la búsqueda de nuevos fármacos, moléculas diagnóstico o vacunas, es imprescindible entender la compleja biología básica de este organismo y los mecanismos que le permiten perpetuarse y sobrevivir tanto en el insecto vector como en el hospedador humano.

El proyecto genoma de *Trityp* reveló una de las características más distintivas de *Trypanosoma cruzi* y es que, en comparación con el resto de kinetoplástidos, este parásito presenta una gran expansión de genes codificantes para proteínas de superficie en su genoma (El-Sayed et al. 2005a). Entre todas ellas, la familia de genes “mucin-associated surface proteins” (MASPs), específica de *T. cruzi* y fuera, por tanto, de la zona de sintenia con *Leishmania spp.*

y *Trypanosoma brucei*, destaca por su abundante presencia en el genoma y por unos altos niveles de co-expresión simultánea (Atwood et al. 2005; Atwood et al. 2006; dos Santos et al. 2012). Si a esto le sumamos que la mayoría de sus proteínas se localizan en la membrana del parásito mientras que algunas de ellas son capaces de ser secretadas, la familia MASP puede ser un factor clave en los procesos de infección por *T. cruzi*, pudiendo participar tanto en la interacción parásito-hospedador como en la evasión de la respuesta inmune (Bartholomeu et al. 2009; De Pablos et al. 2011; dos Santos et al. 2012; Garcia et al. 2007).

Evidencias experimentales recientes demuestran la importancia de considerar la existencia de heterogeneidades fenotípicas entre células isogénicas de una misma población (organismos unicelulares) o de un mismo organismo (organismos pluricelulares) como un mecanismo para interactuar con su medio externo y responder a éste (Balazsi, van Oudenaarden, Collins 2011; Pelkmans 2012; Seco-Hidalgo, Osuna, Pablos 2015). Así, en protozoos pertenecientes tanto al orden Apicomplexa como Kinetoplastida, las variaciones clonales entre células individuales pueden determinar aspectos cruciales como la infectividad o la capacidad de supervivencia en el hospedador (Kafsack et al. 2014; Laffitte et al. 2014; Rovira-Graells et al. 2012a; Seco-Hidalgo, Osuna, Pablos 2015; Souza et al. 2011; Ubeda et al. 2014). Adicionalmente, se ha visto una fuerte relación entre la heterogeneidad fenotípica y las familias génicas que intervienen en la interacción parásito-hospedador (Amambua-Ngwa et al. 2012; Niang, Yan Yam, Preiser 2009; Recker et al. 2011; Rovira-Graells et al. 2012b; Seco-Hidalgo, Osuna, Pablos 2015).

En este contexto y con el objetivo de obtener una mejor comprensión del grado de variación entre poblaciones celulares de un mismo estadio del ciclo de vida de *T. cruzi*, este trabajo de tesis se ha centrado en determinar tanto la variabilidad de las secuencias de la familia MASP entre cepas de diferentes orígenes filogenéticos como en caracterizar la heterogeneidad transcripcional de esta familia entre diferentes líneas isogénicas pertenecientes a una única población de *T. cruzi*.

Así, se ha podido observar como, a diferencia de lo que ocurre con los

pseudogenes, donde existe una mayor variabilidad entre cepas, las secuencias de genes de la familia MASP conservan el orden filogenético. Al respecto, también se ha descrito como los miembros de esta familia, genes y pseudogenes, evolucionan a diferentes velocidades evolutivas, encontrándose determinados miembros de esta familia (incluyendo pseudogenes) que están sometidos a una mayor presión positiva de la selección natural, lo que hace que sus secuencias se mantengan altamente conservadas entre las diferentes cepas de *T. cruzi* analizadas.

Por otra parte, este trabajo supone la primera caracterización de la heterogeneidad transcripcional de la familia MASP dentro de una única población de *T. cruzi*. Para ello se han aplicado varias técnicas moleculares, que han permitido realizar los análisis tanto sobre el ARN como sobre las proteínas de un total de 8 líneas isogénicas de epimastigotes de la cepa Pan4. Todos los resultados muestran el alto grado de variabilidad clonal en la expresión de la familia MASP, donde todas las líneas celulares fueron transcripcionalmente diferentes. Adicionalmente, entre los diferentes clones de *T. cruzi*, se han analizado otros aspectos fenotípicos como son el crecimiento, la metaciclogenesís espontánea o la producción y excreción de diferentes catabolitos al medio, revelándose de nuevo un comportamiento heterogéneo, con marcadas diferencias entre cada una de las líneas isogénicas de *T. cruzi* (Seco-Hidalgo, De Pablos, Osuna 2015a).

Aunque que se desconoce cual es la estrategia evolutiva que mantiene el polimorfismo en la expresión de la familia MASP y la heterogeneidad fenotípica (apuesta diversificada o plasticidad adaptativa), la variabilidad clonal descrita en este trabajo se postula como una importante ventaja para *T. cruzi*, permitiendo una rápida respuesta de las formas epimastigotes del parásito ante las presiones y cambios ambientales que existen en el tracto digestivo del triatomino.













# **I. INTRODUCCIÓN GENERAL**

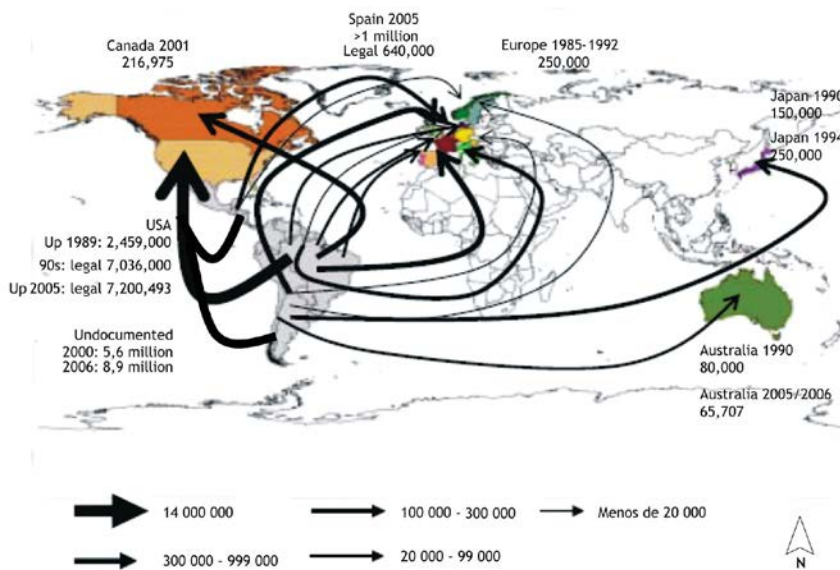






El agente etiológico de esta parasitemia lo constituye un protozoo flagelado conocido como *Trypanosoma cruzi*, perteneciente al orden Kinetoplástida, Familia Trypanosomatidae, Género *Trypanosoma*, y dentro de éste, al grupo denominado Stercolaria.

Pese a su carácter endémico, debido a fenómenos migratorios, esta parasitemia se halla en plena fase de globalización (Figura 1.2), con cerca de 100 millones de personas en riesgo de ser infectadas (Cura et al. 2015; Di Girolamo et al. 2015; Jackson, Varcher Herrera, Gascon 2014; Manne-Goehler, Reich, Wirtz 2015; Pinto, Pett, Jackson 2014; Requena-Mendez et al. 2015; Schmunis and Yadon 2009; World Health Organization 2010). Entre 2005 y 2006 se estimaron valores de 16 por 1,000 inmigrantes infectados en Australia, 9 por 1,000 en Canadá, 25 por 1,000 en España y de 8 a 50 por 1,000 en Estados Unidos (Schmunis 2007). De esta manera, la Tripanosomiasis americana incapacita laboralmente a unas 670,000 personas a la vez que, de forma directa o indirecta, produce alrededor de 11,000 fallecimientos anuales (Tarleton et al. 2007; World Health Organization 2010), constituyendo un problema sanitario a nivel mundial y siendo reconocida desde 2005 por la Organización Mundial de la Salud (OMS), como una de las 17 enfermedades tropicales olvidadas (neglected tropical diseases) priorizadas por dicho organismo (World Health Organization 2010).



**Figura 1.2. Flujo migratorio desde países endémicos para la enfermedad de Chagas.** El tamaño de las flechas indica el volumen de personas desplazadas (Schmunis 2007).

## 1.2 Historia y descubrimiento

En 1909 el Dr. Carlos Chagas (1879-1934) se encontraba en el poblado de Lassance dentro del estado de Minas Geiras, en Brasil, realizando unas investigaciones sobre algunos problemas sanitarios de la región. Fue entonces cuando encontró en las heces de los “barbeiros” (nombre popular de los triatóminos en Brasil) unos protozoos que identificó como parásitos pertenecientes al género *Trypanosoma*, del que ya se conocían algunas especies en otras partes del mundo. Así fue como se aisló antes el parásito que la propia enfermedad. A esta nueva especie la denominó *Trypanosoma cruzi* en homenaje al que fuera su maestro, Oswaldo Cruz. Después de esto, el propio Carlos Chagas siguió estudiando hasta acabar describiendo la enfermedad (Abad-Franch, Santos, Schofield 2010).

Así mismo, esta tripanosomiasis fue estudiada por el Dr. Salvador Mazza, quien contribuyó enormemente a describir y a difundir la importancia sanitaria de la enfermedad, llegando a caracterizar diferentes formas clínicas de la misma.

En honor a estos investigadores, por toda la aportación y trascendencia de sus trabajos, a esta parasitemia se la denominó como enfermedad o mal de Chagas-Mazza.

La existencia de triatóminos en América fue previamente descrita por Charles Darwin a principios del siglo XIX. El siguiente es un extracto de lo que él escribió en su diario de notas “El viaje del Beagle” (*The Voyage of the Beagle*) el 25 de marzo de 1835:

*"Cruzamos el Luján, el cual es un río de talla considerable, si bien su curso hacia las costas del mar es poco conocido: es doblemente dudoso como, pasando sobre sendas planicies, no se evapora y se pierde. Dormimos en la ciudad Luján, la cual es un pequeño lugar rodeado por jardines, y constituye el distrito más sureño en la Provincia de Mendoza; está a cinco leguas al sur de la capital. Por la noche experimenté un ataque (no merece un nombre menor) de vinchucas, una especie de reduviideo, el gran chinche negro de las Pampas. Es de lo más molesto sentir estos insectos sin alas, de aproximadamente una pulgada de largo, trepando sobre el cuerpo. Antes de chupar se mantienen algo delgadas, pero luego se vuelven redondas e hinchadas de sangre, y en*

*este estado se pueden aplastar fácilmente. Uno de los que llevé a Iquito, (también se consiguen en Chile y Perú) estaba muy vacío. Cuando se colocaba en una tabla, aun rodeado de gente, si se le presentaba un dedo, el insecto gallardo extendía inmediatamente su aparato chupador, atacaba, y si se le permitía, absorbía sangre. Ningún dolor causó la herida. Fue curioso ver su cuerpo durante el acto de succión, dado que en menos de 10 minutos este cambió de ser tan plano como una ostia hasta tomar una forma globular. Este banquete, el cual la vinchuca debió a uno de los oficiales, la mantuvo repleta durante cuatro meses enteros; pero la noche cuando se cumplían dos semanas, estaba ya casi lista para picar otra vez."*

Al respecto, se ha especulado sobre la posibilidad de que la patología de la enfermedad y las recaídas que Darwin sufrió tras su estancia en Argentina se produjeran por el contacto con los triatóminos, lo que pudo ocasionarle una enfermedad de larga duración como es el mal de Chagas (Teive and Arruda 2009; Woodruff 1965).

### **1.3 Modos de transmisión**

La enfermedad de Chagas se transmite por insectos triatóminos cuya relación con su reservorio se establece de distintas formas: zoonótica, antropozoonótica, zooantroponótica o zoonosis-amphixenosis (Coura 2013). Estos insectos hematófagos pertenecen a la subfamilia Triatominae, familia Reduviidae, orden Heteroptera/Hemiptera y presentan más de 150 especies descritas hasta el momento, la mayoría de las cuales se distribuyen a lo largo de América, con sólo algunas de ellas presentes en lugares como Asia, África o Australia. Su ciclo de vida se compone de cinco estados ninfales que se suceden hasta alcanzar la fase adulta. Este proceso requiere alrededor de 6 meses para completarse (Noireau, Diosque, Jansen 2009).

Los triatóminos pueden clasificarse según el origen del hospedador que parasitan como triatóminos silvestres o triatóminos domésticos. El primer grupo engloba la mayor parte de las especies y se encuentran en madrigueras, nidos y rocas asociadas a roedores, armadillos, pájaros y a una gran variedad de especies del orden Marsupialia, a un elevado número de especies del orden Chiroptera e incluso a primates y a murciélagos (Coura et al. 2002; Garcia et al. 2007; Noireau,

Diosque, Jansen 2009). Por su parte, el segundo grupo está conformado tan sólo por el 5% de las especies las cuales, debido a la reciente presencia del hombre en sus ecótopos selváticos, han sido capaces de ampliar el abanico de hospedadores, parasitando y transmitiendo el *T. cruzi* a humanos y animales domésticos como perros, gatos y cobayas (Miles, Feliciangeli, de Arias 2003). Así son muchas las especies selváticas de triatóminos que están en proceso de domiciliación (peridomésticas). Estos redúvidos aprovechan la similitud existente entre sus nichos ecológicos y las construcciones domésticas para infestar éstas últimas (Lent, H., Wygodzinsky, P. 1979). De esta manera, además de al ser humano, *T. cruzi* es capaz de infectar a más de 100 especies de mamíferos (domésticos y salvajes) (World Health Organization 2015).

La transmisión de la enfermedad tiene lugar durante la picadura del triatómino infectado. Después de succionar la sangre, el vector deposita sus heces/orina infectadas (con alta carga parasitaria), siendo capaz de transmitir la infección a través de microheridas existentes en la piel, de las membranas mucosas o conjuntivas (Miles, Feliciangeli, de Arias 2003; World Health Organization 2015).

Se considera que todas las especies de la subfamilia Triatominae son vectores potenciales del “Mal de Chagas”, sin embargo las siguientes constituyen, epidemiológicamente, los vectores más importantes: *Rhodnius prolixus*, *Rhodnius pallescens*, *Triatoma infestans*, *Triatoma brasiliensis*, *Triatoma dimidiata*, *Panstrongylus megistus* y *Mepraiia spinolai*.

La transmisión de la enfermedad también puede ocurrir por transfusión sanguínea o trasplante de órganos, por vía oral a través de la ingesta de comida o zumos contaminados por triatóminos infectados, o de forma vertical, es decir, de una manera congénita en la que los parásitos son transmitidos desde la placenta de la madre al hijo (Prata 2001). El riesgo de que un recién nacido adquiera la infección de su madre varía entre el 1.33% y el 5% (Yadon and Schmunis 2009). Otra forma de transmisión de la enfermedad la constituyen los accidentes de laboratorio.

## 1.4 Manifestaciones clínicas

El cuadro clínico que desencadena *T. cruzi* en los seres humanos es variable, va a depender del número de parásitos que constituyan el inóculo inicial así como de las características biológicas de las cepas y clones de *T. cruzi* que establezcan la infección y de la respuesta inmunológica del hospedador (Coura 2007). No obstante, las manifestaciones clínicas pueden encuadrarse en tres fases. Tras la infección por el parásito ocurre la fase aguda con una duración corta, entre 20 y 30 días, caracterizada por un difícil diagnóstico sintomatológico. Aquí se registran bajos porcentajes de mortalidad (<10%) (Urbina 2009) con altos niveles de parasitemia acompañados de síntomas poco específicos como fiebre y malestar general, dolor de cabeza, e inapetencia. El Signo de Romaña (chagoma ocular), es un síntoma que revela la puerta de entrada de la infección, permitiendo así un diagnóstico y tratamiento de la enfermedad más temprano y eficaz, aunque sólo aparece en el 4% de los pacientes. Este signo se caracteriza por la inflamación indolora del párpado superior e inferior de un solo ojo así como del lado facial donde se encuentre el ojo afectado. Dicha inflamación ocular desaparece lentamente en el transcurso de la fase aguda.

La respuesta inmune celular, llevada a cabo también por los linfocitos Th1 CD4+ (Brener and Gazzinelli 1997), frena la proliferación del parásito en fase aguda pero debido a la inmunomodulación que el propio parásito ejerce (Acosta Rodriguez et al. 2007) es incapaz de erradicar la infección, la cual persiste y avanza dando lugar a la fase crónica de la enfermedad caracterizada por una baja parasitemia. En esta fase el paciente puede permanecer durante muchos años completamente asintomático, lo que se conoce como fase indeterminada de la enfermedad, o bien pueden aparecer diferentes manifestaciones tardías en diferentes tejidos como colón o corazón, desarrollándose patologías gastrointestinales y/o cardíacas, las cuales pueden llegar a aparecer hasta 10 años después del primer contacto con el parásito. La patología cardíaca se caracteriza por una miocarditis crónica que va a desembocar en cardiomegalia, insuficiencias cardíacas y arritmias (Barbosa et al. 2015; Moncayo and Silveira 2009; Rassi, Rassi, Little 2000). En su forma digestiva, en casos avanzados del mal de Chagas se observan dilataciones del esófago y/o colon (megaesófago y

megacolon, respectivamente) (Macedo et al. 2004). La cardiopatía se ve acompañada por la pérdida de células del miocardio (Palacios-Pru et al. 1989) mientras una degeneración neuronal y denervación caracterizan al megaesófago y al megacolon respectivamente (Adad et al. 1991).

Entorno al 30% de los afectados va a desarrollar alguno de estos tipos de complicación clínica (The World Health Organization 2002a). Se ha demostrado que existen importantes variaciones geográficas en la morbilidad y prevalencia de las diferentes manifestaciones clínicas. En Brasil, lo más común es encontrar la forma asintomática de la enfermedad (60-70%) seguida de las formas cardíacas y digestivas (20-30% y 8-10%, respectivamente) (Macedo et al. 2004). En la región central de Brasil y Chile la forma digestiva de la enfermedad predomina mientras que se muestra prácticamente inexistente en Venezuela y América Central. Aunque las razones que producen esta heterogeneidad geográfica en las manifestaciones clínicas de la enfermedad todavía no están completamente descritas, se cree que las diferencias genéticas de la población humana junto con el polimorfismo genético, e incluso fenotípico, de las poblaciones de *T. cruzi*, son factores que condicionan fuertemente este fenómeno. Adicionalmente, algunos aspectos inmunológicos, el entorno y la dieta del hospedador no deben ser descartados (Andrade et al. 1999; Macedo et al. 2004; Seco-Hidalgo, Osuna, Pablos 2015).

## **1.5 Diagnóstico**

Una pieza fundamental en la lucha contra la enfermedad de Chagas es el diagnóstico. La efectividad de cualquier campaña de control ya sea actuando sobre el insecto vector, bloqueando la transmisión de la enfermedad o bien vacunando a los individuos, recae sobre un diagnóstico competente capaz de detectar a los individuos infectados.

Como se ha comentado anteriormente, la detección de la enfermedad de Chagas en la fase aguda supone un punto clave para un tratamiento eficaz. No obstante, debido a su sintomatología poco específica y solapada con otras muchas enfermedades, el diagnóstico de la enfermedad en esta fase es difícil en la mayoría de ocasiones.

Los diagnósticos serológicos de esta tripanosomiasis se fundamentan principalmente en tests inmunológicos como los basados en “enzyme-linked immunosorbent assays” (ELISAs), ensayos de inmunofluorescencia indirecta (IIF) y ensayos de hemaglutinación indirecta (IHAs). Todos suelen valerse de extractos o antígenos extraídos a partir de las formas epimastigotes de *T. cruzi*, no siendo ésta la fase del ciclo de vida más adecuada porque, a pesar de encontrarse en el insecto vector, no está presente en los hospedadores mamíferos lo cual puede comprometer la sensibilidad de los resultados. No obstante, por su simplicidad, bajo coste y buen rendimiento genérico, se utilizan kits comerciales basados en estas técnicas consiguiendo sensibilidades entre el 95 y 99% (Caballero et al. 2007). La OMS recomienda usar paralelamente al menos dos tests que se basen en principios antígenos diferentes (The World Health Organization 2002b); si ambos tests son eficaces la sensibilidad asciende llegando incluso al 100% (Caballero et al. 2007; Oelemann et al. 1998); si en cambio los resultados de ambos tests son discordantes se requiere de un tercer test (The World Health Organization 2002b). En los hospitales y bancos de sangre en los que se han empleado estos kits comerciales, se ha observado rendimientos por debajo del 98% (Salles et al. 1996). Recientemente Afonso *et al.* (2012) realizaron una revisión sistemática de la calidad en el diagnóstico de la enfermedad de Chagas; sobre los 61 ensayos analizados se obtuvo una sensibilidad global del 90% y una especificidad del 98%, concluyendo que los ensayos serológicos presentan una sensibilidad y especificidad menor de lo que se había pensado.

De esta manera, los ensayos serológicos para el diagnóstico de la enfermedad de Chagas están sujetos a ciertos inconvenientes como las variaciones en la reproducibilidad de los resultados (da Silveira, Umezawa, Luquetti 2001), reactividad cruzada frente a anticuerpos inducidos por otros patógenos (principalmente *Leishmania*) y descenso de la eficacia frente al análisis de los sueros de pacientes con un título de anticuerpos anti-*T. cruzi* bajo (Marcon et al. 2002). La aplicación de péptidos sintéticos y/o antígenos recombinantes de diferentes cepas y origen geográfico, ha sido propuesta como una mejora de la especificidad y sensibilidad, disminuyendo el número de falsos positivos y falsos negativos (da Silveira, Umezawa, Luquetti 2001).

El radioinmunoensayo (RIA) es una técnica que originalmente fue diseñada para medir la respuesta de anticuerpos frente a *T. cruzi* en ratones y posteriormente adaptada para su uso en el inmunodiagnóstico de la enfermedad de Chagas. Esta técnica que emplea anticuerpos biotinilados y <sup>3</sup>H-avidina, permite distinguir entre *T. cruzi* y *Trypanosoma rangeli*. Tarleton *et al.* (1984) utilizaron epimastigotes fijados e intactos como antígenos en el RIA, obteniendo resultados comparables a los conseguidos con la fase tripomastigote.

Otra técnica que ha surgido con el fin de diagnosticar el mayor porcentaje de individuos infectados es la basada en el antígeno TESA (*Trypomastigote Excreted-Secreted Antigens*), la cual se basa en el empleo de antígenos procedentes de la fase infectiva de *T. cruzi* que están presentes tanto en los triatóminos como en los hospedadores mamíferos. Esta técnica se puede combinar con ELISA (TESA ELISA) o con western blot (TESA blot), ofreciendo unas mayores especificidades y sensibilidades respecto a los tests de los kits comerciales, tanto en la fase aguda como en la fase crónica de la infección (Umezawa *et al.* 1996; Zarate-Blades *et al.* 2007).

La PCR (Polymerase Chain Reaction) y algunas de sus variantes, como *nested*-PCR (N-PCR) y real-time PCR se muestran como métodos moleculares y confirmatorios ideales como complemento de los kits comerciales (Cummings and Tarleton 2003; Marcon *et al.* 2002; Qvarnstrom *et al.* 2012; Ramirez *et al.* 2015). Mediante el uso de estas técnicas de biología molecular se ha podido demostrar la existencia de individuos infectados (ADN de *T. cruzi* circulante) sin rastro serológico del parásito (Salomone *et al.* 2003). Estas técnicas también están sujetas a variaciones como son la intermitencia de las poblaciones de parásitos circulantes en sangre, los métodos de extracción de ADN o las secuencias del parásito utilizadas para realizar las amplificaciones (Schijman *et al.* 2011).

## 1.6 Tratamiento

Se dispone de dos fármacos para el tratamiento específico de la infección por *T. cruzi* (Coura and de Castro, 2002). Nifurtimox, un nitrofurano desarrollado por Bayer en 1967 y comercializado bajo el nombre de Lampit®, actúa reduciendo el grupo nitro dando nitro-aniones los cuales van a reaccionar con el



oxígeno molecular para producir los radicales tóxicos superóxido y peróxido. El segundo compuesto, el Benzinidazol, un nitroimidazol desarrollado por Roche en 1972 y comercializado como Rochagan® o Radanil® (aunque ahora producido por el Laboratorio Farmacéutico del Estado de Pernambuco, Brasil) actúa de forma diferente, produciendo metabolitos que reaccionan con macromoléculas como el ADN, ARN, proteínas y posiblemente lípidos. En ambos casos, la actividad antiparasitaria se encuentra íntimamente ligada a una inherente toxicidad de los compuestos, produciendo de forma frecuente una serie de efectos adversos durante el tratamiento, especialmente en aquellos pacientes de mayor edad. Ambos fármacos pueden ocasionar malestar, dolores de cabeza y pérdida de la concentración; Nifurtimox, concretamente, protagoniza casos de pérdida de apetito, pérdida de peso e incluso anorexia; por su parte Benzinidazol está asociado a dermatitis alérgicas y en ocasiones neuritis periféricas. Todos estos efectos secundarios pueden llegar a ser tan severos que se requiera la suspensión del tratamiento. A pesar de ello, en pocos casos la intolerancia de un fármaco se ha visto asociada a la intolerancia del otro (Abad-Franch, Santos, Schofield 2010).

Otro inconveniente es la eficacia variable de estos fármacos. La OMS determinó que éstos únicamente se mostraban activos ante aquellos enfermos chagásicos que se encontraban en la fase aguda (0 - 4 meses siguientes al inicio de la infección) o en los primeros años de la fase crónica (Guedes et al. 2011; The World Health Organization ). La larga duración del tratamiento por estos fármacos y su elevado coste son otro hándicap a tener en cuenta. Por ejemplo, un tratamiento de Nifurtimox requiere 10 mg por Kg de peso de 3 a 4 veces al día y durante periodos de entre 60 a 120 días, por lo que muchas veces el tratamiento es incompleto y en consecuencia se generan resistencias a dichos fármacos (Wilkinson et al. 2008).

En estudios recientes publicados por Morillo y colaboradores, dentro del proyecto de "BENEFIT trials", se realizó un seguimiento durante una media de 5.4 años a 2,854 pacientes chagásicos afectados de cardiomiopatía y que fueron tratados con Benzinidazol durante 80 días. Los resultados concluyeron que el fármaco ayuda a disminuir significativamente la parasitemia pero no permite una mejora en el deterioro cardíaco del paciente (Morillo et al. 2015). Continuar

investigando hacia una terapia mejorada constituye un reto. Al margen de otros fármacos y compuestos químicos que se están desarrollando, como inhibidores de la biosíntesis de ergosterol (EBI) (Urbina 2009), son muchos los esfuerzos destinados a la búsqueda de una quimioterapia efectiva frente a *T. cruzi*.

## **1.7 Programas de control y prevención de la enfermedad**

Desarrollar un programa de control para esta parasitosis supone un reto complejo tanto por la pobreza que contextualiza a esta endemia, como por la limitación de los medios disponibles para tratar la enfermedad. A pesar de ello, mientras se sigue investigando en el desarrollo de vacunas y en la mejora de los fármacos, son numerosos los hallazgos y muchas las intervenciones sanitarias y sociales que se pueden realizar para paliar los efectos de la Tripanosomiasis americana.

Una primera medida efectiva centrada en la prevención primaria es el control del insecto vector en las zonas endémicas. En este caso se trata de mejorar las características y aislamiento de las viviendas, originalmente hechas con barro, maderas, pajas y hojas de palma; así se consigue evitar la presencia del triatómino en hábitats domésticos. Este hecho combinado con rastreos mediante tests serológicos en bancos de sangre y bancos de órganos que permitan el diagnóstico y el posterior tratamiento del paciente, a la vez que evitan el contagio, son medidas que disminuyen drásticamente el número de infectados. Dichas medidas ya se han puesto en práctica a través de programas como el que sostiene la iniciativa del Cono Sur de Latinoamérica actuando en las regiones de Uruguay en 1997, Chile en 1999, en el centro y sur de Brasil en 2001, en cuatro provincias de Argentina en 2002 y en un distrito de Paraguay en 2003.

A raíz de esta intervención se crearon tres programas más: el programa de control de la Comunidad Andina (Venezuela, Colombia, Perú y Ecuador), el programa de control de América Central y el programa de vigilancia de la cuenca Amazónica (Coura et al. 2002; Dias, Silveira, Schofield 2002). Dichos programas tratan de desarrollar el control en sus respectivas regiones.

También en España se iniciaron dichos tests serológicos en el año 2005 a través del centro de transfusiones de Valencia, siendo el primer país de la Unión Europea en hacerlo (Real Decreto 1088/2005).

Pero la realidad de la situación es que las regiones endémicas siguen necesitando medios y una mejor coordinación de estos programas. Ante las dudas económicas que esta labor humanitaria puede generar, hay que resaltar que el coste derivado de la enfermedad por morbilidad, mortalidad, hospitalización, tratamientos de arritmias, provisión de marcapasos, etc, superan con creces los gastos que supone el desarrollo de un programa de control integrado (Miles, Feliciangeli, de Arias 2003).

## ***2. Trypanosoma cruzi***

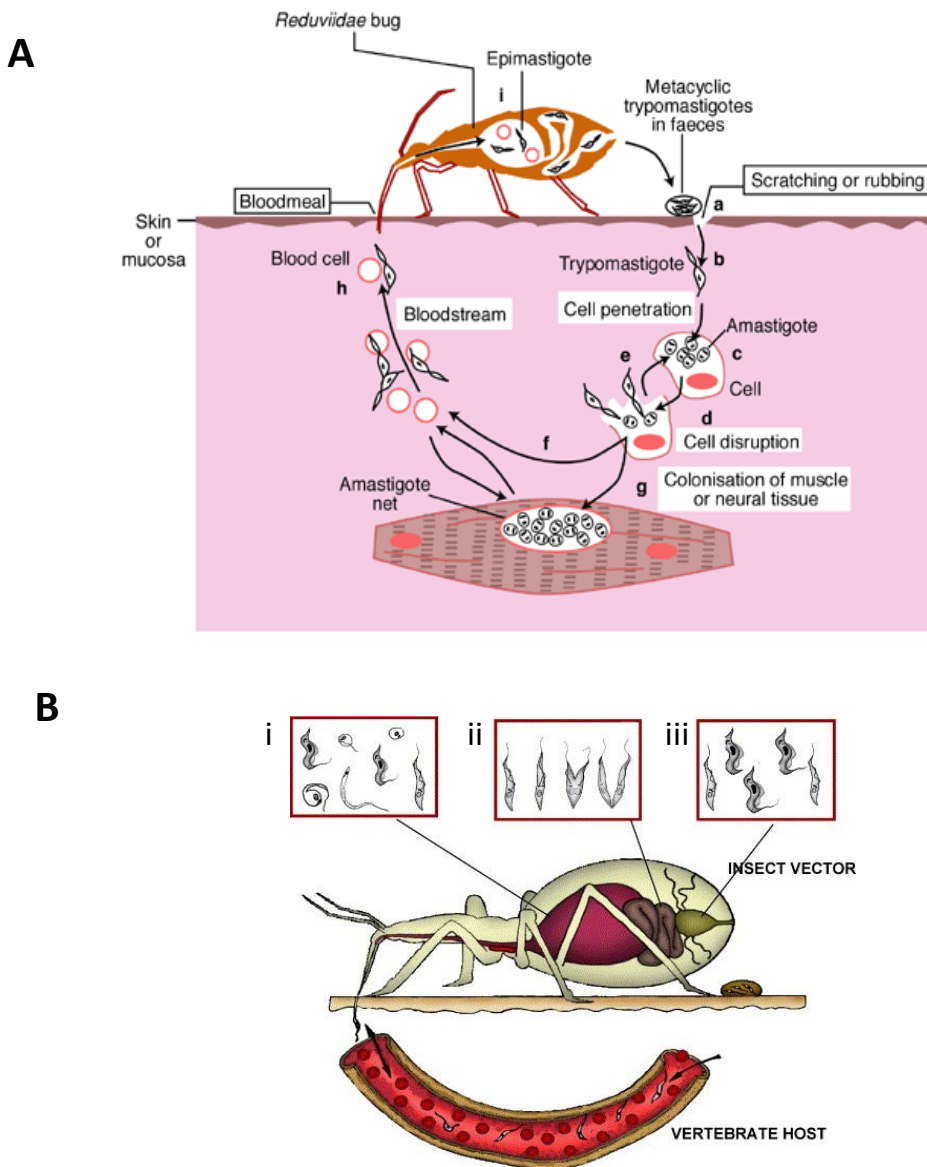
### **2.1 Ciclo de vida**

El ciclo de vida de *T. cruzi* es altamente complejo y supone la presencia de un amplio repertorio de poblaciones de este protozoo que circulan entre humanos, vectores, reservorios silvestres y animales domésticos (García et al. 2007). Para ello, *T. cruzi* es capaz de alternar entre cuatro principales estadios morfológicos y funcionales (si bien existen una mayor variedad de formas intermedias en el tracto digestivo del insecto vector): epimastigote, tripomastigote metacíclico, tripomastigote sanguíneo y amastigote (De Souza 2002). El ciclo se inicia en el momento en que un triatómino infectado se alimenta succionando la sangre de un mamífero a la vez que deposita sus heces y orina, conteniendo las formas tripomastigotes metacíclicas, las cuales entran en el hospedador a través de la escarificación que se produce tras el rascado de la picadura, de las membranas mucosas o conjuntivas (Figura 1.3 Aa). Una vez se encuentran en el interior del hospedador, los tripomastigotes metacíclicos van a invadir las células epiteliales nucleadas cercanas al lugar de la inoculación, diferenciándose a amastigotes intracelulares (Figura 1.3 Ab). Estos amastigotes constituyen la única forma replicativa del parásito en el hospedador mamífero, así mediante fisión binaria, los amastigotes se dividen y ocupan el citoplasma de

la célula hospedadora y, tras varias rondas de replicación, se transforman en tripomastigotes que son liberados a los espacios intercelulares, la linfa y a la circulación sanguínea como tripomastigotes sanguíneos (Figura 1.3 Ac-h). A diferencia de lo que sucede en *Trypanosoma brucei*, los tripomastigotes sanguíneos de *T. cruzi* no tienen capacidad de división y por tanto necesitan infectar otras células nucleadas y replicarse como amastigotes para cerrar un nuevo ciclo intracelular. Estos tripomastigotes sanguíneos presentan la habilidad de infectar a un gran repertorio de células, desde macrófagos (células fagocíticas) hasta células no fagocíticas de diferentes tejidos como intestino, células del músculo liso, tejido adiposo o incluso del sistema nervioso central (Figura 1.3 Ag) (Stuart et al. 2008). El ciclo se cierra cuando un nuevo insecto vector se alimenta de la sangre del mamífero infectado, succionando con ella una carga de tripomastigotes sanguíneos (Figura 1.3 Ah y B). Unos días después, en el estomago del insecto, la mayor parte de los tripomastigotes sanguíneos se transforman en epimastigotes y sólo unos pocos en esferomastigotes (Figura 1.3 Ai y B). Una vez establecidos en el intestino medio, los epimastigotes se dividen rápidamente por fisión binaria. Finalmente, una gran proporción de epimastigotes se adhieren a la cutícula rectal, donde realizan el proceso de transformación a la forma tripomastigote metacíclica del parásito en un proceso conocido como metacicloogénesis (Figura 1.3 B) (García et al. 2007).

Es interesante resaltar también la existencia de dos tipos morfológicos de tripomastigotes sanguíneos: el primero es conocido como “slender”, presenta un núcleo elongado, un kinetoplasto subterminal y un corto flagelo; el segundo tipo morfológico, denominado “broad” o formas “stout” que se caracterizan por tener un núcleo ovalado, también un kinetoplasto terminal y un largo flagelo. La predominancia de una forma u otra depende de la cepa de *T. cruzi* y del tiempo de infección. Al respecto, se ha sugerido que las formas “slender” son las principales responsables de la infección en vertebrados, mientras las formas “broad” son más infectivas en los hospedadores invertebrados (De Souza 2002).

Por último, cabe destacar que existen autores que no restringen la capacidad infectiva de *T. cruzi* a las formas tripomastigote sanguínea y tripomastigote metacíclica, como los amastigotes que son capaces de establecer infecciones en el hospedador definitivo (Mortara et al. 1999).



**Figura 1.3. Representación esquemática del ciclo de vida de *T. cruzi*.** **A:** Se muestra el ciclo completo de *T. cruzi* a través de sus dos hospedadores, insecto vector y mamífero (Macedo, Oliveira, Pena 2002). **B:** Ciclo biológico de *T. cruzi* en el insecto vector: (i) tras succionar la sangre infectada, las formas tripomastigotas se transforman en epimastigotes dando también algunos esferomastigotes en el estómago; (ii) en el intestino, los epimastigotes se multiplican incrementando la población de parásitos; (iii) en el recto, los epimastigotes se transforman a tripomastigotes metacíclicos los cuales son eliminados por las heces y la orina (Garcia et al. 2007).

## 2.2 Biología y estructura celular

La composición de la superficie celular juega un papel muy importante en los parásitos intracelulares, siendo clave en el proceso de infección que éstos desencadenan. Por ello es importante entender su organización estructural, composición y propiedades funcionales. La superficie celular de *T. cruzi* está formada por la bicapa lipídica y por una serie de componentes azucarados, como glicolípidos y glicoproteínas (mucinas, transialidasas, etc.) que son enfrentados al medio extracelular, constituyendo el glicocálix (Souza 2009).

Al analizar la membrana plasmática del parásito mediante criofractura, se pudieron observar tres zonas sobre ésta: el cuerpo celular, el flagelo y la bolsa flagelar (Souza 2009).

El flagelo de este kinetoplastido presenta una estructura típica con nueve pares de dobletes de microtúbulos periféricos y un par central. Incluso en la forma amastigote puede observarse un corto flagelo. Una estructura característica del flagelo de los tripanosomátidos es el bastón paraflagelar (PFR), que se encuentra unido al axonema a través de una matriz de filamentos. El PFR está compuesto por un gran número de proteínas, destacando dos de ellas, "Paraflagellar rod proteins 1 y 2", caracterizadas por ser altamente antigénicas por lo que constituyen una diana para diagnóstico y vacunación (De Souza 2002).

Si hablamos del citoesqueleto de *T. cruzi*, es muy importante resaltar la presencia de los microtúbulos sub-peliculares, los cuales se distribuyen por todo el cuerpo del protozoo (excepto en la bolsa flagelar) presentándose fuertemente unidos a la membrana plasmática y al retículo endoplasmático (Souza 2009). Dichas estructuras están relacionadas con la ruta endocítica de absorción (Souza 2008).

Los tripanosomátidos son células altamente polarizadas donde el 85% de la actividad endocítica está restringida a la bolsa flagelar. El resto de la ruta se relaciona con el citostoma, un dominio del parásito formado por una invaginación de la membrana plasmática (Souza 2009). Siguiendo la unión existente entre el citostoma y la bolsa flagelar, las macromoléculas son rápidamente internalizadas originando unas pequeñas vesículas endocíticas. El pH ácido sugiere que estas vesículas podrían corresponderse con los endosomas

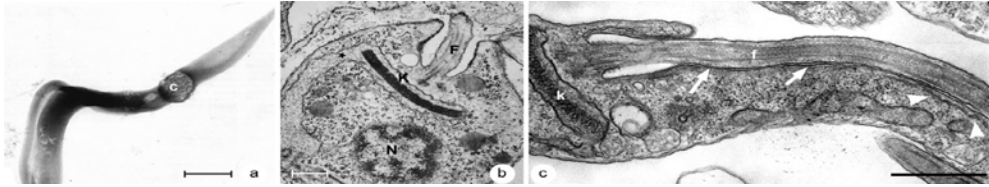
tempranos (EE). El destino final de todas las macromoléculas capturadas por el parásito en el medio extracelular son los reservosomas, unos orgánulos con un gran contenido en proteasas en donde destaca la presencia de la principal proteasa de *T. cruzi*, la cruzipaina. Por ello y porque nunca se han identificado lisosomas en *T. cruzi*, se piensa que el reservosoma puede desempeñar funciones lisosomales. Este orgánulo es particularmente interesante ya que es exclusivo de los organismos *Trypanosoma vespertilionis*, *Trypanosoma dionisii* y *T. cruzi* (Sant'Anna et al. 2009; Souza 2009). Así, el reservosoma es considerado también como el sitio principal de degradación y regulación proteicas. Los reservosomas han sido descritos como una estructura exclusiva de la forma epimastigote, donde cada epimastigote presenta varios reservosomas ubicados en la parte posterior de la célula. En amastigotes y tripomastigotes se han descrito unos orgánulos intracelulares que comparten muchas características con los reservosomas (Sant'Anna et al. 2009; Souza 2009).

Por otra parte, la vía secretora en *T. cruzi* está compuesta por el retículo endoplasmático (RE), el aparato de Golgi (AG) y unas vesículas procedentes de éste último. Estudios recientes describen la capacidad de *T. cruzi* para secretar vesículas (exosomas, micropartículas, ectosomas, etc) de una forma alternativa, demostrando incluso la importancia que éstas tienen en las interacciones parásito-hospedador (Bayer-Santos et al. 2013; Mantel and Marti 2014).

Mediante microscopía óptica y electrónica se ha diferenciado una estructura electrodensa cercana al núcleo, el kinetoplasto. Éste se compone de un tipo de ADN conocido como ADN-kinetoplastídico (ADN-k). Concretamente, el kinetoplasto presenta dos tipos de ADN circulares, los minicírculos y los maxicírculos. Existen miles de minicírculos, cuyo tamaño varía entre 0.5-2.5 kb (según la especie) y algunas docenas de maxicírculos, variando entre 20-40 kb. Todo este ADN va a codificar para ARNs guías que participan en el proceso de "RNA editing" de la mitocondria (Stuart et al. 2005). Por su parte, los maxicírculos son compartimentos análogos, tanto en función como en estructura, al ADN mitocondrial de los eucariotas superiores y codifican ARNr y subunidades de la cadena transportadora de electrones (Souza 2008).

El kinetoplasto se localiza en una porción especial de la mitocondria, siempre cercana al cuerpo basal con el que se conecta mediante estructuras

filamentosas. No obstante, la figura y organización estructural del kinetoplasto varían según el estado morfológico del protozoo (Souza 2009) (Figura 1.4).

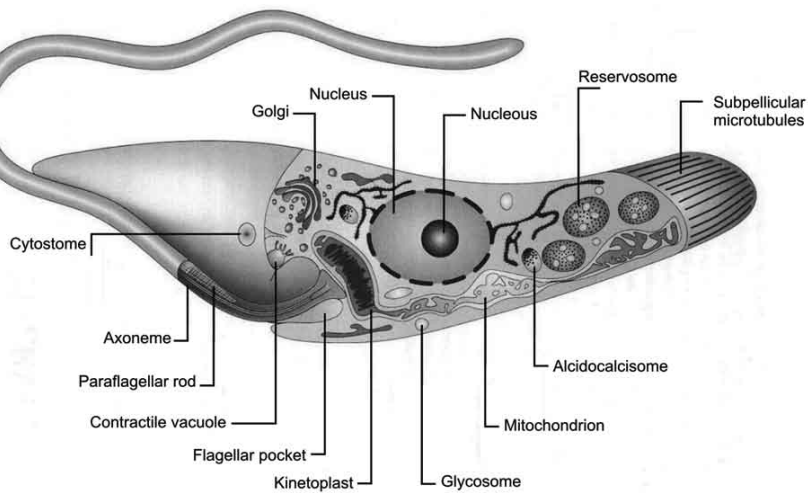


**Figura 1.4. Diferentes vistas del kinetoplasto y de la organización general de las formas de *T. cruzi*.** (a) Tripomastigote, (b) amastigote y (c) epimastigote. F: flagelo; K: kinetoplasto; N: núcleo. Las flechas blancas indican el área de adhesión del flagelo al cuerpo celular. Las cabezas de flechas blancas señalan los perfiles del retículo endoplasmático. Bars = 1  $\mu\text{m}$  (Rocha et al. 2006).

Otros elementos típicos de la biología estructural de *T. cruzi* son el glicosoma, el acidocalcisoma y la vacuola contráctil. El primero de ellos, el glicosoma, consiste en una estructura esférica, carente de ADN, en el que se pueden encontrar proteínas codificadas por genes nucleares, y por el que pasa la mayor parte de la ruta glucolítica del parásito (Opperdoes 1990; Souza 2009). Por su parte, el acidocalcisoma es un orgánulo capaz de transportar protones y calcio desarrollando funciones de almacenaje de calcio, magnesio, sodio, potasio, zinc, hierro y componentes fosforilados. También se le vincula con la homeostasis del pH y con funciones de osmoregulación (Docampo et al. 2013; Moreno and Docampo 2009; Souza 2009). Mientras, la vacuola contráctil, que se compone por algunos túbulos conectados a la vacuola central y se ubica cerca de la bolsa paraflagelar, participa también en funciones de osmoregulación (Souza 2009).

Observaciones iniciales revelaron que el núcleo del parásito está envuelto por una doble membrana porosa, parecida a la de otros organismos eucariotas, con cromatina no condensada a lo largo del nucleoplasma y nucleolos típicos descritos en la fase epimastigote, pero no en la amastigote o tripomastigote (Souza 2009). En la Figura 1.5 se indica la localización de los principales elementos de la estructura celular de la forma epimastigote *T. cruzi*.





**Figura 1.5. Diagrama representativo de las principales estructuras y orgánulos encontrados en la forma epimastigote de *T. cruzi*.** Dichas estructuras fueron detectadas mediante microscopía electrónica de transmisión (Souza 2008).

## 2.3 Genómica y control de la expresión génica

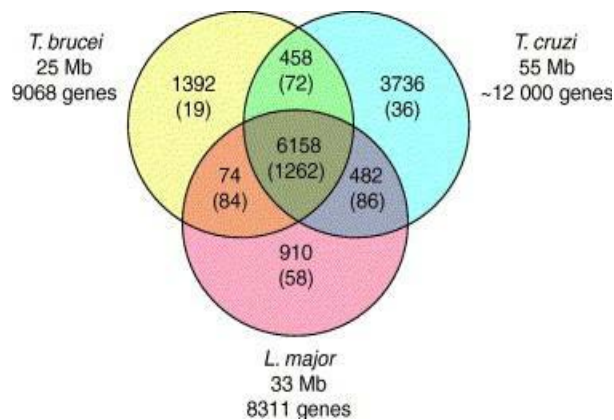
### 2.3.1 Genoma

La iniciativa de secuenciar el genoma de *T. cruzi* surgió como parte del proyecto denominado “Tritryps” en el que se secuenció conjuntamente el genoma de los kinetoplástidos *T. cruzi*, *Trypanosoma brucei* y *Leishmania major* (El-Sayed et al. 2005a; Kissinger 2006).

La secuenciación del genoma de *T. cruzi* se realizó sobre la cepa CL Brener, una cepa híbrida con dos haplotipos (Esmeraldo y no Esmeraldo). De esta manera, cada cromosoma presentaba dos homólogos diferentes entre sí, dificultando el estudio. Para obtener la composición genómica de *T. cruzi* se requirió el empleo de la técnica WGS (secuenciación completa del genoma por técnica de *shotgun*, del inglés “whole-genome shotgun”). El genoma haploide de *T. cruzi* se ensambló en 8,740 “contigs” presentando un total de 67 megabases (Mb), con unos 12,000 genes codificantes (22,570 proteínas codificadas teniendo en cuenta el genoma diploide), donde aproximadamente cada uno de los dos haplotipos aportó la mitad de los alelos. Por otra parte, se estima que el genoma

diploide de *T. cruzi* oscila entre 106.4 y 110.7 Mb, distribuido en 41 pares de cromosomas homólogos cuyos tamaños oscilan entre 74 kb a 2.4 Mb (El-Sayed et al. 2005b; Weatherly, Boehlke, Tarleton 2009). Mientras el genoma haploide de otros kinetoplástidos como *T. brucei* y *L. major* se compone de 25 Mb y 33 Mb, con 8,100 y 8,300 genes codificantes respectivamente (El-Sayed et al. 2005a).

Los genomas solapados de estos tres kinetoplástidos fueron agrupados en clusters de genes ortólogos (COGs). El alineamiento de la secuencia de aminoácidos entre COGs reveló una media de identidad del 57% entre *T. cruzi* y *T. brucei*, y un 44% de identidad entre *L. major* y los otros dos tripanosomátidos. Los parásitos intracelulares *T. cruzi* y *L. major* comparten una mayor proporción de COGs que *T. cruzi* y *T. brucei* y considerablemente más que *L. major* y *T. brucei*. La proporción del genoma exclusiva de cada kinetoplástido se compone de miembros específicos, donde *T. cruzi* (32%) y *T. brucei* (26%) tienen una proporción mucho mayor de la que presenta *L. major* (12%) (Figura 1.6) (El-Sayed et al. 2005a).



**Figura 1.6. Distribución de genes ortólogos (COGs) entre los tres kinetoplástidos** (El-Sayed et al. 2005a).

Aproximadamente, el 50% del genoma de *T. cruzi* está compuesto por ADN repetido, como retrotransposones y genes pertenecientes a familias multigénicas codificantes para proteínas de superficie, incluyendo transialidasas (TS), mucinas, la glicoproteína de superficie gp63 y la recién descrita familia MASP

(*mucin-associated surface protein*) que con 1,377 miembros se sitúa como la segunda familia de mayor tamaño en este tripanosomátido, tras la familia de las transialidasas (TSs) (Tabla 1.1) (El-Sayed et al. 2005b).

A pesar de haber divergido de 200 a 500 millones de años atrás (antes incluso de la aparición de los mamíferos), los genomas de estas tres especies de kinetoplastidas poseen porciones cromosómicas altamente sinténicas (conservadas). El examen de los puntos en los que se rompe la sintenia reveló que el 40% estaba precisamente asociado a expansiones de las mencionadas familias multigénicas, retroelementos y ARNs estructurales (El-Sayed et al. 2005a). De los 1,617 dominios proteicos identificados en los genomas de *Tritryp*s, menos del 5% son únicos de una especie (El-Sayed et al. 2005a). Así, en *L. major* la sintenina se rompe en regiones relativamente pequeñas que se localizan en zonas internas de los cromosomas. En *T. brucei*, las regiones no sinténicas se componen de grandes bloques de *arrays* de genes y pseudogenes en donde destacan los pertenecientes a la familia antigénica VSG (*variant surface glycoprotein*) (El-Sayed et al. 2005a). Por su parte, en *T. cruzi*, estas islas no sinténicas suelen localizarse en las regiones subteloméricas de los cromosomas y están constituidas por *arrays* entrelazados de genes pertenecientes a las familias MASP, TSs, mucinas, DGF-1 (*Dispersed gene family-1*), genes y pseudogenes codificantes para RHS (*retrotransposon hot spot*), así como los retrotransposones VIPER (*vestigial interponed retroelements*), SIRE (*short interspersed repetitive elements*) L1Tc, NARTc (*T. cruzi nonautonomous non-LTR retrotransposons*), y/o los retroelementos DIRE (Tabla 1.1) (El-Sayed et al. 2005a).

**Tabla 1.1. Familias multigénicas de la cepa CL Brener de *T. cruzi*.** En paréntesis se indica el número de pseudogenes (De Pablos 2010).

<b>Producto génico</b>	<b>Miembros</b>	<b>Ortólogos <i>Tritryp</i></b>
<b>Transialidasa (TS)</b>	1,430 (693)	<i>T. brucei</i>
<b>MASP</b>	1,377 (433)	No
<b>Mucinas</b>	863 (201)	No
<b>Retrotransposon hot spot proteins (RHS)</b>	752 (557)	<i>T. brucei</i>
<b>Dispersed gene family-1 (DGF-1)</b>	565 (136)	No
<b>Proteasa gp63</b>	425 (251)	<i>T. brucei</i> + <i>L. major</i>

### 2.3.1.a Genómica comparativa

La genómica comparativa supone una herramienta muy útil para el estudio de la biología funcional ya que permite identificar elementos clave en la biología de un organismo.

Previo a la secuenciación del genoma de otras cepas de *T. cruzi*, Minning *et al.* (2011) hicieron uso de la técnica “whole genome oligonucleotide tiling array” sobre 16 cepas de *T. cruzi* y, aprovechando la hibridación competitiva de los cebadores, fueron capaces de identificar más de 4,000 *copy number variations* (CNVs) de segmentos de ADN, utilizando para ello a CL Brener como cepa de referencia. Los resultados mostraron una clara dicotomía entre las cepas correspondientes al linaje tipo I y tipo II a la vez que se observaron similitudes y diferencias independientes a las Unidades discretas de tipificación (DTUs) al que pertenecían (Minning *et al.* 2011) (la filogenia y las DTUs de *T. cruzi* se explican en el apartado 2.4).

En la actualidad, además de CL Brener, se dispone del genoma de otras cepas de *T. cruzi*, como Marinkellei o Sylvio X10/1. Este hecho ha permitido llevar a cabo estudios comparativos que muestran diferencias significativas en el genoma de los diferentes linajes analizados (Franzen *et al.* 2011; Franzen *et al.* 2012). Así, Franzen y colaboradores, hallaron que el genoma de Marinkellei, una subespecie de *T. cruzi* asociada a murciélagos, era un 11% más pequeño que el de la cepa Sylvio X10/1 (Franzen *et al.* 2012). Esta diferencia fue atribuida a las CNVs de secuencias codificantes y no codificantes. Además, es importante destacar como en la comparación se identificaron 7 secuencias únicas de Marinkellei. También se encontró un 7.5% de divergencia de las regiones codificantes entre ambas cepas, así como una expansión de la familia DGF en Marinkellei respecto a Sylvio X10/1, que posee 1.5 veces más secuencias relacionadas con los elementos VIPER y L1TC (Franzen *et al.* 2011).

Todo esto refleja las variaciones en las trayectorias evolutivas descritas y la plasticidad del genoma existentes entre las diferentes cepas de *T. cruzi*.

### **2.3.2 Control de la expresión génica en tripanosomátidos**

En la mayoría de células eucariotas, los genes codificantes para proteínas se procesan como pre-ARNms monocistrónicos que posteriormente sufren un proceso de maduración, denominado “cis-splicing”, para obtener el ARNm final. La ARN polimerasa II (ARN pol II) es la enzima responsable de la transcripción de genes codificantes para proteínas, mientras que la ARN pol I lleva a cabo la transcripción de genes ribosomales y proteínas de superficie VSG en *T. brucei*, y la ARN pol III de los genes codificantes para ARNs de transferencia.

Por su parte, los tripanosomátidos exhiben determinadas características que los distinguen de otras células eucariotas como la transcripción policistrónica, el trans-splicing 5' y la poliadenilación 3' del pre-ARNm, la edición del ARN mitocondrial y la transcripción de los genes codificantes para proteínas mediante la acción de la ARN polimerasa I (Donelson and Zeng 1990; Smircich et al. 2015).

En los tripanosomátidos se han descrito unidades policistrónicas con más de 100 marcos abiertos de lectura (Fernandez-Moya and Estevez 2010). La transcripción policistrónica implica que todos los genes incluidos en el cluster genético son transcritos a la vez. Sin embargo, es sabido que en los kinetoplastos también hay casos de transcripción monocistrónica en algunos genes (De Gaudenzi et al. 2011).

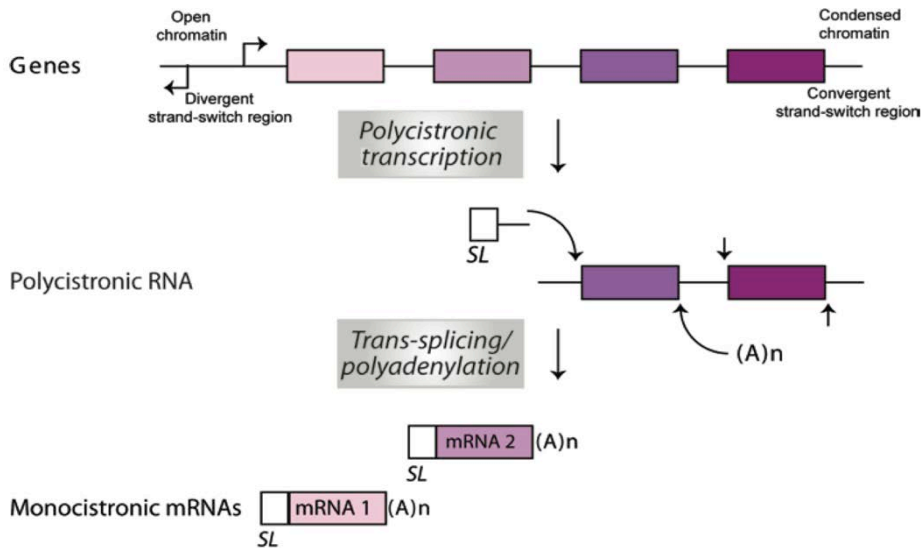
Debido a la ausencia generalizada de un control de inicio de la transcripción, la regulación postranscripcional se convierte en uno de los mecanismos responsables de la mayoría de los cambios en la expresión génica de tripanosomátidos. Es por ello que, respecto a otros eucariotas, los tripanosomas presentan una alta variabilidad en las secuencias de las histonas, principalmente en aquellas regiones sujetas a modificaciones postransduccionales (De Gaudenzi et al. 2011). Pese a todo, la posición genómica dentro de las unidades policistrónicas de transcripción puede ser determinante en la expresión génica temporal en respuesta a situaciones de estrés.

Durante la maduración del ARNm en el núcleo celular se produce la rotura de los ARNm policistrónicos en ARN maduros monocistrónicos. En el proceso de trans-splicing se lleva a cabo la adición de una secuencia hipermetilada de 39 nucleótidos, conocida como “spliced leader” o SL, en el extremo 5'. Por otra

parte, el extremo 3' es poliadenilado (De Gaudenzi et al. 2011; LeBowitz et al. 1993). Una zona localizada entre dos regiones codificantes y rica en bases pirimidínicas parece ser la señal que dirige ambos procesos (De Gaudenzi et al. 2011). De esta manera el trans-splicing y la poliadenilación se muestran como dos mecanismos no independientes que forman parte de un mecanismo de corte y empalme que ocurre concomitantemente o inmediatamente antes de la transcripción (LeBowitz et al. 1993). Debido a que los 5'UTRs son relativamente cortos en los tripanosomátidos, la secuencia SL se añade 30-100 nucleótidos aguas arriba del codón ATG, el cual marca el inicio de la transcripción (De Gaudenzi et al. 2011). Hasta la fecha, el proceso de trans-splicing mejor descrito es en tripanosomátidos; sin embargo otras especies del filo *Euglenozoa* y varios organismos como rotíferos, cnidarios, cordados, nematodos, planelmintos y dinoflagelados también tiene este proceso (Bitar et al. 2013).

Mediante secuenciación de alto rendimiento de ARN de *T. brucei* se ha demostrado que la mayoría de los genes tienen entre uno y tres sitios aceptores alternativos para la secuencia SL y varios susceptibles para la poliadenilación (Siegel et al. 2010).

Por otra parte, estudios recientes sobre las variantes de histonas y el remodelaje de la cromatina indican que los mecanismos epigenéticos llevan a cabo un papel determinante en la regulación postranscripcional de estos parásitos (Fernandez-Moya and Estevez 2010). A las regiones separadoras de los clusters genéticos policistrónicos se las conoce como "strand-switch regions" (SSR). Se ha demostrado que algunas SSR son enriquecidas con sitios de inicio de la transcripción y se las denomina "divergent strand-switch regions" (dSSR), mientras que existen otras que preferiblemente contienen sitios de terminación de la transcripción y son conocidas como "convergent strand-switch regions" (cSSR). En este contexto, las variantes de las histonas y sus modificaciones (acetilaciones y metilaciones) en "divergent strand-switch regions" están relacionados con el inicio bidireccional de la transcripción (Figura 1.7) (De Gaudenzi et al. 2011; Fernandez-Moya and Estevez 2010).



**Figura 1.7. Representación esquemática del proceso de maduración del ARNm por trans-splicing y poliadenilación típico de los tripanosomátidos.** La transcripción se inicia en el punto en que la cromatina está abierta (“divergent strand-switch regions”) y finaliza donde la cromatina está condensada (“convergent strand-switch regions”). El cluster de varios genes codificantes es transcrito en un solo transcrito policistónico, el cual es procesado por los mecanismos de trans-splicing y poliadenilación que dan lugar a ARNm monocistónicos. Los rectángulos representan cuatro secuencias codificantes para proteínas diferentes. (A)n: cola poli(A) (De Gaudenzi et al. 2011).

Las agrupaciones de genes direccionales (DGCs), son característicos de los tripanosomátidos. Sobre estas DGCs existe una fuerte presión selectiva que las mantiene intactas y que, a pesar de las divergencias genéticas, respeta el orden génico. Las singularidades del procesamiento del ARN en los tripanosomátidos se reflejan en la correlación que se establece de forma frecuente entre los largos DGCs característicos de los *Trypanosoma* y los bloques conservados de sintenia (El-Sayed et al. 2005a). El-Sayed *et al.* (2005a) postularon que el proceso de transcripción de genes codificantes para proteínas se inicia sólo en determinados sitios de cada cromosoma. Así, los autores sugieren la acción de una presión selectiva en los grupos de genes policistónicos aguas abajo de dichos sitios, ayudando a mantener dicha sintenia (El-Sayed et al. 2005a).

### 2.3.2.a *Transcriptoma*

El transcriptoma es el conjunto de moléculas de ARN transcritas en una célula/organismo y refleja los genes que están siendo activamente expresados en un tiempo dado. Algunos estudios del transcriptoma muestran que únicamente una pequeña porción de los ARNm están siendo regulados a lo largo del ciclo de vida del parásito, como en *T. brucei*, en donde mediante ARN-seq se ha descrito que el 5-6% del genoma está diferencialmente regulado entre los estadios sanguíneo y procíclico del parásito (Siegel et al. 2010). Sin embargo, otros sugieren que entre el 25-50% de los genes pueden ser diferencialmente expresados en algún estadio del tripanosoma, fenómeno que depende en gran medida de la abundancia de familias multigénicas y de secuencias repetidas en el genoma del parásito (De Gaudenzi et al. 2011; Ouellette and Papadopoulou 2009). La presencia de un ARNm en un estadio del ciclo de vida del parásito no garantiza su traducción a proteína. Una vez procesado, el ARNm maduro necesita ser exportado al citoplasma para ser traducido, protegido hasta su traducción o degradado (De Gaudenzi et al. 2011; Ouellette and Papadopoulou 2009). En este nivel, las “RNA-binding proteins” (RBPs) juegan un papel trascendente, siendo capaces de regular la abundancia de ARNm y su traducción a proteínas (Fernandez-Moya and Estevez 2010).

### 2.3.3 *Pseudogenes*

#### 2.3.3.a *Aspectos generales*

Los pseudogenes originalmente se han definido como secuencias no funcionales de ADN genómico, originados a partir de genes funcionales (parentales). Se caracterizan por exhibir algunas características degenerativas que, teóricamente, van impedir su propia expresión, como lo son la aparición de codones de parada prematuros, la pérdida del codón de inicio, mutaciones que afectan y desplazan el marco de lectura y/o la falta o modificación de las regiones flanqueantes y/o regulatorias (Balakirev, E.S. and Ayala, F.J. 2003; D'Errico, Gadaleta, Saccone 2004; Salmena 2014).



Si la presencia de una única copia de un gen parental es suficiente para cubrir las necesidades de un organismo, se asume que las mutaciones en el pseudogén derivado (discapacitantes o no) son neutras. Así los pseudogenes tenderían a degenerar por la rápida acumulación de mutaciones recurrentes, fenómeno que ha sido observado en los organismos procariotas donde los pseudogenes representan el paso previo a la muerte de los genes funcionales. No obstante, el genoma de los eucariotas contiene un gran número de pseudogenes que parecen haber evitado una degeneración completa, pudiendo evolucionar hasta adquirir un papel funcional como reguladores postranscripcionales de la expresión génica o generadores de diversidad genética (anticuerpos, antígenos, entre otros) (Balakirev, E.S. and Ayala, F.J. 2003; D'Errico, Gadaleta, Saccone 2004; Salmena 2014; Wen et al. 2011). Los alelos mutantes que incorporan nuevas funciones van a ser favorecidos por la selección natural y por lo tanto fijados en la población, así los pseudogenes son considerados “protogenes”, secuencias de ADN con la capacidad potencial de convertirse en nuevos genes, naciendo funcionalmente activos o no, y siendo capaces incluso de revertir su estado hasta volver a ser el gen del que proceden (Balakirev, E.S. and Ayala, F.J. 2003).

En base a como se originan los pseudogenes pueden clasificarse en pseudogenes no procesados, pseudogenes procesados y pseudogenes unitarios:

- **Pseudogenes no procesados (duplicados)**

Los pseudogenes procesados se originan por mecanismos de duplicación génica o por fenómenos inusuales de entrecruzamiento. En cualquier caso se va a producir una copia extra del gen funcional de partida (Figura 1.8). Esta copia, innecesaria y por lo tanto libre de la acción de la selección natural, puede acumular mutaciones sin dañar al organismo. Esta clase de pseudogenes normalmente se encuentra en el mismo cromosoma, a menudo adjuntos a los genes parálogos de los que proceden. En la mayoría de eucariotas, los pseudogenes no procesados, van acompañados de intrones y regiones regulatorias flanqueantes, igual que el gen funcional parental (D'Errico, Gadaleta, Saccone 2004; Wen et al. 2012); esto es algo que no sucede en los pseudogenes no procesados de *T. cruzi* donde, como se ha comentado anteriormente, la transcripción es policistónica y ningún elemento génico

presenta intrones (Donelson and Zeng 1990; Smircich et al. 2015). La mayoría de los genes duplicados van a terminar siendo secuencias no funcionales. Alternativamente, algunos pseudogenes pueden continuar evolucionando y adquirir mutaciones que les lleven a poseer nuevas funciones o modos de expresión (Balakirev, E.S. and Ayala, F.J. 2003; D'Errico, Gadaleta, Saccone 2004; Salmena 2014; Wen et al. 2011).

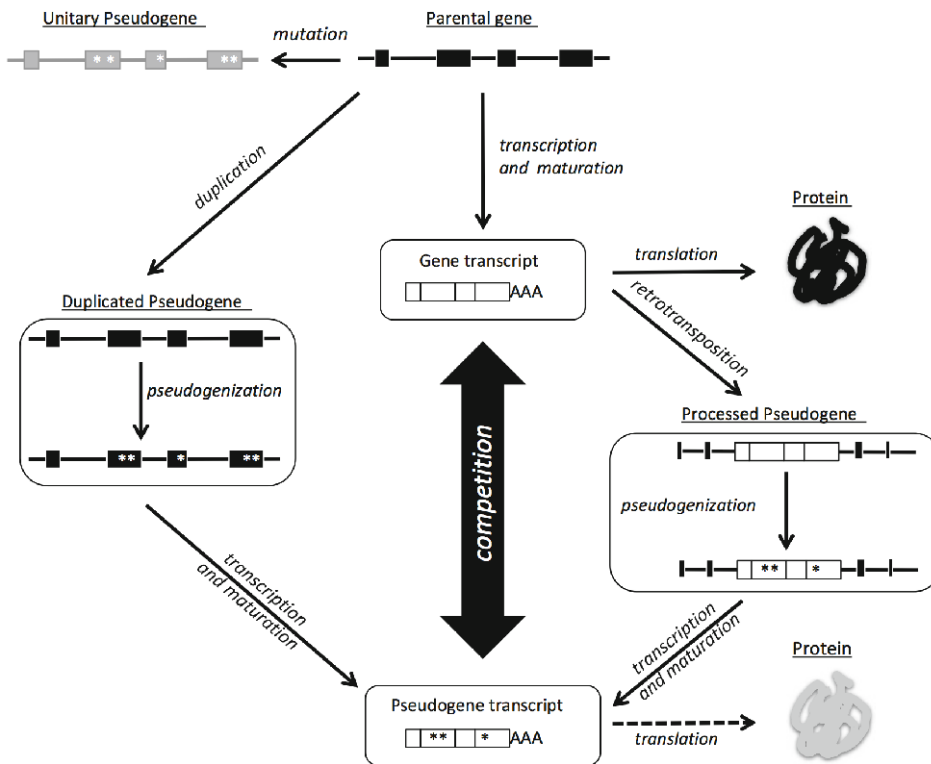
- **Pseudogenes procesados (retrotranscritos)**

Este tipo de pseudogenes surge mediante un proceso de retrotransposición que tiene lugar a partir de ARNm, originándose una secuencia de ADN de doble hélice que puede ser insertada en cualquier punto del genoma (Figura 1.8). Tras la retrotranscripción, los pseudogenes tienden a adoptar una posición inapropiada para su expresión a proteína (D'Errico, Gadaleta, Saccone 2004; Salmena 2014; Wen et al. 2012). Adicionalmente, en el proceso de retrotranscripción tienden a acumularse bastantes errores, produciéndose muchos cambios entre el ARN molde y el ADN complementario (ADNc) (D'Errico, Gadaleta, Saccone 2004).

Los pseudogenes procesados son más comunmente estudiados, probablemente, debido a su gran presencia en los genomas, ya que esta clase de pseudogenes son entre 3 y 4 veces más abundantes que los no procesados (Salmena 2014). Al originarse a partir de ARNm, estas secuencias pseudogénicas van carecer de promotores e intrones (D'Errico, Gadaleta, Saccone 2004; Salmena 2014). Este hecho facilita su identificación en la mayoría de los organismos eucariotas, pero no en *T. cruzi*, caracterizado por su transcripción policistronica y por carecer de intrones (Donelson and Zeng 1990; Smircich et al. 2015). Así, incluso careciendo de regiones controladoras, los pseudogenes procesados pueden llevar a cabo la transcripción haciendo uso de otros elementos transcripcionales que no tienen porque estar próximos necesariamente. Estas secuencias también se caracterizan por estar flanqueadas por repeticiones directas y la presencia de una cola poliA en su extremo 3' (D'Errico, Gadaleta, Saccone 2004; Wen et al. 2012).

- **Pseudogenes unitarios (discapitados)**

En la actualidad, se ha considerado la existencia de una tercera clase de pseudogenes, conocidos como pseudogenes discapacitados o unitarios, que a diferencia de las otras dos categorías, procesados y no procesados, no son producto de una amplificación genómica o de un evento de inserción, sino que se originan por la acumulación de una o varias mutaciones no sinónimas sobre el mismo gen parental (Figura 1.8) (Salmena 2014).



**Figura 1.8. Mecanismos para la formación de pseudogenes.** \*mutaciones, discapacitantes o no, que se producen en la secuencia pseudogénica respecto a la secuencia del gen parental (Salmena 2014).

### 2.3.3.b Expresión y funcionalidad

Recientemente, numerosos estudios han demostrado la capacidad de transcripción de los pseudogenes, teniendo gran presencia no sólo en el genoma sino también en el transcriptoma de varios organismos (Kalyana-Sundaram et al. 2012; Korneev, Park, O'Shea 1999; Poliseno et al. 2010; Salmena 2014). Así, estos

transcritos deben cumplir alguna función en la célula que le compense el gasto de energía que supone realizar el proceso de transcripción.

Como se esperaba, en un pseudogén no funcional y por lo tanto no sujeto a la selección natural, los índices de mutaciones sinónimas y no sinónimas son iguales (Balakirev, E.S. and Ayala, F.J. 2003). Al respecto, en *Drosophila melanogaster*, se ha descrito un pseudogén no procesado llamado  $\psi$  *Adh*. La evolución molecular de  $\psi$  *Adh* presenta características atípicas: a) la velocidad de evolución es sustancialmente más pequeña en los exones que en las regiones intergénicas, siendo sólo ligeramente mayor que en los exones funcionales del resto de genes *Adh*; b) el uso de los mismos codones es conservado para la mayoría de las especies; c) el índice de sustituciones silenciosas (Ks) es mayor que el índice de sustituciones reemplazadoras (Ka), situación que suele ser inversa para un pseudogén (asumiendo que esté libre de constricciones selectivas) y que en este caso el Ks/Ka es sólo un poco menor que en el resto de genes *Adh* (Sullivan et al. 1994). Todos estos datos reflejan la probable funcionalidad de este pseudogén.

El sentido biológico de la transcripción desde secuencias pseudogénicas radica en el papel regulatorio que estos transcritos pueden desempeñar sobre la expresión de otros genes, siendo una fuente de pequeños ARNs reguladores (Guo et al. 2009; Wen et al. 2011). Así, se ha demostrado como los pseudogenes cubren un importante papel como estructuras codificantes de ARN antisentido capaces de hibridar con el ARN sentido de aquellos genes con los que mantienen una alta homología; de esta manera se establece un dúplex ARN-ARN capaz de bloquear la expresión del gen, presentándose así como una fuente muy importante de pequeños ARNs reguladores (Balakirev, E.S. and Ayala, F.J. 2003; Korneev, Park, O'Shea 1999).

Se han podido identificar un gran número transcritos codificados a partir de pseudogenes en un gran número de organismos. En humanos, entre el 5-20% de los pseudogenes han mostrado evidencias de transcripción (Yano et al. 2004; Zheng et al. 2005). Algunos de los muchos pseudogenes funcionales descritos en el ser humano son  $\Psi$  *Little ROCK* (ROCK1P1),  $\Psi$  *GS*,  $\Psi$  *P450c21A*,  $\Psi$  *DRD5-1* o  $\Psi$  *PTEN* (Balakirev, E.S. and Ayala, F.J. 2003; Bristow et al. 1993; Montefusco et al. 2010; Nguyen et al. 1991; Poliseno et al. 2010). En el reino vegetal, además de

$\Psi$ TGG3, descrito en *Arabidopsis thaliana* (Zhang et al. 2002), se ha visto la expresión de  $\Psi$  *rps19* en *Oenothera berteriana* (Balakirev, E.S. and Ayala, F.J. 2003). En protozoos, el ejemplo más claro lo constituye el pseudogén que codifica para la proteína de unión a eritrocitos  $\Psi$  *EBA165* de *Plasmodium falciparum* (Balakirev, E.S. and Ayala, F.J. 2003). En *T. brucei* destaca también los *siRNAs* (RNAs small interference) generados a partir de la transcripción de pseudogenes. Estos pequeños ARNs no codificantes van a ser capaces de regular la expresión y traducción de diferentes genes en este parásito (Wen et al. 2011).

Adicionalmente, se han descrito varios casos de proteínas traducidas desde pseudogenes, aunque esto es algo menos común que los fenómenos de transcripción desde pseudogenes (Li, Yang, Wang 2013; Pink et al. 2011; Poliseno 2012). Al respecto, uno de los primeros casos descritos fue *PGAM3*, cuyo pseudogen se genera por retrotransposición a partir del gen *PGAM1* y se localiza en un intrón del gen de la enfermedad de Menkes (Betran et al. 2002). Otro ejemplo de pseudogen traducido a proteína lo constituye  $\Psi$ *BRAF*, el cual codifica para una proteína truncada de 244 aa capaz de activar la vía MAP-kinasa específica del cáncer de tiroides (Zou et al. 2009).

Por otra parte, se ha descrito la gran variabilidad que existe en el número de pseudogenes presentes en un genoma según el organismo analizado. Esta cifra es incluso más variable que la que se encuentra en genes y no presenta una relación directa ni con el número total de genes, ni con el tamaño del genoma (Tabla 1.2). El número de pseudogenes va a depender de los índices relativos de duplicación génica y de pérdida de pseudogenes (Balakirev, E.S. and Ayala, F.J. 2003; D'Errico, Gadaleta, Saccone 2004; Guo et al. 2009; Podlaha and Zhang 2010).

### 2.3.3.c *Pseudogenes en T. cruzi*

Es interesante observar como el organismo en el que se han detectado un mayor número de pseudogenes es el ser humano, con más de un 33% respecto al número total de genes. Dato que da una idea del complejo papel funcional que pueden desempeñar los pseudogenes. Por su parte, en la cepa CL Brener de *T. cruzi*, los pseudogenes representan un 16% del número total de genes, siendo el tripanosomátido que contiene un mayor porcentaje de éstos (Tabla 1.2).

Tabla 1.2. Porcentaje de pseudogenes en diferentes organismos, incluido *T. cruzi*.

Organismo	Tamaño del genoma diploide	Nº genes (G)	Nº pseudogenes procesados ( $\Psi G$ )	% $\Psi G/G$
<i>C. elegans</i>	100Mb	22,941	104	0.5%
<i>D. melanogaster</i>	128Mb	13,525	34	0.25%
<i>M. musculus</i>	2.6Gb	26,762	5,000	18.38%
<i>H. sapiens</i>	2.8Gb	23,531	8,000	33.41%
<i>T. brucei</i>	50Mb	~18,136	~1,808	9.96%
<i>L. major</i>	66Mb	~16,622	~68	0.04%
<i>T. cruzi</i>	<b>106.4-110.7Mb</b>	<b>22,570 (1,377)*</b>	<b>3,590 (433)*</b>	<b>15.9% (1.92%)*<sup>a</sup> (31.45%)*<sup>b</sup></b>

\*Datos referentes a la familia multigénica MASP donde <sup>a</sup> representa el porcentaje de pseudogenes MASP respecto al número total de genes en *T. cruzi*; <sup>b</sup> es el porcentaje de pseudogenes MASP respecto al número total de genes MASP.

Es en las familias multigénicas de *T. cruzi* donde se encuentran la mayor parte de los pseudogenes presentes en este parásito. La familia TS, con un mayor número de genes, es también la que presenta un mayor número de pseudogenes (693), seguida por la familia Retrotrasposon “hot spot RHS proteins” (557), la familia MASP (433), la proteasa de superficie gp63 (251), Mucinas (201) y la familia DGF-1 (136) (El-Sayed et al. 2005b). Estas secuencias pseudogénicas pueden ser transcritas gracias a la naturaleza policistrónica de *T. cruzi* a la vez que van a contribuir al aumento del repertorio genético del parásito mediante recombinación. Otro motivo por el que los pseudogenes se mantienen en el genoma y forman parte de estas familias multigénicas, es por la diversidad antigénica que son capaces de generar (Vihinen 2014).

Los retrotransposones constituyen un mecanismo generador de pseudogenes. Como se describió anteriormente, las familias multigénicas de *T. cruzi* se ubican en zonas subteloméricas, fuera de la zona de sintenia. Estas zonas subteloméricas son puntos calientes para la inserción de retrotransposones. La

inserción de fragmentos de ADN en estas zonas tiende a originar nuevos pseudogenes en las familias multigénicas descritas. Al respecto, es importante remarcar la asociación existente entre el retrotransposon no autónomo TcTREZO y los genes MASP (Souza et al. 2007).

#### **2.3.4 Retrotransposones**

Los retrotransposones son elementos ubicuos de ADN capaces de transponerse por el genoma de un organismo a través de un ARN intermediario. En cambio, los transposones se mueven por el genoma basándose únicamente en el propio ADN (Fujiwara 2015). Los retrotransposones se clasifican en elementos que presentan largas repeticiones terminales en ambos extremos (LTRs), y en aquellos que carecen de esas repeticiones, no-LTR. Estos dos tipos de retroelementos presentan diferentes mecanismos de integración en el genoma. Los retrotransposones LTR son similares, en estructura y mecanismo de retrotransposición, a los retrovirus, utilizando un mecanismo más sencillo de integración (Bringaud et al. 2006; Fujiwara 2015). Por su parte los retrotransposones no-LTR van a codificar una endonucleasa, que va a romper el ADN diana, generando un extremo libre 3' hidroxilo el cual es utilizado como cebador para iniciar un proceso de retrotranscripción (transcripción reversa). La síntesis de la hebra complementaria de la nueva copia de ADN del retrotransposon se lleva a cabo mediante una proteína reverso transcriptasa codificada por él mismo. Algunos pares de bases aguas debajo de la primera rotura, la misma endonucleasa se va a encargar de realizar una segunda rotura de ADN en la otra hebra. Esto va a suponer otro cebador disponible para la síntesis de la segunda cadena del retroelemento. De esta manera, los retrotransposones están flanqueados por una repetición directa, denominada sitio diana de duplicación, que se corresponde con la secuencia situada entre los dos sitios de rotura de la cadena sencilla de ADN (Bringaud, Berriman, Hertz-Fowler 2009; Fujiwara 2015).

Mientras que no se ha descrito la presencia de transposones en *T. cruzi*, los retroelementos LTR y no-LTR constituyen los elementos móviles más abundantes descritos en *T. cruzi*, representando el 5% del genoma haploide de

este parásito y el 2% del genoma haploide de *T. brucei*. Algunas copias de los no-LTR retrotransposones CZAR y SLACS, con sitio de inserción específico, se han visto en los loci codificantes para péptido señal en *T. cruzi* y *T. brucei* pero no en *L. major*. No obstante, los elementos no-LTR más abundantes son L1Tc y NARTc, con 320 copias de L1Tc y 133 de NARTc por genoma haploide (El-Sayed et al. 2005b; Souza et al. 2007).

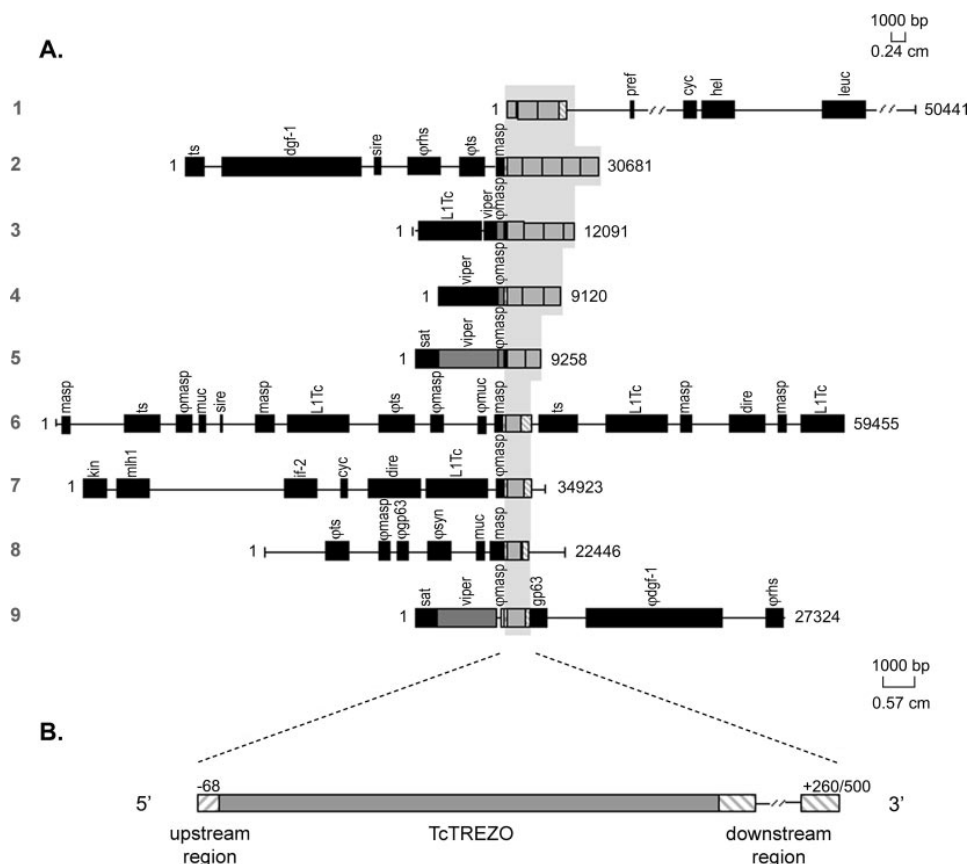
Las repeticiones representan una fuente extraordinaria de información acerca de los procesos biológicos que acontecen, funcionando como un archivo molecular que narra todos los eventos evolutivos que se dan en el genoma. Se ha sugerido que los elementos de ADN repetitivo pueden estar involucrados en reordenamientos genéticos y control de la expresión genética en tripanosomátidos. La mayoría de las secuencias repetidas son específicas de una especie (Souza et al. 2007). Estudiando el ADN repetitivo de *T. cruzi* es como se identificó recientemente el retroelemento "sitio-específico" de estructura repetida denominado TcTREZO (*Trypanosoma cruzi tandem repetitive element ZO*). Con un tamaño de ~ 1.6kb, TcTREZO consiste en una secuencia derivada de diferentes regiones del genoma, que se caracteriza estructuralmente por la presencia de un núcleo central compuesto por la repetición en tándem de una secuencia de 55 pb y que a su vez está flanqueado por dos dominios diferentes de 995 pb y 254 pb, que se muestran altamente conservados. Algunas evidencias sugieren que TcTREZO sería un retrotransposon no autónomo, no-LTR, que emigra a través de un ARN intermediario. Análisis northern blot y RT (Reverse Transcriptase)-PCR muestran que los transcritos de TcTREZO son expresados como transcritos con una cola oligoA. Los transcritos de estas secuencias no codifican ninguna proteína, sugiriendo que no se traducen. El transcrito oligoA obtenido del extremo 3' de TcTREZO podría generar un cebador sobre el que actuara una transcriptasa inversa produciendo una copia de ADNc de TcTREZO que se insertaría en el genoma. El número de copias de secuencias de TcTREZO estimado en el genoma haploide de *T. cruzi* es de 173 (Souza et al. 2007). Dicho retrotransposon parece haberse originado por inserciones de secuencias sobre un elemento progenitor (Souza et al. 2007).

Los elementos TcTREZO se han encontrado en largas regiones no sinténicas del genoma de *T. cruzi*, intercalándose principalmente en el extremo 3' de genes



y pseudogenes MASP además de en TS, Gp63, mucinas, DGF-1, los genes RHS y en retrotransposones como VIPER, L1Tc, SIRE y DIRE (Figura 1.9), sugiriéndose que existe un punto caliente de recombinación homóloga en todos ellos. Así, Souza *et al.* (2007) demostraron que TcTREZO se encuentra flanqueado por secuencias conservadas (ver Materiales y Métodos, apartado 9), indicando que efectivamente se trata de un elemento con sitios de inserción específicos en el genoma de *T. cruzi*. De esta manera, el sitio de unión específico de TcTREZO en los genes MASP podría dejar “arrays” en tándem de este elemento.

Una relación muy parecida a la de TcTREZO con la familia multigénica MASP se describió en *T. brucei* entre los retrotransposones no-LTR ingi/RIME y los genes RHS (Bringaud *et al.* 2004).



**Figura 1.9. Representación esquemática y localización de un retrotransposon TcTREZO.** (A) Detalle de los contigs de *T. cruzi* que contienen TcTREZO (cajas grises). La longitud (en pb) se indica a la derecha de cada contig. Las cajas negras son los diferentes locus comprendidos en cada contig. (B) Esquema representativo de la estructura típica de TcTREZO (Souza *et al.* 2007).

## 2.4 Filogenia y diversidad genética

Varios análisis bioquímicos, biológicos y moleculares muestran que *T. cruzi* posee un gran número de polimorfismos genéticos, siendo un taxón altamente heterogéneo. Por ello la clasificación de este parásito en grupos discretos acorde con sus variantes moleculares y eco-epidemiológicas supone un reto complejo que ha evolucionado desde sus primeros pasos (Brisse, Dujardin, Tibayrenc 2000; Miles et al. 1977; Miles et al. 1978; Miles et al. 1980; Souto and Zingales 1993; Souto et al. 1996; Tibayrenc, Kjellberg, Ayala 1990) hasta la clasificación actual (Zingales et al. 2009; Zingales et al. 2012).

La genotipificación por multilocus permitió clasificar a *T. cruzi* en distintas DTUs que en un principio fueron englobadas en dos subgrupos principales, denominados DTU I y DTU II. Los DTUs son definidos como “conjuntos de stocks genéticamente más relacionados entre ellos que con otro stock e identificables por marcadores genéticos, moleculares e inmunológicos comunes” (Tibayrenc 1998). De esta manera, cepas de *T. cruzi* que estén incluidas en un mismo DTU no tienen porque ser genéticamente idénticas.

Como se ha observado, desde 1999 hasta la actualidad, la comunidad científica ha adquirido un mayor conocimiento acerca de la diversidad de *T. cruzi*. Por ese motivo, en 2009 se celebró el “Second Satellite Meeting” en el que el comité reestructuró los DTUI y DTUII en seis nuevos DTUs, desde *T. cruzi* I hasta *T. cruzi* VI (Tabla 1.3).

**Tabla 1.3. Nomenclatura actual para las divisiones de *T. cruzi* (Zingales et al. 2009).**

DTU designation	Abbreviation	Equivalence to former <i>T. cruzi</i> grouping schemes
<i>T. cruzi</i> I	TcI	<i>T. cruzi</i> I <sup>a,b</sup> and DTU I <sup>c</sup>
<i>T. cruzi</i> II	TcII	<i>T. cruzi</i> II <sup>a</sup> and DTU IIb <sup>c</sup>
<i>T. cruzi</i> III	TcIII	Z3/Z1 ASAT <sup>d</sup> , Z3-A <sup>e</sup> , DTU IIc <sup>e</sup> and <i>T. cruzi</i> III <sup>f</sup>
<i>T. cruzi</i> IV	TcIV	Z3 <sup>d</sup> , Z3-B <sup>e</sup> and DTU IIa <sup>e</sup>
<i>T. cruzi</i> V	TcV	Bolivian Z2 <sup>d</sup> , rDNA 1/2 <sup>e</sup> , clonet 39 <sup>h</sup> and DTU II <sup>d</sup> <sup>e</sup>
<i>T. cruzi</i> VI	TcVI	Paraguayan Z2 <sup>i</sup> , Zymodeme B <sup>j</sup> and DTU II <sup>e</sup> <sup>c</sup>

a: Anonymous 1999; b: Falla et al. 2009; c: Brisse et al. 2000; d: Miles et al. 1981; DTU: discrete typing units; e: Mendonça et al. 2002; f: Freitas et al. 2006; g: Souto et al. 1996; h: Tibayrenc and Ayala 1991; i: Chapman et al. 1984; j: Carneiro et al. 1990.

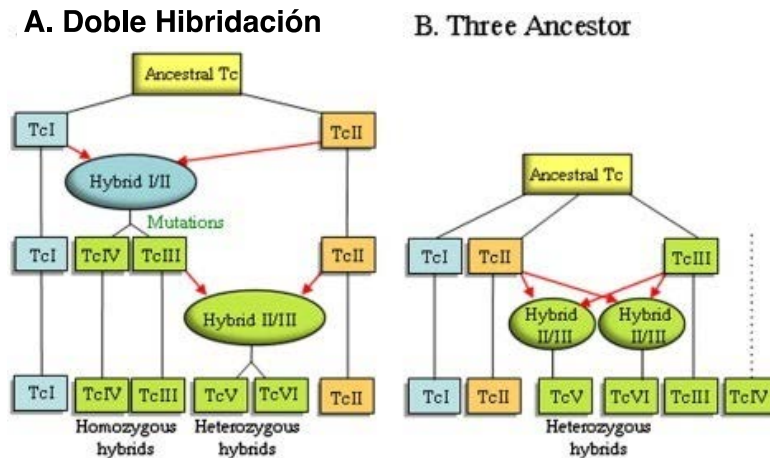
Dentro de esta estructuración intraespecífica de *T. cruzi*, CL Brener, originalmente clasificada como *T. cruzi* IIe, ha sido reclasificada como *T. cruzi* VI,

siendo confirmada su naturaleza híbrida (Pena, Machado, Macedo 2009; Zingales et al. 2009).

Briones y colaboradores relacionaron a TcI con la fauna nativa de Sudamérica mientras TcII queda vinculada con los mamíferos placentarios de Norteamérica. Así se estableció la idea de que los linajes principales de *T. cruzi* permanecieron separados entre 88 y 37 millones de años atrás, divergiendo hasta considerarse dos especies diferentes, o por lo menos dos subespecies, con características epidemiológicas y ecológicas particulares. La mezcla entre estos dos grupos ancestrales no se produjo hasta la unión de las dos Américas. Durante el Oligoceno y el Plioceno, TcII fue capaz de dispersarse por Sudamérica gracias a roedores, primates y fauna de mayor tamaño. Mientras que TcI hizo lo mismo por Norteamérica, en su caso, mediante la difusión de los marsupiales por este continente. Todo esto explicaría la asociación del linaje 2 con los marsupiales y del linaje 1 con la enfermedad en humanos (Briones et al. 1999).

Es importante mencionar como han surgido crecientes evidencias de que en la variabilidad genética de *T. cruzi*, inicialmente atribuida a la evolución clonal, el intercambio genético de los parásitos ha contribuido a formar la estructura poblacional presente (Tibayrenc et al. 1986; Zingales et al. 2009; Zingales et al. 2012). Por definición, un clado es un concepto que representa unidades evolutivas que están estrictamente aisladas. Sin embargo, la consideración de los fenómenos de intercambio genético en la evolución de *T. cruzi* hizo que fuera necesario desarrollar el concepto de unidades discretas de tipificación (Zingales et al. 2012). La heterocigosis de algunas cepas de *T. cruzi* ha sido detectada gracias al análisis de genes individuales. Así se observó una heterocigosis pronunciada en aislados naturales de TcV y TcVI, lo que sugiere la naturaleza híbrida de estas DTUs, originadas a partir de TcII y TcIII (Sturm and Campbell 2010; Zingales et al. 2012). Los análisis alélicos realizados sobre el resto de DTUs, TcI, TcII, TcIII, TcIV mostraron un alto grado de homocigosis (Zingales et al. 2012). Al respecto, se han descrito dos modelos posible que explican los sucesos de hibridación sobre la estructura clonal de las poblaciones de *T. cruzi*: i) Doble Hibridación (Westenberger et al. 2005) y ii) Árbol Ancestro (de Freitas et al. 2006). Aunque, en ambos modelos se incorporan dos eventos de hibridación, el modelo de Doble Hibridación defiende que la existencia de un antiguo

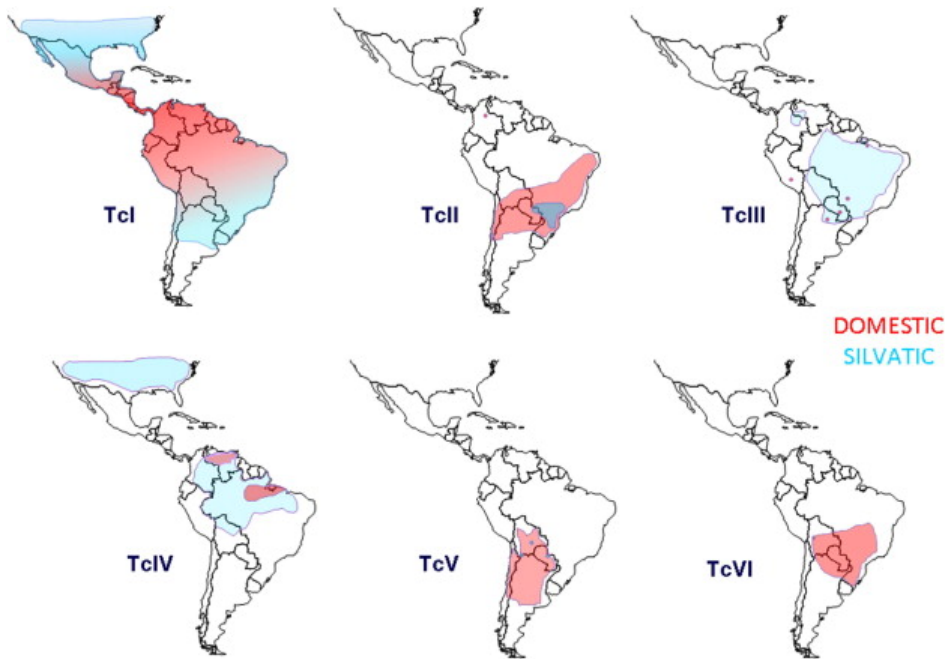
intercambio genético entre TcI y TcII y la pérdida de heterocigosidad en la progenie produjeron la aparición de TcIII y TcIV; seguidamente, una segunda y más reciente hibridación entre TcII y TcIII originó TcV y TcVI (Figura 1.10A). Por otra parte, el modelo de Árbol Ancestro indica que estos dos sucesos de intercambio genético fueron recientes y que ambos se produjeron entre TcII y TcIII originando TcV y TcVI (Figura 1.10B) (Zingales et al. 2012).



**Figura 1.10. Comparación de los modelos de (A) Doble Hibridación y (B) Árbol Ancestro para explicar los mecanismos de intercambio genético durante la evolución clonal de *T. cruzi*.** Los rectángulos indican diferentes DTUs. La fusión de dos células y el consecuente intercambio genético es indicada mediante óvalos, con la contribución parental indicada por las flechas rojas. Los clados mitocondriales están rellenos de diferentes colores: Azul = clado A; Verde = clado B; naranja = clado C (Zingales et al. 2012).

Existen numerosas evidencias experimentales que muestran como las DTUs de *T. cruzi* exhiben un comportamiento diferencial respecto a su virulencia (capacidad de multiplicarse en el hospedador) y patogenicidad (habilidad para producir lesiones en los tejidos y desencadenar una respuesta inmunológica) (Andrade 1974; Andrade et al. 1999; Andrade, Brodskyn, Andrade 1983; Postan, McDaniel, Dvorak 1987). Pese a todo no existe un marcador genético ligado a la severidad de la infección. TcI, TcII, TcV y TcVI constituyen los principales agentes de la enfermedad de Chagas para los humanos en Latinoamérica, siendo

cada uno de estos DTUs capaz de producir cardiomiopatías, mientras sólo TcII, TcV y TcVI han sido también asociados a los síndromes digestivos de la enfermedad (Zingales et al. 2012). Dentro de este marco, resalta TcI como el DTU más abundante en el continente, el cual puede estar asociado con los ciclos salvajes y domésticos de la enfermedad (Figura 1.11) (Zingales et al. 2012).



**Figura 1.11. Distribución geográfica aproximada de las diferentes DTUs de *T. cruzi*, con los dos ciclos de transmisión, selvático y doméstico (Zingales et al. 2012).**

Incluso dentro de un mismo DTU puede haber un alto grado de polimorfismo tal y como se demostró en TcI. Para ello, Cura *et al.* (2010) caracterizaron un fragmento de 350pb incluido en la región intergénica del péptido señal sobre 105 muestras y aislados de *T. cruzi* I procedentes desde Argentina hasta USA, permitiendo la caracterización de 76 genotipos y 54 sitios polimórficos. Así, establecieron la existencia de cuatro subgrupos TcI SL-IR: Ia (11 genotipos), Ib (11 genotipos), Id (35 genotipos) y Ie (19 genotipos).

Por último y como comentamos en apartados anteriores, en la actualidad se ha propuesto también la existencia de una séptima rama en la filogenia y taxonomía de *T. cruzi*, nombrada Tcbat y procedente de murciélagos en la que se encuentra la cepa Marinkellei (Marcili et al. 2009).

## 2.5 Bases moleculares y mecanismos de entrada en la célula hospedadora

### 2.5.1 Aspectos generales

La mencionada variabilidad genética que existe en *T. cruzi* se traduce en un alto grado de heterogeneidad en el proceso de invasión celular llevado a cabo por el parásito, donde cada paso de este proceso puede variar teniendo en cuenta la combinación única que se establece entre la cepa del parásito, su estado morfológico y el tipo de célula hospedadora (Epting, Coates, Engman 2010).

La capacidad de infectar y replicarse en el interior de una amplia variedad de tipos celulares en el hospedador mamífero, es una característica esencial de *T. cruzi* que reside en el amplio espectro de proteínas presentes en su superficie procedentes del genoma altamente complejo (Burleigh and Woolsey 2002). Tan sólo las células sin núcleo (Osuna et al. 1983) o células nucleadas que hayan sido pretatadas con actinomicina D (Mester et al. 1987), se van mostrar resistentes a la invasión por las formas infectivas de *T. cruzi*.

Aislar las moléculas que intervienen en el establecimiento de la infección del parásito es fundamental para comprender la patogénesis de la enfermedad de Chagas. Los modelos de infección *in vitro* (invasión y replicación intracelular) son muy útiles para el estudio de las bases moleculares y celulares de la interacción parásito-célula hospedadora (Burleigh and Woolsey 2002). Así, al contrario que determinados parásitos intracelulares, parece que *T. cruzi* no internaliza en la célula a través de un proceso de fagocitosis que dependa exclusivamente de actina (Andrade and Andrews 2004).

Una vez iniciada la infección en el citoplasma de la célula hospedadora, se forma una vacuola, en el interior de la cual el parásito va a permanecer de forma transitoria (Ley et al. 1990). Se ha demostrado que *T. cruzi* invade con mayor eficiencia aquellas células que se encuentren en la fase S del ciclo celular; durante este período la célula hospedadora presenta un potencial de membrana más adecuado y expresa un mayor nivel de componentes necesarios para el reconocimiento, unión y entrada del parásito (Osuna et al. 1984a).

### ***2.5.2 Modos de internalización: Movilización de Calcio y lisosomas y despolimerización de la actina***

El estadio tripomastigote de *T. cruzi* es relativamente largo (~20  $\mu\text{m}$  de longitud) y sin capacidad de división, siendo el encargado de establecerse en el interior del citoplasma de la célula hospedadora y diferenciándose posteriormente en un amastigote que va a llevar a cabo la división.

En células fagocíticas, *T. cruzi* induce su endocitosis al interior de los macrófagos. Aquí el proceso de internalización se produce gracias a la interacción de las proteínas de superficie del parásito con proteínas de la superficie del macrófago y depende de la acción de proteínas receptoras de tipo *Toll Like Receptor* (TLR) (Epting, Coates, Engman 2010; Kuhn 1994; Maganto-Garcia et al. 2008). *T. cruzi* interactúa con TLR2 a través de moléculas ricas en glicosilfosfatidilinositol (GPI) regulando su entrada en el macrófago a través de la activación de fosfatidilinositol 3-kinase (PI3K). Tras esto, PI3K va a activar a la proteína de membrana del macrófago Rab5, una proteína perteneciente a la familia de las Rab GTPasas que se va a unir a endosomas tempranos, constituyendo un paso clave para la entrada de *T. cruzi* en los macrófagos (Maganto-Garcia et al. 2008). En la actualidad se han descrito 13 TLRs, 10 de los cuales han sido descritos en células humanas. Además del TLR2, el TLR4 y TLR9 también se asocian con el proceso de reconocimiento de *T. cruzi* y posterior respuesta inmune (Epting, Coates, Engman 2010; Kuhn 1994; Maganto-Garcia et al. 2008).

Por otra parte, durante el proceso de entrada en células no fagocíticas, el tripomastigote induce un aumento de Calcio ( $\text{Ca}^{2+}$ ) libre debido a la liberación del catión desde sus depósitos intracelulares como son el retículo endoplasmático, el núcleo y la mitocondria (Osuna et al. 1990; Scharfstein et al. 2000; Tardieux, Nathanson, Andrews 1994). La salida de  $\text{Ca}^{2+}$  del retículo endoplasmático, depende de diversas señales inducidas por el parásito que van dirigidas a activar la enzima Fosfolipasa C (PLC) de la célula hospedadora, lo que conlleva un aumento de los niveles de IP3 en el interior de la misma. Entre todas estas señales procedentes de *T. cruzi*, también se ha visto como las proteínas Cruzipaína (CZ) y la Oligopeptidasa B (OB), participan activamente en

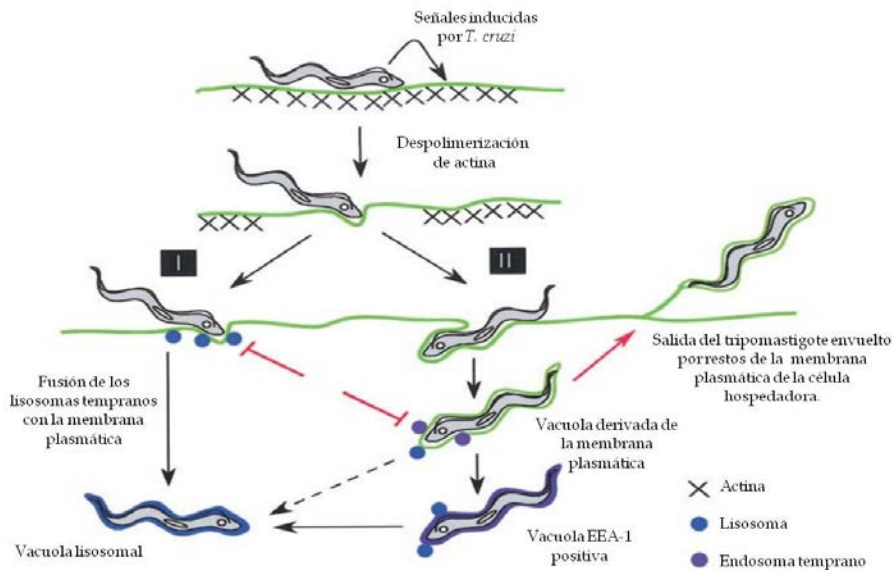
el proceso. Una vez secretada, la CZ, se encarga de romper Kininógeno de alto peso molecular para obtener Kininas las cuales se van a unir a la célula hospedadora a través de sus receptores de Bradikinina para estimular la salida de  $\text{Ca}^{2+}$  dependiente de  $\text{IP}_3$  (Scharfstein et al. 2000). Por su parte, la OB citosólica actúa sobre un precursor del parásito agonista de  $\text{Ca}^{2+}$  (Caler et al. 1998). Finalmente, también se ha demostrado la participación de la glicoproteína gp82 y de proteínas ricas en Serina-Alanina-Prolina (SAP) como moléculas capaces de estimular la liberación de  $\text{Ca}^{2+}$  en la célula hospedadora (Baida et al. 2006; Ruiz et al. 1998).

Tras la liberación del  $\text{Ca}^{2+}$ , *T. cruzi* en su fase infectiva, utiliza dos estrategias para penetrar en la célula y constituir la vacuola parasitófora en el transcurso de la invasión. El primer mecanismo que se describió se basa en el reclutamiento dirigido de lisosomas de la célula hospedadora hacia el punto exacto en el que el parásito va a realizar su entrada (Tardieux et al. 1992). A parte, se ha descrito una vía alternativa lisosoma-independiente para el tripomastigote de *T. cruzi* (Woolsey et al. 2003). Este mecanismo de entrada implica una invaginación inducida de la membrana plasmática de la célula hospedadora y la posterior formación de una vacuola derivada de la misma membrana, la cual en un primer momento se muestra desprovista de marcadores lisosomales (Woolsey and Burleigh 2004), asegurando así una entrada inicial ajena a los lisosomas. Las vacuolas de *T. cruzi* resultantes de la vía de entrada lisosoma-independiente sufren un proceso de maduración en el que tiene lugar la adquisición gradual de marcadores endosomales y/o lisosomales, originando el nacimiento de la vacuola parasitófora.

Es importante resaltar que los dos modos de internalización del tripomastigote, bien a través de la membrana plasmática del hospedador (vía lisosoma-independiente) o mediante el reclutamiento lisosomal (vía lisosoma-dependiente), coexisten produciéndose simultáneamente sobre diferentes tipos celulares (Woolsey et al. 2003). Por ello, una herramienta para distinguir la vía de entrada que toma el parásito es el análisis de marcadores específicos de los compartimentos vinculados a cada una de las dos rutas. La cuantificación de la asociación relativa de *T. cruzi* con esos marcadores revela que solo el 20-30% de los parásitos recién internalizados están asociados con los lisosomas, mientras



~50% presentan marcadores de la membrana plasmática de la célula hospedadora. El 20-30% restante fue encontrado dentro de vacuolas enriquecidas por marcadores endosomales tempranos, EEA1 (Woolsey et al. 2003). El hecho de que los EEA1 nunca se encontraron asociados con parásitos parcialmente internalizados, hace pensar que estos marcadores endosomales tempranos son reclutados por las vacuolas derivadas de la membrana plasmática después de su formación. Por ello, las vacuolas EEA1 positivas se consideran una extensión del proceso de entrada lisosoma-independiente (Figura 1.12) (Woolsey et al. 2003).



**Figura 1.12. Modelo integrado de la invasión de *T. cruzi* en células no fagocíticas.**

Los tripomastigotes de *T. cruzi* se unen a la superficie de la célula del hospedador mamífero activando las vías de señalización para la despolimerización transitoria de la actina del citoesqueleto cortical. La entrada del parásito se inicia por la invaginación de la membrana plasmática seguido de la fusión con lisosomas tempranos (vía I: entrada lisosoma-dependiente) o forma vacuolas derivadas de la membrana plasmática que posteriormente se fusionan con los endosomas o lisosomas (vía II: entrada lisosoma-independiente). También se representa la salida de los tripomastigotes cuando se impide el reensamblamiento de los microfilamentos de actina, por tratamiento con citocalasina D, bloqueando la fusión de lisosomas tempranos y endosomas con los tripomastigotes recién internalizados. Finalmente, se puede observar la convergencia entre la vía I y II en el compartimento lisosomal (Woolsey and Burleigh 2004).

En cualquier caso, sea cual sea la ruta por la que el parásito accede a la célula, la presencia transitoria de éste en el compartimiento lisosomal se considera imprescindible para el establecimiento de la infección intracelular (Ley et al. 1990) y para la diferenciación del tripomastigote en amastigote (Tomlinson et al. 1995). Así, la supervivencia de un parásito que invade una célula a través de la vía lisosoma-independiente va a depender de su capacidad para unirse a los lisosomas de la célula hospedadora (Woolsey and Burleigh 2004). Woolsey y colaboradores (2003) demostraron que, tras 60 minutos de internalización, el 90% de los parásitos estaban cubiertos por marcadores lisosomales. De esta manera, aunque la mayoría de los tripomastigotes de *T. cruzi* penetran en las células no fagocíticas del mamífero mediante el mecanismo lisosoma-independiente, éste, va a converger con la vía lisosoma-dependiente, produciéndose la fusión de las vacuolas derivadas de la membrana plasmática (con el tripomastigote en su interior) con los lisosomas de la célula hospedadora (Woolsey et al. 2003). Como se ha descrito, la exocitosis de lisosomas regulada por el calcio citosólico liberado desde los depósitos del retículo endoplasmático y de la mitocondria (Osuna et al. 1990), va a ser un fenómeno clave en la entrada del *T. cruzi*. El reclutamiento de lisosomas tiene lugar mediante un sistema activo kinesina-dependiente (Rodríguez et al. 1996; Tardieux et al. 1992).

En la célula hospedadora tiene lugar una marcada reorganización del citoesqueleto como respuesta al tripomastigote (Rodríguez et al. 1995; Tyler et al. 2005). Al penetrar en la célula, la dinámica normal de los microfilamentos y microtúbulos de la célula hospedadora se ve alterada y dirigida hacia la sinapsis célula-parásito. Para ello las moléculas de  $\alpha$  y  $\gamma$ -tubulina se asocian con la vacuola parasitófora que se acaba de formar (Tyler et al. 2005). El reordenamiento de estas moléculas ha sido relacionado con la despolimerización de los microtúbulos en la cara citoplasmática de la membrana de la célula hospedadora y también en las membranas de los lisosomas reclutados, permitiendo que se produzca la invaginación en la vacuola parasitófora (Tardieux et al. 1992). De esta manera, el citoesqueleto cortical de actina actuaría como una barrera para la fusión de los lisosomas con la membrana plasmática, por lo que su despolimerización y reordenamiento, inducidos por el aumento de  $\text{Ca}^{2+}$  citosólico, promueven la entrada de parásitos debido a un incremento en la

presencia de lisosomas asociados al punto exacto de la membrana donde el parásito se encuentra adherido, listos para fusionarse con el tripomastigote (Rodríguez et al. 1995; Tardieux et al. 1992; Woolsey and Burleigh 2004).

Por el contrario y sorprendentemente, existen varios estudios que asocian la despolimerización de los microfilamentos de actina a una mayor dificultad de *T. cruzi* para infectar la célula hospedadora (Ferreira et al. 2006; Schenkman, Robbins, Nussenzweig 1991; Tardieux et al. 1992). Tal discrepancia puede ser explicada por la propia variabilidad genética e incluso fenotípica de *T. cruzi*, la cual, puede traducirse en un proceso de invasión celular altamente heterogéneo. De esta manera, dependiendo de la cepa de *T. cruzi*, pueden variar algunas de las moléculas del parásito que interaccionan con la célula hospedadora (gp35/50 o gp82) (Ferreira et al. 2006) o incluso el mecanismo de entrada (vía lisosoma dependiente o independiente) (Epting, Coates, Engman 2010; Osuna et al. 1984a; Woolsey and Burleigh 2004). Por todo ello se explica como, pese a someter a las células hospedadoras a un pretratamiento con Citocalasina D (agente antagonista a la asociación de los tripomastigotes con los lisosomas) el número de parásitos intracelulares aumenta, demostrando la existencia de un segundo mecanismo de entrada, la vía lisosoma-independiente (Osuna et al. 1984a; Woolsey and Burleigh 2004).

Otro hecho inesperado se ha observado cuando los tripomastigotes permanecen en células con el citoesqueleto de actina comprometido; en esa situación se produce la salida de los parásitos desde las células (visualizada mediante un marcado de la membrana plasmática de la célula hospedadora por GFP (Woolsey and Burleigh 2004). La capacidad de que los tripomastigotes internalizados permanezcan en el interior celular está asociada con el reensamblamiento de la actina en la célula hospedadora (Woolsey and Burleigh 2004).

La invasión celular de *T. cruzi* mediada por la despolimerización de los microfilamentos de actina y el reclutamiento de lisosomas va a depender de fenómenos de señalización celular por parte de la célula hospedadora donde se ven implicadas proteínas como las PI-3 kinasas, proteínas Akt (tipo serin Threonin/kinasa) (Burleigh and Woolsey 2002; Todorov et al. 2000) y proteínas tipo Rho-GTPasa (Woolsey and Burleigh 2004).

## 2.6 Proteínas de superficie e invasión celular

### 2.6.1 Aspectos generales

Una de los aspectos más característicos de *T. cruzi* es que, en comparación con los otros kinetoplastidos, este parásito presenta una gran expansión de genes codificantes para proteínas de superficie, los cuales se encuentran fuera de la isla sinténica. Así, una gran parte de las interacciones de *T. cruzi* y del resto de tripanosomátidos con la célula hospedadora y la matriz extracelular ocurren a través de un grupo diverso de glicoproteínas de superficie y proteasas (Epting, Coates, Engman 2010). Muchas de estas glicoproteínas comparten el motivo GPI de unión a membrana. La adición de la molécula GPI consiste en una modificación postransduccional que se realiza en el lumen (interior) del retículo endoplasmático con el objetivo de adherir moléculas a la bicapa lipídica de las células eucariotas (Canepa et al. 2012). Como muchas otras, las proteínas que se unirán al motivo GPI, contienen un péptido señal que las conduce al retículo endoplasmático donde se procesa la unión de este glicolípido a la proteína. Posteriormente, la proteína con el anclaje GPI es llevada mediante transporte vesicular al aparato de Golgi, y finalmente a la membrana plasmática, resultando una proteína extracelular asociada a la membrana (Brown and Waneck 1992). Las estructuras y funciones de estas proteínas son increíblemente diversas participando en procesos de adhesión celular, comunicación célula-célula, señalización paracrina y diferenciación celular (Canepa et al. 2012; Epting, Coates, Engman 2010; Fujita and Jigami 2008). En tripanosomátidos se ha observado un gran número de proteínas con anclaje GPI (Bartholomeu et al. 2009; Ferguson et al. 1994; Field, Medina-Acosta, Cross 1991). Un claro ejemplo de proteínas con anclaje GPI lo constituyen las VSG de *T. brucei*, que ejercen un papel crucial en la evasión de la respuesta inmune (Ferguson et al. 1994). Por otra parte, en *T. cruzi* se ha vinculado a un gran número de proteínas con anclaje GPI con el fenómeno de interacción del parásito con la célula hospedadora (Bartholomeu et al. 2009; Maganto-Garcia et al. 2008; Ropert et al. 2002).

Entre las proteínas de superficie de *T. cruzi* se encuentra la familia de las mucinas, cuya importancia queda reflejada al determinarse que constituyen entorno al 1% del genoma del parásito y que presentan hasta  $4 \times 10^6$  de estas

moléculas por parásito (Buscaglia et al. 2006). *T. cruzi* es capaz de llevar a cabo la expresión de dos grupos distintos de mucinas: TcSMUG, presente en la fases típicas del insecto vector; y TcMUC, que se expresa en los estadios del hospedador mamífero (Buscaglia et al. 2006; Canepa et al. 2012; Gonzalez et al. 2013; Tarleton 2005). Las mucinas son proteínas altamente O-glicosiladas (mediante mecanismos diferentes según sean TcSMUG o TcMUC) que tienen motivos GPI, con los que se anclan a la membrana de *T. cruzi* (Buscaglia et al. 2006; Canepa et al. 2012; Di Noia et al. 1996). Así, esta familia multigénica va a recubrir la superficie celular de *T. cruzi* desempeñando un papel muy importante en el que proporciona (i) protección frente los mecanismos de defensa del vector y del mamífero (Canepa et al. 2012; Garcia et al. 2007), (ii) interviene en el reconocimiento previo a la entrada en la célula hospedadora (Buscaglia et al. 2006) y (iii) participa en la adhesión de los epimastigotes al lumen de la parte posterior del intestino medio del insecto, aspecto clave para la metaciclologénesis (Gonzalez et al. 2013).

Una característica distintiva de los genes pertenecientes a la familia TcMUCII es que presentan una región central variable, posiblemente relacionada con la evasión de la respuesta inmune (Buscaglia et al. 2006). Las TcMUC aparecen relacionadas con las transialidasas (TS), otra familia proteica esencial en el proceso de invasión celular. Las TS actúan tanto transfiriendo el ácido siálico de la célula hospedadora a las glicoproteínas superficiales del parásito, así como ligandos durante el proceso de interacción con la célula hospedadora (Ming et al. 1993; Villalta et al. 2001). Las mucinas se muestran como los principales aceptores de los residuos de ácido siálico.

Ruiz y colaboradores (1998) describieron el papel que algunas glicoproteínas podían tener en el proceso de invasión. En la superficie de las formas metacíclicas de las cepas G y CL Brener de *T. cruzi*, encontraron respectivamente, las glicoproteínas gp35/50 y gp82. Aunque gp82 se expresaba en ambas cepas, se demostró que los tripomastigotes metacíclicos de la cepa G utilizan preferentemente gp35/50 para interaccionar con la célula hospedadora, posiblemente debido a que la capacidad de adhesión de gp35/50 es mucho mayor que la que posee gp82 (Ferreira et al. 2006; Ruiz et al. 1998). En cualquier caso, ambas moléculas inducían la movilización de calcio desde los depósitos

intracelulares de la célula hospedadora, aunque mediante mecanismos opuestos. Mientras CL Brener utiliza gp82 para inducir la despolimerización de la actina en la célula hospedadora (Cortez, Atayde, Yoshida 2006), las mucinas gp35/50 de la cepa G, promueven la entrada del parásito reclutando fibras de actina (Ferreira et al. 2006; Ruiz et al. 1998).

También se ha identificado como una porción de las mucinas que carecen de motivo GPI son secretadas al medio. Entre estos productos secretados se encuentran mucinas de ambos tipos, tanto TcSMUG como TcMUC, que pese a no tener la modificación postransduccional en su extremo N- terminal que supone la adición de GPI, son mucinas maduras y altamente O-glicosiladas. Se sugiere incluso la posibilidad de que entre las mucinas secretadas, se encuentren los productos codificados por los pseudogenes de esta familia, especialmente de TcMUC II. Todos estos productos de secreción pueden moldear y afectar a la interacción parásito/hospedador (Canepa et al. 2012).

### **2.6.2 “*Mucin-associated surface proteins*” (MASPs)**

Dentro de las proteínas de superficie de *T. cruzi*, la familia multigénica codificante para las denominadas proteínas *mucin-associated surface proteins* (MASPs), constituye uno de los principales hallazgos que tuvo lugar durante el proyecto genoma de este parásito, incluido dentro de la iniciativa Tritryp (El-Sayed et al. 2005b). En la cepa CL Brener esta familia está formada por 1,377 miembros, siendo la segunda familia más grande descrita en *T. cruzi* y representando entorno al 6% del genoma diploide del parásito (El-Sayed et al. 2005b). Los genes de esta familia MASP se distribuyen en largos clusters, fuera de la zona de sintenia con *L. major* y *T. brucei*. Concretamente, los genes MASPs se encuentran flanqueados aguas arriba y aguas abajo por los genes de las mucinas y los de la familia de las TS respectivamente. Esta organización genómica, lejos de ser azarosa, parece evitar la homogenización de las secuencias contribuyendo a mantener la diversidad dentro de la familia MASP (Bartholomeu et al. 2009). La expresión de proteínas MASPs ocurre preferentemente durante la fase tripomastigote del parásito (Bartholomeu et al. 2009; dos Santos et al. 2012), aunque también se ha observado su expresión en el resto de estadios de *T. cruzi* (De Pablos et al. 2011). Por otra parte, la estructura

proteica de las MASPs es característica, con dominios N- y C- terminales altamente conservados que codifican una péptido señal y un sitio de adición GPI respectivamente, lo cual sugiere su localización en la superficie celular del tripomastigote, aunque también se ha descrito como algunos miembros de la familia MASP son excretados al medio (Bartholomeu et al. 2009; Bayer-Santos et al. 2013; De Pablos et al. 2011). La región central se muestra hipervariable tanto en longitud como en secuencia (Bartholomeu et al. 2009; El-Sayed et al. 2005b). Analizando las proteínas MASPs se pueden observar varios sitios susceptibles de ser N- y O- glicosilados mostrando también sitios aptos para la fosforilación (Bartholomeu et al. 2009). Es interesante señalar que en contraste con los genes VSG de *T. brucei*, en los que la expresión se restringe a un solo miembro de la familia en cada momento, Atwood y colaboradores (2005; 2006) describieron 14 genes de la familia MASP expresándose simultáneamente en las fases tripomastigote y/o amastigote de *T. cruzi*; mientras que recientemente dos Santos y colaboradores (2012) fueron capaces de elaborar librerías de ADNc donde detectaron desde 60 hasta 90 miembros de la familia MASP expresándose simultáneamente en una población de formas tripomastigotas del parásito, determinando además que el repertorio de genes MASPs expresado oscilaba según el número de pases de cultivo y la célula hospedadora sobre la que se realizaba la infección *in vitro*. Adicionalmente, los autores mostraron el potencial inmunogénico de las proteínas de esta familia, siendo capaces de detectarse en sueros provenientes de fase aguda en ratones infectados. Al respecto, el análisis mediante *arrays* de péptidos y ELISA de 64 MASPs expresadas demostró la amplia variabilidad en la intensidad del reconocimiento de la librería, así como un reconocimiento diferencial tanto por anticuerpos IgM como IgG presentes en los sueros (dos Santos et al. 2012). Este polimorfismo de expresión de genes MASPs y de reconocimiento inmunológico refleja un mecanismo por el cual el parásito evade la respuesta inmune del hospedador contribuyendo a la hipergammaglobulinemia asociada a la fase aguda de la enfermedad de Chagas.

Por lo tanto, debido a su abundante presencia en el genoma, los altos niveles de co-expresión, la localización en la superficie de sus proteínas, su reconocimiento inmunológico e incluso la habilidad para ser secretadas, la familia MASP de *T. cruzi*, se postula como un factor clave en las interacciones

célula-parásito y en la evasión de la respuesta inmune del hospedador (Bartholomeu et al. 2009; Buscaglia et al. 2006; De Pablos et al. 2011; dos Santos et al. 2012; Garcia et al. 2007).

## **2.7 Poblaciones celulares: mosaicos de variaciones entre células**

Recientes avances en tecnología de análisis celular individual revelan la importancia de considerar las poblaciones celulares como complejos mosaicos de células que, pese a estar sujetas a un mismo programa genético, pueden manifestar un alto grado de variación transcripcional y/o fenotípica (Pelkmans 2012; Seco-Hidalgo, Osuna, Pablos 2015; Xia et al. 2014). Se ha observado que este mosaicismo celular puede constituir una ventaja selectiva actuando como uno de los mecanismos principales en las decisiones celulares que tienen lugar dentro de una población de células isogénicas, tanto en organismos unicelulares como pluricelulares (Balazsi, van Oudenaarden, Collins 2011; Pelkmans 2012; Seco-Hidalgo, Osuna, Pablos 2015).

En la actualidad, se dispone de varias técnicas que permiten trabajar sobre las células aisladas a partir de poblaciones celulares, como son: PCR de colonias, clonación de células mediante dilución seriada, plataformas basadas en microfluidos, amplificación completa del genoma (WGA) o transcriptoma (WTA) sobre células individuales, hibridación fluorescente *in situ* sobre moléculas individuales de ARN (Seco-Hidalgo, Osuna, Pablos 2015). De esta manera, el análisis sobre células individuales canaliza sus esfuerzos en caracterizar y comprender los procesos biológicos cambiando el enfoque tradicional, basado en mediciones promedio sobre poblaciones celulares, por un punto de vista más complejo, donde se considera a cada célula de una manera individual, analizándose su variabilidad dentro de la población.

Aunque el principal objetivo de este trabajo de Tesis es el de indagar en la importancia de las variaciones celulares en una población de *T. cruzi*, es importante resaltar como estas heterogeneidades son capaces de determinar el destino de células cancerígenas y células madre (Kumar et al. 2014; Navin 2014; Torres-Padilla and Chambers 2014; Xia et al. 2014), así como también se ha evidenciado la presencia de variabilidad clonal en poblaciones de bacterias



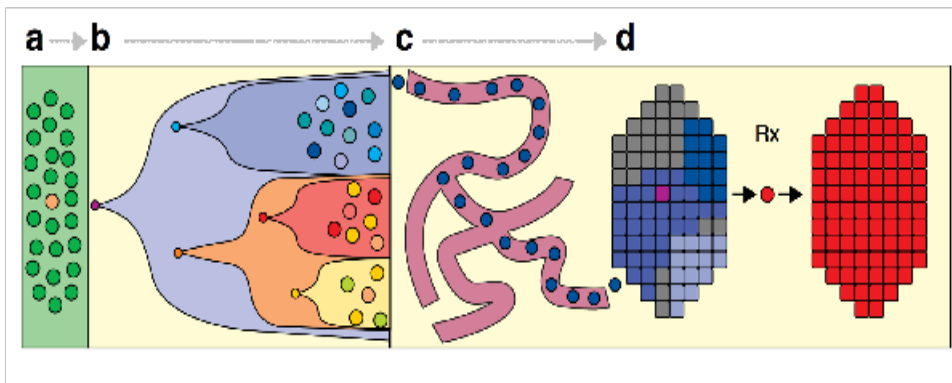
patogénicas (Balazsi, van Oudenaarden, Collins 2011; Casadesus and Low 2013) y en hongos (Levy, Ziv, Siegal 2012; Morrow and Fraser 2013).

### ***2.7.1 Variabilidad clonal en poblaciones de células cancerígenas y células madre***

La muerte celular en todas sus formas, apoptosis y necrosis, es un aspecto clave para el buen funcionamiento y desarrollo de los tejidos. En exceso puede desencadenar enfermedades neurodegenerativas mientras que, un déficit o falta de la misma pueden traducirse en cáncer o trastornos autoinmunes.

Un estudio reciente que muestra la heterogeneidad entre células cancerígenas ha sido llevado a cabo por Fallahi-Sichani y colaboradores (2013). En este trabajo, centrado en la muerte celular y el cáncer, los autores observaron como la vida media de células cancerígenas, tratadas con un inhibidor de la vía PI3K, mostraba una respuesta clonal. Por su parte, Nicholas Navin, llevó a cabo la secuenciación del genoma y del transcriptoma de células individuales “single-cell sequencing (SCS)” a partir de diversas muestras de tejidos cancerígenos (Navin et al. 2011; Wang et al. 2014), proponiendo un modelo en el que desde la aparición hasta la propagación y metastasis derivadas del cáncer se explican por la variabilidad célula-célula (Figura1.13) (Navin 2014; Navin 2015).

Así la dinámica y el desarrollo de un proceso cancerígeno están altamente influenciados, no sólo por la heterogeneidad genómica derivada de mutaciones, sino también por la heterogeneidad fenotípica que se produce como consecuencia de modificaciones epigenéticas o cambios transcripcionales (Navin 2014).



**Figura 1.13. Evolución de un proceso cancerígeno desde la singularidad de una única célula.** Aunque una célula individual va a interactuar con el resto de células del tejido conjuntivo de un órgano, numerosas partes del proceso cancerígeno se desarrollan por la acción de una única célula cancerígena: (a) transformación de una célula somática normal en una célula cancerígena; (b) evolución clonal que se produce por la acumulación de mutaciones que hace que células individuales diverjan hacia una heterogeneidad intratumoral; (c) liberación de células individuales desde el tumor primario a otros órganos a través del torrente sanguíneo, originándose tumores metastásicos; y (d) adquisición de quimioresistencia mediante la acumulación de mutaciones (que confieren dicha resistencia) en células tumorales individuales. Esto va a permitir la reconstitución de la masa tumoral (Navin 2014).

Diversos análisis sobre células madre también demuestran que las células pluripotentes son capaces de manifestar una expresión génica heterogénea en ambientes de cultivo uniformes.

La pluripotencia celular depende de la acción coordinada de redes regulatorias de genes (GRN) ensambladas por factores de transcripción y sujetas a modulación por vías de señalización capaces de responder a señales externas (Torres-Padilla and Chambers 2014).

La somatogénesis, o conversión de células pluripotentes en células somáticas, se produce por un descenso en algunos de los factores de transcripción que determinan dicha pluripotencia, como son NANOG u OCT4 (Torres-Padilla and Chambers 2014). Trabajando sobre células madre embrionarias (ESCs), diferentes autores han sido capaces de detectar variaciones célula-célula en algunos de estos factores de transcripción, como son NANOG

(Abranches, Bekman, Henrique 2013; Chambers et al. 2007), REX1 (Toyooka et al. 2008), STELLA (Hayashi et al. 2008) ZSCAN4 (Zalzman et al. 2010), entre otros.

Aunque los mecanismos que sostienen este mosaico de variaciones entre células pluripotentes no están bien descritos, se cree que además de la regulación transcripcional, participan procesos a nivel postranscripcional tales como la síntesis proteica o incluso la dinámica del ciclo celular (Torres-Padilla and Chambers 2014).

### **2.7.2 Heterogeneidad celular en poblaciones de protozoos parasitarios**

Nuevos hallazgos han mostrado como determinados protozoos parasitarios, ya sea en un contexto de regulación transcripcional (como el que presenta *Plasmodium falciparum*) o en ausencia del mismo (Orden *Kinetoplastida*), basan en la heterogeneidad celular algunos de los aspectos más importantes de sus fenotipos, como son el repertorio de antígenos expuestos en sus superficies, su diferenciación sexual o incluso la resistencia a fármacos (Kafsack et al. 2014; Laffitte et al. 2014; Lima et al. 2013; Rovira-Graells et al. 2012a; Ubeda et al. 2014).

Trabajando sobre *P. falciparum*, Rovira-Graells et al. (2012a) describieron una enorme variabilidad transcripcional entre subclones en el estado merozoito, fase responsable de la multiplicación asexual del parásito en el interior de los glóbulos rojos. Así observaron como, en contraste con genes esenciales o de copia única, dicha heterogeneidad clonal se centraba principalmente en las familias multigénicas responsables de llevar a cabo la interacción célula hospedadora-parásito. Además, identificaron y asociaron diferentes estados de condensación de la heterocromatina con aquellos genes que habían mostrado una respuesta clonal, indicando que los mecanismos epigenéticos son los que estaban detrás de esta variación clonal (Rovira-Graells et al. 2012a; Seco-Hidalgo, Osuna, Pablos 2015). A su vez, propusieron una naturaleza estocástica en la condensación de la cromatina como medio para compensar la habilidad limitada de *P. falciparum* para responder rápida y directamente a estímulos externos.

Otro aspecto del agente etiológico de la malaria (*P. falciparum*) que recientemente se ha visto ligado a la heterogeneidad celular individual es su gametogénesis, la habilidad que tiene este protozoo de diferenciarse desde un estado asexual a otro sexual. Así, Kafsack y colaboradores (2014) demostraron

como este cambio morfológico y funcional se lleva a cabo únicamente en un conjunto de parásitos dentro de la población total. Para ello, Kafsack *et al.* analizaron los niveles de expresión del gen *pfap2-g*, cuyos productos proteicos (PfAP2-G) se relacionan con la formación de gametocitos, observando como dichos niveles variaban entre células isogénicas de la misma población. Así los autores proponen que *P. falciparum* mantiene una expresión del gen *pfap2-g* regulada epigenética y estocásticamente, con una línea basal de conversión (Kafsack *et al.* 2014; Seco-Hidalgo, Osuna, Pablos 2015). Sin embargo, otros autores aportan evidencias firmes de que este protozoo es capaz de detectar estímulos ambientales y responder ajustando con precisión la proporción de células dedicadas a la gametogénesis, lo que supone un ejemplo de plasticidad fenotípica adaptativa (Pollitt *et al.* 2011; Seco-Hidalgo, Osuna, Pablos 2015). De esta manera, en presencia de fármacos anti-malaria, el parásito se declina hacia un aumento claro de la producción de gametocitos, forma transmisible del protozoo, asegurando así su progenie y posterior propagación (Schneider *et al.* 2012; Seco-Hidalgo, Osuna, Pablos 2015).

Por otra parte, se ha demostrado como los kinetoplástidos pese a la ausencia de un control transcripcional, también van a ser capaces de generar variaciones entre células de una misma población. Recientes estudios muestran como *Leishmania spp.* cuando sufre cambios ambientales o estrés es capaz de promover la amplificación reversible de segmentos de ADN generando CNVs que confieren resistencia a fármacos y que, por lo tanto, favorecen el éxito de un conjunto de células de la población (Laffitte *et al.* 2014; Seco-Hidalgo, Osuna, Pablos 2015; Ubeda *et al.* 2014).

Adicionalmente, en la actualidad hay evidencias que muestran la existencia de CNVs y otras variaciones clonales en *T. cruzi* (Bartholomeu *et al.* 2009; Lima *et al.* 2013). Precisamente, estos autores mediante ensayos de inmunofluorescencia (Bartholomeu *et al.* 2009) y análisis del cariotipo (tanto en número como en tamaño de las bandas cromosómicas analizadas) (Lima *et al.* 2013). Pese a ello, este fenómeno no está bien caracterizado. Por el importante papel funcional que se le atribuye a la familia MASP, son necesarias más evidencias experimentales que permitan comprender el modo de expresión de esta familia dentro de un población de *T. cruzi*.







## **II. OBJETIVOS/OBJECTIVES**





*T. cruzi* es una taxón genéticamente muy diverso que muestra en su genoma una gran expansión de familias codificantes para proteínas de superficie, entre las que destaca la familia MASP. Esta familia multigénica se postula como un factor determinante en las interacciones parásito-hospedador.

Evidencias experimentales recientes han demostrado la importancia de la variabilidad clonal en la dinámica poblacional, la supervivencia y la evolución de los parásitos.

Bajo estas premisas, el objetivo principal de este trabajo de investigación ha sido el de estudiar y caracterizar la variabilidad transcripcional de la familia MASP entre diferentes clones de *T. cruzi*, así como la heterogeneidad fenotípica de algunos aspectos básicos de la biología de este protozoo. Para lograrlo se planteron los siguientes objetivos específicos:

1. Desarrollar un método que permita una amplificación específica y de alto rendimiento de transcritos MASP para llevar a cabo un análisis comparativo tanto a nivel de secuencia como de expresión de esta familia multigénica.
2. Determinar los niveles de expresión y el tiempo de vida del ARN MASP entre las fases epimastigote y tripomastigote de *T. cruzi*.
3. Caracterizar el grado de variabilidad de poblaciones clonales isogénicas de *T. cruzi*, mediante el estudio de la expresión de la familia MASP a nivel de ARN y proteína, así como mediante el análisis de varios aspectos fenotípicos como crecimiento, metaciclogénesis espontánea y secreción de diferentes catabolitos.



*T. cruzi* is a taxon genetically diverse which genome shows a dramatic expansion of several families of surface molecules where is highlighted the MASP family. This multigene family is postulated as an important factor in the host-parasite specific interactions.

Recent experimental evidences have demonstrated the importance of clonal variations for parasite population dynamics, survival and evolution

Thus, the main aim of the current work is to study and characterize the transcriptional variability in the expression of the MASP family among different clonal lines of *T. cruzi*, as well as the phenotypic heterogeneity of some basic characters. With this purpose, several specific objectives were proposed:

1. To develop a method that allow for an specific and high efficiently amplification of the MASP transcripts to perform a comparative analysis at sequence and expression level of this multigene family
2. To determine the level expression and the lifespan of the MASP RNA between the stages epimastigote and trypomastigote of *T. cruzi*
3. To characterize the degree of variability among different clonal isogenic populations of *T. cruzi* by the study of the MASP family expression at RNA and protein level, as well as by the analysis of phenotypic aspects such as growth, spontaneous metacyclogenesis and secretion of different catabolites.













### **III. MATERIALES Y MÉTODOS**



## 1. Cultivo y manejo de *T. cruzi*

El material biológico de *T. cruzi* usado en este trabajo de tesis se corresponde con la cepa Pan4 de *T. cruzi*. Esta cepa, perteneciente al DTUI a-d (Cura et al. 2010), fue aislada en el año 2006 a partir de un paciente de 32 años (varón) en Panamá (distrito de Arraiján, comunidad de Burunga) y donada a nuestro laboratorio por la Dra. A. Ying de la Universidad de Panamá en ese mismo año.

### 1.1 Mantenimiento del cultivo de *T. cruzi*

Las formas epimastigotas de Pan4 fueron cultivadas a 28°C en medio RPMI suplementado con un 10% v/v de Suero Bovino Fetal inactivado (SBFi) mediante incubación a 56°C durante 30 min, y colocados en frascos de cultivo Roux (Nunc) de diferentes tamaños según los requerimientos experimentales. Los cultivos de epimastigotes se mantuvieron en su fase exponencial de crecimiento, realizando resiembras periódicas (2-3 veces por semana) de  $1 \times 10^6$  epimastigotes/ml a partir de los stocks de parásitos.

Todos los medios de cultivos utilizados para *T. cruzi* fueron elaborados en un ambiente estéril (cámara de flujo laminar vertical, Glatt), siendo posteriormente esterilizados por filtración bajo presión negativa mediante filtros estériles Millex- GP de 0.22  $\mu$ M (Millipore).

- **Medio RPMI:** 16.4 g RPMI-1640 Medium (Sigma), 2.2 g Carbonato Sódico, 0.134 g Penicilina, 0.060 g Estreptomicina, 0.134 g Kanamicina, 100 ml Suero Bovino Fetal (Gibco) previamente inactivado mediante incubación a 56°C durante 30 min; agua bidestilada hasta 1 litro; el pH de la solución se ajustó a 7.2-7.4.

### 1.2 Criopreservación de *T. cruzi*

La criopreservación fue utilizada para el mantenimiento de material biológico, lo que permitió tanto disponer de la cepa original de *T. cruzi* como conocer el número de pases en cultivo de cada una de las líneas de parásitos. Con

este fin, 15 ml de un cultivo de parásitos en fase exponencial de crecimiento ( $6 \times 10^6$  parásitos/ml) se centrifugaron a  $1,000 \times g$  durante 10 min a  $4^\circ\text{C}$ . El botón obtenido se resuspendió en 2 ml de medio de congelación y se alicuotó en viales de congelación (Nalgene Cryoware). Estos crioviales fueron introducidos en un contenedor Nalgene® Cryo 1 que contenía alcohol isopropílico, y que a su vez fue enfriado en un arcón de  $-80^\circ\text{C}$ . De esta manera, los parásitos fueron sometidos a un proceso de congelación progresiva, con un descenso de temperatura de  $-1^\circ\text{C}/\text{min}$ . Finalmente, tras alcanzar los  $-80^\circ\text{C}$ , los viales fueron almacenados en un tanque de nitrógeno líquido a  $-170^\circ\text{C}$ .

La descongelación de parásitos se realizó de manera rápida incubando los viales de congelación a  $37^\circ\text{C}$  en un baño maria. Seguidamente, con el objetivo de retirar el medio de congelación, la suspensión de parásitos fue centrifugada a  $1,000 \times g$  durante 5 min a  $4^\circ\text{C}$ , seguido de tres lavados con PBS (tampón fosfato salino). Finalmente, los parásitos fueron resuspendidos en su medio de cultivo habitual e incubados a  $28^\circ\text{C}$ .

- **Medio de congelación de *T. cruzi*:** Medio RPMI suplementado con 10% de SBFi y 15% de Glicerol estéril.
- **Tampón fosfato salino (PBS):** 10 mM  $\text{KH}_2\text{PO}_4$  y 10 mM  $\text{Na HPO}_4$ . La solución se ajustó pH 7.2.

### 1.3 Obtención *in vitro* de formas tripomastigotes metacíclicas

Para aislar la fase metacíclica de *T. cruzi*, se siguió la metodología descrita por Osuna *et al.* (Osuna Carrillo, Jimenez Ortiz, Lozano Maldonado 1979; Osuna *et al.* 1986). Dicho proceso de metaciclologénesis se llevó a cabo a partir de las formas epimastigotas en fase exponencial de crecimiento. Los epimastigotes fueron cultivados hasta alcanzar su fase logarítmica de crecimiento, siendo posteriormente centrifugados a  $1,000 \times g$ , a  $4^\circ\text{C}$  durante 10 min. Tras varios lavados con PBS, el botón de parásitos fue resuspendido en medio Grace modificado, ajustando la densidad a  $1 \times 10^7$  flagelados/ml. Los cultivos se realizaron en frascos Roux de 25 ml (Nunc), con un volumen de 5 ml de medio e incubados a  $27^\circ\text{C}$ . Tras 9-12 días del inicio del cultivo se consiguieron las

mayores tasas de transformación. La tasa de transformación fue determinada a través del recuento de tripomastigotes metacíclicos teñidos con solución de Giemsa o mediante conteo en cámara hemocitométrica de Neubauer.

En la realización de todos los experimentos posteriores, únicamente se emplearon formas de *T. cruzi* con más de un 90% de pureza.

- **Medio Grace modificado** (hasta 100 ml agua bidestilada): El medio se realizó disolviendo 5.2 g de Grace's Insect Medium (Sigma) y 0.250 g de Hemoglobina. Posteriormente, se filtró la solución y se añadió un 10% de SBFi. A dicho medio se añadieron 100 mEq/l de NaCl que se obtuvieron tras la acidificación del medio con HCl 1M hasta alcanzar un pH de 1.5, y posteriormente se neutralizó con NaOH 1M hasta obtener un pH final de 5.8. Finalmente, el medio se filtró a través de filtros de policarbonato de 0.22  $\mu\text{m}$  de tamaño de poro (Osuna Carrillo, Jimenez Ortiz, Lozano Maldonado 1979; Osuna et al. 1986).

#### **1.4 Curvas de crecimiento y metaciclologénesis espontánea**

Las curvas de crecimiento se realizaron a partir de  $1 \times 10^6$  epimastigotes de cada una de las líneas celulares, las cuales fueron cultivadas por triplicado de manera independiente, siguiendo las condiciones habituales que se describen en el apartado 1.1. El recuento de parásitos se realizó cada 24 h durante un total de 11 días, utilizando para ello una cámara hematocitométrica de Neubauer. El porcentaje de formas metacíclicas, entre los días 9 y 11 del estudio, fue también obtenido mediante recuento en la cámara hematocitométrica de Neubauer.

## **2. Cultivo y manejo de células Vero**

Se utilizaron cultivos de células Vero para diferenciar el parásito a fase amastigote intracelular y posteriormente aislar la fase tripomastigote sanguíneo derivado de estas células. Esta línea celular, procedente del riñón de mono Verde Africano adulto, fue obtenida en el Servicio de Líneas Celulares del Centro de

Instrumentación Científica de la Universidad de Granada (CIC) con Código Internacional: ECACC 84113001.

## 2.1 Mantenimiento del cultivo de células Vero

Las células Vero se cultivaron en medio DMEM suplementado con 10% SBFi a 37°C y en una atmósfera con 5% de CO<sub>2</sub>. Las células crecieron adheridas en monocapa a la superficie de frascos de cultivo (Nunc) dispuestos en posición horizontal. Una vez que las células alcanzaron un estado de confluencia en su crecimiento, se procedió a su tripsinización. Para ello, se realizó un primer lavado con la solución EDTA-Tripsina cubriendo la monocapa. A continuación, se volvió a añadir la solución EDTA-Tripsina cubriendo las células e incubando a 37°C durante 5-10 min. De esta manera se consiguió el desprendimiento de la monocapa celular, que tras lavarla con medio de cultivo fue centrifugada a 600 x g, a 4°C durante 10 min. Tras tres lavados con solución de Hank's (Sigma), el botón celular fue resuspendido en medio de cultivo realizando una nueva siembra con 1 x 10<sup>4</sup> células/ml.

Al igual que para *T. cruzi*, todos los medios de cultivos utilizados sobre células Vero, fueron elaborados en un ambiente estéril (cámara de flujo laminar vertical, Glatt), siendo posteriormente esterilizados por filtración bajo presión negativa mediante filtros estériles Millex- GP de 0.22 µM (Millipore).

- **Medio DMEM (*Dulbecco's Modified Eagle's medium*):** 9.53 g de medio de Eagle en polvo conteniendo aminoácidos Esenciales (MEM), 2.2 g de Bicarbonato Sódico, 4 g de HEPES, 0.056 g de Penicilina, 0.1 g de Estreptomina, 0.048 g de Kanamicina, 100 ml SBFi; agua bidestilada hasta 1 litro; el pH se ajustó a 7.2-7.4.
- **Solución EDTA-Tripsina:** 0.1 g EDTA, 0.5 g Tripsina; hasta 200 ml de PBS; pH ajustado a 7.2.

## 2.2 Criopreservación de células Vero

La criopreservación de células Vero se realizó sobre cultivos de células en fase exponencial. Tras decantar el medio de cultivo, las células Vero fueron tripsinizadas y la suspensión celular obtenida fue centrifugada a  $300 \times g$ , a  $4^{\circ}\text{C}$  durante 5 min. A continuación, el botón celular se resuspendió en medio de congelación y se alicuotó en viales para una congelación progresiva, como se ha descrito en el apartado 1.2, hasta  $-80^{\circ}\text{C}$ , los cuales fueron mantenidos en tanques de nitrógeno líquido a  $-170^{\circ}\text{C}$ .

La descongelación de células Vero se realizó siguiendo el mismo proceso que se llevó a cabo para descongelar *T. cruzi*. Finalmente, las células Vero fueron resuspendidas en medio DMEM + 10% SBFi.

- **Medio de congelación Vero:** 80% de medio de cultivo DMEM, 10% de SBFi adicional, y 10% de DMSO.

## 3. Infección de células Vero con tripomastigotes metacíclicos y obtención de tripomastigotes sanguíneos

Cultivos de células Vero semiconfluentes fueron infectados con las formas metacíclicas de la cepa Pan4 de *T. cruzi*. Para ello, previamente se procedió a sincronizar las células Vero con Timidina 25 mM durante 9 h acorde con la metodología descrita por Osuna *et al.* (Osuna *et al.* 1984b), permitiendo obtener una mayor eficiencia en la infección. A continuación, las células se lavaron tres veces con PBS y fueron infectadas incubando durante 8h con una suspensión de tripomastigotes metacíclicos en medio DMEM sin SBFi, empleando una relación parásito:célula de 5:1.

Seguidamente, las células se lavaron tres veces con PBS a  $4^{\circ}\text{C}$ , para retirar los parasitos no-adheridos o no-internalizados, y se añadió medio DMEM suplementado con un 10% SBFi atemperado a  $37^{\circ}\text{C}$ . De esta manera, en el cultivo



sólo permanecieron los tripomastigotes metacíclicos que fueron capaces de adherirse y de penetrar en las células, dando lugar a las formas amastigotes que, tras replicarse durante 96 h en el interior celular, terminaron por transformarse en las formas tripomastigotes sanguíneos derivadas del cultivo celular, las cuales salen rompiendo la célula para infectar otras células del cultivo.

Para aislar las formas tripomastigotes derivadas de cultivo celular, se procedió a recoger el medio de células Vero 4-5 días después de que éstas fueran infectadas con las formas metacíclicas. Se dejó decantar dicho medio a 37°C durante 4 h para descartar posibles restos celulares. Tras esto, se recogió cuidadosamente el sobrenadante (con los tripomastigotes sanguíneos), se centrifugó a 1,000 x g, a 4°C durante 10 min, lávándose tres veces en PBS.

#### **4. Clonación de la cepa Pan4 de *T. cruzi***

El proceso de clonación de la cepa Pan4 se llevó a cabo mediante dilución seriada. Para ello, se recogieron los tripomastigotes sanguíneos resultantes de tres rondas consecutivas de infección sobre células Vero. Estas formas flageladas fueron centrifugadas a 500 x g, a 4°C durante 10 min y, tras tres lavados con PBS, cultivadas como epimastigotes en medio RPMI. Sobre estos epimastigotes, en fase exponencial de crecimiento ( $6 \times 10^6$  parásitos/ml), se realizó una dilución de 1,000 parásitos/ml a partir de la cual se extrajeron alícuotas de 1  $\mu$ l, que teóricamente contenían un solo organismo, y que fueron añadidas de manera individual en los pocillos de una microplaca de cultivo de 96 pocillos (Nunc) para aislar células individuales. Para evitar posibles contaminaciones cruzadas entre los clones, se dejó un pocillo vacío entre cada pocillo utilizado para el proceso de clonación. Únicamente se agregaron 100  $\mu$ l de medio RPMI sobre aquellos pocillos en los que había un solo parásito (comprobado microscópicamente), siendo incubados en condiciones de cultivo durante una semana. Posteriormente, cada una de las poblaciones clonales isogénicas aisladas

fue traspasada a un frasco de cultivo Roux (Nunc) para proceder con su cultivo. Una vez obtenida una masa de  $6 \times 10^6$  parásitos/ml en 15 ml, cada una estos clones fue centrifugado, alicuotado y criopreservado a fin de mantener sus propiedades biológicas.

Todos los experimentos posteriores se llevaron a cabo utilizando el mismo número de generaciones de los parásitos y las mismas densidades celulares en las diferentes líneas clonales.

## 5. Cultivo y manejo de *Escherichia coli*

Se utilizó la cepa DH5 $\alpha$  de *Escherichia coli*. Su genotipo cromosómico es el siguiente: F-  $\Phi$ 80*lacZ* $\Delta$ M15  $\Delta$ (*lacZYA-argF*) U169 *recA1 endA1 hsdR17* (rK-, mK+) *phoA supE44*  $\lambda$ - *thi-1 gyrA96 relA1*.

### 5.1 Cultivo de *E. coli*

Las bacterias de la cepa DH5 $\alpha$  fueron cultivadas en medio LB líquido en agitación constante (240 rpm) a 37°C. A su vez, el medio de cultivo fue suplementado con los antibióticos de selección correspondientes. También se realizaron cultivos en placas Petri con medio LB Agar, incubándose a 37°C durante una noche para el crecimiento de colonias individuales. El medio LB fue suplementado con los inductores de selección y antibióticos correspondientes que se describen en el apartado 6.9.3.b.

- **Medio LB líquido:** 1 g Triptona, 0.5 g Levadura, 1 g NaCl; hasta 100 ml de agua bidestilada. Esterilizado mediante autoclave (121°C durante 20 min).
- **Medio LB-Agar:** 2 g de Agar en 100 ml de medio LB líquido. Esterilizado mediante autoclave (121°C durante 20 min).

## 5.2 Preparación de células competentes

Para este procedimiento, partiendo de cultivos en placa de la cepa DH5 $\alpha$ , se realizó una siembra de bacterias en 5 ml de medio LB líquido, siendo incubada en agitación (240rpm) a 37°C durante 12-14 h. Transcurrido este tiempo, se añadieron 4 ml del cultivo bacteriano sobre 100 ml de medio TYM precalentado volviéndose a someter a incubación en agitación (240rpm) a 37°C entre 1 h 15 min a 2 h, hasta que el cultivo alcanzó un densidad celular/ml equivalente a una densidad óptica de 0.5 medida espectrofotométricamente a 600 nm.

Seguidamente, se enfrió el cultivo 5 min en hielo para detener el crecimiento bacteriano y se centrifugó a 3,500 x g, a 4°C durante 5 min, tras lo cual se resuspendió cuidadosamente el botón bacteriano en 20 ml de la solución TFB-I fría. Esta solución de bacterias en TFB-I fue incubada en hielo durante 15 min, volviéndose a centrifugar a 3,500 x g, a 4°C durante 5 min. El botón fue resuspendido en 2 ml de solución fría TFB-II y nuevamente incubado en hielo durante 15 min. A partir de aquí se realizaron las alícuotas de bacterias competentes (de 50 o 100  $\mu$ l) que fueron rápidamente almacenadas a -80°C.

- **Medio TYM:** 2% Triptosa, 0.5% Extracto de levadura, 0.1% NaCl, 0.01% MgCl<sub>2</sub>. Esterilizado mediante autoclave (121°C por 20 min).
- **Medio TFB-I:** Acetato de potasio 30 mM, MnCl<sub>2</sub> 50 mM, KCl 100 mM, CaCl<sub>2</sub> 10 mM. Esterilizado mediante filtración.
- **Medio TFB-II:** Na MOPS 10 mM, CaCl<sub>2</sub> 75 mM, 15% Glicerol; pH ajustado a 7. Esterilizado mediante filtración.

## 5.3 Criopreservación de *E. coli*

Las bacterias de interés fueron seleccionadas, resembradas en 3-5 ml de LB (medio LB suplementado con los antibióticos de selección correspondientes) y cultivadas a 37°C en agitación (240 rpm), durante toda la noche. Seguidamente, los cultivos bacterianos fueron centrifugados a 4,000 x g, a 4°C durante 10 min. Posteriormente, los botones bacterianos obtenidos fueron resuspendidos en

medio de congelación, alicuotándose 1 ml de éstos en viales de congelación, los cuales fueron almacenados hasta su uso a  $-80^{\circ}\text{C}$ .

- **Medio congelación bacteriana:** LB líquido suplementado con un 15% de Glicerol estéril.

## 6. Técnicas de Biología Molecular

### 6.1 Extracción de ADN genómico de *T. cruzi*

La purificación de ADN genómico se realizó a partir de un cultivo de *T. cruzi* en su estadio epimastigote. Para ello se empleó el kit DNeasy Blood and Tissue (Qiagen) siguiendo el protocolo de purificación para células en cultivo que se detalla en el manual del fabricante. Tras la extracción, el ADN genómico se cuantificó, como se detalla a continuación (apartado 6.4), y se almacenó a  $-20^{\circ}\text{C}$  hasta su posterior uso.

### 6.2 Extracción de ARN total de *T. cruzi*

Este proceso se realizó de manera independiente sobre  $1 \times 10^8$  de las formas epimastigote y tripomastigote. Previamente los cultivos fueron lavados por centrifugación ( $1,000 \times g$  durante 10 min) en tres volúmenes de PBS. Para la extracción del ARN total se utilizó el kit comercial RNeasy (Qiagen), siguiendo el protocolo para células animales que se describe en el manual del fabricante con leves modificaciones. El botón de células obtenido de los lavados fue lisado añadiendo  $600 \mu\text{l}$  de tampón RLT. La suspensión fue pipeteada durante 1 min. Se añadió un volumen de etanol al 70% y se homogenizó por pipeteo. Se transfirió a una columna del kit con su tubo colector y se centrifugó a  $8,000 \times g$  durante 30 s. El eluido del tubo colector fue descartado. Se añadió  $350 \mu\text{l}$  de tampón RW1 y se centrifugó a  $8,000 \times g$  durante 30 sec. Se mezcló  $15 \mu\text{l}$  de DNase I en  $70 \mu\text{l}$  de tampón RDD y esta mezcla fue añadida a la columna e incubada a  $25^{\circ}\text{C}$  durante 15 min. Transcurrido este tiempo,  $700 \mu\text{l}$  de tampón

RW1 fueron añadidos a la columna y centrifugados a 8,000 x g durante 30 sec. Para lavar la membrana de la columna se la colocó en un nuevo tubo colector y se añadieron 500 µl de tampón RPE, se centrifugó a 8,000 x g durante 30 sec y se descartó el eluido. Este último paso fue repetido pero aumentando el tiempo de centrifugación a 2 min. Nuevamente se cambió el tubo colector y la membrana fue centrifugada durante 1 min para eliminar los posibles restos de los tampones. Para eluir el ARN de la membrana de la columna se añadieron 30 µl de agua ultrapura libre de RNasas.

Este ARN total fue directamente procesado a ARNm o almacenado a -80°C hasta su posterior uso, dependiendo del interés experimental.

### **6.3 Extracción de ARNm de *T. cruzi***

Se procedió al aislamiento de ARNm a partir del ARN total de las fases epimastigote y tripomastigote.

Para llevar a cabo esta purificación se utilizó el kit Oligotex (Quiagen) siguiendo el protocolo de purificación ARN poli A<sup>+</sup> a partir de ARN total que se describe en el manual del fabricante.

Tras determinar la cantidad de ARN total de la muestra (tal y como se describe en el apartado 6.4), se procedió a añadir la cantidad adecuada de tampón OBB, incubándose a 70°C durante 3 min y, seguidamente, a 25°C durante 10 min (paso donde se produce la hibridación entre los oligo dT<sub>30</sub> y la cola de poli A). A continuación, el complejo formado fue centrifugado a 15,000 rpm durante 2 min, descartándose cuidadosamente el sobrenadante. La muestra fue resuspendida en 400 µl de tampón OW2, añadida sobre una columna que fue centrifugada a máxima velocidad durante 1 min, descartándose el eluido (este paso se repitió una vez más). Por último, la muestra fue transferida a un nuevo tubo colector, añadiéndose 50 µl de tampón OEB previamente calentado a 70°C. Con el objetivo de eluir el ARNm Poli A, se centrifugó a 14,000 rpm durante 1 min. Este último paso se repitió una vez más, obteniéndose 100 µl finales de ARNm Poli A.

El ARNm fue directamente retrotranscrito a ADNc, tal y como se describe a continuación en el apartado 6.5.

## 6.4 Cuantificación de ADN y ARN

La pureza y concentración de las muestras de ADN y ARN extraídas se determinaron espectrofotométricamente con el equipo Nanodrop ND-1000 (Thermo Scientific) midiendo la absorbancia de la muestra a una dilución adecuada a 260 nm frente a un blanco de agua ultrapura. Las relaciones  $A_{260}/A_{280}$  y  $A_{260}/A_{230}$  fueron utilizadas para estimar el grado de pureza de las muestras. Se consideró que la muestra tenía buena calidad con valores de  $A_{260}/A_{280}$  de ~ 1.8 para ADN y ~ 2.2 para ARN; y con valores de  $A_{260}/A_{230}$  de 2.0-2.2.

## 6.5 Retrotranscripción del ARNm

La retrotranscripción se realizó con el kit iScript cDNA Synthesis (Bio-Rad) utilizando como cebador la solución oligo(dT)<sub>20</sub>. Para esta reacción se mezclaron los siguientes reactivos:

- 4 µl 5x iScript Reaction Mix
- 2 µl oligo(dT)<sub>20</sub>
- 1 µg de ARN total
- 1 µl iScript Reversotranscriptasa
- Agua ultrapura hasta 20 µl

Las condiciones empleadas para retrotranscripción se describen a continuación:

- 42°C durante 90 min
- 85°C durante 5 min

El ADNc obtenido fue almacenado a -20°C hasta su posterior uso.

## **6.6 Análisis del tiempo de vida del ARNm en *T. cruzi***

Para inhibir la transcripción y poder analizar el tiempo de vida media de los transcritos en *T. cruzi* se utilizó Actinomicina D (AD). Este compuesto, aislado a partir de la actinobacteria *Streptomyces spp.*, es capaz de inhibir el proceso de transcripción uniéndose al ADN en el complejo de iniciación de la transcripción. Esta unión va a evitar que la enzima ARN polimerasa realice el proceso de elongación. La unión de la AD al ADN se establece a través de los residuos deoxiguanosina (Davis et al. 1998).

La AD se empleó de manera independiente sobre un total de  $6 \times 10^7$  células de las formas epimastigotes y tripomastigotes para cada tratamiento. En primer lugar, las células se lavaron mediante centrifugación ( $1,000 \times g$  durante 10 min) y el botón de parásitos fue resuspendido en un frasco Roux de 25 cm<sup>2</sup> (Nunc) con 10 ml de medio cultivo habitual (RPMI 10% SBFi), que en el caso de las formas tipomastigotes fueron tamponados con Hepes 25 mM y ajustados a un pH de 7.4. Sobre cada cultivo se añadió AD (Sigma) a una concentración de 10 µg/ml de AD, incubando a diferentes tiempos (0 min, 15 min, 30 min o 2 h). Transcurrido el tiempo de cada tratamiento y para eliminar los restos de AD, los flagelados se lavaron tres veces con PBS mediante centrifugación ( $1,000 \times g$  durante 10 min) y fueron directamente procesados para la extracción de su ARN total.

## **6.7 Electroforesis de ADN en geles de agarosa**

La electroforesis de ADN permitió comprobar el tamaño y la integridad de las moléculas de ADN con las que se trabajó en el transcurso de esta tesis. Esta técnica se fundamenta en la migración diferencial de moléculas en función de su tamaño mediante la acción de un campo eléctrico sobre una matriz tamponada compuesta de agarosa.

Las concentraciones de agarosa del gel variaron entre el 1-2% según el tamaño de los fragmentos analizados. Para la obtención del gel, la agarosa se disolvió en Tampón TAE 1X, añadiendo 5 µl de SYBR Green (10,000X)

(Invitrogen) por cada 50 ml de gel. Como marcador de peso molecular se utilizó el HyperLadder I® con bandas de tamaños comprendidos entre los 200 y 10,000 pb (Bioline). Durante la electroforesis, las muestras fueron sometidas a un voltaje constante de 120 V durante 30-40 min en un sistema electroforético MAXI-VG (VWR) acoplado a una fuente de alimentación Power Pac®HV (Bio-Rad). El resultado fue visualizado en un equipo transiluminador y fotodocumentador Benchtop 2UV® Transilluminator (UVP).

- **Tampón TAE 10X:** 0.4 M Tris, 20 mM EDTA, 11.14 ml Ácido acético; se ajustó a pH 8.2.
- **Tampón de electroforesis:** TAE 1X (diluir 10 veces TAE 10X con agua bidestilada).

## 6.8 Purificación productos de PCR desde geles de agarosa

Para este proceso se realizó una separación electroforética de la muestra en geles de agarosa de bajo punto de fusión (USBiological). A continuación, se recortaron las bandas de interés, extrayendo el ADN contenido en ellas mediante incubación a 50°C durante 10 min con el Tampón de solubilización del Kit QIAquick Gel Extraction (Qiagen). La purificación se completó mediante centrifugación de la muestra en columnas con membranas de sílica, tal y como se especifica en el manual del fabricante. Este protocolo se basa en la purificación de fragmentos de ADN entre 70 pb y 10 kb mediante la retención en membrana de sílica, los tampones de captura dan a la muestra la correcta concentración de sales y el pH adecuado para ser retenidos en la columna y luego poder ser eluidos en agua ultrapura estéril.

Una vez extraído el ADN de la banda, se midió su concentración y pureza mediante espectrofotometría tal y como se describió en el apartado 6.4.

## 6.9 Elaboración de librería de expresión MASP (*3-step MASP library*)

La clonoteca de miembros MASP se realizó a partir del ARNm del estadio tripomastigote tras tres pases de infección en células Vero; este ARNm fue



retrotranscrito a ADNc, como se detalla en el apartado 6.5, el cual sirvió de molde para la librería de expresión MASP.

### 6.9.1 Amplificación de (pseudo)genes MASP (3- step nested PCR)

El ADNc de tripomastigote fue amplificado mediante una PCR anidada de tres pasos denominada *3-step nested PCR*. Este proceso se basó en el descrito por dos Santos *et al.* (2012) en el cual realizaron una PCR semi-anidada con una primera reacción donde emplearon la combinación de cebadores SL y 3'UTR1; y una segunda PCR donde mantuvieron el cebador SL combinándolo con un cebador más interno, el 3'UTR2. A partir de esta segunda reacción de PCR nosotros desarrollamos un tercer paso de amplificación (detallado a continuación) que permitió obtener la librería MASP. Las combinaciones de cebadores empleados en cada uno de los pasos de la *3-step nested PCR* se recogen en la tabla 3.1.

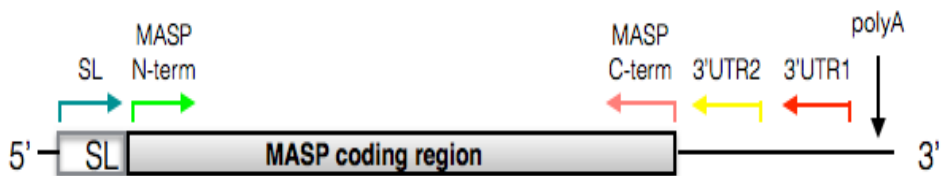
**Tabla 3.1. Combinación de cebadores empleados para la realización de la *3-step nested PCR*.**

Tipo de PCR	Nombre del Cebador	Secuencia Cebador R (5' -3')	Temperatura de annealing (°C)
1ª: PCR	SL	AACGCTATTATTGATACAGTTTC TGTACTATATTG	61.2
	3'UTR1	GTGTGCTTCGTGGGGTGAGGTG	
2ª: Semi-nested PCR	SL	AACGCTATTATTGATACAGTTTC TGTACTATATTG	63
	3'UTR2	CTCACTCTCACGCGGCCACCAC CACCG	
3ª: Nested PCR	MASP N-term_F	ATGGCGATGATGATGACCGGC	56
	MASP C-term_R	CCACCACCGCAGTAGCAG	

Los cebadores MASP N-term\_F y MASP C-term\_R se diseñaron con el objetivo de amplificar el mayor número de transcritos MASP posible. Para ello se realizó una búsqueda de motivos conservados sobre 250 miembros MASP en

la cepa CL Brener de *T. cruzi* utilizando el software MEME (Bailey and Elkan 1994). Los motivos más conservados fueron seleccionados para diseñar las secuencias de cada cebador. Por último, previo a la síntesis de los cebadores, se analizaron sus secuencias mediante el algoritmo “BLASTN” (Altschul et al. 1990) para confirmar el alto grado de identidad con la familia MASP y descartar inespecificidades.

A continuación, la figura 3.1 permite visualizar la posición de los cebadores en el prototipo de un transcrito MASP.



**Figura 3.1. Representación de un transcrito MASP estándar indicando la posición de los cebadores que se han empleado para la amplificación de MASP mediante 3-step nested PCR.**

Cada uno de los pasos de esta combinación de PCRs se elaboraron en un volumen total de 25  $\mu$ l a partir de 600 ng de ADNc para la primera reacción y 3.3  $\mu$ l de producto amplificado procedente de la primera y segunda reacción como molde para la segunda y tercera reacción respectivamente. Se añadieron 200  $\mu$ M dNTPs, 2 mM  $MgCl_2$ , 1 mM de cada cebador y 1 unidad de Taq ADN polimerasa (MBL) en 1X “MBL PCR Buffer”.

Las condiciones en las que se realizaron cada una de las tres reacciones de la 3-step nested PCR fueron las siguientes:

- 95°C - 5 min
  - 95°C - 1 min
  - de 56 a 63°C\* - 1 min
  - 72°C - 1 min
  - 72°C - 5 min
- } 30 ciclos

\* Las temperaturas de *annealing* variaron para cada uno de los tres pasos de la *3-step nested PCR*, según la combinación de cebadores que se empleó (tabla 3.1).

## **6.9.2 Purificación de los productos de PCR**

Antes de ser introducido en un vector de clonación, el producto resultante de la 3ª *nested PCR* fue sometido al Kit de “Qiaquick PCR Purification” (Qiagen) para eliminar restos de cebador y otros reactivos. El proceso se hizo siguiendo el protocolo para microcentrifuga incluido en el manual del fabricante y finalmente eluyendo el ADN en agua ultrapura estéril. Este kit está basado en el mismo principio que el kit QIAquick Gel Extraction Kit explicado en el apartado 6.8.

## **6.9.3 Clonación en pGEM®-T Easy**

### **6.9.3.a Reacción de ligación**

Tras su purificación descrita en el apartado anterior, los distintos amplicones MASP fueron ligados al vector de clonación pGEM-T® Easy (Promega) mediante el uso T4 DNA ligasa (Promega). Esta enzima cataliza la ligación de dos hebras de ADN entre los grupos 5´-fosfato del residuo timidina terminal del vector de clonación y el 3´-hidroxilo que se encuentra en la adenina terminal de los productos de PCR. En esta reacción se utilizó una relación molar inserto-vector de 3:1, añadiéndose tres unidades de enzima a la reacción en un volumen total de 10 µl. Para obtener una mayor eficiencia en la reacción, la mezcla se incubó a 4°C durante toda la noche. El producto obtenido de este proceso de ligación fue utilizado para la transformación de bacterias competentes.

### **6.9.3.b Transformación de bacterias competentes**

La reacción de transformación se llevó a cabo mezclando en agitación ligera 50 µl de células competentes de *E. coli* cepa DH5α, con 2 µl del producto de ligación. Se incubó 20 min en hielo y 2 min a 42°C en baño María. Rápidamente se volvió a incubar en hielo durante 2 min; todo esto produce un choque térmico

que aumenta la fluidez de la membrana bacteriana, facilitando la entrada del ADN. A continuación, se añadieron 950  $\mu\text{l}$  de medio SOC líquido sobre las células transformadas, incubándolo a 37°C en agitación (240 rpm) durante 1 h. Por último 100  $\mu\text{l}$  de las bacterias sometidas a la reacción de transformación fueron sembradas sobre placas de LB agar, conteniendo Ampicilina como antibiótico de selección, IPTG 0.5 mM como inductor del gen de la  $\beta$ -Galactosidasa, y X-gal (100  $\mu\text{g}/\text{ml}$ ). Las placas se incubaron a 37°C durante toda la noche, siendo las bacterias transformadas, con el vector ligado al ADN de interés, capaces de crecer en dicha placa formando colonias blancas, que fueron seleccionadas sobre el fondo de colonias azules, las cuales carecen de dicho inserto en el vector de clonación.

- **Medio SOC (100 ml):** 2 g Triptosa, 0.5 g de extracto de levadura, 1 ml 1M NaCl, 0.25 ml 1M KCl. Todo ello fue esterilizado mediante autoclave (121°C durante 20 min). A continuación se añadió, 1 ml 2M Glucosa y 1 ml de una solución 2M  $\text{Mg}^{2+}$  compuesta por 2.033 g  $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ , 2.465 g  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ . Finalmente, se ajustó el pH a 7 y se esterilizó mediante filtración.
- **Para 100 ml de LB:**
  - Ampicilina (100  $\mu\text{g}/\text{ml}$ ): 100  $\mu\text{l}$  (100 mg/ml)
  - IPTG (0.5 mM): 500  $\mu\text{l}$  (35g mg/ml (1 M))
  - X-Gal (100  $\mu\text{g}/\text{ml}$ ): 100  $\mu\text{l}$  (100mg/ml)

### 6.9.3.c Rastreo de clones transformantes

Para determinar la presencia de un inserto MASP, los clones transformantes fueron analizados mediante PCR. Con este fin, se seleccionaron el máximo número de colonias positivas (blancas) que fueron resembradas en placas LB agar con ampicilina e incubadas a 37°C durante toda la noche. Al día siguiente, se picaron colonias que fueron suspendidas en 25  $\mu\text{l}$  de agua ultrapura estéril y calentadas a 92°C durante 2 min para producir su lisis. Dicha suspensión de bacterias lisadas fue centrifugada a 10,000 x g durante 5 min, utilizando 1  $\mu\text{l}$  de la parte superior del sobrenadante obtenido como ADN molde para la realización de PCRs independientes con los cebadores del inserto clonado, así como también con los cebadores del vector (abajo detallados). Dichas PCRs

permitieron confirmar la presencia del inserto de interés en el interior de los clones bacterianos transformados.

- **Cebadores del inserto:** MASP N-term\_F y MASP C-term\_R (descritos anteriormente en el apartado 5.5.1 de esta sección)
- **Cebadores del vector:**
  - T7(F): 5'-TAATACGACTCACTATAGGG-3'
  - SP6(R): 5'-TATTTAGGTGACACTATAG-3'

Las condiciones a las que se realizaron las PCR para el rastreo de clones bacterianos transformados:

- 95°C - 10 min
  - 95°C - 20 segundos
  - 55°C - 20 segundos
  - 72°C - 1 min
  - 72°C - 5 min
- } 30 ciclos

Finalmente, los productos amplificados fueron visualizados mediante electroforesis en gel de agarosa al 1%.

#### 6.9.3.d Purificación del ADN plasmídico

La extracción y obtención de ADN plasmídico se realizó sobre clones bacterianos transformados que fueron identificados como positivos por PCR. Para ello se empleó el kit comercial de purificación de ADN plasmídico PureYield™ Plasmid Miniprep System (Promega) siguiendo el protocolo de aislamiento que se especifica en el manual del fabricante.

La concentración y pureza de los plásmido fueron comprobadas empleando el espectrofotómetro Nanodrop ND-1000 (Thermo Scientific) a través de la relación de las mediadas de absorbancia a 280/260 y 230/260 respectivamente (ver apartado 6.4).

### **6.9.3.e Secuenciación de ADN**

Este proceso se llevó a cabo en el servicio de secuenciación de ácidos nucleicos del Instituto de Parasitología y Biomedicina “López-Neyra” (CSIC) Granada.

La reacción de secuenciación se realizó en un volumen final de 12 µl utilizando entre 100 y 600 ng de ADN molde según si éste procedía de productos de PCR o era ADN plasmídico. Se emplearon también 6.4 pg de cada cebador, dNTPs marcados con fluorescencia BigDye® pertenecientes al kit BigDye® Terminator v 1.1 cycle sequencing (Applied Biosystems, CA, USA).

La secuenciación se obtuvo mediante un equipo “3130XL Genetic Analyzer” (Applied Biosystems).

### **6.9.4 Anotación de genes MASP**

Las secuencias obtenidas en esta tesis fueron depositadas en GenBank, a través del software Bankit, obteniéndose los números de acceso KT222275-KT222324 (Seco-Hidalgo, De Pablos, Osuna 2015b).

## **6.10 Análisis de los niveles de expresión MASP**

Los niveles de expresión MASP fueron analizados tanto sobre la línea celular parental, así como sobre cada una de las líneas clonales isogénicas de *T. cruzi*. Para ello se emplearon diversas técnicas que se describen a continuación.

### **6.10.1 RT-qPCR**

La cuantificación de la expresión MASP mediante RT-qPCR se realizó a partir del ADNc de las distintas líneas celulares de *T. cruzi* en su fase epimastigote. Para ello se amplificó un fragmento específico y conservado, que se localiza en el extremo 5' de la región codificante de la familia MASP, utilizando la pareja de cebadores MASP N-term\_F (anteriormente descrito) y MASP N-term\_R cuyas secuencias se describen a continuación:

- MASP N-term\_F: 5' ATGGCGATGATGATGACCGGC 3'
- MASP N-term\_R: 5' ACCACAGCACGCACAGGGC 3'

Estos oligonucleótidos fueron diseñados siguiendo la metodología descrita en el apartado 5.9.1 de esta sección.

La reacción de PCR para MASP N-term se realizó en un volumen total de 10  $\mu$ l:

- 5  $\mu$ l Ssofast 2X
- 400 nM MASP N-term\_F
- 400 nM MASP N-term\_R
- 50ng ADNc
- Agua ultrapura hasta 10  $\mu$ l

La expresión MASP fue normalizada utilizando Gliceraldehido 3-fosfatodehidrogenasa (GAPDH) como gen de referencia. Para ello se amplificó un fragmento de este gen mediante la siguiente combinación de cebadores:

- GAPDH\_F: 5' AGCGCGCGTCTAAGACTTACA 3'
- GAPDH\_R 5' TGGAGCTGCGGTTGTCATT 3'

La reacción de PCR para GAPDH se realizó en un volumen total de 10  $\mu$ l:

- 5  $\mu$ l Ssofast 2X
- 500 nM GAPDH\_F
- 500 nM GAPDH\_R
- 50ng ADNc
- Agua ultrapura hasta 10  $\mu$ l

La amplificación se realizó bajo las siguientes condiciones:

- 95°C - 30 segundos
  - 95°C - 10 segundos
  - 60°C - 10 segundos
- } 39 ciclos

- 95°C - 1 min
- Curva de melting 65°C a 95°C (incremento 0.5°C; 5 segundos/paso)

La eficiencia de los cebadores se comprobó mediante una curva de calibración a partir diluciones seriadas de concentraciones de ADN molde conocidas. El cálculo de la expresión relativa de la familia MASP se determinó según el método de  $\Delta CT$ : referencia/prueba =  $2^{CT\ referencia - CT\ prueba}$ ; donde la referencia es el gen GAPDH y la prueba es el gen de interés.

Todas las reacciones se realizaron por triplicado usando SYBR® Green Supermix (Bio-Rad) en un termociclador C-1000 conectado a un módulo de Real-time CFX96 (Bio-Rad) bajo las condiciones recomendadas por el fabricante.

### **6.10.2 Electroforesis capilar**

El ADNc de cada uno de los clones también fue amplificado mediante *3-step nested PCR*, descrita previamente en el apartado 5.9.1 de esta sección. Los amplicones obtenidos fueron resueltos mediante electroforesis capilar en un 3130xl Genetic Analyzer (Applied Biosystems) en el servicio de secuenciación de ácidos nucleicos del Instituto de Parasitología y Biomedicina “López-Neyra” (CSIC) de Granada. La identificación de picos y la medida del tamaño de los fragmentos obtenidos se realizaron usando el programa “Peak Scanner” (Applied Biosystems). El gel *in silico* y el dendograma fueron realizados usando el software “Fingerprinting II” (Bio-Rad), empleando el coeficiente de Cosine, que tiene en cuenta tanto la presencia y ausencia de la banda como la intensidad del pico de cada banda.

### **6.10.3 Northern blot**

#### **6.10.3.a Obtención de las sondas**

Para el análisis de expresión relativa por northern blot se utilizaron dos sondas que fueron diseñadas en base a la librería MASP: i) sonda *masp20*: sonda aislada a partir de la secuencia única del clon *masp20* de la librería de expresión



MASP (número de acceso GenBank KT222294). ii) Sonda *masp19*: esta sonda se obtuvo a partir de la secuencia conservada del clon *masp19* de la librería de expresión MASP (número de acceso GenBank KT222293). Así, la sonda *masp19*, presenta altos porcentajes de identidad (tanto en secuencia como en longitud) con 23 de los 50 transcritos MASP (Tabla 4.5 de resultados). Ambas sondas fueron obtenidas mediante PCR con los cebadores del inserto (MASP N-term\_F and MASP C-term\_R) utilizando como molde los vectores pGEM-T® Easy donde se clonaron dichas secuencias. Antes de ser utilizados como sondas, los productos resultantes de la amplificación fueron sometidos a un proceso de purificación sobre geles de agarosa (descrito en el apartado 5.8 de esta sección), seleccionando exclusivamente el amplicón de interés correspondiente a los dos genes *masp* mencionados.

Adicionalmente se amplificaron los genes 18S y  $\beta$ -tubulina para su utilización como sondas en la normalización de la expresión del ARN MASP.

#### 6.10.3.b Electroforesis, hibridación y cuantificación de la señal

La separación electroforética se realizó a partir de 3  $\mu$ g de ARN total de cada una de las líneas celulares isogénicas de *T. cruzi*. Se mezcló cada ARN con 7.5  $\mu$ l de tampón de carga y con agua ultrapura estéril hasta un volumen de 17.5  $\mu$ l. Esta mezcla fue incubada a 65°C durante 5 min y rápidamente depositada en hielo. Se añadieron 2.5  $\mu$ l de tinte de carga y se procedió a la separación mediante electroforesis en un gel de formaldehído al 0.9% agarosa (80V durante 2 h). Seguidamente se transfirieron las muestras desde el gel a una membrana de nylon (MSI 0.45 micron). Para ello, en la base de una bandeja se colocaron varios papeles de *blotting* (Schleicher and Schuell GB 002), cortados al tamaño del gel y completamente humedecidos con la solución 20x SSC. Sobre estos papeles y tras eliminar cualquier posible burbuja que pueda interferir en la transferencia, se colocó el gel. Encima de el gel se dispuso la membrana, también cortada al tamaño del gel, y sobre ésta 6 papeles de *blotting* más. De nuevo, se eliminaron posibles burbujas, se colocó una esponja y se procedió a sellar la bandeja para evitar la evaporación del 20x SSC. Finalmente, se añadió un peso en la parte superior y se incubó a temperatura ambiente durante toda la noche. Una vez las

muestras fueron transferidas a la membrana, se realizó un proceso de *crosslinking* de las moléculas de ARN transferidas a la membrana mediante el *Stratalinker UV crosslinker* (Thermo), utilizándose para ello luz UV a 254nm durante 20-25 segundos con una irradiación de 120,000 microjulios/cm<sup>2</sup>. Tras ello, la membrana fue incubada en agitación orbital constante a 42°C durante toda la noche en el tampón de pre-hibridación suplementado con 1.25 µg/µl ácido deoxiribonucleico procedentes de testículos de salmón (Sigma) previamente desnaturalizados a 95°C durante 5 min y almacenados posteriormente en hielo. Este paso sirve para bloquear la membrana y evitar inespecificidades en el northern blot. Al día siguiente, sobre el mismo medio de pre-hibridación se añadieron de manera independiente cada una de las sondas que también fueron previamente desnaturalizadas mediante incubación a 95°C durante 5 min, y marcadas con dCTP-α[P<sup>32</sup>] (Perkin-Elmer).

Antes del revelado, las membranas fueron sometidas a un proceso secuencial de lavados estrictos con las soluciones de lavado 1 y 2, incubándose a 65-68°C durante 20 min en cada lavado. La cuantificación de cada uno de los northern se llevó a cabo empleando un "Typhoon imager" y el software "ImageQuant" (GE Healthcare Life Sciences).

- **Tampón de electroforesis northern blot 10X:** 50 mM NaOAc, 0.2 M MOPS, 10 mM EDTA. Agua ultrapura estéril (autoclavada 121°C durante 20 min) hasta 1 litro.
- **Gel de formaldehído** (200 ml): 0.9% agarosa en 1X tampón de electroforesis, 10 ml formaldehído (37%), 20 ml 10X MOPS.
- **1.5X tampón de carga:** 750 µl formamida, 75 µl formaldehído (37%), 150 µl 10X tampón de electroforesis, 15 µl Et Br, 10 µl agua ultrapura estéril.
- **Tinte de carga:** Azul de bromophenol blue en 35% glicerol y 1X de tampón de electroforesis.
- **20X SCC:** 3 M NaCl, 0.3 M Citrato de sodio, 800 ml agua ultrapura; el pH se ajustó a 7. Se añadió agua ultrapura hasta un volumen final de 1 litro. Esterilizado por autoclave (121°C durante 20 min).

- **Tampón de pre-hibridación:** 50% Formamida, 5X Tampón SSC, 5X solución Denhart y 0,1% dodecil sulfato sódico (SDS).
- **Solución de lavado 1:** 2X SSC y 0.1% SDS. Hecho en agua ultrapura estéril.
- **Solución de lavado 2:** 0.2X SSC y 0.1% SDS. Hecho en agua ultrapura estéril.

## 7. Técnicas de análisis proteico e inmunológico

### 7.1 Técnicas de análisis proteico

#### 7.1.1 *Electroforesis de proteínas. SDS-PAGE*

##### 7.1.1.a *Obtención de homogenados proteicos*

La obtención de proteínas de *T. cruzi* se realizó a partir de 50 ml de cultivo de las formas epimastigote ( $1 \times 10^6$  células/ml) en su día 5 (fase exponencial de crecimiento). También se extrajeron las proteínas totales de  $5 \times 10^7$  tripomastigotes sanguíneos derivados de cultivo que fueron aislados tal y como se describe en el apartado 3. Los parásitos fueron lavados por centrifugación ( $1,000 \times g$  durante 10 min) en tres volúmenes de PBS. El botón se congeló a  $-80^\circ\text{C}$  y posteriormente se descongeló en hielo. Se añadieron 3 v/p de tampón de lisis y se incubó durante 10 min en hielo. Finalmente, las muestras fueron sonicadas durante 2 min, alternando 10 seg de sonicación con 10 seg de reposo durante tres ciclos y almacenadas a  $-20^\circ\text{C}$  hasta su posterior uso.

- **Tampón de lisis:** 10 mM fosfato potásico a pH 7.4, 0.25 mM sacarosa, 1 mM EDTA, 0.145 mM KCl, 1% Tritón X-100. Finalmente, por cada 25ml de tampón de lisis se añadió un inhibidor de proteasa “Complete Mini inhibitor protease cocktail” (Roche Molecular Biochemicals).

### *7.1.1.b Cuantificación de la concentración de proteínas*

La determinación de la concentración de proteínas se realizó mediante el método de Bradford.

Para llevar a cabo este proceso se elaboró una curva estándar de concentraciones conocidas de albúmina bovina (Sigma), desde 0 a 10 µg/ml, las cuales fueron obtenidas a partir de una solución madre de 1 mg/ml. Las diferentes diluciones de albúmina, realizadas en PBS, se mezclaron con el reactivo Bradford (Sigma) en una proporción 1:1 para un volumen final de 2 ml, dejándose interactuar a temperatura ambiente durante 20 min y en oscuridad. A continuación, se midió la absorbancia de cada una de las diluciones a una longitud de onda de 595 nm utilizando un espectrofotómetro de Mitton Roy (Spectronic 3000). A partir de los valores de densidad óptica de la curva estándar se elaboró una recta de regresión que, posteriormente, permitió extrapolar y conocer la concentración de proteínas en las muestras de interés. Para ello las muestras proteicas fueron procesadas de igual manera que las diluciones de albúmina y medidas a 595 nm.

### *7.1.1.c Elaboración de geles y condiciones de electroforesis*

El análisis de proteínas se realizó mediante la técnica de electroforesis en geles de poliacrilamida, en presencia de dodecil sulfato sódico (SDS-PAGE), utilizando el equipo Mini-protean system (Bio-Rad). La composición del gel separador y del gel concentrador se detalla en la tabla 3.2. Se mezclaron 7 µl de la muestra proteica con otros 7 µl de tampón de carga 2X y se calentaron a 95°C durante 10 min antes de que fueran cargadas en el gel para su migración a un voltaje constante de 120 V, durante 1 h. El marcador de peso molecular usado fue Precision Plus Protein® Standards (Bio-Rad) de 10 kDa a 250 kDa.

- **Tampón de la muestra 2X** (10 ml): 2 ml Tris-HCl 0.5 M a pH 6.8, 2 ml SDS 10%, 0.2 ml EDTA 0.2M a pH 8, 4 ml Glicerol 50%, 1.8 ml agua ultrapura, azul de bromofenol, y 2% de β-mercaptoetanol.
- **Tampón de electroforesis 5X** (1,000 ml agua ultrapura): 15.1 g Trizma, 72 g Glicina, 5 g SDS a pH 8.3.

- **Solución gel concentrador y del gel separador:** La composición de ambos geles se describe en la Tabla 3.2.

**Tabla 3.2. Composición de las soluciones de gel separador y gel concentrador para SDS-PAGE.**

Reactivos	Gel Separador 12%	Gel Concentrador
Acrilamida 30%/ bisacrilamida 0.8% (ml)	3.9	0.8
Agua ultrapura (ml)	3.45	3.06
Tampón de gel* (ml)	2.5	0.625
SDS 10% (ml)	0.1	0.05
Persulfato de amonio 10% (APS) (ml)	0.2	0.05
TEMED ( $\mu$ l)	10	10

\* Tampón de gel separador: 1.5 M Tris-HCl pH 8.8. Tampón de gel concentrador: 1 M Tris- HCl pH 6.8.

#### 7.1.1.d Tinción de los geles

Tras la electroforesis se procedió a la tinción sumergiendo el gel en la solución de tinción durante 45 min. Para desteñir el gel se retiró la solución anterior y se sumergió en una solución de lavado. Una vez desteñido, el gel fue conservado en agua destilada, capturándose su imagen con el fotodocumentador ChemiDoc® MP System (Bio-Rad)

- **Solución de tinción:** Agua ultrapura 50% (v/v), Metanol 40% (v/v), Ácido Acético 10% (v/v), Azul de Comassie 0.025% (p/v).
- **Solución de lavado:** Agua ultrapura 80% (v/v), Ácido Acético 10% (v/v), Propanol 10% (v/v).

## 7.2 Técnicas inmunológicas

### 7.2.1 Diseño de péptidos inmunogénicos

El diseño de péptido sintético MASP-Ag se basó en la secuencia proteica de los miembros MASP expresados en nuestra librería de expresión. Para ello se identificaron los mejores motivos aminoacídicos en base a la posición dentro de la zona hipervariable de los genes MASP y su presencia en la clonoteca utilizando el programa MEME (Bailey and Elkan 1994). A continuación, sobre los motivos seleccionados, se determinó la hidrofobicidad, utilizando el programa ProtScale (<http://web.expasy.org/protscale/>), y la antigenicidad, mediante la herramienta bioinformática Antigenic (<http://emboss.bioinformatics.nl/cgi-bin/emboss/antigenic>). Antes de ser sintetizado, el péptido fue analizado mediante el programa BLASTP para confirmar la especificidad de sus secuencias. Finalmente el péptido fue sintetizado por Genscript.

- **Secuencia del péptido sintético MASP-Ag:** AESQPAGVSVQDATG

### 7.2.2 Acoplamiento del péptido antigénico al carrier

Como *carrier* se utilizó la Hemocianina de cangrejo (KLH), la cual permite un aumento en el reconocimiento antigénico del péptido de interés. Para ello se mezclaron 3 mg de MASP-Ag con 3 mg de KLH (10 mg/ml) en PBS hasta un volumen final de 2 ml. Seguidamente, se agregaron 2 ml de Glutaraldehído 0.2%, como *crosslinker*. Esta mezcla fue incubada en agitación (agitador magnético) constante a 4°C durante toda la noche. Al día siguiente, con el objetivo de bloquear los grupos aldehído libres, se añadió 1 ml de Glicina 1 M a pH 7.2 y se volvió a incubar en agitación a 4°C durante 2 h. Seguidamente, se procedió a la diálisis introduciendo el polímero en un membrana con un tamaño de corte de 4,000 Da (Spectra Pore) e incubándolo en agitación con acetato de amonio 0.1 M a 4°C durante 48 h, con cambios cada 12 h (Harlow, E., Lane, D. 1988). Por último, se liofilizó el polímero empleando un equipo liofilizador LyoQuest (Telstar). El resultado de la polimerización fue evaluado mediante electroforesis SDS- PAGE.

### ***7.2.3 Inmunización y obtención de sueros policlonales***

El suero anti-MASP-Ag se obtuvo mediante inyección intraperitoneal en ratones BALB/c de 50  $\mu$ g correspondientes al péptido sintético MASP-Ag acoplado al carrier KLH y mezclado con adyuvante completo de Freund (Sigma) para la primera inoculación. Posteriormente, se realizaron nuevas inoculaciones cada 15 días hasta un total de 5 inmunizaciones en las que siguió el mismo proceso pero esta vez con el polímero MASP-Ag-KLH que fue inyectado con adyuvante incompleto de Freund (Sigma). Durante este proceso se hizo un seguimiento de la respuesta inmune del ratón valorando el título de anticuerpos mediante la técnica ELISA (Enzyme Linked Immunosorbent Assay), que se describe a continuación. Dos semanas después de la última inmunización, se procedió al sangrado de los ratones depositando la sangre en tubos (BD Microtainer® tubes). Finalmente, la sangre se incubó durante 30 min a 37°C tras lo que se mantuvo a 4°C durante 3 h para favorecer la retracción del coágulo y obtener la cantidad máxima de suero. El suero recogido fue centrifugado a 5,000 x g durante 5 min y almacenado a -20°C.

### ***7.2.4 Titulación de los sueros policlonales (ELISA)***

Esta técnica permitió cuantificar el nivel de antígenos o anticuerpos específicos que existen frente al inmunógeno anti-MASP-Ag.

El procedimiento comenzó sensibilizando un placa de microtitulación de poliestireno de 96 pocillos tipo Maxisorp (Nunc) añadiendo 10  $\mu$ g/pocillo del antígeno MASP-Ag resuspendido en tampón carbonato. A continuación, con el objetivo de fijar el antígeno uniformemente en el pocillo, las placas fueron incubadas en agitación orbicular a 4°C durante toda la noche cámara húmeda. Al día siguiente, los pocillos fueron lavados tres veces con tampón de lavado y se procedió al bloqueo agregando 300  $\mu$ l/pocillo de tampón de bloqueo e incubando a 37°C durante 1 h. Tras el bloqueo, se procedió a tres lavados con tampón de lavado y se añadió, por triplicado, 100  $\mu$ l/pocillo del suero anti-MASP-Ag en diluciones seriadas desde 1:100 hasta 1:6,400 (en PBS); las placas fueron incubadas a 37°C durante 2 h en agitación. Se lavó de nuevo tres veces y

se añadieron 100  $\mu$ l/pocillo de anticuerpo secundario "Polyclonal Goat Anti-Mouse Immunoglobulins/ HRP" (Sigma) a dilución 1:2,000 en PBS, incubando a 37°C durante 1 h 30 min en agitación. Transcurrido este tiempo, se lavó como se ha descrito y se reveló agregando 100  $\mu$ l/pocillo de solución sustrato peroxidasa, incubándose 15-20 min en oscuridad a 37°C y en agitación. Finalmente, con el fin de parar la reacción, se añadieron 50  $\mu$ l/pocillo de solución de parada e inmediatamente después la placa se leyó en un lector Elisa Multiskan® Spectrum a una longitud de onda de 492 nm.

- **Tampón de Bloqueo:** Tween 0.1% en PBS y Leche en polvo 2%.
- **Tampón de lavado:** Tween 0.3% en PBS.
- **Tampón Carbonato (0.1 M):** Carbonato sódico 0.1 M, Bicarbonato sódico 0.1 M. La solución se ajustó a pH 9.3.
- **Tampón citrato:** 0.1 M Acido Cítrico , 0.1 M Citrato sódico, pH: 5.
- **Solución sustrato peroxidasa:** 12.5 ml agua ultrapura, 12.5 ml Tampón citrato y una pastilla de OPD (O-fenil diaminobenzidina) (Sigma), resultando una concentración final de 0.4 mg/ml OPD. Justo antes de ser utilizada se añadieron 20  $\mu$ l H<sub>2</sub>O<sub>2</sub> a la solución.
- **Solución de parada:** HCl 3M.

#### **7.2.4 Western blots**

Esta técnica inmunoenzimática combina la capacidad analítica, que ofrece la electroforesis, con la sensibilidad del inmunoanálisis, permitiendo así la identificación de proteínas (Towbin, Staehelin, Gordon 1992).

Así tras la electroforesis (apartado 7.1.1.c de esta sección), las muestras de proteínas fueron transferidas a una membrana de PVDF (Polivinilideno difluorido) empleando para ello un Trans-blot® turbo transfer system (Bio-Rad). Para ello, se preparó un sándwich con la membrana y el gel con la siguiente distribución: cuatro hojas de papel de filtro del tamaño del gel, la membrana de nitrocelulosa, el gel, y cuatro trozos más de papel de filtro. Este proceso de transferencia se llevó a cabo a 1.3 A, 25 V durante 10 min. Seguidamente, la membrana fue incubada en tampón de bloqueo a temperatura ambiente durante



2 h y lavada tres veces mediante incubación en agitación con tampón de lavado a temperatura ambiente durante 10 min. A continuación, se usó el suero anti-MASP-Ag, descrito en el apartado 7.2.3, como anticuerpo primario (Ac 1ario) a una dilución 1:200 en tampón de lavado, incubándose en agitación a 4°C durante toda la noche. La membrana se volvió a lavar tres veces y se incubó en agitación a 37°C durante 2h con el anticuerpo secundario (Ac 2ario) Polyclonal goat anti-mouse immunoglobulins/HRP (Dako) conjugado con peroxidasa a dilución 1:1,000 en tampón de lavado. Para el revelado, la membrana fue lavada tres veces e incubada con la solución substrato peroxidasa a temperatura ambiente durante 15-20 min y en agitación.

Como controles de carga las membranas fueron incubadas, repitiendo el proceso descrito anteriormente, con el Ac 1ario Polyclonal sheep anti-alpha/beta tubulin (Cytoskeleton) a dilución 1:1,000 y con el Ac 2ario Anti-Sheep IgG whole molecule (Sigma) , conjugado con peroxidasa, a dilución 1:2,000.

Para la cuantificación de la señal obtenida en los western blot se realizó un análisis densitométrico mediante el software ImageJ 1.48 (Schneider, Rasband, Eliceiri 2012).

- **Tampón de transferencia:** 25 mM Tris, 192 mM Glicina, Metanol 20% en agua ultrapura y ajustado a pH 8.3.
- **Tampón de bloqueo:** Tween 0.1% en PBS y Leche en polvo 2%.
- **Tampón de lavado:** Tween 0.1% en PBS.
- **Solución substrato peroxidasa** (para 100 ml agua ultrapura): 1.21 g Tris-HCl 0.1 M a pH 7.4, 0.05 g 3-3' Diaminobencidina tetrahidroclorhídrica. Justo antes de ser utilizada, se añadieron 50 µl H<sub>2</sub>O<sub>2</sub> a la solución.

## **8. Hierarchical Clustering y estimación de las distancias entre la librería MASP y diversas cepas de *T. cruzi***

Con el objetivo de identificar secuencias ortólogas a la librería MASP de

Pan4 en el genoma de las cepas CL Brener (TcVI), Dm28c (TcI), Sylvio X10/1 (TcI) y Marinkellei (Tcbat) de *T. cruzi* (<http://tritrypdb.org/tritrypdb/>), se utilizaron los algoritmos BLASTN (para pseudogenes) y BLASTP (para genes) (Altschul et al. 1990). El e-valor probabilístico mínimo utilizado, para considerar una secuencia como ortóloga en las búsquedas de BLAST, fue de  $10^{-08}$ . Seguidamente, las secuencias ortólogas identificadas en cada cepa fueron alineadas frente a las secuencias de nuestra librería en Pan4 mediante el algoritmo ClustalW2 (Goujon et al. 2010; Larkin et al. 2007). Posteriormente, los alineamientos fueron editados en el software Jalview (Waterhouse et al. 2009). Basándose en los resultados de los alineamientos, se empleó la opción “pairwise distances” del programa Mega6.06 (Tamura et al. 2013) para obtener las distancias entre cada uno de los miembros de nuestra clonoteca y sus respectivos ortólogos, utilizándose 1,000 replicas de *bootstrap resampling* como método de varianza. Las distancias obtenidas se utilizaron para realizar el *hierarchical clustering* con el que analizar la distribución de dichas distancias entre los miembros de la librería de expresión así como entre ortólogos en las distintas cepas. También se utilizaron las distancias para elaborar *boxplots* con los que representar las distancias medias de las cepas. Para todo ello se utilizó el paquete informático *gplots* mediante sus funciones *heatmap.2* y *boxplots* respectivamente, siendo visualizados en la plataforma R (R Core Team 2014).

Adicionalmente, los miembros de la librería MASP fueron alineados frente a otras tres librerías MASP obtenidas previamente en nuestro laboratorio (De Pablos 2010) y frente a siete librerías MASP elaboradas por dos Santos *et al.* (dos Santos et al. 2012) utilizando el algoritmo ClustalW2. Esto permitió identificar variaciones genéticas (SNPs) entre las diferentes librerías que fueron visualizadas mediante GeneDoc (Nicholas and Nicholas 1997).

## **9. Búsqueda de sitios de inserción TcTREZO en la librería MASP**

La búsqueda de sitios de inserción para el retroelemento TcTREZO se llevó a cabo sobre todos los miembros de la librería MASP. Para ello, mediante

ClustalW2 (Goujon et al. 2010; Larkin et al. 2007), se realizaron dos alineamientos múltiples enfrentando por separado a los motivos de inserción TcTREZO *upstream* (*up*) y *downstream* (*down*) (Souza et al. 2007) que se describen a continuación, con todos los miembros de la librería MASP. Posteriormente los resultados fueron editados mediante el programa Genedoc (Nicholas and Nicholas 1997).

- **TcTREZO *Up*:**

```
GGCGACAGTGACGGCAGCACCGCGGTCTCCCACACCACCTCCCC  
TCTTTTGCTTCTTCTTCTTGTTGC
```

- **TcTREZO *Down*:**

```
GTGIGTGCGGCTGCGGCTGCGGTGGTGGCCGCGTGAGAGTGAGT  
GAGAGAGCGCCATGCACCGC
```

## 10. Resonancia magnética nuclear de protones (H NMR)

El análisis H NMR se realizó en la Unidad de Resonancia Magnética Nuclear perteneciente al Centro de Instrumentación Científica de la Universidad de Granada.

Esta técnica fue utilizada para analizar los metabolitos finales de la glicólisis que se producen y se excretan durante un cultivo de epimastigotes. Con este propósito, cada clon de Pan4 fue cultivado partiendo de  $1 \times 10^6$  epimastigotes/ml en 25 ml de medio fresco. En el día 5 de su crecimiento (fase exponencial), el mismo número de parásitos de cada cultivo fue centrifugado a  $300 \times g$  a  $4^\circ\text{C}$  durante 10 min. Se recogieron los sobrenadantes, que fueron procesados a través de filtros estériles Millex- GP de  $0.22 \mu\text{M}$  (Millipore) para su posterior análisis por H NMR. El espectro H NMR se obtuvo en un espectrómetro "VARIAN INOVA UNITY" de 300 MHz (2 canales) operando a 300.13 MHz con una sonda dual de banda ancha  $^1\text{H}$  con un gradiente Z de 5 mm. La temperatura de la sonda se mantuvo a  $27^\circ\text{C}$ . La técnica de pulso y la transformación de Fourier fueron utilizadas con pulsos de  $90^\circ$  y un espectro de 3287.5 Hz. Para eliminar marcas de agua se utilizó el método de presaturación, irradiando selectivamente a la frecuencia del agua durante 2.5 segundos, con

intervalos entre los pulsos de 7.5 segundos y acumulando 160 FIDs (*Free Induction Decay*), el total de los cuales fue multiplicado exponencialmente, antes de la transformación de Fourier, con una línea de ampliación de 0.2. Las señales obtenidas fueron expresadas en partes por millón (ppm).

## **11. Análisis estadístico**

La significancia de las variaciones en las curvas de crecimiento y metaciclo­génesis espontánea fueron estadísticamente analizadas usando el test “Tukey-Kramer” del software “Graph Pad InStat” version 3.05 (32 bit). Las diferencias entre líneas de parásitos isogénicas fueron consideradas significativas cuando  $p \leq 0.05$  y extremadamente significativas cuando  $p \leq 0.001$ .







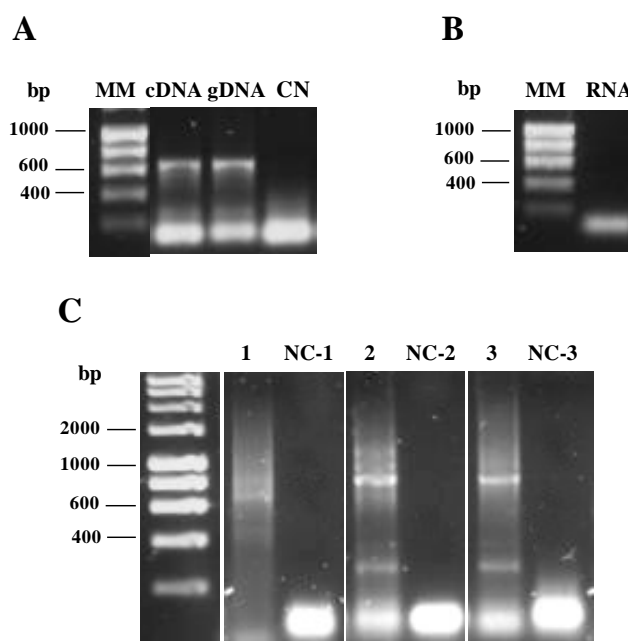
## **IV. RESULTS**





## 1. MASP expression library

Firstly, a MASP expression library was performed *cruzi* by 3-step nested PCR using the cDNA from the trypomastigotes of the Pan4 strain of *T. cruzi*. As quality control, a PCR of the AHADH2 gene (L-alpha-hydroxy acid dehydrogenase) was performed using trypomastigote gDNA, RNAm and cDNA samples (Figure 4.1A and 4.1B). The AHADH2 gene was only amplified in cDNA and gDNA samples but not in the RNAm, indicating that the RNAm was free of DNA contaminants (Figure 4.1B). On the other hand, the amplification profile in each step of the library construction presented a smear (Figure 4.1C), indicating the co-expression of several MASP transcripts with different lengths.



**Figure 4.1. Construction of the MASP expression library from trypomastigotes of the Pan4 strain by 3-step nested PCR: agarose electrophoresis gels (1%) for RNA, cDNA (quality controls) and for the 3-step nested PCR for MASP amplification. (A) PCR for AHADH2 gene using cDNA, gDNA and water as negative control. (B) Negative control PCR for AHADH2 gene using total RNA as template. (C) Three steps of the nested PCR reaction: 1- first PCR reaction for MASP family using cDNA as template**

with primers SL and 3'UTR1. 2- second PCR reaction for MASP family using the first PCR product as template with primers SL and 3'UTR2. 3- third PCR reaction for the amplification of MASP specific transcripts using the second PCR product as template with primers MASP N-term and MASP C-term. gDNA: PCR using 10 ng *T. cruzi* genomic DNA as template; NC-1- NC-2- NC-3: water control DNA of the first, the second and the third PCR reactions respectively.

To analyze the sequence of those amplified transcripts, the products of the third PCR reaction were cloned into pGEM-T vector and a total of 62 clones were sequenced. Based on the percentage of valid sequences (complete sequences without indeterminations), 93.5%, the library was considered of a good quality.

**Table 4.1. Sequencing results of the 3-step Pan4 library.**

No. clones sequenced	62
Valid sequences	58
% Valid sequences	93.5
Masp hits <sup>a</sup>	57
% Masp hits	98.3
Masp complete gene sequence	45
Masp pseudogene sequence	12
Unique Masp complete gene <sup>b</sup>	37
Unique Masp pseudogene <sup>b</sup>	12
% Unique Masp complete gene <sup>b</sup>	82.2
% Unique Masp pseudogene <sup>b</sup>	100

<sup>a</sup> Valid sequences identified as MASP members. <sup>b</sup> Unique sequences (genes and pseudogenes) are those which are not repetitive in the library. The e-value (expected value) cutoff for BLAST searches used was  $10^{-8}$ .

The *T. cruzi* (pseudo)genes corresponding to each valid sequence, as well as its closest orthologs in other *T. cruzi* strains, are shown in the Tables 4.2 and 4.3.

**Table 4.2. List of MASP genes from the 3-step MASP library and their respective orthologs in the different strains.**

GENES	bp	aa	Genbank ID	CL Brener				Dm28c				Sylvio X10/1			
				Annotation	Best Hit	Score	E-value	Annotation	Best Hit	Score	E-value	Annotation	Best Hit	Score	E-value
<i>masp1</i>	648	219	KT222275	MASP	<a href="#">TcCLB.505025.130</a>	340	2E-118	MASP	<a href="#">TCDM_10876</a>	342	1E-118	MASP	<a href="#">TCSYLVIO_009410</a>	322	1E-110
<i>masp2</i>	648	219	KT222276	MASP	<a href="#">TcCLB.505025.130</a>	343	2E-119	MASP	<a href="#">TCDM_10876</a>	344	2E-119	MASP	<a href="#">TCSYLVIO_009410</a>	325	1E-111
<i>masp3</i>	648	220	KT222277	MASP	<a href="#">TcCLB.508871.60</a>	347	7E-121	MASP	<a href="#">TCDM_10324</a>	356	2E-124	MASP	<a href="#">TCSYLVIO_009410</a>	330	1E-113
<i>masp4</i>	651	221	KT222278	MASP	<a href="#">TcCLB.508871.60</a>	358	1E-125	MASP	<a href="#">TCDM_10220</a>	358	7E-125	MASP	<a href="#">TCSYLVIO_009410</a>	317	1E-108
<i>masp5</i>	392	136	KT222279	MASP	<a href="#">TcCLB.510629.350</a>	189	1E-58	MASP	<a href="#">TCDM_10436</a>	186	2E-58	MASP	<a href="#">TCSYLVIO_002084</a>	154	5E-46
<i>masp6</i>	651	216	KT222280	MASP	<a href="#">TcCLB.508871.60</a>	351	4E-123	MASP	<a href="#">TCDM_10220</a>	356	6E-124	MASP	<a href="#">TCSYLVIO_009410</a>	313	5E-107
<i>masp7</i>	648	220	KT222281	MASP	<a href="#">TcCLB.508871.60</a>	347	7E-121	MASP	<a href="#">TCDM_10324</a>	356	2E-124	MASP	<a href="#">TCSYLVIO_009410</a>	330	1E-113
<i>masp8</i>	432	149	KT222282	MASP	<a href="#">TcCLB.511255.280</a>	201	1E-62	MASP	<a href="#">TCDM_09622</a>	281	1E-93	MASP	<a href="#">TCSYLVIO_009298</a>	137	4E-38
<i>masp9</i>	852	288	KT222283	MASP	<a href="#">TcCLB.504239.420</a>	368	4E-127	MASP	<a href="#">TCDM_10569</a>	484	8E-173	MASP	<a href="#">TCSYLVIO_002050</a>	249	5E-81
<i>masp10</i>	780	259	KT222284	MASP	<a href="#">TcCLB.506001.14</a>	276	2E-91	MASP	<a href="#">TCDM_10436</a>	330	7E-113	MASP	<a href="#">TCSYLVIO_002084</a>	154	5E-44
<i>masp11</i>	432	149	KT222285	MASP	<a href="#">TcCLB.511255.280</a>	204	8E-64	MASP	<a href="#">TCDM_09709</a>	228	5E-73	MASP	<a href="#">TCSYLVIO_009298</a>	137	4E-38
<i>masp12</i>	201	71	KT222286	MASP	<a href="#">TcCLB.510629.70</a>	122	6E-34	MASP	<a href="#">TCDM_10436</a>	114	5E-32	MASP	<a href="#">TCSYLVIO_002084</a>	115	4E-32
<i>masp13</i>	285	94	KT222287	MASP	<a href="#">TcCLB.511173.460</a>	141	2E-41	MASP	<a href="#">TCDM_12580</a>	63.5	1E-12	MASP	<a href="#">TCSYLVIO_006684</a>	62.4	3E-12
<i>masp14</i>	645	214	KT222288	MASP	<a href="#">TcCLB.506757.50</a>	326	6E-113	MASP	<a href="#">TCDM_08758</a>	321	2E-110	MASP	<a href="#">TCSYLVIO_008640</a>	324	3E-112
<i>masp15</i>	639	217	KT222289	MASP	<a href="#">TcCLB.506757.50</a>	327	3E-113	MASP	<a href="#">TCDM_08758</a>	324	6E-112	MASP	<a href="#">TCSYLVIO_008640</a>	325	2E-112
<i>masp16</i>	652	223	KT222290	MASP	<a href="#">TcCLB.506757.50</a>	322	2E-111	MASP	<a href="#">TCDM_08758</a>	316	1E-108	MASP	<a href="#">TCSYLVIO_008640</a>	318	6E-110
<i>masp17</i>	927	308	KT222291	MASP	<a href="#">TcCLB.510013.290</a>	278	4E-91	MASP	<a href="#">TCDM_08850</a>	404	2E-140	MASP	<a href="#">TCSYLVIO_006646</a>	406	2E-141
<i>masp18</i>	648	215	KT222292	MASP	<a href="#">TcCLB.506757.50</a>	329	5E-114	MASP	<a href="#">TCDM_08758</a>	322	3E-111	MASP	<a href="#">TCSYLVIO_008640</a>	326	7E-113

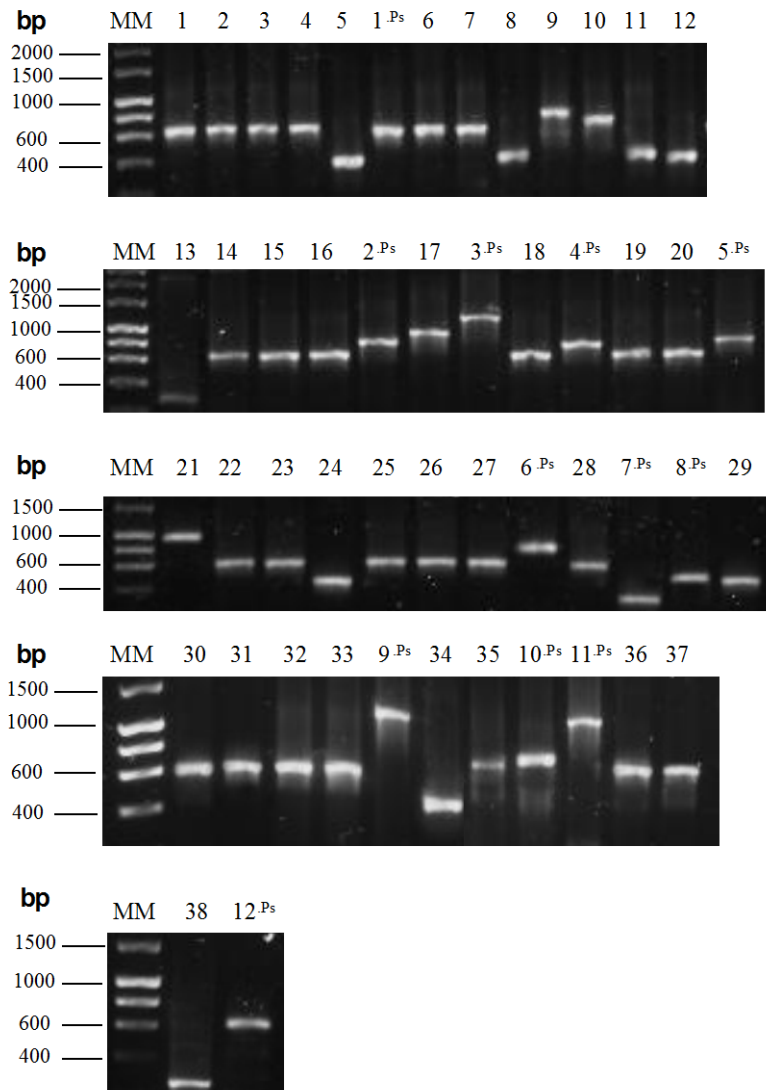
GENES	bp	aa	Genbank ID	CL Brener				Dm28c				Sylvio X10/1			
				Annotation	Best Hit	Score	E-value	Annotation	Best Hit	Score	E-value	Annotation	Best Hit	Score	E-value
<i>masp19</i>	648	215	KT222293	MASP	TcCLB.506757.50	328	2E-113	MASP	TCDM_08758	319	6E-110	MASP	TCSYLVIO_008640	323	1E-111
<i>masp20</i>	798	265	KT222294	MASP pseudogene	TcCLB.504261.100	234	2E-75	MASP	TCDM_10163	285	6E-95	MASP	TCSYLVIO_005163	264	9E-87
<i>masp21</i>	945	319	KT222295	MASP	TcCLB.511797.167	294	8E-97	MASP	TCDM_10436	301	1E-100	MASP	TCSYLVIO_002084	214	2E-66
<i>masp22</i>	648	215	KT222296	MASP	TcCLB.508871.60	347	7E-121	MASP	TCDM_10324	356	2E-124	MASP	TCSYLVIO_009410	330	1E-113
<i>masp23</i>	639	212	KT222297	MASP	TcCLB.506757.50	327	4E-113	MASP	TCDM_08758	321	1E-110	MASP	TCSYLVIO_008640	324	3E-112
<i>masp24</i>	432	144	KT222298	MASP	TcCLB.511255.280	204	1E-63	MASP	TCDM_10606	283	3E-94	MASP	TCSYLVIO_009298	138	2E-38
<i>masp25</i>	648	221	KT222299	MASP	TcCLB.506757.50	320	3E-110	MASP	TCDM_08758	310	3E-106	MASP	TCSYLVIO_008640	306	5E-105
<i>masp26</i>	648	215	KT222300	MASP	TcCLB.506757.50	329	5E-114	MASP	TCDM_08758	322	3E-111	MASP	TCSYLVIO_008640	326	7E-113
<i>masp27</i>	645	220	KT222301	MASP	TcCLB.508873.210	318	4E-109	MASP	TCDM_08758	329	6E-114	MASP	TCSYLVIO_008640	347	5E-121
<i>masp28</i>	612	204	KT222302	MASP	TcCLB.510961.10	248	1E-79	MASP	TCDM_11069	54.7	1E-08	MASP	TCSYLVIO_009300	53.1	3E-08
<i>masp29</i>	432	144	KT222303	MASP	TcCLB.511255.280	199	8E-62	MASP	TCDM_09709	239	4E-77	MASP	TCSYLVIO_009298	136	7E-38
<i>masp30</i>	638	219	KT222304	MASP	TcCLB.506757.50	328	9E-114	MASP	TCDM_08758	322	5E-111	MASP	TCSYLVIO_008640	326	8E-113
<i>masp31</i>	648	215	KT222305	MASP	TcCLB.508873.210	312	8E-107	MASP	TCDM_08758	315	2E-108	MASP	TCSYLVIO_008640	327	4E-113
<i>masp32</i>	648	215	KT222306	MASP	TcCLB.506757.50	329	7E-114	MASP	TCDM_08758	322	5E-111	MASP	TCSYLVIO_008640	326	7E-113
<i>masp33</i>	651	216	KT222307	MASP	TcCLB.506757.50	303	1E-103	MASP	TCDM_08758	297	4E-101	MASP	TCSYLVIO_008640	307	1E-105
<i>masp34</i>	410	137	KT222308	MASP	TcCLB.510629.350	189	1E-58	MASP	TCDM_10436	186	1E-58	MASP	TCSYLVIO_002084	154	5E-46
<i>masp35</i>	648	215	KT222309	MASP	TcCLB.506757.50	328	9E-114	MASP	TCDM_08758	322	6E-111	MASP	TCSYLVIO_008640	326	7E-113
<i>masp36</i>	642	213	KT222310	MASP	TcCLB.508871.60	338	9E-118	MASP	TCDM_10876	355	2E-123	MASP	TCSYLVIO_009410	324	1E-111
<i>masp37</i>	648	215	KT222311	MASP	TcCLB.506757.50	328	1E-113	MASP	TCDM_08758	321	1E-110	MASP	TCSYLVIO_008640	325	1E-112
<i>masp38</i>	219	73	KT222312	Hypothetical protein*	TcCLB.509205.50	26.9	1.40	Hypothetical protein	TCDM_09521	57.8	2E-11	Hypothetical protein*	TCSYLVIO_008612	38.9	1E-04

\* Orthologs with e-value higher than  $10^{-8}$  (the cutoff established for BLAST searches).

**Table 4.3. List of MASP pseudogenes from the 3-step MASP library and their respective orthologs in the different strains.**

PSEUDOGENES	bp	aa (to 1st stop)	Genbank ID	CL Brener (esmeraldo, no-esm, CL)				Dm28c				Sylvio X10/1			
				Annotation	Best Hit	Score	E-value	Annotation	Best Hit	Score	E-value	Annotation	Best Hit	Score	E-value
<i>masp1.Ps</i>	645	155	KT222313	MASP	<a href="#">TcCLB.510359.600</a>	940	0.00	MASP	<a href="#">TCDM_08758</a>	996	0.00	MASP	<a href="#">TCSYLVIO_008640</a>	1009	0.00
<i>masp2.Ps</i>	773	4	KT222314	MASP pseudogene	<a href="#">TcCLB.506459.299</a>	1272	0.00	MASP	<a href="#">TCDM_09835</a>	168	8E-41	MASP	<a href="#">TCSYLVIO_008115</a>	168	8E-41
<i>masp3.Ps</i>	1164	188	KT222315	MASP pseudogene	<a href="#">TcCLB.511643.73</a>	1790	0.00	MASP	<a href="#">TCDM_11700</a>	253	4E-66	MASP	<a href="#">TCSYLVIO_010405</a>	798	0.0
<i>masp4.Ps</i>	778	59	KT222316	MASP pseudogene	<a href="#">TcCLB.506967.140</a>	1011	0.00	MASP	<a href="#">TCDM_12282</a>	672	0.0	MASP	<a href="#">TCSYLVIO_006664</a>	356	3E-97
<i>masp5.Ps</i>	846	237	KT222317	MASP	<a href="#">TcCLB.508871.60</a>	933	0.00	MASP	<a href="#">TCDM_10220</a>	1050	0.00	MASP	<a href="#">TCSYLVIO_009410</a>	1050	0.00
<i>masp6.Ps</i>	854	18	KT222318	MASP pseudogene	<a href="#">TcCLB.511605.5</a>	601	5E-171	MASP	<a href="#">TCDM_12197</a>	136	5E-31	MASP	<a href="#">TCSYLVIO_010758</a>	122	1E-26
<i>masp7.Ps</i>	231	23	KT222319	MASP	<a href="#">TcCLB.508221.894</a>	311	3E-84	MASP	<a href="#">TCDM_10614</a>	264	4E-70	MASP	<a href="#">TCSYLVIO_010808</a>	284	4E-76
<i>masp8.Ps</i>	454	26	KT222320	MASP	<a href="#">TcCLB.511255.280</a>	634	0.00	MASP	<a href="#">TCDM_09622</a>	726	0.00	MASP	<a href="#">TCSYLVIO_009298</a>	246	2E-64
<i>masp9.Ps</i>	1063	39	KT222321	MASP pseudogene	<a href="#">TcCLB.511081.41</a>	1420	0.00	MASP	<a href="#">TCDM_11444</a>	168	1E-40	MASP	<a href="#">TCSYLVIO_009181</a>	178	2E-43
<i>masp10.Ps</i>	776	130	KT222322	MASP pseudogene	<a href="#">TcCLB.506459.299</a>	1274	0.00	MASP	<a href="#">TCDM_13341</a>	134	2E-30	MASP	<a href="#">TCSYLVIO_007735</a>	140	4E-32
<i>masp11.Ps</i>	1092	45	KT222323	MASP	<a href="#">TcCLB.507957.150</a>	1112	0.0	MASP	<a href="#">TCDM_10560</a>	1458	0.00	MASP	<a href="#">TCSYLVIO_006721</a>	1276	0.00
<i>masp12.Ps</i>	645	155	KT222324	MASP	<a href="#">TcCLB.506765.50</a>	928	0.00	MASP	<a href="#">TCDM_08758</a>	933	0.00	MASP	<a href="#">TCSYLVIO_005169</a>	957	0.00

The different sequences of the 3-step MASP library obtained had a range of ~1000bp length, varying from ~200 up to ~1200 bp as shown in the figure 4.2.



**Figure 4.2. Agarose electrophoresis gels (1%) with the different MASP members of the 3-step MASP library.** The numbers on the top of the gels correspond to the name of each *masp* clone. <sup>Ps</sup>, indicates pseudogene.

From the 62 amplicons sequenced, 7 MASP sequences were repetitive in the library, with a total of 50 unique sequences: 49 MASP sequences, 37 genes (75.5%) and 12 pseudogenes (24.5%), plus a gene sequence codifying for a hypothetical protein (not included in the hierarchical cluster analysis). These repetitive sequences might be a consequence of the highest expression for this sequences in the trypomastigote stage. It is interesting to note that within the 3-step MASP library, it has been found several expressed members of the same size but with different numbers and positioning of SNPs (single nucleotide polymorphisms) in their sequences (Figure 4.3). The SNPs were found all across the sequences with no special distribution or preference for insertion.



## A

```

      *           20           *           40           *           60           *           80           *           100           *
masp29/1-4 : ATGGCGATGATGATGACTGGCCGTGTGCTGCTGCTGTGTGTGCCCTCTGCGTGTGTGGTGGCGGTGCTGTGGGAGTTTATGCAAGGTATCTCGACAACAACACAGTGGGTGGCTGCATG : 117
masp11/1-4 : ATGGCGATGATGATGACTGGCCGTGTGCTGCTGCTGTGTGTGCCCTCTGCGTGTGTGGTGGCGGTGCTGTGGGAGTTTATGCAAGGTATCTCGACAACAACACAGTGGGTGGCTGCCTG : 117
masp8/1-43 : ATGGCGATGATGATGACTGGCCGTGTGCTGCTGCTGTGTGTGCCCTCTGCGTGTGTGGTGGCGGTGCTGTGGGAGTTTATGCAAGGTATCTCGACAACAACACAGTGGGTGGCTGCCTG : 117
masp24/1-4 : ATGGCGATGATGATGACTGGCCGTGTGCTGCTGCTGTGTGTGCCCTCTGCGTGTGTGGTGGCGGTGCTGTGGGAGTTTATGCAAGGTATCTCGACAACAACACAGTGGGTGGCTGCATG : 117
      ATGGCGATGATGATGACTGGCCGTGTGCTGCTGCTGTGTGTGCCCTCTGCGTGTGTGGTGGCGGTGcCqG GG GTTTATGC AGGTATCTCGACAACAAC CAGTGGGTGGCTGC TG

      120           *           140           *           160           *           180           *           200           *           220           *
masp29/1-4 : CCGTCTGGAGGTTTTGGAGCCAAAAATCGTATGCGAAGAGCGGCTGTGGGAAAACCTGCGCTTACGCTGCCTTTGCGGTCCATTTTGCCCATCCCGCCATATAAAGCCGAGGATGAG : 234
masp11/1-4 : CCGTCTGGAGGTTTTGGAGCCAAAAATCGTATGCGAAGAGCGGCTGTGGGAAAACCTGCGCTTACGCTGCCTTTGCGGTCCATTTTGCCCATCCCGCCATATAAAGCCGAGGATGAG : 234
masp8/1-43 : CCGTCTGGAGGTTTTGGAGCCAAAAATCGTATGCGAAGAGCGGCTGTGGGAAAACCTGCGCTTACGCTGCCTTTGCGGTCCATTTTGCCCATCCCGCCATATAAAGCCGAGGATGAG : 234
masp24/1-4 : CCGTCTGGAGGTTTTGGAGCCAAAAATCGTATGCGAAGAGCGGCTGTGGGAAAACCTGCGCTTACGCTGCCTTTGCGGTCCATTTTGCCCATCCCGCCATATAAAGCCGAGGATGAG : 234
      GCGTCTGGAGGTTTTGGAGCCAAAAATCGTATGCGAAGAGCGGCTGTGGGAAAACCTGCGCTTACGCTGCCTTTGCGGTCCAT TTGCCCATCCc GCcaTAcAAGCCGAGGATGAG

      240           *           260           *           280           *           300           *           320           *           340           *
masp29/1-4 : CAAGAAGCACGTGAGTACAGGGAAGGACCCCTTTGGGTTTCAGTTGCGACTACAGGTCGGGCTCTGTGTTTCTGGTGGAGGTCACGGTACTGTTGGTCTGGAGGTGCTGGTGTATTC : 351
masp11/1-4 : CAAGAAGCACGTGAGTACAGGGAAGGACCCCTTTGGGTTTCAGTTGCGACTACAGGTCGGGCTCTGTGTTTCTGGTGGAGGTCACGGTACTGTTGGTCTGGAGGTGCTGGTGTATTC : 351
masp8/1-43 : CAAGAAGCACGTGAGTACAGGGAAGGACCCCTTTGGGTTTCAGTTGCGACTACAGGTCGGGCTCTGTGTTTCTGGTGGAGGTCACGGTACTGTTGGTCTGGAGGTGCTGGTGTATTC : 351
masp24/1-4 : CAAGAAGCACGTGAGTACAGGGAAGGACCCCTTTGGGTTTCAGTTGCGACTACAGGTCGGGCTCTGTGTTTCTGGTGGAGGTCACGGTACTGTTGGTCTGGAGGTGCTGGTGTATTC : 351
      CAAGAAGCACGTGAGTACAGGGAAGGAGACC CTTTGGGTTTCAGTTGCGACTACAGGtCCGGG TCTGTTTCTGGTGGAGGtCACGGTACTGTTGGTCTGGAGGTGCTGGTGTATTC

      360           *           380           *           400           *           420           *
masp29/1-4 : GGTCTGGGCCCAATTTGAATCACCTTCTTCTGTTGATACGACTGTAACTCGCGGTGGTGGTGGCCGCGTGAGAGTGAGA : 432
masp11/1-4 : GGTCTGGGCCCAATTTGAATCACCTTCTTCTGTTGATACGACTGTAACTCGCGGTGGTGGTGGCCGCGTGAGAGTGAGA : 432
masp8/1-43 : GGTCTGGGCCCAATTTGAATCACCTTCTTCTGTTGATACGACTGTAACTCGCGGTGGTGGTGGCCGCGTGAGAGTGAGA : 432
masp24/1-4 : GGTCTGGGCCCAATTTGAATCACCTTCTTCTGTTGATACGACTGTAACTCGCGGTGGTGGTGGCCGCGTGAGAGTGAGA : 432
      GGTG TGGGCCCAATTTG AATCACCTtCTTCTGTTGATACGACTGTAACTC CGGTGGTGGTGGCCGCGTGAGAGTGAGA

```



## C

```

      *           20           *           40           *           60           *           80           *           100           *
masp26/1-6 : ATGGCGATGATGATGACTGGCCGTGTGCTGCTGGTGTGTGCCCTCTGCGTGTGTGGTTCGGTGCCTGGCCGGTGTGTTATGCCAACAAAGTGGTGGAAAGTTCCTGCCGGTGTGTTG : 115
masp32/1-6 : ATGGCGATGATGATGACTGGCCGTGTGCTGCTGGTGTGTGCCCTCTGCGTGTGTGGTTCGGTGCCTGGCCGGTGTGTTATGCCAACAAAGTGGTGGAAAGTTCCTGCCGGTGTGTTG : 115
masp35/1-6 : ATGGCGATGATGATGACTGGCCGTGTGCTGCTGGTGTGTGCCCTCTGCGTGTGTGGTTCGGTGCCTGGCCGGTGTGTTATGCCAACAAAGTGGTGGAAAGTTCCTGCCGGTGTGTTG : 115
masp19/1-6 : ATGGCGATGATGATGACTGGCCGTGTGCTGCTGGTGTGTGCCCTCTGCGTGTGTGGTTCGGTGCCTGGCCGGTGTGTTATGCCAACAAAGTGGTGGAAAGTTCCTGCCGGTGTGTTG : 115
masp37/1-6 : ATGGCGATGATGATGACTGGCCGTGTGCTGCTGGTGTGTGCCCTCTGCGTGTGTGGTTCGGTGCCTGGCCGGTGTGTTATGCCAACAAAGTGGTGGAAAGTTCCTGCCGGTGTGTTG : 115
masp18/1-6 : ATGGCGATGATGATGACTGGCCGTGTGCTGCTGGTGTGTGCCCTCTGCGTGTGTGGTTCGGTGCCTGGCCGGTGTGTTATGCCAACAAAGTGGTGGAAAGTTCCTGCCGGTGTGTTG
      ATGGCGATGATGATGACTGGCCGTGTGCTGCTGGTGTGTGCCCTCTGCGTGTGTGGTTCGGTGCCTGGCCGGTGTGTTATGCCAACAAAGTGGTGGAAAGTTCCTGCCGGTGTGTTG

      120           *           140           *           160           *           180           *           200           *           220           *
masp26/1-6 : CCAGTAACAGAACGGACGACGAGAATTTGATCTTGAATTGGTATTTTTTGGTGCGGGAAGAATGCGCGAATGAAAGCACGACAGGAGGGAAGTTGAATGTGACAGCTGAGAGGAT : 230
masp32/1-6 : CCAGTAACAGAACGGATGACGAGAATTTGATCTTGAATTGGTATTTTTTGGTGCGGGAAGAATGCGCGAATGAAAGCACGACAGGAGGGAAGTTGAATGTGACAGCTGAGAGGAT : 230
masp35/1-6 : CCAGTAACAGAACGGATGACGAGAATTTGATCTTGAATTGGTATTTTTTGGTGCGGGAAGAATGCGCGAATGAAAGCACGACAGGAGGGAAGTTGAATGTGACAGCTGAGAGGAT : 230
masp19/1-6 : CCAGTAACAGAACGGATGACGAGAATTTGATCTTGAATTGGTATTTTTTGGTGCGGGAAGAATGCGCGAATGAAAGCACGACAGGAGGGAAGTTGAATGTGACAGCTGAGAGGAT : 230
masp37/1-6 : CCAGTAACAGAACGGATGACGAGAATTTGATCTTGAATTGGTATTTTTTGGTGCGGGAAGAATGCGCGAATGAAAGCACGACAGGAGGGAAGTTGAATGTGACAGCTGAGAGGAT : 230
masp18/1-6 : CCAGTAACAGAACGGATGACGAGAATTTGATCTTGAATTGGTATTTTTTGGTGCGGGAAGAATGCGCGAATGAAAGCACGACAGGAGGGAAGTTGAATGTGACAGCTGAGAGGAT
      CCAGTAACAGAACGGATGACGAGAATTTGATCTTGAATTGGTATTTTTTGGTGCGGGAAGAATGCGCGAATGAAAGCACGACAGGAGGGAAGTTGAATGTGACAGCTGAGAGGAT

      240           *           260           *           280           *           300           *           320           *           340
masp26/1-6 : CTGTATTCAAAAGTATGAAGGGGGTTTGTGACGCTTTTTATGATAAAACGTCGGGTGTAACCTCATGACCGTGACGTTGAGTTTATTTGTACGTATTACGCTACCATTCCAGAG : 345
masp32/1-6 : CTGTATTCAAAAGTATGAAGGGGGTTTGTGACGCTTTTTATGATAAAACGTCGGGTGTAACCTCATGACCGTGACGTTGAGTTTATTTGTACGTATTACGCTACCATTCCAGAG : 345
masp35/1-6 : CTGTATTCAAAAGTATGAAGGGGGTTTGTGACGCTTTTTATGATAAAACGTCGGGTGTAACCTCATGACCGTGACGTTGAGTTTATTTGTACGTATTACGCTACCATTCCAGAG : 345
masp19/1-6 : CTGTATTCAAAAGTATGAAGGGGGTTTGTGACGCTTTTTATGATAAAACGTCGGGTGTAACCTCATGACCGTGACGTTGAGTTTATTTGTACGTATTACGCTACCATTCCAGAG : 345
masp37/1-6 : CTGTATTCAAAAGTATGAAGGGGGTTTGTGACGCTTTTTATGATAAAACGTCGGGTGTAACCTCATGACCGTGACGTTGAGTTTATTTGTACGTATTACGCTACCATTCCAGAG : 345
masp18/1-6 : CTGTATTCAAAAGTATGAAGGGGGTTTGTGACGCTTTTTATGATAAAACGTCGGGTGTAACCTCATGACCGTGACGTTGAGTTTATTTGTACGTATTACGCTACCATTCCAGAG
      CTGTATTCAAAAGTATGAAGGGGGTTTGTGACGCTTTTTATGATAAAACGTCGGGTGTAACCTCATGACCGTGACGTTGAGTTTATTTGTACGTATTACGCTACCATTCCAGAG

      *           360           *           380           *           400           *           420           *           440           *           460
masp26/1-6 : GAGCCCGTCGAACCCCAACACCCGCAAGGCCACCATCAGATTTCAGCTGTGAGCAATACCACCAGTAATGAGGGAACGCCGGAAGATGCACCCGGAATCCGATGCTGCAGGGAGAG : 460
masp32/1-6 : GAGCCCGTCGAACCCCAACACCCGCAAGGCCACCATCAGATTTCAGCTGTGAGCAATACCACCAGTAATGAGGGAACGCCGGAAGATGCACCCGGAATCCGATGCTGCAGGGAGAG : 460
masp35/1-6 : GAGCCCGTCGAACCCCAACACCCGCAAGGCCACCATCAGATTTCAGCTGTGAGCAATACCACCAGTAATGAGGGAACGCCGGAAGATGCACCCGGAATCCGATGCTGCAGGGAGAG : 460
masp19/1-6 : GAGCCCGTCGAACCCCAACACCCGCAAGGCCACCATCAGATTTCAGCTGTGAGCAATACCACCAGTAATGAGGGAACGCCGGAAGATGCACCCGGAATCCGATGCTGCAGGGAGAG : 460
masp37/1-6 : GAGCCCGTCGAACCCCAACACCCGCAAGGCCACCATCAGATTTCAGCTGTGAGCAATACCACCAGTAATGAGGGAACGCCGGAAGATGCACCCGGAATCCGATGCTGCAGGGAGAG : 460
masp18/1-6 : GAGCCCGTCGAACCCCAACACCCGCAAGGCCACCATCAGATTTCAGCTGTGAGCAATACCACCAGTAATGAGGGAACGCCGGAAGATGCACCCGGAATCCGATGCTGCAGGGAGAG
      GAGCCCGTCGAACCCCAACACCCGCAAGGCCACCATCAGATTTCAGCTGTGAGCAATACCACCAGTAATGAGGGAACGCCGGAAGATGCACCCGGAATCCGATGCTGCAGGGAGAG

      *           480           *           500           *           520           *           540           *           560           *
masp26/1-6 : GAGAACAATAATGCAAGGAGCACGGCAATACGAAGGAaAAAGCAGTCGAAACCCGCCGTGTGAAAAGTATCCCGAGGACTACCGACAGTGACGGCAGCACCCGGTCTCCACAC : 575
masp32/1-6 : GAGAACAATAATGCAAGGAGCACGGCAATACGAAGGAaAAAGCAGTCGAAACCCGCCGTGTGAAAAGTATCCCGAGGACTACCGACAGTGACGGCAGCACCCGGTCTCCACAC : 575
masp35/1-6 : GAGAACAATAATGCAAGGAGCACGGCAATACGAAGGAaAAAGCAGTCGAAACCCGCCGTGTGAAAAGTATCCCGAGGACTACCGACAGTGACGGCAGCACCCGGTCTCCACAC : 575
masp19/1-6 : GAGAACAATAATGCAAGGAGCACGGCAATACGAAGGAaAAAGCAGTCGAAACCCGCCGTGTGAAAAGTATCCCGAGGACTACCGACAGTGACGGCAGCACCCGGTCTCCACAC : 575
masp37/1-6 : GAGAACAATAATGCAAGGAGCACGGCAATACGAAGGAaAAAGCAGTCGAAACCCGCCGTGTGAAAAGTATCCCGAGGACTACCGACAGTGACGGCAGCACCCGGTCTCCACAC : 575
masp18/1-6 : GAGAACAATAATGCAAGGAGCACGGCAATACGAAGGAaAAAGCAGTCGAAACCCGCCGTGTGAAAAGTATCCCGAGGACTACCGACAGTGACGGCAGCACCCGGTCTCCACAC
      GAGAACAATAATGCAAGGAGCACGGCAATACGAAGGAaAAAGCAGTCGAAACCCGCCGTGTGAAAAGTATCCCGAGGACTACCGACAGTGACGGCAGCACCCGGTCTCCACAC

      580           *           600           *           620           *           640
masp26/1-6 : CACCTCCCCTCTTTTGTCTTCTTCTTTGTTGTTGCGTGTGCGGCTGCTGCTGCGGTGGTGGTGGCCCGGTGA : 648
masp32/1-6 : CACCTCCCCTCTTTTGTCTTCTTCTTTGTTGTTGCGTGTGCGGCTGCTGCTGCGGTGGTGGTGGCCCGGTGA : 648
masp35/1-6 : CACCTCCCCTCTTTTGTCTTCTTCTTTGTTGTTGCGTGTGCGGCTGCTGCTGCGGTGGTGGTGGCCCGGTGA : 648
masp19/1-6 : CACCTCCCCTCTTTTGTCTTCTTCTTTGTTGTTGCGTGTGCGGCTGCTGCTGCGGTGGTGGTGGCCCGGTGA : 648
masp37/1-6 : CACCTCCCCTCTTTTGTCTTCTTCTTTGTTGTTGCGTGTGCGGCTGCTGCTGCGGTGGTGGTGGCCCGGTGA : 648
masp18/1-6 : CACCTCCCCTCTTTTGTCTTCTTCTTTGTTGTTGCGTGTGCGGCTGCTGCTGCGGTGGTGGTGGCCCGGTGA
      CACCTCCCCTCTTTTGTCTTCTTCTTTGTTGTTGCGTGTGCGGCTGCTGCTGCGGTGGTGGTGGCCCGGTGA

```

## D

```

      *          20          *          40          *          60          *          80          *          100          *
masp4/1-65 : ATGGCGATGATGATGACGGCCGCTGTGCTGCTGGTGTGTGCCCTCTGCGTGTGTGGTGC GGCGTGGTTTTGCCAAGGAGGCGGAAGCTTCTGCCAATGGTGTGCCAGT : 117
masp6/1-65 : ATGGCGATGATGATGACGGCCGCTGTGCTGCTGGTGTGTGCCCTCTGCGTGTGTGGTGC GGCGTGGTTTTGCCAAGGAGGCGGAAGCTTCTGCCAATGGTGTGCCAGT : 117
      ATGGCGATGATGATG C GGCCGCTGTGCTGCTGGTGTGTGCCCTCTGCGTGTGTGGTGC GGCGTGGTTTTGCCAAGGAGGCGGAAGCTTCTGCCAATGGTGTGCCAGT

      120          *          140          *          160          *          180          *          200          *          220          *
masp4/1-65 : AAGACAACACCTGTTGACAGAATCATCTTGAATTGGCATAAGCTGACGCACGAAGAATGCGAGACGGATAATACGAAGAATGGGACGGTGAATGTGCCAGCGATGAAGAGATGCATA : 234
masp6/1-65 : AAGACAACACCTGTTGACAGAATCATCTTGAATTGGCATAAGCTGACGCACGAAGAATGCGAGACGGATAATACGAAGAATGGGACGGTGAATGTGCCAGCGATGAAGAGATGCATA : 234
      AAGACAACACCTGTTGACAGAATCATCTTGAATTGGCATAAGCTGACGCACGAAGAATGCGAGACGGATAATACGAAGAATGGGACGGTGAATGTGCCAGCGATGAAGAGATGCATA

      240          *          260          *          280          *          300          *          320          *          340          *
masp4/1-65 : CAAGTTGCTATGAAGGGGCTTTGTGACACTTTTATAATAAGACGCCGAGTGAATTCATGGCCCTGGGTTACGGATATGTGTACTATTACGCTGCCATTCCAGAGGAGCCCGTC : 351
masp6/1-65 : CAAGTTGCTATGAAGGGGCTTTGTGACACTTTTATAATAAGACGCCGAGTGAATTCATGGCCCTGGGTTACGGATATGTGTACTATTACGCTGCCATTCCAGAGGAGCCCGTC : 351
      CAAGTTGC ATGAAGGGG TTTGTGACACTTTTATAATAAGACGCCGAGTGAATTCATGGCC G GGTACGGATATGTGTAC TATTACGCTGC ATTCCAGAGGAGCCCGTC

      360          *          380          *          400          *          420          *          440          *          460
masp4/1-65 : GAACCCCAACACCCGAAGGGCCACACCAGATTCAGTGTGTCAGCAATACCCCACTAATGAGGGGACGCCGAGAAGATGCACCGGAATTCGATCTGCAGGGAGAGAAGGAAAAACA : 468
masp6/1-65 : GAACCCCAACACCCGAAGGGCCACACCAGATTCAGTGTGTCAGCAATACCCCACTAATGAGGGGACGCCGAGAAGATGCACCGGAATTCGATCTGCAGGGAGAGAAGGAAAAACA : 468
      GAACCCCAACACCCGAAGGGCCAC ACCAGATTCAG TGTGTCAGCAATACCCCACTAATGAGGGGACGCCGAGAAGATGCACCGGAATTCGAT CTGCAGGGAGAGAAGGAAAAACA

      *          480          *          500          *          520          *          540          *          560          *          580
masp4/1-65 : GATGAAATGAGCAGCGCAATACGAAGGAAAAAGCATTCGAAAACGCCGCTGTGAAAGTTATCACCAAGAATACCGACAGTGACACCGACAGCAGCACCGGGTCTCCACCCACC : 585
masp6/1-65 : GATGAAATGAGCAGCGCAATACGAAGGAAAAAGCATTCGAAAACGCCGCTGTGAAAGTTATCACCAAGAATACCGACAGTGACACCGACAGCAGCACCGGGTCTCCACCCACC : 585
      GATG AAATGAGCAGCGCAATACGAAGGAAAAAGCA TCGAAAACGCCGCTGTGAAAGTTATCAC AAGAATACCGACAGTGACA CGACAGCAGCACCG GGTCTCCAC CCACC

      *          600          *          620          *          640          *
masp4/1-65 : TCCCCTCTTTTGCTTCTTCTGTTGTTGTTGTTCGGCTCTGCTGCGGTGGTGGTGGCCGCGTGA : 651
masp6/1-65 : TCCCCTCTTTTGCTTCTTCTGTTGTTGTTGTTCGGCTCTGCTGCGGTGGTGGTGGCCGCGTGA : 651
      TCCCCTCTTTTGCTTCTTCTGTTGTTGTTGTTCGGCTCTGCTGCGGTGGTGGTGGCCGCGTGA

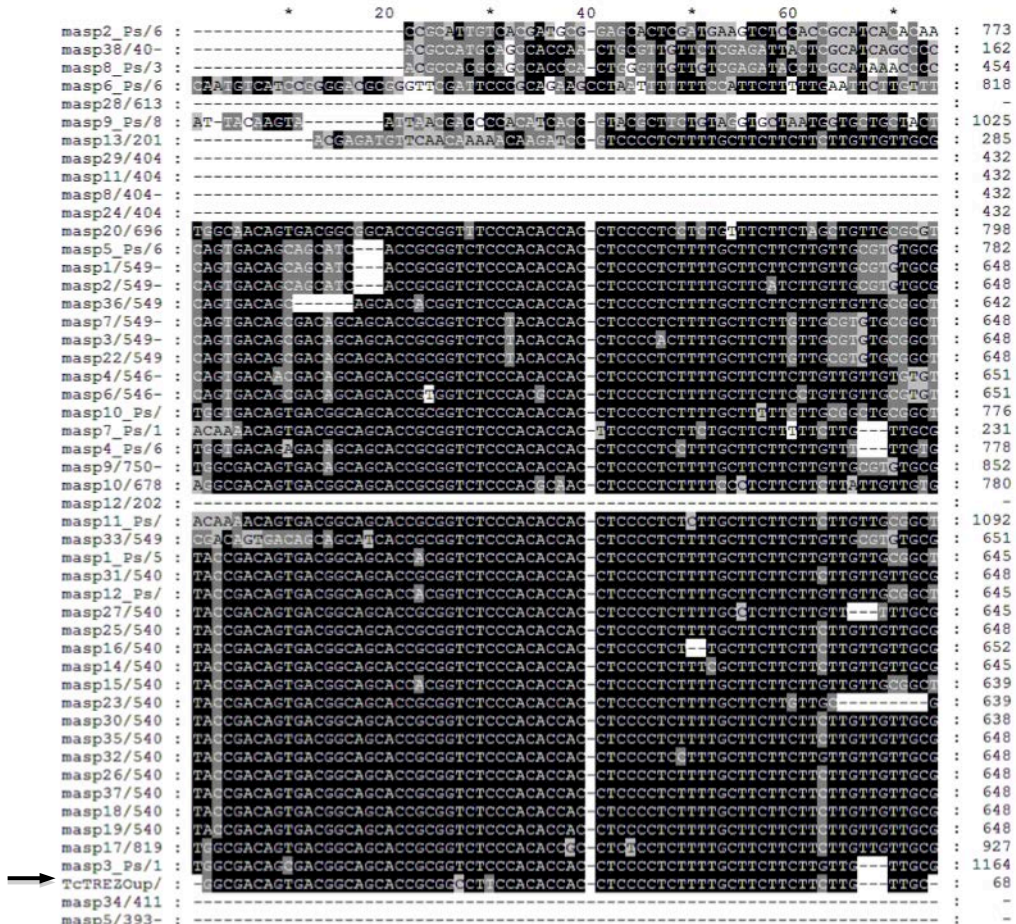
```

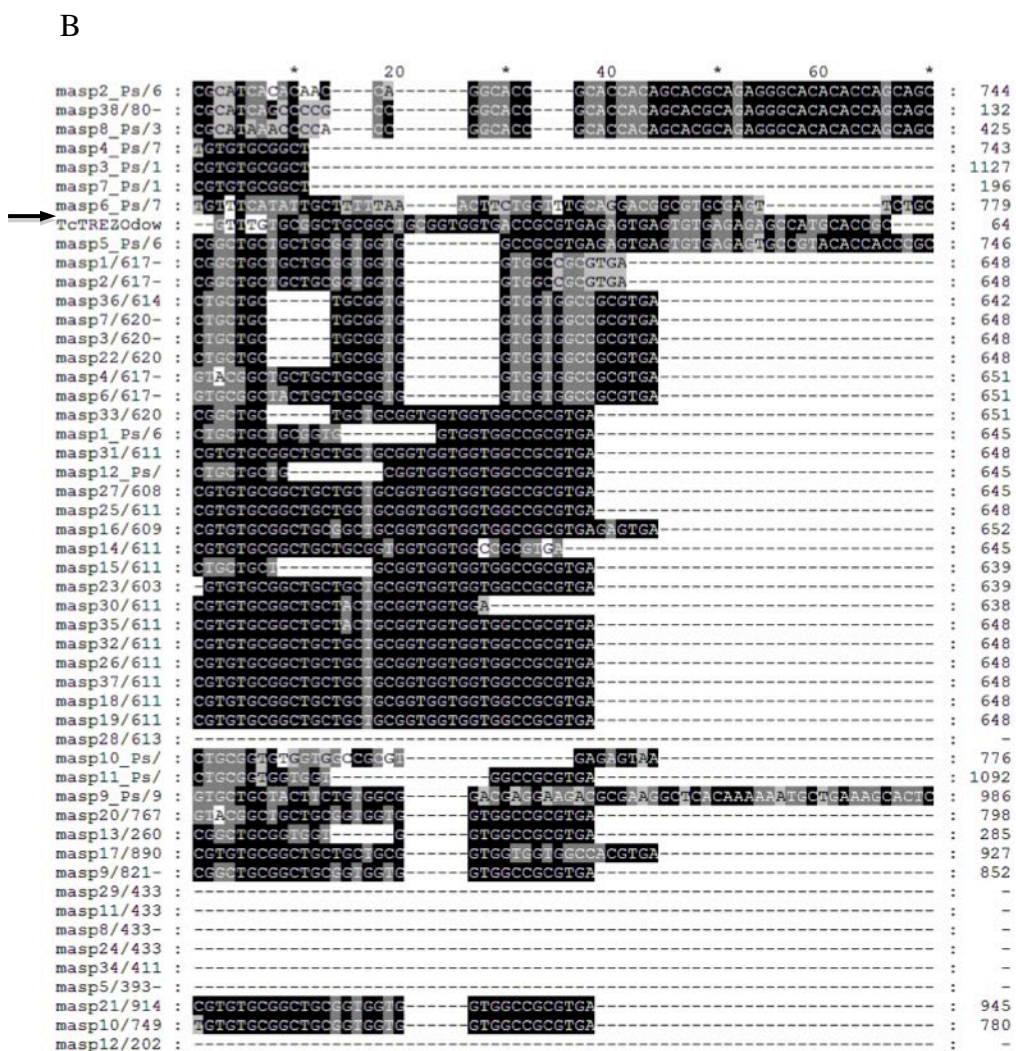
**Figure 4.3.** SNPs analysis by multiple alignment among different MASP sequences from the 3-step MASP library with lengths of 432 bp (A), 638 bp and 639 bp (B), 648 bp (C) and 651 bp (D). Complete homology among the sequences is represented with black areas whereas nucleotides in grey represent areas with less than 80% of homology.

## 2. Search of TcTREZO specific insertion sites in the 3-step MASP library

The figure 4.4 shows the results of the alignment and the location of the two insertion sites of TcTREZO respect with the members of the 3-step MASP library. This analysis revealed that the extreme 3' of the 73.5% of the MASP sequences (36/49) showed more than 80% of identity with the TcTREZO *upstream* specific site (Figure 4.4A). On the other hand, the specific motif TcTREZO *downstream* showed lower levels of identity with the members of the MASP library. Only 11 members of the library conserved, in their extreme 3', a part (~50%) of this insertion specific sequence (figure 4.4B).

A

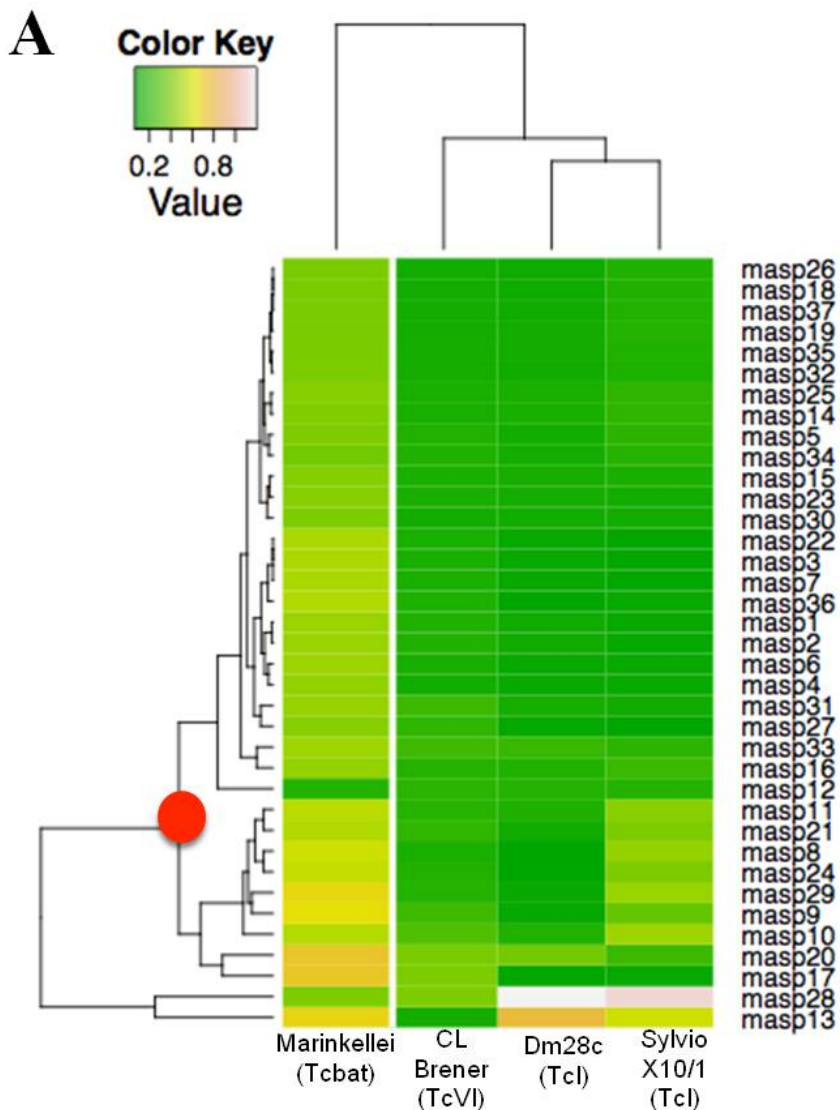


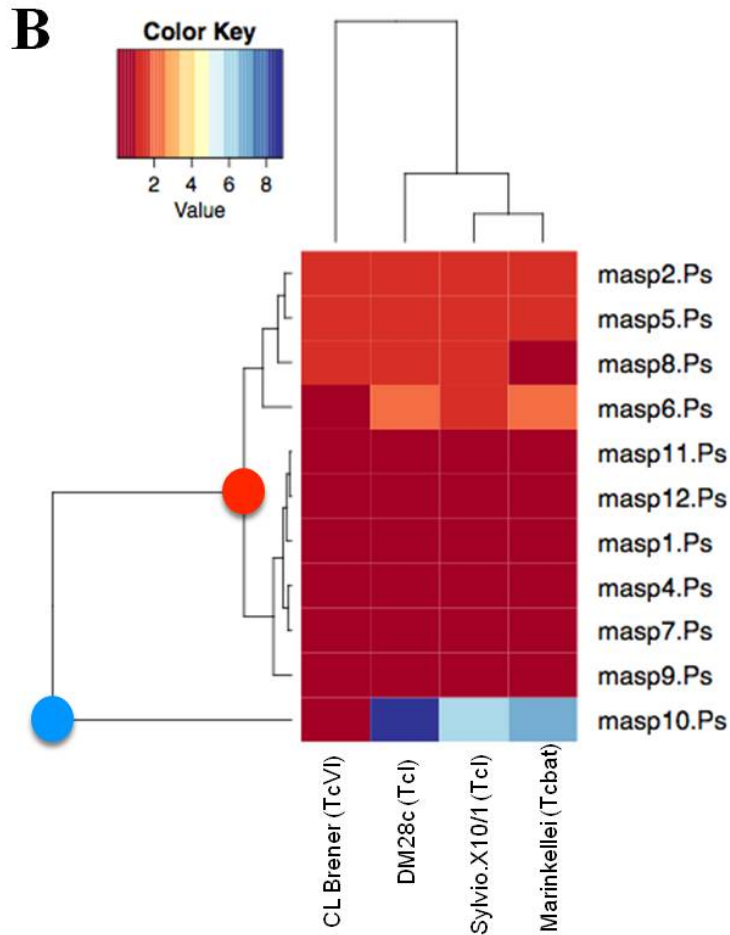


**Figure 4.4. Multiple alignment for the TcTREZO specific insertion sites respect with the 3-step MASP library.** (A) Multiple alignment for TcTREZO upstream. (B) Multiple alignment for TcTREZO downstream. The arrows indicate the position of the TcTREZO motifs in the alignment. Sequences corresponding to the conserved residues are shaded in black (>80% conservation) and gray (60 to 80% conservation); no shading denotes residues with <60% conservation.

### 3. 3-step MASP library hierarchical cluster analysis

The hierarchical cluster analysis showed as heatmaps, allowed to represent the distances of the 3-step MASP library with the closest MASP orthologs identified in CL Brener, Dm28c, Sylvio X10/1 and Marinkellei genomes at the gene and pseudogene level (Figure 4.5; Table 4.4). In this manner, was possible to study the range of conservation of the different MASP sequences not only among the outlined *T. cruzi* strains, but also within the own 3-step MASP library.





**Figure 4.5. Hierarchical cluster analyses based on the 3-step MASP library according to the pairwise distances with the ortholog sequences in CL Brener, Dm28c, Sylvio X10/1 and Marinkellei genomes.** Heatmaps were splitted on gene (A) and pseudogene (B) sequences. The horizontal nodes and their length represent the distance among the MASP members of the 3-step MASP library while the vertical nodes and their length represent the distances among the outlined strains. In the heatmap of genes, the green colors show the minimum distance while white colors indicate maximum distance. In the heatmap of pseudogenes, the minimum and maximum distances are represented with red and blue colors respectively. The red dots show the point of divergence between the slow and fast evolving groups of MASP genes and pseudogenes. The blue dot indicates the point of divergence for *masp10.Ps*, which has the highest mean distance pairwise, being an exception to the pseudogene distribution in the two groups outlined.



The analysis at gene level showed a correlation between the MASP genes and the phylogenetic origin of the different strains (vertical nodes). As was expected, the strain Marinkellei (outgroup), which is a bat-associated subspecies of *T. cruzi* (Tcbat) (Cosentino and Aguero 2012; Franzen et al. 2012; Marcili et al. 2009), showed the most distanced MASP pattern, whereas Sylvio X10/1 and Dm28c strains (Tcl) were the closest to the Pan4 strain (Tcl) (Figure 4.5A). The clustering of the individual MASP genes (horizontal nodes) also showed two main groups of MASP sequences that, according with their distances, can be grouped as slow ( $0.22 \pm 0.13$ ) and fast evolving ( $0.40 \pm 0.27$ ) (Figure 4.5A). The mean distances of all MASP genes in each strain, represented as boxplots, showed similar values for CL-Brener, Sylvio X10/1 and Dm28c strain, being Marinkellei again the most distanced strain to the Pan4 library (Figure 4.5; Table 4.4).

At the pseudogene level, the cluster analysis showed a loss of correlation with the phylogenetic origin of the sequences (vertical nodes). A singular member in the 3-step MASP library was the pseudogene *masp3.ps*, which presented a sequence such divergent with respect to the rest of the strains that was not possible to estimate the distance values and was therefore excluded from the hierarchical study. Thus, 11 MASP pseudogenes were analyzed and, according to the distances, were clustered into two different groups as slow ( $0.24 \pm 0.19$ ) and fast evolving ( $1.15 \pm 0.37$ ), with the exception of the member *masp10.ps* ( $5.52 \pm 3.82$ ) (horizontal nodes) (Figure 4.5B; Table 4.4). This pseudogene showed high distance values with respect to the closest orthologs in Sylvio X10/1, Dm28c and Marinkellei strains, but at the same time, *masp10.ps* presented an enormous degree of conservation in the CL Brener strain, being detected in the MASP libraries performed by dos Santos *et. al* from the cDNA of bloodstream trypomastigotes after 2 and 10 passages in rhesus-monkey epithelial cells (LLC-MK2) and from trypomastigotes isolated from rat mioblast cells (L6) after 14 passages (dos Santos et al. 2012). Additionally, *masp10.ps* also showed high percentages of identity with members of other preliminary MASP libraries previously performed in our research group in CL Brener, Pan4 and Maracay strains of *T. cruzi* (De Pablos 2010), with eventual recovering of the complete

open reading frame in the case of the Maracay (Figure 4.6). Although in general the mean pairwise distances were smaller among genes compared to pseudogene MASP sequences, it is interesting to note the distance similarity of the pseudogenes among the five strains of *T. cruzi* analyzed (Figure 4.5 and 4.7; Table 4.4).

masp10.Ps\*..... **ATGGCGATGATGATGACTGGCCGTGTGCTGCTGGTGTGTGCCCTCTGCCTGCTGTGGTGCCTGCCTGGTTGTGTGATGCG** : 81  
 masp6Maracay\*\*..... **ATGGCGATGATGATGACTGGCCGTGTGCTGCTGGTGTGTGCCCTCTGCCTGCTGTGGTGCCTGCCTGGTTGTGTGATGCG** : 81  
 masp3CL-Brener\*\*..... **ATGGCGATGATGATGACTGGCCGTGTGCTGCTGGTGTGTGCCCTCTGCCTGCTGTGGTGCCTGCCTGGTTGTGTGATGCG** : 81  
 masp3PS3Pan4\*..... **ATGGCGATGATGATGACTGGCCGTGTGCTGCTGGTGTGTGCCCTCTGCCTGCTGTGGTGCCTGCCTGGTTGTGTGATGCG** : 81  
 TcCLB.506459.299\*\*\*. **ATGGCGATGATGATGACTGGCCGTGTGCTGCTGGTGTGTGCCCTCTGCCTGCTGTGGTGCCTGCCTGGTTGTGTGATGCG** : 81

masp10.Ps\*..... **GTTGGAGACTTCATCGAGTGTCTCCGCATCGTGACAAATGCGGAGGAAACGACGACTCCAGAGGAGGTGACTGAGGCTATCGAA** : 162  
 masp6Maracay\*\*..... **GTTGGAGACTTCATCGAGTGTCTCCGCATCGTGACAAATGCGGAGGAAACGACGACTCCAGAGGAGGTGACTGAGGCTATCGAA** : 162  
 masp3CL-Brener\*\*..... **GTTGGAGACTTCATCGAGTGTCTCCGCATCGTGACAAATGCGGAGGAAACGACGACTCCAGAGGAGGTGACTGAGGCTATCGAA** : 162  
 masp3PS3Pan4\*..... **GTTGGAGACTTCATCGAGTGTCTCCGCATCGTGACAAATGCGGAGGAAACGACGACTCCAGAGGAGGTGACTGAGGCTATCGAA** : 162  
 TcCLB.506459.299\*\*\*. **GTTGGAGACTTCATCGAGTGTCTCCGCATCGTGACAAATGCGGAGGAAACGATGACTCCAGAGGAGGTGACTGAGGCTATCGAA** : 162

masp10.Ps\*..... **AAGCTCGCAGAGGGAGTGTGGACATGATTCGAGTATCTTGTGCGAATGTCCTTCATTACGAGTGTAATGCCACGTTACCC** : 243  
 masp6Maracay\*\*..... **AAGCTCGCAGAGGGAGTGTGGACATGATTCGAGTATCTTGTGCGAATGTCCTTCATTACGAGTGTAATGCCACGTTACCC** : 243  
 masp3CL-Brener\*\*..... **AAGCTCGCAGAGGGAGTGTGGACATGATTCGAGTATCTTGTGCGAATGTCCTTCATTACGAGTGTAATGCCACGTTACCC** : 243  
 masp3PS3Pan4\*..... **AAGCTCGCAGAGGGAGTGTGGACATGATTCGAGTATCTTGTGCGAATGTCCTTCATTACGAGTGTAATGCCACGTTACCC** : 243  
 TcCLB.506459.299\*\*\*. **AAGCTCGCAGAGGGAGTGTGGACATGATTCGAGTATCTTGTGCGAATGTCCTTCATTACGAGTGTAATGCCACGTTACCC** : 243

masp10.Ps\*..... **GCGCACAACGATACAGAAGTTGAGAAATGCGTAAACAAAACGGCGGTTGGATTTCCGGCGTCACCCCTCGTGGCCGGCCT** : 324  
 masp6Maracay\*\*..... **GCGCACAACGATACAGAAGTTGAGAAATGCGTAAACAAAACGGCGGTTGAGATTTCCGGCGTCACCCCTCGTGGCCGGCCT** : 324  
 masp3CL-Brener\*\*..... **GCGCACAACGATACAGAAGTTGAGAAATGCGTAAACAAAACGGCGGTTGAGATTTCCGGCGTCACCCCTCGTGGCCGGCCT** : 324  
 masp3PS3Pan4\*..... **GCGCACAACGATACAGAAGTTGAGAAATGCGTAAACAAAACGGCGGTTGGATTTCCGGCGTCACCCCTCGTGGCCGGCCT** : 324  
 TcCLB.506459.299\*\*\*. **GCGCACAACGATACAGAAGTTGAGAAATGCGTAAACAAAACGGCGGTTGGATTTCCGGCGTCACCCCTCGTGGCCGGCCT** : 324

masp10.Ps\*..... **GCAGCCGCAACCACAGACGCCACTTTGCGCTCGCCATTGCCGTCCTCAAGTTCCTGCTCCAGCGCATAGCGGGGCAGGCG** : 405  
 masp6Maracay\*\*..... **GCAGCCGCAACCACAGACGCCACTTTGCGCTCGCCATTGCCGTCCTCAAGTTCCTGCTACCAGCGCCTTGCAAGGGGCAGGCG** : 405  
 masp3CL-Brener\*\*..... **GCAGCCGCAACCACAGACGCCACTTTGCGCTCGCCATTGCCGTCCTCAAGTTCCTGCTACCAGCGCCTTGCAAGGGGCAGGCG** : 405  
 masp3PS3Pan4\*..... **GCAGCCGCAACCACAGACGCCACTTTGCGCTCGCCATTGCCGTCCTCAAGTTCCTGCTCCAGCGCATAGCGGGGCAGGCG** : 405  
 TcCLB.506459.299\*\*\*. **GCAGCCGCAACCACAGACGCCACTTTGCGCTCGCCATTGCCGTCCTCAAGTTCCTGCTCCAGCGCATAGCAAGGGGCAGGCG** : 405

```

masp10.Ps*.....CAAGGATCACCAGATGCTGCACCAGAGCCTGTGAAACCCCGCAGACAGATCCGGTCAATAATGAACAGCTGTCTGAAACG : 486
masp6Maracay**.....CAAGGATCACCAGGCTGCTGCACCAGAGCCTGTGAAACCCCGCAGACAGATTCAGTCAATACCGAACAGCTGTCTGAAACG : 486
masp3CL-Brener**.....CAAGGATCACCAGCTGCTGCACCAGAGCCTGTGAAACCCCGCAGACAGACCCGGTCAATAGTGAACAGCTGTCTGAAACG : 486
masp3PS3Pan4**.....CAAGGATCACCAGATGCTGCACCAGAGCCTGTGAAACCCCGCAGACAGATCCGGTCAATAATGAACAGCTGTCTGAAACG : 486
TcCLB.506459.299***..CAAGGATCACCAGCTGCTGCACCAGAGCCTGTGAAACCCCGCAGACAGACCCGGTCAATAGTGAACAGCTGTCTGAAACG : 486

masp10.Ps*.....CCGGCCTCCACAGGCGGTGGTCCCTGGTGGCGGTGGTGAGTTGGCTGACAGGAGCAGCCGTACACCGACCGACGTGCGAAGA : 567
masp6Maracay**.....CCGGCCTCCACAGGCGGTGGT-----GAGTTGGCTGACAGGAGCAGCCATACACCGACCGACGTGCGAAGA : 552
masp3CL-Brener**.....CCGGCCTCCACAGGCGGTGGTCCCTGGTGGCGGTGGTGAGTTGGCTGACAGGAGCAGCCGTACACCGACCGACGTGCGAAGA : 567
masp3PS3Pan4**.....CCGGCCTCCACAGGCGGTGGTCCCTGGTGGCGGTGGTGAGTTGGCTGACAGGAGCAGCCGTACACCGACCGACGTGCGAAGA : 567
TcCLB.506459.299***..CCGGCCTCCACAGGCGGTGGTCCCTGGTGGTGGTGGTGAGTTGGCTGACAGGAGCAGCCGTACACCGACCGACGTGCGAAGA : 567

masp10.Ps*.....ACACATGAAGCAGGCCAGCTGGTAACGCCACTTCCACACCGACGGCACTCAGGGGATTTCGAGAAGTCACCGGCCGAACTC : 648
masp6Maracay**.....ACACATGAAGCAGGCCAGCTGGTAACGCCA ACTTCCACACCGACGGCACTCAGGGGATTTCGAGAAGTCACCGGCCGAACTC : 633
masp3CL-Brener**.....ACACATGAAGCAGGCCAGCTGGTAACGCCACTTCCACACCGACGGCACTCAGGGGATTTCGAGAAGTCACCGGCCGAACTC : 648
masp3PS3Pan4**.....ACACATGAGCAGGCCAGCTGGTAACGCCACTTCCACACCGACGGCACTCAGGGGATTTCGAGAAGTCACCGGCCGAACTC : 648
TcCLB.506459.299***..ACACATGAAGCAGGCCAGCTGGTAACGCCACTTCCACACCGACGGCACTCAGGGGATTTCGAGAAGTCACCGGCCGAACTC : 648

masp10.Ps*.....ACCACCCGAACACGTGAGACAAACGAGACTGGTGACAGTGACGGCAGCACCCGGGTCCTCCACACACCTCCCCTCTTTTG : 729
masp6Maracay**.....ACCACCCGAACACGTGAGACAAACGAGACTGGCGACAGTGACGGCAGCACCCGGGTCCTCCACAGCACCTCCCCTC--TTTG : 712
masp3CL-Brener**.....ACCACCCGAACACGTGAGACAAACGAGACTGGTGACACCAGCAGGGCAGCACCCGGGTCCTCCACTCCACCTCCCCTCTTTTG : 729
masp3PS3Pan4**.....ACCACCCGAACACGTGAGACAAACGAGACTGGTGACAGTGACGGCAGCACCCGGGTCCTCCACACACCTCCCCTCTTTTG : 729
TcCLB.506459.299***..ACCACCCGAACACGTGAGACAAACGAGACTGGTGACACCAGCAGGGCAGCACCCGGGTCCTCCACTCCACCTCCCCTCTTTTG : 729

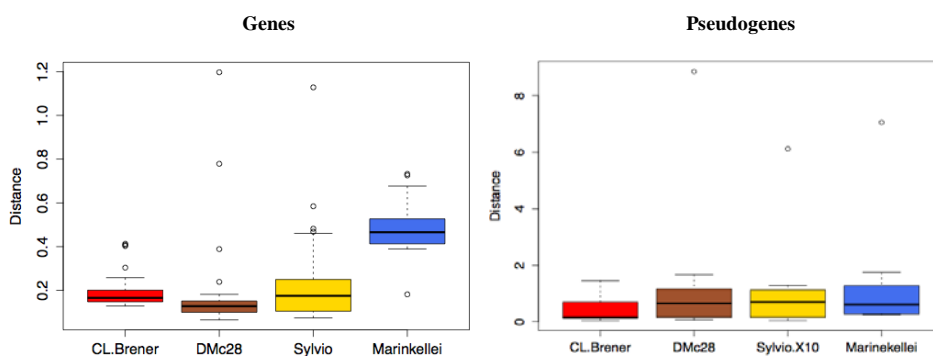
```

**Figure 4.6.** SNP's analysis by Clustal multiple alignment of MASP expressed sequences in different libraries and strains of *T. cruzi*. Complete homology among the sequences is represented with black areas whereas grey nucleotides represent areas with less than 80% of homology. \*3-step MASP library (trypomastigote stage); \*\*preliminary MASP library (metacyclic trypomastigote stage) ; \*\*\* dos Santos et al. MASP library (bloodstream trypomastigote stage) (dos Santos et al. 2012).

**Table 4.4. Distances obtained with MEGA6.06 among the members of the 3-step MASP library and the ortholog sequences identified in CL Brener, Dm28c, Sylvio X10/1 and Marinkellei strains of *T. cruzi*.**

Clone	CL Brener	Dm28c	Sylvio X10/1	Marinkellei
<i>masp1</i>	0.177	0.126	0.098	0.477
<i>masp2</i>	0.177	0.126	0.093	0.477
<i>masp3</i>	0.157	0.100	0.083	0.511
<i>masp4</i>	0.131	0.103	0.103	0.458
<i>masp5</i>	0.178	0.131	0.203	0.411
<i>masp6</i>	0.148	0.098	0.103	0.482
<i>masp7</i>	0.157	0.100	0.083	0.511
<i>masp8</i>	0.156	0.065	0.461	0.582
<i>masp9</i>	0.254	0.097	0.351	0.646
<i>masp10</i>	0.304	0.177	0.482	0.534
<i>masp11</i>	0.190	0.168	0.438	0.542
<i>masp12</i>	0.201	0.182	0.182	0.182
<i>masp13</i>	0.141	0.779	0.585	0.677
<i>masp14</i>	0.150	0.156	0.210	0.421
<i>masp15</i>	0.157	0.145	0.157	0.431
<i>masp16</i>	0.186	0.174	0.242	0.462
<i>masp17</i>	0.407	0.085	0.105	0.726
<i>masp18</i>	0.129	0.123	0.175	0.403
<i>masp19</i>	0.129	0.129	0.181	0.403
<i>masp20</i>	0.403	0.389	0.250	0.732
<i>masp21</i>	0.219	0.128	0.405	0.527
<i>masp22</i>	0.157	0.100	0.083	0.511
<i>masp23</i>	0.152	0.141	0.135	0.436
<i>masp24</i>	0.177	0.065	0.405	0.566
<i>masp25</i>	0.152	0.152	0.212	0.433
<i>masp26</i>	0.129	0.123	0.175	0.403
<i>masp27</i>	0.216	0.096	0.074	0.434
<i>masp28</i>	0.413	1.197	1.128	0.413
<i>masp29</i>	0.181	0.110	0.468	0.667
<i>masp30</i>	0.135	0.129	0.135	0.413
<i>masp31</i>	0.248	0.142	0.125	0.466
<i>masp32</i>	0.139	0.128	0.168	0.408
<i>masp33</i>	0.258	0.239	0.202	0.488

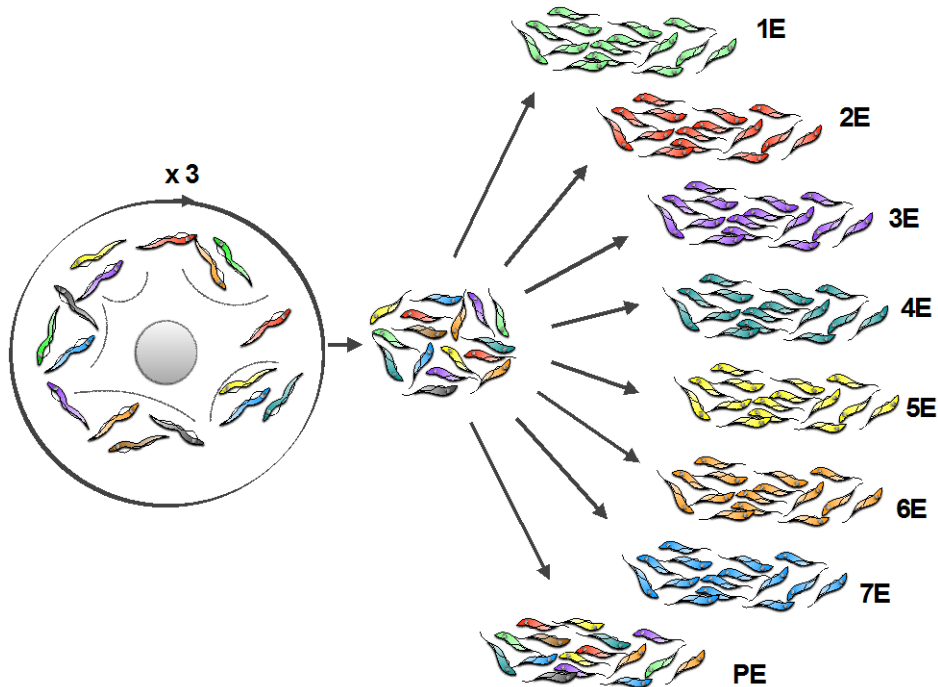
<i>masp34</i>	0.176	0.141	0.188	0.390
<i>masp35</i>	0.140	0.129	0.175	0.403
<i>masp36</i>	0.166	0.082	0.104	0.526
<i>masp37</i>	0.129	0.123	0.175	0.403
<i>masp1.Ps</i>	0.327	0.070	0.048	0.269
<i>masp2.Ps</i>	1.449	1.172	1.137	1.333
<i>masp4.Ps</i>	0.110	0.218	0.401	0.261
<i>masp5.Ps</i>	1.101	1.149	1.113	1.213
<i>masp6.Ps</i>	0.126	1.668	1.285	1.740
<i>masp7.Ps</i>	0.150	0.222	0.217	0.242
<i>masp8.Ps</i>	1.070	1.073	0.909	0.789
<i>masp9.Ps</i>	0.143	0.651	0.693	0.608
<i>masp10.Ps</i>	0.044	8.855	6.117	7.049
<i>masp11.Ps</i>	0.136	0.066	0.089	0.327
<i>masp12.Ps</i>	0.089	0.095	0.077	0.243



**Figure 4.7.** Boxplots representing the pairwise distances between the 3-step MASP library and the ortholog sequences of CL Brener, Dm28c, Sylvio X10/1 and Marinkellei. The boxplots show the distribution of the mean pairwise distances values for the ortholog sequences among the different strains.

#### 4. Analysis of the MASP clonally variant expression

As was outlined about, we performed a cloning process where 7 clonally isogenic populations derived from single cells were isolated from a parental culture of the Pan4 strain (Figure 4.8).

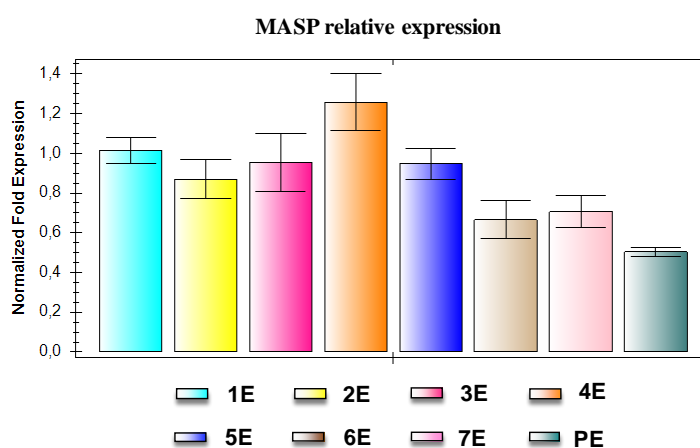


**Figure 4.8. Representation of the cloning process with three passages of the trypanosomes through VERO cells and further cloning procedure to obtain the final clonal independent cell lines from the parental line. [1E-7E], name of each epimastigote clonal population; PE, epimastigote forms of the parental cell line**

To study the clonal heterogeneity in the expression of the MASP family we analyzed, at RNA and protein levels, the expression of all the clonal lines, keeping the same passages and conditions for all the cell cultures in order to make them comparable (see material and methods, section 4).

## 4.1 RT-qPCR

At RNA level, the RT-qPCR analysis allowed to obtain a global quantification of the MASP transcripts by amplifying the conserved 5' extreme region of this family on the epimastigote forms. These results revealed clonal differences with approximately a 2.5 fold change variation in expression among the different cell lines. The clone4 showed the highest MASP expression while the parental strain had the lowest levels (Figure 4.9).



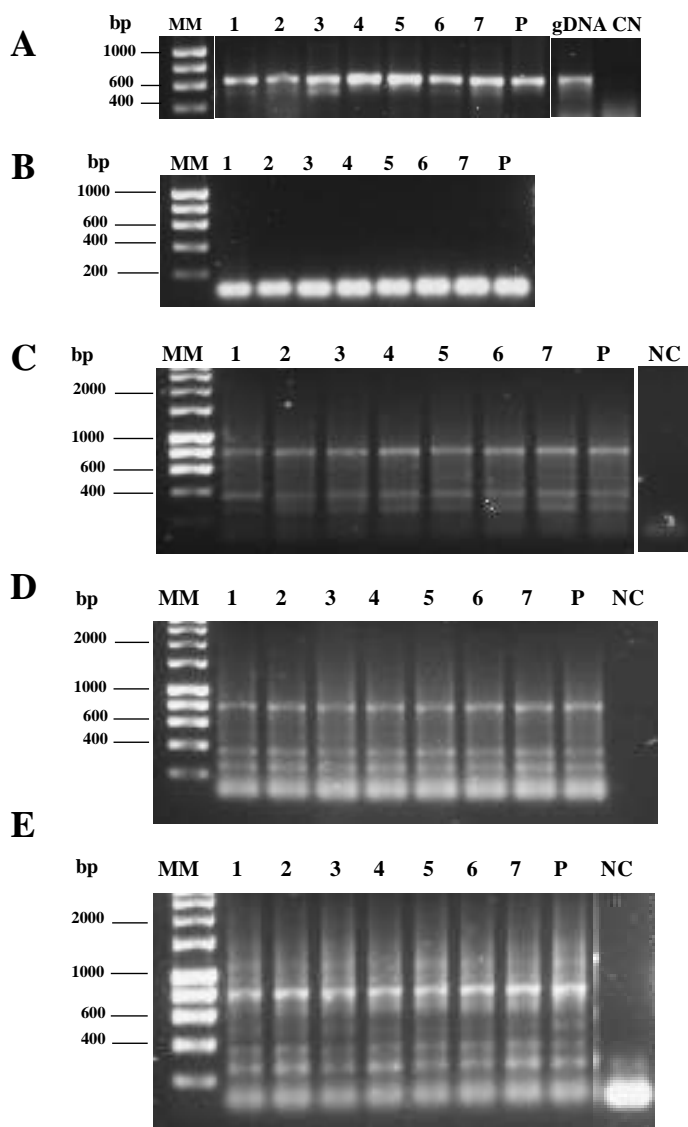
**Figure 4.9.** Quantification of the MASP clonal expression by qRT-PCR in the epimastigotes forms of all the clones and the parental strain. The analysis was normalized using the GAPDH gene. [1E-7E], name of each epimastigote clonal population; PE, epimastigote forms of the parental cell line.

## 4.2 Capillary electrophoresis

With the aim of obtaining specific MASP transcriptional differences between the clones, the 3-step nested PCR strategy used to construct a MASP expression library (previously described in material and methods, section 6.9.1) was also applied to finally get an *in silico* picture of the MASP transcriptional landscape of clonal populations in *T.cruzi*. Thus, the products from each step of the 3-nested

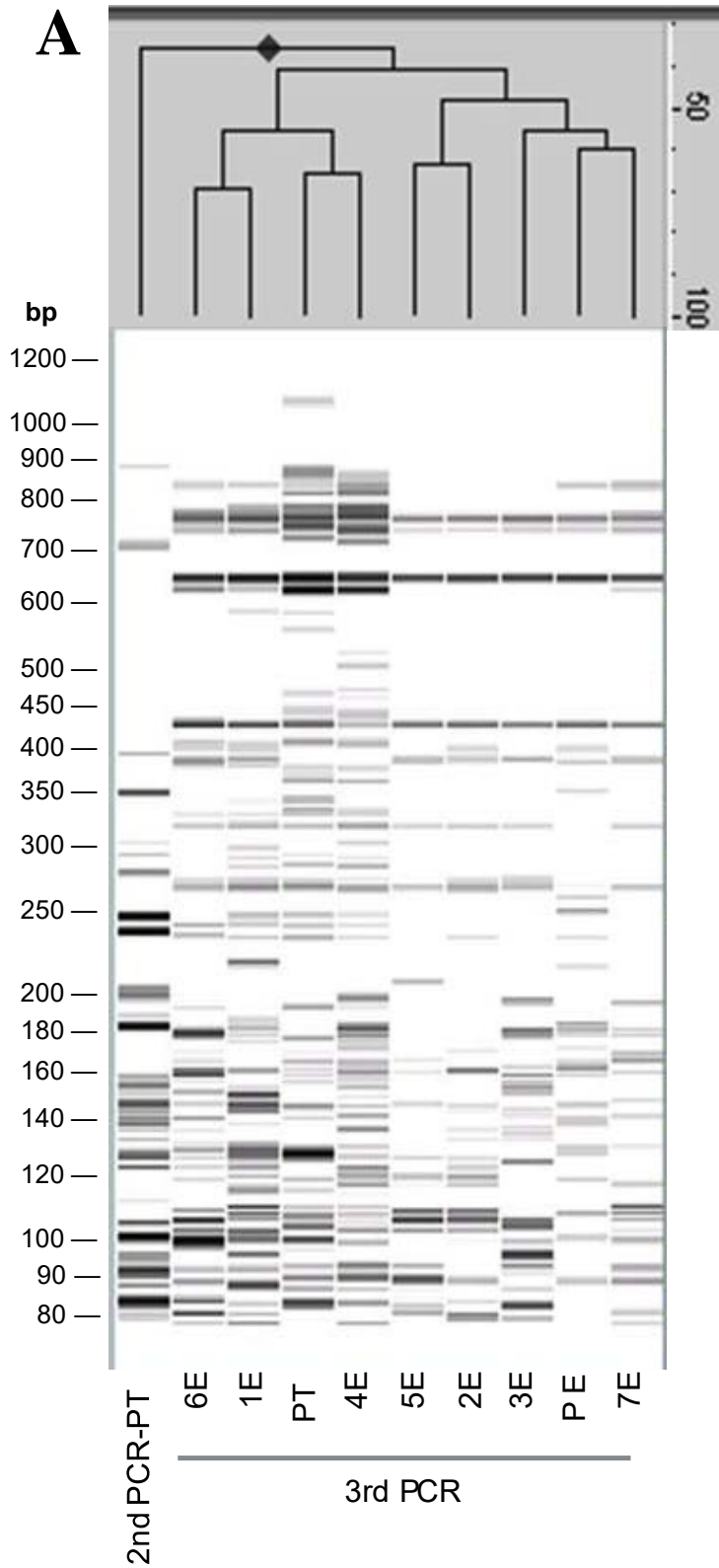


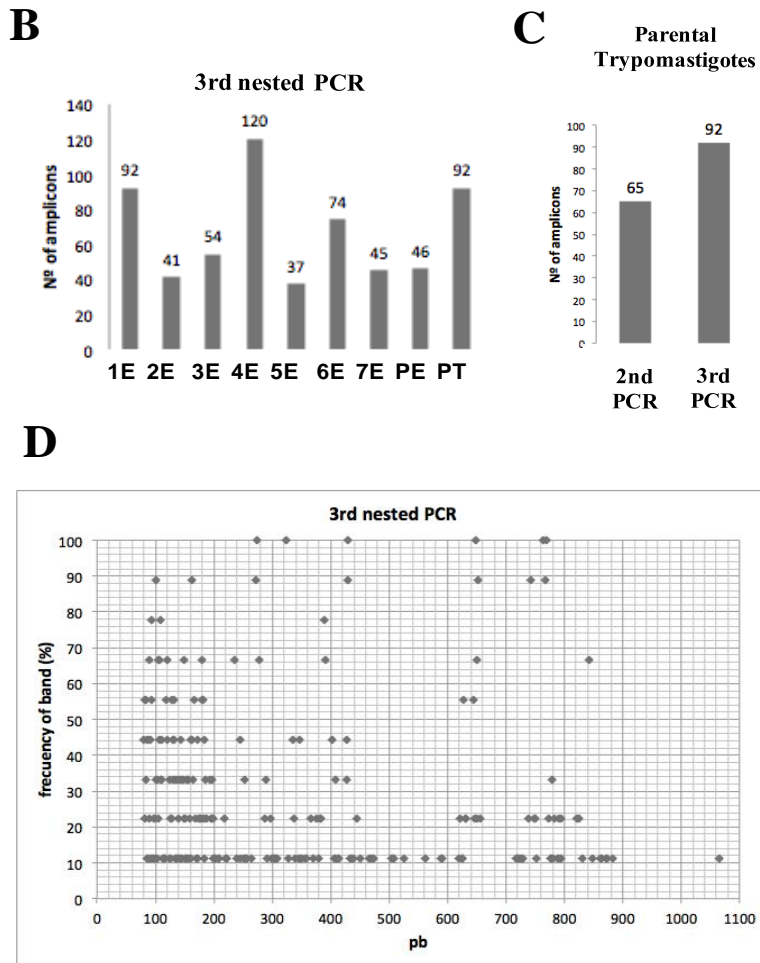
PCR from all the clones and the parental line in the epimastigote form, were electrophoresed in an agarose gel (1%). In the negative control of the PCR, where the RNAm was directly submitted to PCR for AHADH2 gene, there was no any amplicon, indicating the good quality of the RNAm, free of DNA contaminants. On the other hand, in all the samples the results showed a smear which increased after each step of the nested PCR. Nevertheless, was not possible to clearly identify clonal variations among the isogenic lines by agarose electrophorese gels (Figure 4.10).



**Figure 4.10. 3-step nested PCR for clonal MASP expression analysis in the epimastigote form. Agarose electrophoresis gels (1%) for RNA, cDNA (quality controls) and for the 3-step nested PCR for MASP amplification. (A) PCR for AHADH2 gene using the cDNA as template. (B) Negative Control PCR for AHADH2 gene using total each RNA as template. (C) First PCR reaction for MASP family using each cDNA as template with primers SL and 3'UTR1. (D) Second PCR reaction for MASP family using each first PCR product as template with primers SL and 3'UTR2. (E) Third PCR reaction for MASP family using each second PCR product as template with primers MASP N-term and MASP C-term. [1-7], name of each clone; P, parental strain (Pan4); gDNA: PCR using 10ng *T. cruzi* genomic DNA as template; NC, negative control of PCR reaction with no DNA template.**

Afterwards and as outlined above, we analyzed the third nested PCR amplicons, from all the clones and the parental strain, by capillary electrophoresis, elaborating a finger printing analysis. As result, we obtained a high-resolution *in silico* gel where none of the cell lines analyzed were transcriptionally identical. As was described by qRT-PCR, the clone4 showed the highest repertoire of bands in the *in silico* gel with 120 MASP amplicons; a number even higher than that obtained by the parental strain in its trypomastigote form with 92 bands (the stage where the highest MASP repertoire expression was expected) (Bartholomeu et al. 2009; De Pablos et al. 2011; dos Santos et al. 2012). Meanwhile, the epimastigote form of the parental cell line showed one of the poorest profiles with 46 bands. The lowest number of *masp* amplicons was found in the clone5 profile, with 37 bands resolved (Figure 4.11A and B). Another interesting result is the increase in the number of amplicons between the epimastigote, with the outlined 46 bands, and the trypomastigote stage of the parental cell line, with 92 bands (Figure 4.11A and B). Among the 92 bands of the parental trypomastigote, 23 bands (25%) were specific to this stage and were not found in any other parasite line in the epimastigote stage (Figure 4.11A). On the top of the *in silico* gel, the dendrogram based on the similarity of the pattern of the MASP amplicons (absence/presence and intensity of the bands), allowed to cluster all the samples in different nodes. Here, the clone7 showed the most similar pattern to the parental strain in the epimastigote form (Figure 4.11A). Furthermore, the frequency of apparition of the bands was determined according to their size for all the clones in the epimastigote stage and also for the parental line in the epimastigote and trypomastigote form. Unique bands were showed by 97 amplicons of a total of 601 (16.14%) and only 6 bands (0.99%) were present in all the samples. Thus, 82.87% of the bands were repeated in 2 to 8 of the samples (Figure 4.11D). Additionally, we also analyzed the profile of the second semi-nested and third nested PCR product from the trypomastigote of Pan4. We obtained 65 and 92 bands respectively, which confirm the increased sensitivity of the extra 3rd step of PCR (Figure 4.11A and C).





**Figure 4.11. Analysis of MASP clonal expression resolved by capillary electrophoresis.** (A) *In silico* gel obtained by PCR fingerprinting analyzing PCR products of 3-step nested PCR for each isogenic line. On the top, a dendrogram based on the similarity of the MASP amplicons pattern. The distances are measured at the right side of the dendrogram with the percentage of similarity among the samples. (B) Bar graph representing the number of MASP amplicons obtained in the 3rd step of the nested PCR for each clone and the parental strain. (C) Number of amplicons obtained from the 2nd and the 3rd steps of the nested PCR in the trypomastigotes of the parental cell line. (D) Frequency of bands, according to their relative mobility, of products from the 3rd nested PCR of the epimastigote stage for all the clones and from the epimastigote and trypomastigotes forms of the parental cell line. 2nd PCR and 3rd PCR, are the products of the second and third PCRs of the 3-step nested PCR; [1E-7E], name of each epimastigote clonal population; PE, epimastigote forms of the parental cell line; PT, trypomastigote forms of the parental cell line.

### 4.3 Northern blot

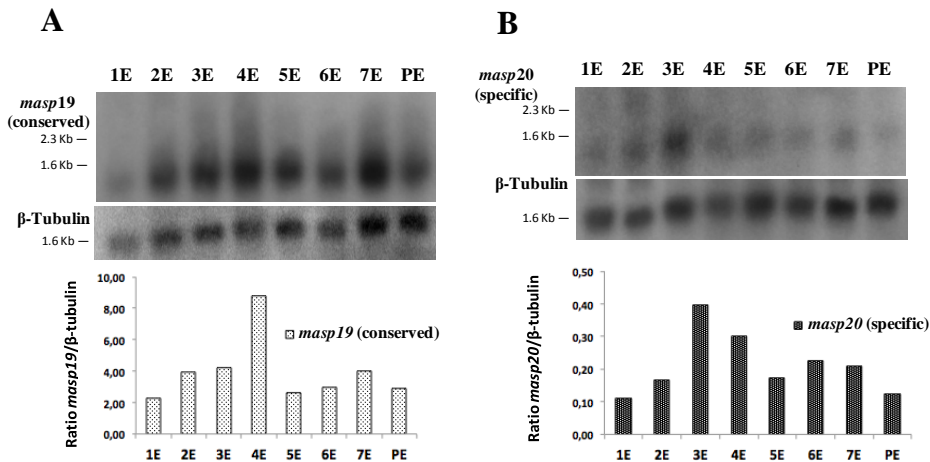
To confirm the clonal differences found in MASP expression by RTqPCR and capillary electrophoresis, two northern blots experiments were performed. The results showed the clonally variant expression of the MASP transcripts. The hybridization of the mRNA with *masp19* probe (sequence highly conserved in 23 of the 50 transcripts of the 3-step MASP library, with percentages of identity ranging from 82.68 to 100%) (Table 4.5), produced a smear, suggesting the co-expression and detection of several MASP genes, as expected by the great degree of homology of this sequence in the 3-step MASP library (Figure 4.12).

**Table 4.5. Percentages of identity of *masp19* probe with 23 transcripts of the 3-step MASP library.**

Members of 3-step MASP library	% identity respect to <i>masp19</i> probe	Members of 3-step MASP library	% identity respect to <i>masp19</i> probe
<i>masp1</i>	84.17	<i>masp22</i>	81.60
<i>masp2</i>	84.48	<i>masp23</i>	99.06
<i>masp3</i>	81.45	<i>masp25</i>	97.38
<i>masp4</i>	83.41	<i>masp26</i>	99.54
<i>masp6</i>	82.94	<i>masp27</i>	96.90
<i>masp7</i>	81.45	<i>masp30</i>	99.53
<i>masp14</i>	97.98	<i>masp31</i>	95.68
<i>masp15</i>	98.59	<i>masp32</i>	99.54
<i>masp16</i>	99.23	<i>masp33</i>	94.39
<i>masp18</i>	99.85	<i>masp35</i>	99.54
<i>masp19</i>	100	<i>masp36</i>	82.70
		<i>masp37</i>	99.69

The ratio *masp19*/β-tubulin changed depending on the clone analyzed. Approximately, 4-fold of variation in expression has been found among clones, with the highest signal corresponding to the clone4 and lowest to the clone1 (Figure 4.12A). To find the variation of single MASP transcripts, the RNA from the different clonal lines was hybridized with the probe *masp20* (specific probe

for the *masp20* transcript). As a whole, the signal was lower than that obtained with *masp19*. However, differences were also detected among the samples with again ~4 fold-range of variation, with the clones 3 and 4 having the highest levels of relative expression. Again, the clone1 showed the lowest expression signal (Figure 4.12B).

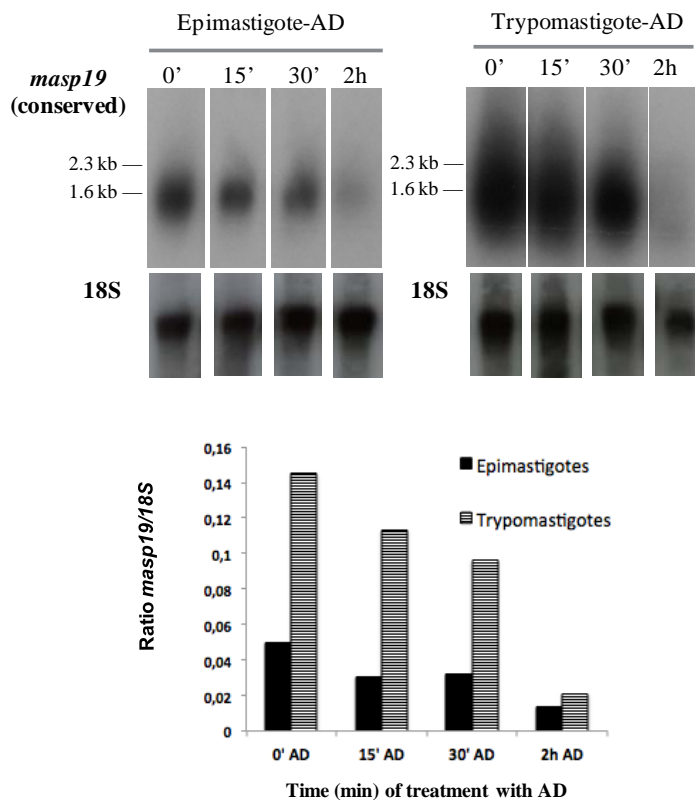


**Figure 4.12** Northern blot analysis of total RNA extracts from the epimastigote forms of all the clones and the parental strain hybridized with radiolabelled *masp19* conserved probe (A) and *masp20* specific probe (B).

#### 4.4. MASP RNA turnover assay

Having demonstrated the different RNA expression patterns among the different cell lines and even between different life cycle stages, we look for differences in the regulation of MASP RNAs between epimastigote and trypomastigotes. Thus, we studied the lifespan of MASP transcripts taking advantage of the inhibitory effects of AD on transcription to follow up the kinetics of RNAm decay. Northern blot analysis shows the relative expression of the MASP genes from the epimastigote and trypomastigote stages detecting by hybridization with conserved *masp19* probe, which, as has been outlined, is potentially able to hybridize with several MASP transcripts simultaneously (Figure 4.12A and 4.13; Table 4.5). The smear obtained mainly in trypomastigotes

but also in epimastigotes, suggests the ability of *masp19* to detect a wide repertoire of MASP transcripts expressed by Pan4 strain, demonstrating again the conserved feature of the detected sequences. At time 0h of the treatment with AD, a greater expression in the MASP recognized by the *masp19* probe in the trypomastigote forms was found, which was approximately 3-fold higher than in epimastigotes. At 2h (end of the AD treatment) there were no major differences in the RNA levels between both *T. cruzi* forms. Despite being difficult to accurately measure the intensity of that probe due to the smeary nature of its signal (specially in the case of 0', 15' and 30' trypomastigote treatment), the lifespan seems to be similar for all the MASP RNAs detected in trypomastigotes and epimastigotes, with the hybridization signal almost lost after 2h of AD treatment in both life cycle stages (Figure 4.13).



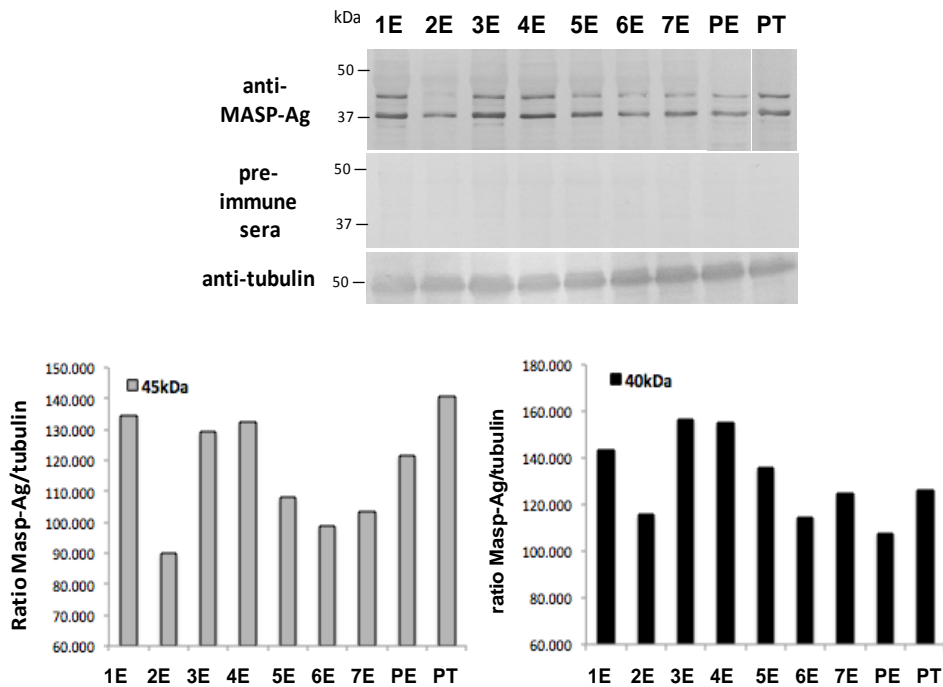
**Figure 4.13.** Northern blot analysis of total RNA extracted from epimastigote and trypomastigote forms of Pan4. Both life cycle stages were treated with 10 $\mu$ g/ml actinomycin D (AD) during different periods of time (0', 15', 30' and 2h).



Subsequently, their RNAs were hybridized with radiolabelled masp19 conserved probe and with radiolabelled 18S probe as loading control. The bars represent the ratio masp19/18S.

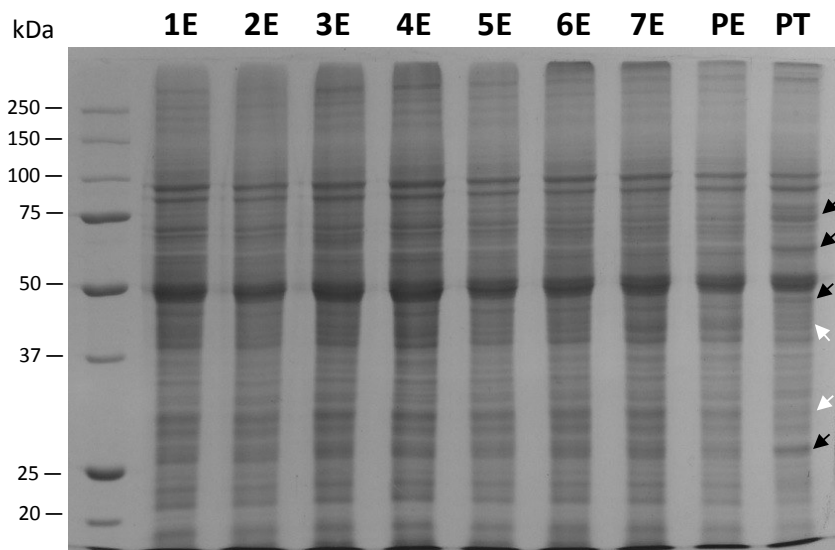
#### **4.5 Western blot**

In order to analyze the presence of MASP expression heterogeneities at protein level in the different cell lines, we used anti-MASP-Ag antibodies, which reacted with two epimastigote MASP proteins of 40 and 45 kDa in a clonally dependent manner. For the band of 45kDa, the highest intensity of signal was shown in the trypomastigotes of parental cell line while for the band of 40 kDa, the clone3 and clone4 showed the highest intensities, even higher than the trypomastigotes of the parental line (Figure 4.14). Moreover, when both bands were compared, we detected higher degree of variation for the 45 kDa band, indicating that MASP protein expression patterns are completely independent. Additionally, both bands were present in epimastigote and trypomastigote stage of *T.cruzi*, indicating the conservation of those proteins between life cycle stages.



**Figure 4.14. Clonally variant MASP expression at protein level.** Western blot analysis of total protein extracts from the epimastigotes forms of all the clones and the epimastigote and trypomastigote stage of the parental strain using anti-MASP-Ag mouse antisera. Anti-tubulin sera was used as loading control and pre-immune sera as negative control. Below, bar graphs representing the ratio 45kDa/tubulin and 40 kDa/tubulin. [1E-7E], name of each epimastigote clonal population; PE, epimastigote forms of the parental cell line; PT, trypomastigote forms of the parental cell line.

On the other hand, as a parallel quality and loading control, the total protein profile of the same samples was stained with comassie blue. Variations among clones and parental strain were not detected among. However, differences in the intensity levels of some bands were found comparing the protein profiles of trypomastigote and epimastigote stages of the parental strain (Figure 4.15).



**Figure 4.15. Protein extracts from the isogenic parasite populations electrophoresed and stained with Coomassie brilliant blue.** The black arrows indicate the bands overexpressed while the white arrows indicate those which are underexpressed in PT respect PE. [1E-7E], name of each epimastigote clonal population; PE, epimastigote forms of the parental cell line; PT, trypomastigote forms of the parental cell line.

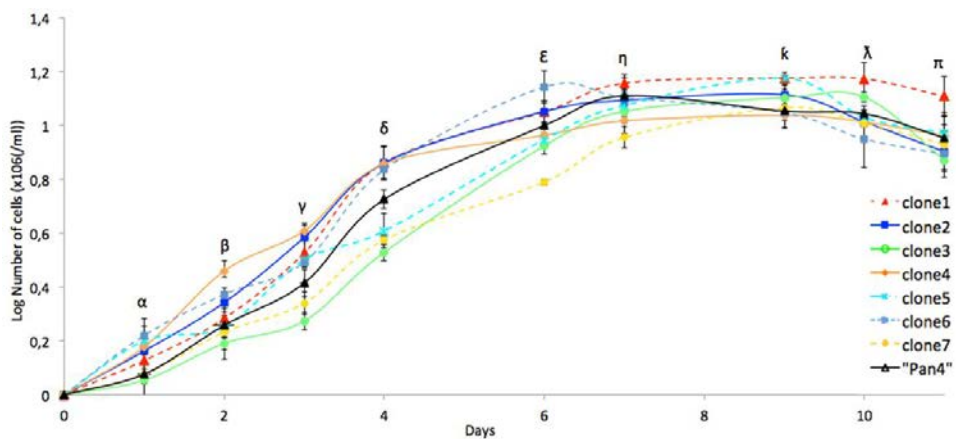
## **5. Analysis of the phenotypic variations of clonally isogenic populations of cells**

With the aim of detecting phenotypic differences, other aspects such as growth, the ability for the metacyclogenesis and the metabolites secreted to the medium were analyzed and compared among the isogenic cultures, which were grown with the same passages and under the same conditions.

## 5.1 Growth curve analysis

In the growth curves analysis, despite describing normal growth kinetics as a general trend, each clone manifested to a greater or lesser extent significant differences from the first day of growth (Figure 4.16).

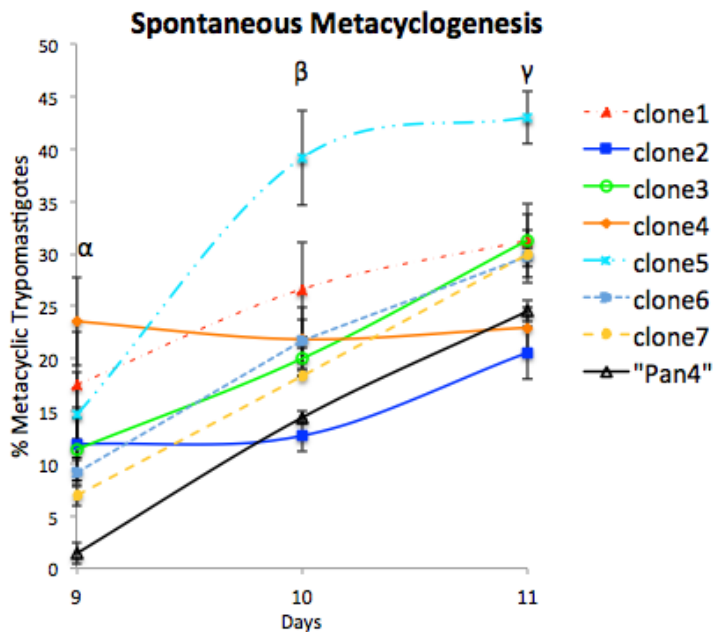
The first clone to reach the stationary phase was clone7 at the 6th day of growth, one day earlier than the parental strain and two days earlier than the rest of the clones. On the other hand, the highest cellular densities were observed in clones1 and 5, both with  $1.175 \times 10^7$  cells. Once all the cultures reached the stationary phase, each of them showed a different kinetics within the death phase, ending the counts in a similar parasite density. Only clone1 presented a significant higher number of epimastigotes at the 11th day (Figure 4.16).



**Figure 4.16. Analysis of the phenotypical clonal variation in the growth curves of epimastigote cell lines.** Letters above the time points mean that cell numbers among cell lines were significantly different ( $p \leq 0.01$ ) with the exception of: Day 1,  $\alpha$  (2vs4, 3vs7, 3vs8, 4vs5, 5vs6, 7vsPan4); Day 2,  $\beta$  (5vs6); Day 3,  $\gamma$  (5vs6), Day 4,  $\delta$  (1vs2, 1vs4, 2vs4); Day 6,  $\epsilon$  (1vs2); Day 7,  $\eta$  (2vs6 and 6vsPan4); Day 9,  $\kappa$  (1vs5, 2vs3, 4vs6, 4vs8, 6vs7, 6vs8, 7vsPan4); Day 10,  $\lambda$  (2vs4, 2vs5, 2vs7, 4vs5, 4vs7, 5vsPan4); Day 11,  $\pi$  (2vs6, 4vs5, 4vsPan4 and 5vsPan4).

## 5.2 Spontaneous metacyclogenesis analysis

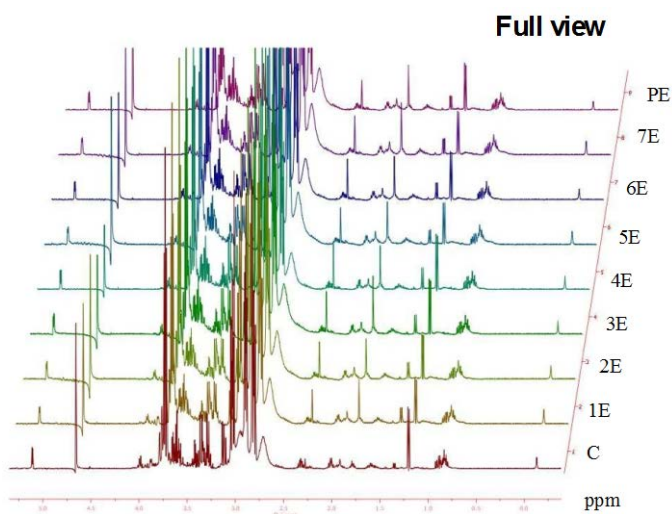
Regarding the spontaneous metacyclogenesis, significative different dynamics in metacyclic transformation were detected among the cell lines. Clone5 showed the best ability to differentiate into metacyclic forms with more than the 40% of the cells in metacyclic stage at the 11<sup>th</sup> day of culture. On the contrary, clone2 presented the lowest metacyclogenesis activity with 20.56% while Pan4 had 24.52% of metacyclic forms in culture at 11th day (Figure 4.17).



**Figure 4.17.** Analysis of phenotypical clonal variations in spontaneous metacyclogenesis yields at day 9, 10 and 11 of the culture. Letters above the time points mean that cell numbers among cell lines were significantly different ( $p \leq 0.01$ ) with the exception of: Day 9,  $\alpha$  (2vs3); Day 10,  $\beta$  (4vs6); Day 11,  $\gamma$  (1vs3, 6vs7).

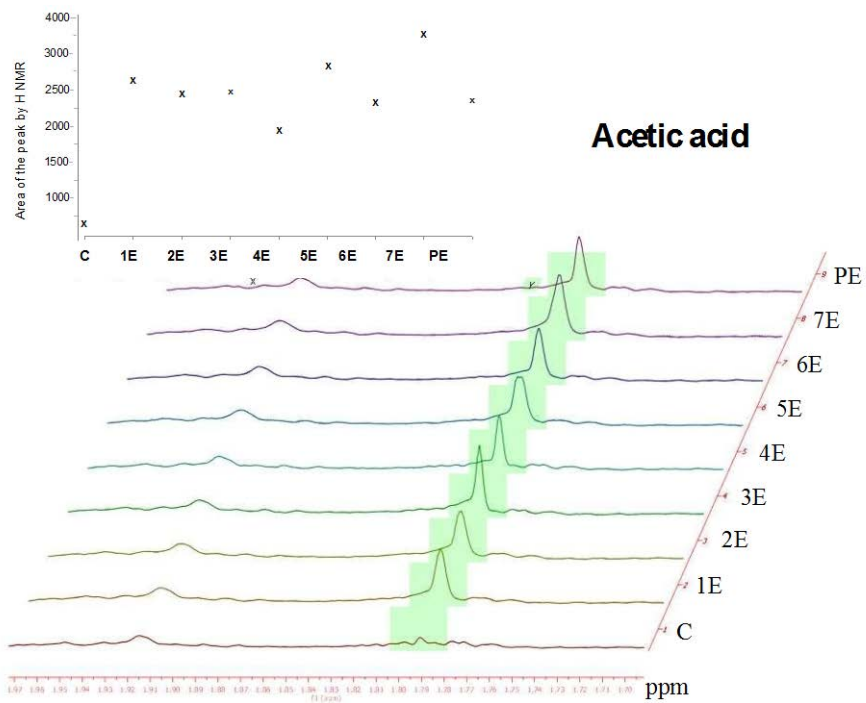
### 5.3 Analysis of the phenotypical clonal variation by H NMR

With the aim of comparing the metabolic profiles of the different epimastigote clonal cell lines at the logarithmic phase of growth (5<sup>th</sup> day of the growth curve), we analyzed a set of secreted metabolites by H NMR. The full view of the H NMR spectrums for the metabolites analyzed during the epimastigote cultures for each isogenic parasite line is showed in the Figure 4.18.

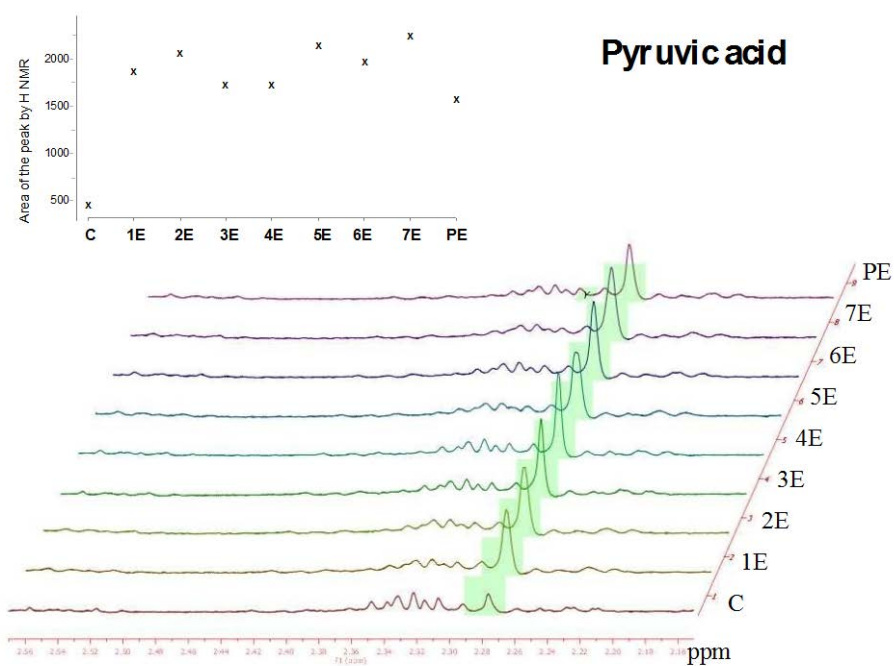


**Figure 4.18. Full view of the H NMR spectrums for the different clonal lines of epimastigotes.** C, RPMI with 10% IFSC (blank); [1E-7E], name of each epimastigote clonal population; PE, epimastigote forms of the parental cell line.

We identified differences in the signal intensities of the proton spectrums obtained from the acetic acid, L-alanine and pyruvic acid. The highest differences, measured as the area under the peak, were found for the acetic acid secretion with ~1.5 fold changes in the production of this metabolite between the clone4 (lowest) and clone7 (highest). Moreover, in the L-alanine and pyruvic acid spectrums, clone7 also showed the highest yields of secreted metabolites, contrary to the parental cell line, which showed the lowest levels for L-alanine and pyruvic acid production (Figure 4.19- 4.21).

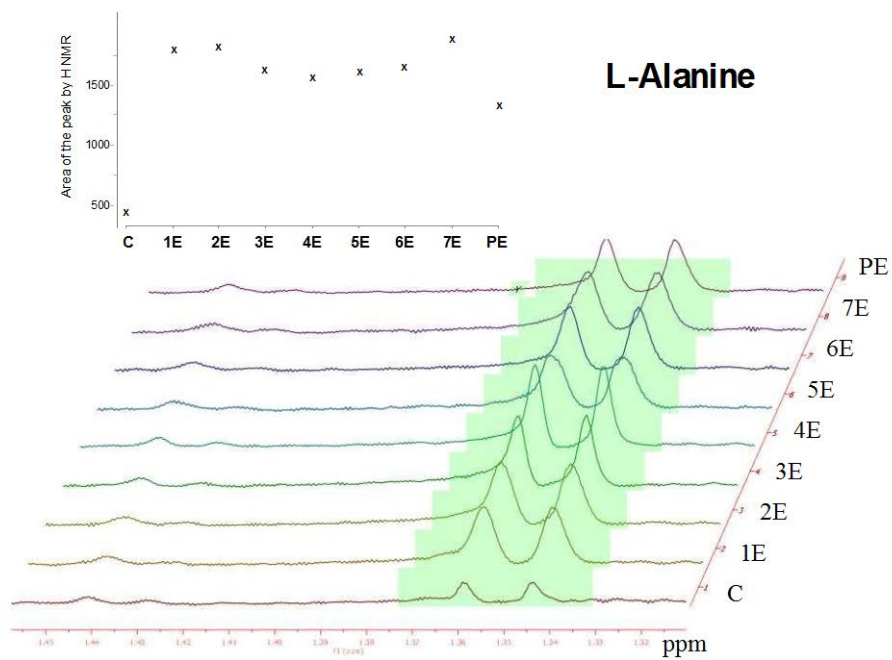


**Figure 4.19.** Full area under the peak obtained from the  $^1\text{H}$  NMR spectrums for acetic acid produced and excreted during the epimastigote cultures of the isogenic lines. C, RPMI with 10% IFSC (blank); [1E-7E], name of each epimastigote clonal population; PE, epimastigote forms of the parental cell line.



**Figure 4.20. Full area under the peak obtained from the H NMR spectrums for pyruvic acid produced and excreted during the epimastigote cultures of the isogenic lines. C, RPMI with 10% IFSC (blank); [1E-7E] name of each epimastigote clonal population; PE, epimastigote forms of the parental cell line.**





**Figure 4.21. Full area under the peak obtained from the H NMR spectrums for L-alanine produced and excreted during the epimastigote cultures of the isogenic lines. C, RPMI with 10% IFSC (blank); [1E-7E] name of each epimastigote clonal population; PE, epimastigote forms of the parental cell line**





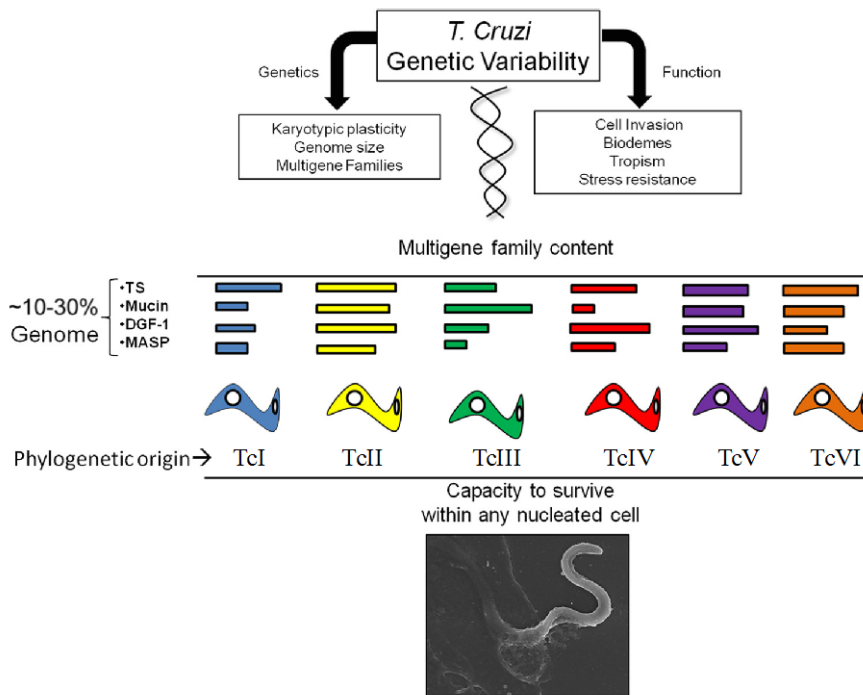
## **V. DISCUSIÓN**



Son muchos los estudios moleculares, bioquímicos y biológicos que muestran a *T. cruzi* como a un taxón altamente heterogéneo, con una estructura poblacional compleja, de naturaleza clonal y con eventos de intercambio genético (Brisse, Dujardin, Tibayrenc 2000; Buscaglia and Di Noia 2003; Miles et al. 1978; Miles et al. 1980; Souto and Zingales 1993; Souto et al. 1996; Zingales et al. 2012). Al respecto, se ha descrito la relación existente entre la mencionada heterogeneidad genética de *T. cruzi* y la variabilidad en las manifestaciones clínicas de la enfermedad de Chagas, que van desde casos asintomáticos hasta severas afecciones cardíacas o digestivas; en este aspecto pueden intervenir otros factores como son la propia genética del hospedador o eventos de coinfección (Andrade et al. 1999; Buscaglia and Di Noia 2003; Macedo, Oliveira, Pena 2002; Macedo et al. 2004; Zingales et al. 2012). Así, tal y como se comentó en la introducción (apartado 2.4), en función de las diferencias moleculares, eco-epidemiológicas y geográficas, se ha podido dividir al taxón de *T. cruzi* en siete subdivisiones intraespecíficas: seis unidades discretas de tipificación (DTUs) nombradas desde TcI-TcVI y la rama Tcbat, que infecta exclusivamente a hospedadores murciélagos (Brisse, Dujardin, Tibayrenc 2000; Marcili et al. 2009; Zingales et al. 2012).

El proyecto genoma de *Trityp* reveló una de las características más distintivas de *T. cruzi* y es que, en comparación con el resto de los kinetoplastidos, este parásito presenta una gran expansión de genes codificantes para proteínas de superficie pertenecientes a las familias multigénicas DGF-1, TSs, Mucinas y MASPs (De Pablos and Osuna 2012a; El-Sayed et al. 2005a; El-Sayed et al. 2005b). Este gran repertorio de proteínas de superficie se postula como una pieza fundamental participando tanto en la evasión de la respuesta inmune del hospedador, como en el proceso de interacción y adhesión que el parásito es capaz de establecer con un amplio repertorio de células, desde fagocitos a células epiteliales y cardíacas en el hospedador mamífero, hasta con las células que conforman el endotelio intestinal en el insecto vector (Bartholomeu et al. 2009; Burleigh and Woolsey 2002; Buscaglia et al. 2006; De Pablos and Osuna 2012a; Garcia et al. 2007).

Recientes estudios de genómica comparativa han revelado como cepas de *T. cruzi* con diferentes orígenes filogenéticos, muestran un alto número de CNVs que se centran principalmente en zonas de ADN repetitivo donde se incluyen las familias multigénicas de este protozoo (Franzen et al. 2012; Minning et al. 2011; Reis-Cunha et al. 2015). De esta manera, la cantidad de genes que conforman cada una de las familias multigénicas varía según el linaje de *T. cruzi*, oscilando en su conjunto entre 10 y 30% del genoma del parásito (Figura 5.1) (De Pablos and Osuna 2012a; Minning et al. 2011). Este polimorfismo en el contenido de las familias multigénicas entre diferentes cepas de *T. cruzi*, puede estar relacionado con el tropismo diferencial del parásito en el organismo del hospedador mamífero, siendo por lo tanto, un fenómeno clave para entender parte de la variación en la manifestación clínica de la enfermedad (Andrade et al. 1999).



**Figura 5.1. Representación esquemática del polimorfismo en el porcentaje del genoma ocupado por las familias multigénicas TS, Mucina, DGF-1 y MASP de *T. cruzi* al comparar cepas con diferentes orígenes filogenéticos.** Las barras horizontales representan el número de genes de cada familia multigénica en cada linaje (De Pablos and Osuna 2012a).

Como ya se ha comentado anteriormente, la familia MASP fue uno de hallazgos más interesantes que se obtuvieron de la secuenciación del genoma de *T. cruzi* (El-Sayed et al. 2005b). Esta familia, específica de *T. cruzi* y fuera, por tanto, de la zona de sintenia con *Leishmania spp.* y *Trypanosoma brucei*, destaca por su abundante presencia en el genoma y por unos altos niveles de co-expresión simultánea (Atwood et al. 2005; Atwood et al. 2006; De Pablos 2010; dos Santos et al. 2012). Pese a que su funcionalidad permanece sin ser caracterizada, se sabe que las proteínas de la familia MASP se localizan en la membrana del parásito, mientras que algunas de ellas pueden ser secretadas al medio. De esta manera, la familia MASP se postula como un factor clave en los procesos de infección por *T. cruzi* participando tanto en la interacción célula-parásito como en la evasión de la respuesta inmune (Bartholomeu et al. 2009; De Pablos et al. 2011; dos Santos et al. 2012). Así, en términos evolutivos, la especificidad y expansión de esta familia en *T. cruzi* es uno de sus aspectos más característicos, el cual, probablemente, intervino en la separación de esta especie del resto de Kinetoplástitos a partir de un antepasado común.

Con el objetivo de profundizar en el conocimiento de la familia MASP se elaboró una librería de expresión MASP (*3-step MASP library*) que, mediante “*hierarchical clustering*”, ha permitido establecer un análisis de la variabilidad de esta familia entre cepas de *T. cruzi* de diferentes orígenes filogenéticos y medir la velocidad evolutiva de los miembros de la propia librería. La estrategia para la realización de la librería de expresión MASP se basó en la estructura típica de sus miembros, con unos extremos N- y C- terminales altamente conservados y una zona central hipervariable (Bartholomeu et al. 2009). Así, se diseñaron cebadores capaces de hibridar en dichas zonas conservadas que, a partir del ADNc de las formas tripomastigotes de la cepa Pan4 (TcIa-d) y mediante una PCR anidada de tres pasos (*3-step nested PCR*), permitieron amplificar y clonar un total de 49 miembros de la familia MASP (37 genes y 12 pseduogenes) que estaban expresándose en un momento dado de la población de tripomastigotes (Tabla 4.2 y 4.3). Según el porcentaje de secuencias válidas clonadas (secuencias completas sin indeterminaciones), 93.5%, la librería de expresión MASP fue considerada de buena calidad.



También es importante resaltar como el 98.3% de los amplicones generados por la PCR anidada de tres pasos, fueron identificados como miembros de la familia MASP (tabla 4.1). Mientras que, en las siete librerías MASP obtenidas por dos Santos *et al.* (2012) mediante una PCR semi-anidada de dos pasos, se observaron porcentajes de identificación de miembros MASP que oscilaron entre el 50%-83.9%. Esto significó que, respecto a la PCR semi-anidada de dos Santos, el tercer paso en la amplificación *3-step nested PCR*, en el que se emplearon unos cebadores más cercanos a la región codificante de los miembros MASP (Figura 3.1), supuso un incremento del 14.4 al 48.3% en el número de amplicones específicos de esta familia multigénica. Este enriquecimiento pudo deberse a que en la PCR semi-anidada de dos Santos se utilizaron cebadores para la región UTR de genes MASP que podrían ser compartidos parcialmente con otras familias de genes; mientras que la inclusión del tercer paso de la PCR anidada desarrollada en este trabajo, se basó en cebadores correspondientes a los extremos de la región codificante de esta familia, lo que, tal y como se ha mostrado, derivó en un aumento de la especificidad de los amplicones generados.

Para el análisis de la variabilidad de la familia MASP entre diferentes cepas de *T. cruzi*, se compararon las secuencias de los miembros de dicha librería de expresión frente a las secuencias de sus respectivos ortólogos en los genomas de las diferentes cepas de *T. cruzi* que se encontraban disponibles en la base de datos Tritryp (<http://tritrypdb.org/tritrypdb/>) (acceso 3 de Febrero 2015)

- CL Brener (TcVI): CL Brener es una de las cepas más estudiadas y mejor caracterizadas, siendo elegida como cepa modelo para la secuenciación del primer genoma completo de *T. cruzi* (El-Sayed et al. 2005b). TcVI es un DTU originado mediante la hibidación de TcII y TcIII, probablemente como consecuencia de un episodio de co-infección en el insecto vector o en algunos de sus hospedadores. Es importante remarcar que este linaje está únicamente asociado con la transmisión doméstica de la infección (Figura 1.11) (Zingales et al. 2012).

- SylvioX10/1 y Dm28c (TcI): El linaje de estas dos cepas coincide con el de la cepa Pan4, en la que se basó la librería de expresión MASP. Tras CL Brener, SylvioX10/1 fue la segunda cepa de *T. cruzi* en ser secuenciada (Franzen et al. 2011). TcI es el DTU más abundante y ampliamente distribuido en América (Figura 1.11) y se caracteriza por presentar un alto grado de diversidad genética con un total de cuatro subgrupos Ia-Id cuyo ciclo de transmisión está asociado tanto a los ciclos domésticos (humanos) como a selváticos (marsupiales, primates, etc) (Cura et al. 2010; Zingales et al. 2012). Así, mientras SylvioX10/1 fue aislada de un paciente en fase aguda de la enfermedad de Chagas, la cepa Dm28c se aisló a partir de la zarigüeya (*Didelphis marsupialis*) (Zingales et al. 2009).
- Marinkellei (Tcbat): A este linaje en realidad se le considera como una subespecie de *T. cruzi* por lo que es nombrado como *T. cruzi marinkellei*. De esta manera a los DTUs TcI-TcVI se les nombra como *T. cruzi cruzi* (Franzen et al. 2012). Este subgrupo está asociado a especies del género *Chiroptera*, que se distribuyen en Centro y Sudamérica. Así, los murciélagos se presentan como su único hospedador conocido y los triatominos del género *Cavernicola* como el insecto vector (Franzen et al. 2012; Zingales et al. 2012). Aunque existe cierta discrepancia acerca de los modelos y tiempos de evolución de los diferentes linajes de este parásito, se cree que *T. cruzi* y *T. cruzi marinkellei* divergieron de un antepasado común hace 200-475 millones de años, mientras que el resto de linajes de *T. cruzi* pudieron separarse hace 33-88 millones de años (Tomasini and Diosque 2015).

Las distancias obtenidas entre todas las cepas analizadas mostraron como las secuencias de los genes MASP se distribuyen de acuerdo a su origen filogenético (Figura 4.3A; Tabla 4.4). Así, los resultados y distribución son coincidentes con los encontrados en un estudio reciente en el que se llevó a cabo el análisis filogenético de los genes V7V8 rRNA y gGAPDH en las mismas cepas (Lima et al. 2015). Adicionalmente, asumiendo la interacción parásito-hospedador, se comprende que la trayectoria evolutiva que describe cada DTU viene determinada, no sólo por la propia genética de la cepa de *T. cruzi*, sino

también por la genética del hospedador y por la presión inmune que éste ejerce. De esta manera, es fácil entender que Marinkellei, con el origen filogenético más alejado, se comporte como la cepa más divergente dentro del análisis de variabilidad MASP.

Por otro lado, se observó una mayor variabilidad entre las secuencias pseudogénicas, produciéndose incluso, la pérdida de la correlación con el origen filogenético de las cepas. No obstante, se encontraron excepciones en donde la distancia entre los pseudogenes MASP de las diferentes cepas fue menor de lo que podría esperarse, especialmente para los pseudogenes 1, 4, 7, 9, 11 y 12 (Figura 4.5 y 4.6; Tabla 4.4). Todo esto podría explicarse por la presencia de una presión selectiva positiva a la que pueden verse sometidos aquellos pseudogenes que presentaron una secuencia conservada entre las diferentes cepas, sugiriendo una posible función biológica para éstos, por ejemplo, actuando como reguladores postranscripcionales de la expresión génica. Al resto de pseudogenes secuenciados, que presentaron mayores distancias evolutivas, se les supone una menor presión selectiva, lo que les va a permitir acumular mutaciones en un proceso de evolución que probablemente les destine a adquirir nuevas funciones, pudiendo actuar también como generadores de diversidad antigénica tal y como ocurre con los pseudogenes VSG en *T.brucei* (Balakirev, E.S. and Ayala, F.J. 2003; D'Errico, Gadaleta, Saccone 2004; Hall, Wang, Barry 2013; Salmena 2014; Wen et al. 2011).

Es interesante resaltar como el *hierarchical clustering* también permitió clasificar a los genes y pseudogenes MASP de la librería en base a su velocidad evolutiva, definiéndose dos grupos en cada uno de ellos (secuencias de evolución lenta y rápida) (Figura 4.3; Tabla 4.4). El pseudogén *masp10.ps*, debido a su comportamiento evolutivo singular, quedó al margen de esta clasificación. Este pseudogén muestra una gran distancia respecto a sus ortólogos en las cepas Sylvio X10/1, Dmc28 y Marinkellei, por lo que es considerado como el miembro que evoluciona a mayor velocidad dentro de la librería MASP. No obstante y de manera sorprendente, *masp10.ps* presenta altos porcentajes de identidad respecto a diferentes MASPs localizadas en el genoma de CL Brener, en las tres librerías

de expresión MASP preliminares (desarrolladas en las cepas Pan4 (TcIa-d), Maracay (TcII) y CL-Brener (TcVI) (De Pablos 2010) y en tres de las siete librerías MASP de dos Santos en la cepa CL-Brener (dos Santos et al. 2012) (Figura 4.3B y 4.4; Tabla 4.4). Este alto grado de conservación, sumado a los SNPs que se observaron al comparar y alinear entre sí las secuencias de determinados miembros pertenecientes a la librería MASP (con un máximo de diferencia de 3pb), sugiere que dentro del amplio repertorio y variabilidad implícitos en la familia MASP, existe un subgrupo de ARNs dentro del transcriptoma de las MASPs que está altamente conservado, no sólo entre diferentes fases del ciclo celular, sino que también entre diferentes cepas de *T. cruzi*. De hecho, como se ha mencionado anteriormente, los pseudogenes MASP pueden conservar su secuencia, siendo positivamente seleccionados para desempeñar alguna función biológica, como competidores naturales de genes MASP, donde podrían actuar como cebo de proteínas de unión a ARN (RBPs) regulando indirectamente la expresión de genes MASP. Otra posible función es su actuación como reguladores directos de la expresión génica a nivel de ARN. A este último respecto, se ha descrito que los pseudogenes son una fuente de ARNs antisentido en *Lymnaea stagnalis*, donde la expresión del gen NO sintasa (NOS) está regulada por la co-transcripción del pseudo-transcrito NOS, el cual forma un dúplex ARN-ARN con NOS, suprimiéndose selectivamente la expresión de este gen (Korneev, Park, O'Shea 1999).

Los retrotransposones constituyen un mecanismo generador de pseudogenes. A excepción de la familia MASP, otras familias multigénicas de *T. cruzi* que están fuera de sintonía, como DGF-1, TSs o mucinas, se ubican en las zonas subteloméricas de los cromosomas, (Moraes Barros et al. 2012), condicionando la presencia de los retrotransposones, los cuales tienden a insertarse sobre los genes que constituyen dichas familias creando nuevos miembros. En este contexto, cabe remarcar la asociación existente entre el retrotransposon no autónomo TcTREZO y el extremo 3' de (pseudo)genes MASP (Souza et al. 2007). La búsqueda de los sitios de inserción específicos del retroelemento TcTREZO sobre la librería MASP demostró una amplia presencia de su dominio *upstream* en el extremo 3' de la mayoría de los miembros MASP

(73.5%) (Figura 4.4A). Este resultado sugiere como la mayor parte de las secuencias (pseudo)génicas MASP clonadas pueden originarse y propagarse gracias a eventos de retrotransposición llevados a cabo por este elemento que, según el producto formado, va a dar lugar a los denominados retrogenes o retropseudogenes (Ding et al. 2006). De esta manera, se confirmaría el importante papel que el retroelemento TcTREZO pudiera desempeñar actuando como uno de los principales mecanismos para la generación específica de diversidad dentro de la familia MASP.

Por último, otro de los mecanismos generadores de esta diversidad y variabilidad puede ser la inestabilidad cromosómica y plasticidad genética de *T. cruzi*. En este sentido, la comparación de 16 cepas mediante “whole genome oligonucleotide tiling arrays” con la cepa CL Brener como referencia, demostró un alto número de CNVs (~4000) entre ellas, que se localizaron especialmente en secuencias repetitivas de familias multigénicas como la familia MASP. De esta manera, los autores de este trabajo sugieren que *T. cruzi* utiliza dichas secuencias para impulsar la recombinación mitótica en el parásito lo que generaría la diversidad de familias multigénicas de proteínas de superficie en *T. cruzi* (Minning et al. 2011). Adicionalmente, Reis-Cunha *et al.* han comparado los 41 cromosomas de CL Brener con cepas pertenecientes a los DTUs TcI, TcII y TcIII, identificando CNV Cromosómicos (CCNV), otro nivel de CNVs en el que bloques enteros de genes se expanden de manera simultánea. Esta amplificación cromosómica, además de generar variabilidad, contribuiría a modificar la expresión génica, lo que puede resultar crucial para un parásito como *T. cruzi*, cuyo control de la expresión depende principalmente de mecanismos postranscripcionales (Reis-Cunha et al. 2015).

El análisis de la variabilidad de la familia MASP entre diferentes cepas de *T. cruzi* ayudó a determinar el grado de heterogeneidad genotípica que existe dentro de este taxón. Tras esto, se pretendió ir un paso más allá, estudiando tanto la heterogeneidad transcripcional de esta familia multigénica (a nivel transcripcional y proteico), como la variabilidad de varios aspectos fenotípicos (crecimiento, metaciclogenesis espontánea y producción y excreción de

metabolitos) que puedan manifestarse entre las células de una única población de *T. cruzi*.

Recientemente, las variaciones clonales han sido caracterizadas como un propiedad intrínseca de las poblaciones tanto de organismos unicelulares como pluricelulares, siendo capaces de determinar aspectos fundamentales de la biología de bacterias patogénicas (Balazsi, van Oudenaarden, Collins 2011; Casadesus and Low 2013), hongos (Levy, Ziv, Siegal 2012; Morrow and Fraser 2013) o incluso definir el destino de células cancerígenas. A este respecto, estudios centrados en la muerte celular y cáncer han propuesto un modelo en el que desde la aparición hasta la propagación y metastasis derivadas del cáncer se explican por la variabilidad célula-célula (Figura 1.13) (Navin 2014; Navin 2015).

En protozoos, la heterogeneidad celular también resulta determinante, facilitando la adaptación a diferentes ambientes o condiciones a lo largo del ciclo de vida del parásito (Balazsi, van Oudenaarden, Collins 2011; Casadesus and Low 2013; Kumar et al. 2014; Levy, Ziv, Siegal 2012; Morrow and Fraser 2013; Navin 2014; Seco-Hidalgo, Osuna, Pablos 2015; Xia et al. 2014). Estas variaciones han sido descritas tanto en genes individuales como en familias multigénicas involucradas en la interacción parásito-hospedador (Amambua-Ngwa et al. 2012; Niang, Yan Yam, Preiser 2009; Recker et al. 2011; Rovira-Graells et al. 2012b; Seco-Hidalgo, Osuna, Pablos 2015). A pesar del importante papel de la familia MASP en este tipo de interacciones (tanto en el hospedador mamífero como en el vector), su modo de expresión entre las células isogénicas de una misma población, permanece sin caracterizar. El repertorio de técnicas utilizadas en este estudio, ha permitido caracterizar, por primera vez, el alto grado de heterogeneidad clonal asociado a la expresión de la familia MASP a diferentes niveles (ARN y proteínas) sobre un total de ocho líneas de parásitos isogénicas (7 clones y el cultivo parental).

La electroforesis capilar de amplicones se ha mostrado como un método con un alto potencial resolutivo, permitiendo una separación en alta definición de los amplicones MASP, comprendidos entre 80 y 1100pb, y obtenidos mediante de el proceso de *3-step nested PCR*. Esto permitió la identificación de diferencias

transcripcionales específicas entre los aislados clonales y la cepa parental, donde cada una de las líneas de parásitos presentó un perfil de amplicones diferente (Figura 4.10A y B). Además, la electroforesis capilar también permitió confirmar la eficiencia de la tercera reacción en la *3-step nested PCR* respecto al protocolo original (dos Santos et al. 2012), suponiendo un incremento del 29.35% en el número de amplicones MASP de tripomastigotes de Pan4 (Figura 4.10C).

Adicionalmente, mediante northern blot (Figura 4.7), RT-qPCR (Figura 4.8) y western blot (Figura 4.11) se pudo confirmar el resultado de la electroforesis capilar, demostrándose de nuevo la existencia de importantes variaciones clonales a nivel de ARN y de proteínas entre las líneas clonales de epimastigotes.

De esta manera, en todas las técnicas utilizadas para el análisis de la heterogeneidad transcripcional, y contrariamente a lo esperado, la línea parental mostró una tendencia común, con los niveles de variabilidad MASP más bajos. Por su parte, el clon4 presentó un comportamiento opuesto, exponiendo siempre uno de los mayores repertorios y niveles de expresión MASP. Una posible explicación a este fenómeno se detalla más adelante, tras discutir los resultados de variabilidad clonal fenotípica.

La co-expresión simultánea de varios miembros de la familia MASP es un hecho previamente documentado en varias formas del ciclo *T. cruzi*, como tripomastigote o amastigote (Atwood et al. 2005; Atwood et al. 2006; dos Santos et al. 2012), aunque dichos estudios recogen la diversidad sobre un *pool* de parásitos provenientes tanto de cultivos *in vitro*, como de infecciones en ratón. A diferencia de estos estudios, en el trabajo desarrollado en este proyecto de tesis, se ha analizado la habilidad del parásito para generar variabilidad y heterogeneidad en una población mediante la obtención y el cultivo de líneas clonales a partir de un cultivo original. Los resultados de la electroforesis capilar (Figura 4.1A) sumados a los 49 miembros MASP obtenidos en la librería MASP (Figura 4.1C) y el *smear* generado en los análisis por northern blot utilizando la sonda conservada *masp19* (Figura 4.7A), confirman el alto grado de co-expresión MASP que ocurre a nivel de ARN en las fases epimastigote y tripomastigote de la cepa parental y en la fase epimastigote de los distintos clones aislados. Así, se

confirma como la co-expresión MASP en un cultivo no es un fenómeno exclusivo de las fases infectivas de *T. cruzi* (De Pablos and Osuna 2012b), detectándose también un amplio repertorio de MASP expresándose de manera simultánea en el ARN de la fase epimastigote de *T. cruzi*, donde deben cumplir algún papel funcional.

Adicionalmente, trabajando sobre el ARN de las formas epimastigotes y tripomastigotes de *T. cruzi*, hemos podido verificar como los niveles de los transcritos reconocidos por la sonda conservada *masp19*, fueron mayores en la fase tipomastigote que en la epimastigote. No obstante, el ARN de ambas fases mostró un tiempo de vida similar, cercano a las 2h tras el tratamiento con el antibiótico AD (Figura 4.19). Este resultado, en el que se establece un proceso de regulación parecido sobre el ARN de ambas fases, puede explicarse por el alto grado de conservación de las regiones MASP 3' UTR (Bartholomeu et al. 2009) y por lo tanto de las secuencias cis-reguladoras. Así, las diferencias en los patrones de expresión entre las diferentes fases del ciclo celular podrían estar motivadas por diferencias en el nivel de factores trans-reguladores (p.ej. proteínas de unión al ARN o ARNs no codificantes), por un uso diferente de codones, por la disponibilidad de ARNs y por la propia plasticidad del genoma mostrada por las formas replicativas del parásito (epimastigotes y amastigotes) (Horn 2008; Lima et al. 2013; Minning et al. 2011; Souza et al. 2011).

A pesar de la gran variabilidad de la familia MASP, la técnica de la electroforesis capilar, también permitió encontrar la presencia de 6 bandas con un tamaño conservado entre los clones en su estado epimastigote y la cepa parental en las formas epimastigote y tripomastigote (Figura 4.1 A y D). Esto refuerza la teoría de la selección positiva a la que pueden estar expuestos determinados miembros de esta familia MASP que, como se ha comentado anteriormente, hace que algunas de sus secuencias se mantengan conservadas debido a alguna especialización funcional. Adicionalmente, el constante remodelaje y expresión diferencial de la familia MASP pueden ser utilizados por el parásito como una fuente para adquirir y fijar nuevas especializaciones en su superficie, pudiendo favorecerse la presencia de algunas proteínas MASP por



encima de otras. En este sentido, las dos bandas reconocidas por el anticuerpo anti-MASP-Ag (40 y 45kDa) tanto sobre las formas epimastigotes como tripomastigotes (aunque a distintos niveles de intensidad), puede indicar que esas dos proteínas MASP deben en efecto proporcionar alguna función específica para ambas fases. Un fenómeno similar de especialización ha sido descrito en *T. brucei*, donde el receptor para la transferrina, *haptoglobin-haemoglobin receptor (HpHbR)*, o el *Serum Resistance Associated (SRA)* han evolucionado a partir de miembros de la familia multigénica VSG, para nutrir el parásito con grupos hemo o para protegerlo de la acción tripanolítica de la Apolipoproteína L presente en el suero humano (Higgins et al. 2013; Jackson et al. 2012; Jackson et al. 2013; Xong et al. 1998).

En *T. cruzi* el fenómeno de co-expresión ha sido detectado en otras familias multigénicas como TS y mucinas (Buscaglia et al. 2006; Campo et al. 2006; De Pablos and Osuna 2012a; Freitas et al. 2011; Jager, Muia, Campetella 2008). Sin embargo, estos estudios fueron realizados trabajando con la media poblacional.

Por otra parte, los análisis de expresión de la familia VSG, cuyos genes codifican para el principal antígeno en la membrana de *T. brucei*, muestran como, en un momento dado, todas las células que conforman la población, van a expresar un sólo gen de esta familia. Este fenómeno se debe a la singularidad de las unidades de transcripción que determinan la expresión de los miembros de esta familia. Dichas unidades, conocidas como sitios de expresión (ESs), se localizan cerca de los telómeros de los cromosomas. Se estima que hay un total de 15 ESs, en los que la ARN polimerasa I lleva a cabo una transcripción policistronica, y es que además de un gen VSG, cada ES contiene genes asociados entre los que se incluyen otros genes codificantes para proteínas de membrana (Cross, Kim, Wickstead 2014; Li 2015; Pays et al. 2001; Stanne and Rudenko 2010). Aunque normalmente cada ES contiene un gen VSG diferente, el promotor ES es una secuencia altamente conservada cuya activación parece depender de fenómenos epigenéticos. Con este mecanismo de expresión monoalélica, *T. brucei* se asegura que el gen VSG no va a estar activo por mucho tiempo en su membrana, lo que le va a permitir alternar en la expresión de VSGs, dificultando

que el hospedador establezca una respuesta inmune eficaz frente a *T. brucei* (Cross, Kim, Wickstead 2014; Li 2015; Pays et al. 2001; Stanne and Rudenko 2010).

La familia MASP y el resto de familias multigénicas de *T. cruzi* carecen de tal control en su expresión pudiendo tratarse de casos reales de co-expresión. Las evidencias experimentales aportadas en este trabajo de tesis, descifrando el carácter ampliamente heterogéneo con el que se expresa la familia MASP entre líneas clonales de epimastigotes, nos anima a sugerir que los SNPs detectados al alinear secuencias pertenecientes a la librería MASP (máximo de 3pb de diferencia), pueden ser el resultado de la propia expresión clonal de la familia MASP dentro de las poblaciones de *T. cruzi*. Sin embargo, nuevos estudios basados en técnicas de análisis sobre células individuales serían necesarios para aclarar si un sólo parásito es capaz de expresar varios genes MASP de manera simultánea, o por el contrario, la co-expresión de MASP es el resultado del modo de expresión clonal de esta familia, en el que en una población de *T. cruzi*, cada parásito expresa un sólo gen MASP.

Además de la variabilidad descrita en la transcripción de la familia MASP, otros aspectos fenotípicos, como el crecimiento, la metacicloogénesis espontánea o la producción y excreción de catabolitos al medio, fueron también analizados en este trabajo, revelándose, de nuevo, un comportamiento heterogéneo, con marcadas diferencias entre las líneas clonales isogénicas de *T. cruzi*. Así, cada una de ellas presentó densidades celulares significativamente diferentes en la fase logarítmica y estacionaria de su curva de crecimiento. Precisamente, el descenso de densidad celular de la fase estacionaria, se ve acompañado de un aumento heterogéneo en el número de formas metacíclicas que aparecen en cada cultivo. La metacicloogénesis, diferenciación de las formas epimastigotes en tripomastigotes metacíclicos, es un proceso fundamental en la biología de *T. cruzi*, ya que va a permitir la propagación del parásito desde el vector hasta los hospedadores vertebrados (Garcia and Azambuja 1991). Se ha descrito como este proceso acontece en la ampolla rectal del triatomino (Figura 5.3), en donde el parásito se va a ver sometido a unas condiciones de estrés nutricional que, como

se describe más adelante, inducen cambios en su expresión génica (de Godoy et al. 2012; Garcia and Azambuja 1991; Parodi-Talice et al. 2007). Así, se comprende como, en condiciones *in vitro*, la metacicloénesis espontánea se produce al final de la fase estacionaria, donde la disponibilidad de nutrientes es menor. Nuestros resultados, demostrando la variabilidad clonal en la metacicloénesis, refuerzan los datos previamente aportados por Lauria-Pires *et al.* (1996) y Avendaño *et al.* (2006). Lauria-Pires y colaboradores trabajaron sobre 2 y 4 clones derivados de los aislamientos hSLU239 y mSLU142 respectivamente, los cuales procedían de casos clínicos de cardiopatía chagásica y megaloesófago. Así, observaron un comportamiento diferente respecto a la metacicloénesis espontánea entre dichos clones. Por su parte, en el estudio de Avendaño *et al.*, se compararon tres clones de la cepa Costa Rica TC-4 con la cepa parental, detectando también el comportamiento heterogéneo de este proceso en *T. cruzi*.

Por otro lado, tal y como se ha comentado, en este trabajo de tesis también se observaron diferencias clonales en los niveles de excreción para los catabolitos ácido acético, ácido pirúvico y L-alanina, todos ellos productos finales del catabolismo de la glucosa (Sanchez-Moreno et al. 1995). Durante el crecimiento exponencial, en las células se produce el consumo de glucosa y la disimilación del acetato. La acetogénesis, excreción del acetato al ambiente, se debe a la necesidad de regenerar el NAD<sup>+</sup> consumido en la glicólisis y de reciclar la coenzima A (CoASH) necesaria para convertir el piruvato en acetyl-CoA (punto central del metabolismo) (Wolfe 2005). Por otra parte, el acetato es también un producto esencial para la biosíntesis de lípidos (Mazet et al. 2013; Riviere et al. 2009), pudiéndose sugerir que la cantidad de proteínas con anclaje GPI, que se expresan de manera clonal en las poblaciones isogénicas de epimastigotes, podría estar relacionada con el procesamiento heterogéneo de ácidos grasos, quedado reflejado en los niveles de los distintos catabolitos analizados. Adicionalmente, el tráfico de proteínas ancladas a GPI (como la familia MASP) hasta su exposición en la membrana plasmática, así como su difusión a través de dicha membrana de *T. cruzi*, es un factor que podría limitar y alterar el crecimiento de una población de células, tal y como se ha demostrado mediante la modificación de los ectodominios de proteínas VSG expresadas en la

membrana plasmática de *T. brucei* (Hartel et al. 2015). De esta manera, aunque aún se desconoce si es una o varias las MASPs expresadas por una única célula al mismo tiempo, es muy probable que la expresión clonal de esta familia multigénica, sumada a la alta diversidad de sus secuencias y la variabilidad en el tamaño de sus ectodominios, puedan afectar al crecimiento de *T. cruzi* explicando parte de las diferencias clonales encontradas en las curvas de crecimiento descritas por las líneas isogénicas de epimastigotes analizadas en este trabajo de tesis (Figura 4.16). Este fenómeno también afectaría al consumo de glucosa y por lo tanto se vería reflejado en los catabolitos analizados

Aunque no se ha podido encontrar una conexión biológica completa entre la heterogeneidad clonal de la expresión MASP y los parámetros fenotípicos analizados, es interesante remarcar como la línea celular parental presentó nuevamente los niveles más bajos de metaciclologénesis espontánea y producción de catabolitos, tal y como ocurrió con los niveles y la variabilidad de expresión MASP. En este sentido, el estrés al que fueron sometidos los parásitos durante el proceso de clonación y las características únicas de cada clon, podrían determinar las diferencias encontradas en la expresión de MASP y en los aspectos fenotípicos mencionados. Alternativamente, es posible que el proceso de clonación haya supuesto un cuello de botella para la población original de epimastigotes, en donde las células mejor adaptadas fueran aquellas que estuvieran en constante cambio de la expresión MASP, presentando también una mayor actividad metabólica.

La concordancia de los resultados así como la metodología empleada en este estudio, en el que todas las líneas celulares fueron sometidas al mismo número de pases y todos los experimentos fueron realizados sobre el mismo número de parásitos cuando éstos se encontraban en la misma fase de crecimiento (fase exponencial), permite descartar que la heterogeneidad hallada a diferentes niveles en poblaciones clonales de epimastigotes, sean producto de un artefacto experimental.

Aunque resulte paradójico, la heterogeneidad transcripcional y fenotípica entre células que son genéticamente idénticas y que se encuentran bajo un

ambiente teóricamente común, es un fenómeno que, gracias a las nuevas tecnologías de análisis de células individuales, empieza a describirse en una gran cantidad de organismos y procesos celulares (Balazsi, van Oudenaarden, Collins 2011; Casadesus and Low 2013; Levy, Ziv, Siegal 2012; Navin 2014; Pelkmans 2012; Rovira-Graells et al. 2012a; Seco-Hidalgo, Osuna, Pablos 2015; Ubeda et al. 2014; Xia et al. 2014). Así, la heterogeneidad celular es una característica implícita en la producción y control de los fenotipos y, aunque no todas las variaciones fenotípicas van a ser seleccionadas, la generación de mosaicos celulares, dentro de una población, ocupa un papel importante en ambientes cambiantes como en los que vive *T. cruzi* (Herman et al. 2014; Seco-Hidalgo, Osuna, Pablos 2015).

La variación fenotípica puede generarse aleatoriamente o en base a una respuesta celular a señales ambientales externas. De esta manera, dependiendo de como pueden predecirse los cambios en el ambiente, las células presentan diferentes estrategias evolutivas: plasticidad adaptativa (AP, del inglés "*adaptive plasticity*"), apuesta diversificada (BH, del inglés "*bet-hedging*") o estrategias fijas (FS, del inglés "*fixed strategies*"). Todas estas estrategias evolutivas que aquí se describen pueden ser consideradas tanto para genes, individuos o poblaciones (Seco-Hidalgo, Osuna, Pablos 2015):

- **AP:** La AP se define como la habilidad que tiene un solo genotipo para producir diferentes fenotipos en respuesta a la variación ambiental (Forsman 2014). La plasticidad del genotipo puede anticiparse (activa) o ser una consecuencia del ambiente (pasiva), lo que implica mecanismos sensoriales de adaptación. Para que dicha plasticidad sea adaptativa debe producirse un fenotipo que esté acorde a los valores óptimos que favorecen la selección en el nuevo ambiente (Ghalambor et al. 2007). Debido a que ningún organismo es infinitamente plástico, existen límites y costes que definen la amplitud de la respuesta (Auld, Agrawal, Relyea 2010; Dewitt, Sih, Wilson 1998; Seco-Hidalgo, Osuna, Pablos 2015). De esta manera, la AP optimiza la esperanza de vida dentro de una misma generación.

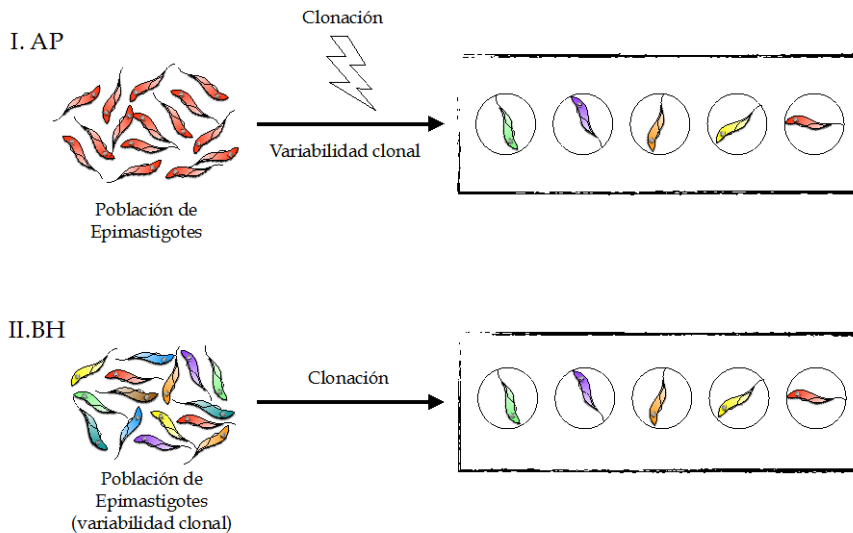
- **BH:** En ambientes que varían de forma impredecible, la supervivencia puede asegurarse mediante un mosaicismo celular mantenido por estrategias de BH (Simons 2011). Este mecanismo maximiza la esperanza de vida geométrica a expensas de la esperanza de vida aritmética (Ghalambor et al. 2007). Así el BH tiende a generar variabilidad constante en determinados rasgos fenotípicos, aún cuando estas variantes no sean las más óptimas con los requerimientos concretos del ambiente externo. Esto produce un descenso de la esperanza de vida en la generación siguiente, pareciendo algo perjudicial en períodos cortos. Sin embargo, la producción basal de variabilidad mantenida por BH va a evitar el riesgo de extinción ante fluctuaciones imprevistas del ambiente, maximizando la esperanza de vida de generaciones futuras (Ghalambor et al. 2007; Seco-Hidalgo, Osuna, Pablos 2015; Viney and Reece 2013). Igual que en AP, un fenotipo tampoco puede adoptar un número infinito de formas mediante BH, limitándose a un conjunto de estados que se producen en proporciones aleatorias. Para demostrar que la variabilidad de un rasgo se debe a la acción de BH, los diferentes estados del fenotipo deben encontrarse en proporciones similares y ser ajenos a señales (externas o internas) del ambiente (Viney and Reece 2013). Una dificultad con la que se encuentran los investigadores es la de definir los límites de la aleatoriedad, los cuales no siempre resultan evidentes (Seco-Hidalgo, Osuna, Pablos 2015)
- **FS:** En ambientes impredecibles en los que además la respuesta al estímulo se produce con retraso, los costes biológicos van a ser cuantiosos. En estos casos, los organismos pueden optar por estrategias fijas (FS) en detrimento de las estrategias plásticas (AP y BH). FS se refiere a aquellas situaciones en las que un organismo adopta un fenotipo determinado que no va a variar a pesar del ambiente (Seco-Hidalgo, Osuna, Pablos 2015; Tilmon 2008).

Teniendo en cuenta lo complejo y diverso del ciclo de vida de los parásitos como *T. cruzi* (con varios hospedadores), las estrategias de adaptación y variación deben de estar activas a lo largo de su ciclo vital (Kochin, Bull, Antia 2010). Así, acorde con lo revisado en el apartado 2.7.2 de la Introducción de este

trabajo, pese a la ausencia de un mecanismo de regulación transcripcional como tal, los kinetoplastidos *T. cruzi* y *Leishmania spp.* son capaces de producir y mantener heterogeneidad fenotípica a través de diferentes medios (Laffitte et al. 2014; Lima et al. 2013; Seco-Hidalgo, Osuna, Pablos 2015; Ubeda et al. 2014)

En el caso descrito en este trabajo, el mosaico transcripcional de la familia MASP podría sustentarse a través de cualquiera de las dos estrategias plásticas mencionadas: bien por mecanismos de plasticidad adaptativa (AP), que requieren de sensores ambientales y que permitan responder a *T. cruzi* frente a cambios externos; o mediante expresión estocástica clonal (BH), estrategia por la cual, dentro de unos límites, las parásitos van a producir constante y aleatoriamente diferentes MASP.

Profundizando en esta idea, se sugiere que si la diversidad clonal detectada fuera consecuencia del AP, el proceso de clonación (llevado a cabo para la obtención de la diferentes líneas clonales) y el estrés derivado del mismo, podrían actuar como un estímulo ambiental frente al cual *T. cruzi* generaría la heterogeneidad celular observada. Mientras que, si el BH fuera el mecanismo evolutivo responsable, la variación encontrada en la población de epimastigotes no se vería afectada por el proceso de clonación, siendo la heterogeneidad clonal la consecuencia de la selección de células cuya variación está predeterminada de una forma constante y azarosa (Figura 5.2).



**Figura 5.2. Representación esquemática del proceso de clonación en el que se visualiza la manera en que AP o BH actuarían sobre la variabilidad clonal de *T. cruzi*.** En AP (I), el proceso de clonación sería un estímulo abiental que fomentaría la heterogeneidad celular. Mientras que BH (II), actuaría produciendo dicha heterogeneidad sin verse afectado por el proceso de clonación.

De cualquier manera, la heterogeneidad clonal en la naturaleza pudiera ser una respuesta a las numerosos factores que *T. cruzi* tiene que enfrentarse durante su trayecto a través del tracto digestivo del insecto vector (factores que no han sido analizados en este trabajo), permitiéndole asegurarse su propia supervivencia dentro del vector.

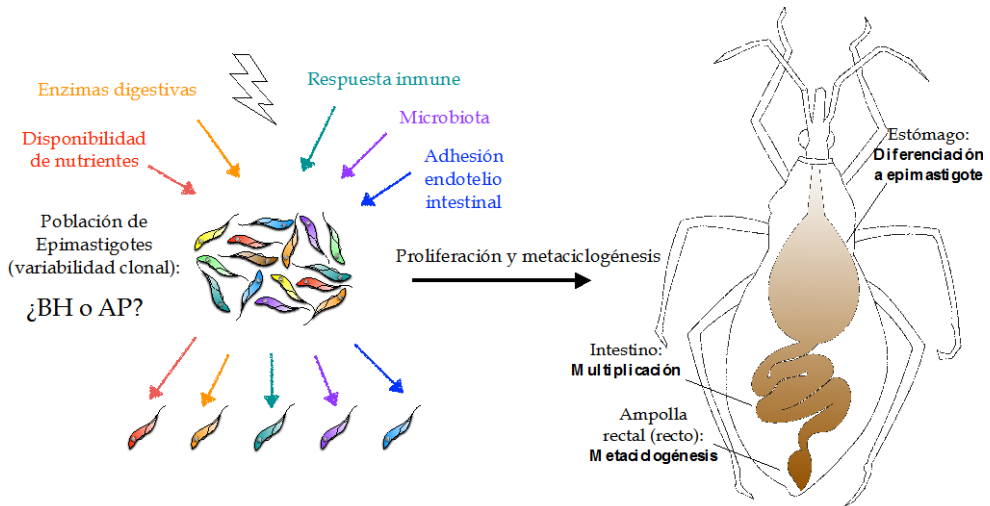
Nuestro estudio de mosaicismo fenotípico se realizó sobre la fase mayoritaria de *T. cruzi* en el insecto vector, el estado epimastigote. En gran medida, la enfermedad de Chagas y por tanto la propagación de *T. cruzi*, van a depender de la interacción que se establezca entre el protozoo y el insecto vector.

En la co-evolución de parásitos e insectos se establece una carrera evolutiva en la que los mecanismos del parásito y del insecto compiten para facilitar o impedir el desarrollo del parásito en el hospedador invertebrado, habiendo



múltiples factores que pueden determinar dicho proceso (Azambuja, Feder, Garcia 2004; Azambuja, Garcia, Ratcliffe 2005; Garcia et al. 2007; Noireau, Diosque, Jansen 2009). Al respecto, existen estudios que demuestran como *T. cruzi* presenta una infectividad variable sobre el insecto vector, lo que va a depender tanto de la susceptibilidad de las diferentes especies de triatóminos como de la capacidad de adaptación de las cepas de *T. cruzi* al hospedador invertebrado (Carvalho-Moreira et al. 2003; Garcia and Dvorak 1982; Perlowagora-Szumlewicz and Moreira 1994). Pese a que los aspectos de interacción parásito-vector nunca han sido abordados desde la singularidad de una única célula, algunos autores ya han atribuido la variabilidad del comportamiento de *T. cruzi* en el triatómino a la estructura poblacional altamente heterogénea del protozoo (de Lana et al. 1998; Lima et al. 1999).

Un total de tres estadios morfológicos de *T. cruzi* ( tripomastigotes sanguíneo → epimastigote → tripomastigotes metacíclico) deben de sucederse en el interior del triatómino para que posteriormente el protozoo pueda cerrar su ciclo en el hospedador vertebrado. La interacción entre en el parásito y el insecto hematófago comienza cuando este último ingiere las formas tripomastigotes sanguíneas de *T. cruzi* circulantes en la sangre del hospedador vertebrado. Unos días después de la ingesta, esta forma del parásito va a transformarse en epimastigotes, los cuales van a estar presentes desde el estómago hasta el recto del triatómino. En este trayecto, los epimastigotes son responsables de la multiplicación (en el intestino delgado) y de la metacicloogénesis (en el recto) (Figura 5.3).



**Figura 5.3. Representación esquemática de los estímulos ambientales que pueden incidir sobre la variabilidad clonal de una población de epimastigotes de *T. cruzi* durante su trayecto por el aparato digestivo del triatómino.**

Analizando la superficie celular de *T. cruzi* se ha visto como los epimastigotes presentan una composición lipídica, proteica y de carbohidratos diferente a la de la forma tripomastigote (Ferguson 1999; Garcia et al. 2007; Previato et al. 2004). Sin embargo, se han observado denominadores comunes en la membrana de ambas fases morfológicas; en este respecto, se enmarcarían las dos proteínas MASP, de 40 y 45 kDa, reconocidas por el anticuerpo anti-MASP-ag en epimastigotes y tripomastigotes de *T. cruzi*. Así, al igual que las mucinas, las proteínas de la familia MASP, tal y como se ha mostrado en este trabajo, se expresan también en la membrana de los epimastigotes (Buscaglia et al. 2006; De Pablos et al. 2011; El-Sayed et al. 2005b) donde pueden tener un papel clave evitando la respuesta del insecto o incluso protegiendo al parásito de la acción de enzimas digestivas (Garcia et al. 2007).

El establecimiento de las poblaciones de *T. cruzi* en el tracto digestivo del triatómino va a depender de diferentes aspectos bioquímicos y fisiológicos. Por una parte, se ha descrito como *T. cruzi* compete con el vector por los nutrientes

(Garcia et al. 2007). El triatómino obtiene los nutrientes succionando grandes cantidades de sangre seguido de una rápida diuresis y una lenta digestión, con episodios poco frecuentes de inanición. Todo esto supone un cambio en el ambiente intestinal en el que vive *T. cruzi*, afectando tanto a su densidad poblacional como al porcentaje de las diferentes fases morfológicas del mismo (Kollien and Schaub 2000). Por otra parte, para metabolizar el alimento y obtener los nutrientes la mayoría de los insectos hacen uso de las enzimas digestivas del tipo serín proteasa (tripsina y quimiotripsina); en cambio, los triatominos se basan en la acción de las cisteín proteasas para digerir las proteínas de la sangre (Borges et al. 2006). Dentro de esta familia de enzimas destacan las catepsinas que son activas a pH ácido y se expresan en la parte posterior del intestino medio. En triatóminos infectados con *T. cruzi* se observó un aumento en la actividad de la catepsina D entre 1-3 días después de la ingesta de sangre (Borges et al. 2006), aunque en *R. prolixus* tratados con pepstatin, un inhibidor de cisteín proteasas, no se vió ningún efecto sobre los índices de infección (Garcia and Gilliam 1980).

La respuesta inmune del triatomino constituye otro obstáculo de carácter fisiológico que debe superar el protozoo para consagrar la infección en el insecto vector. En esta respuesta de tipo humoral destacan moléculas como la defensina, un péptido antimicrobiano de 43 aminoácidos que se produce principalmente en el tejido graso y, en menor medida, en el intestino de los triatóminos (Lopez et al. 2003). Sin embargo, Lopez *et al.* (2003) sugieren que esta molécula producida a altos niveles podría eliminar las bacterias simbiotas que ayudan al insecto a conseguir vitaminas y nutrientes esenciales. Es por ello, que el triatómino no produce defensina en altos niveles y no ejerce efecto alguno sobre el desarrollo del parásito.

Ursic-Bedoya y Lowenberger (2007) fueron capaces de identificar otras moléculas relacionadas con el sistema inmune del insecto infectando a *R. prolixus* con *Escherichia coli*, *Micrococcus luteus* y *T. cruzi*. Entre éstas destacan las *mucin/peritrophic-like molecules*, a las que se relaciona con la interacción específica con *T. cruzi* (Azambuja, Ratcliffe, Garcia 2005). También describieron la presencia

de transferrinas que forman parte de la respuesta innata y están implicadas en el metabolismo del hierro; de  $\beta 1-3$  *glucan recognition protein*, relacionadas con el reconocimiento de patógenos en insectos como ocurre en *Manduca sexta*, entre otros; de la proteinasa de la hemolimfa, que participa en la activación de la profenoloxidasa; y también nitroporinas (Ursic-Bedoya and Lowenberger 2007). Las nitroporinas son un conjunto de moléculas relacionadas con el transporte del óxido nítrico (NO). Durante la picadura, las nitroporinas liberan NO desde las glándulas del triatómino hacia el hospedador vertebrado para asegurar la vasodilatación y así la ingesta de sangre. El NO es sintetizado por el gen óxido nítrico sintasa (NOS) (Gazos-Lopes et al. 2012). Inyectando en glándulas salivales de *R. prolixus*, glicoinositolfosfolípidos derivados de *T. rangeli* (Tr GiPL) o de *T. cruzi* (Tc GiPL), Gazos-Lopes *et al.* (2012) observaron un bloqueo en la producción de NO. Estos resultados sugieren que los glicolípidos de los tripanosomátidos pueden modular la producción de NO en el insecto vector para favorecer su propia transmisión al hospedador vertebrado.

También, es interesante resaltar como, en el tracto digestivo del triatómino se encuentran microorganismos que pueden afectar al desarrollo y supervivencia del parásito en el insecto vector, frente a los cuales *T. cruzi* deberá convivir o incluso competir (Azambuja, Feder, Garcia 2004; Caballero et al. 2015; Garcia et al. 2007; Soares et al. 2015). Azambuja *et al.* (2004) analizaron el efecto que tienen las bacterias simbiotes residentes en *R. prolixus* sobre la infección por *T. cruzi*. Para ello se alimentó al insecto vector con sangre infectada por las cepas de *T. cruzi* Y y Dm28c y observaron como la bacteria *Serratia marcescens* proliferaba en detrimento de los parásitos de la cepa Y, mientras que la cepa Dm28c permanecía estable. En otros estudios llevados a cabo por Espino y colaboradores (2009), también se determinó la presencia de la alfa proto bacteria intracelular *Wolbachia* en diferentes órganos (gónadas y glándulas salivares) e incluso en las heces del triatómino *Rhodnius pallescens*, en presencia o ausencia de *T. cruzi* y de *T. rangeli*. Además, estudios recientes realizados mediante genómica comparativa han llegado a sugerir que la continua exposición e interacción de *T. cruzi* con la microbiota intestinal podrían tener como consecuencia la transferencia horizontal de material genético desde las bacterias al parásito, pudiendo explicar los

distintos factores de virulencia que *T. cruzi* tiene en su genoma. Si bien también pudiera tratarse de una convergencia evolutiva, se ha encontrado que una fracción de las secuencias de genes MASP poseen numerosas analogías con genes SipD de *Salmonella thypy*, relacionadas también con mecanismos de invasión y patogenicidad en la bacteria (Silva et al. 2013). Otros ejemplos recientes apoyan la importancia de estos eventos de transferencia, como el caso del gen de la prolina Racemasa de *T. cruzi* que actúa como mitógeno para células B durante la infección por *T. cruzi* y que se postula como transferido horizontalmente desde bacterias del género *Gemella* (Caballero et al. 2015)

Otra de las principales interacciones que ocurre entre el parásito y el insecto vector es la adhesión de *T. cruzi* a la superficie de las células epiteliales que recubren el intestino. Este proceso resulta clave tanto para establecer la infección, en el triatomino (permitiendo la multiplicación de epimastigotes) como para la metacicloénesis (Oliveira et al. 2012). A diferencia de los epimastigotes, la fase tripomastigote no es capaz de adherirse a la superficie intestinal (Garcia et al. 2007). Por su parte, los epimastigotes se adhieren a la superficie del lumen en el estómago e intestino. En el estómago, la adhesión de los epimastigotes se produce a través del cuerpo celular y del flagelo mientras que en el intestino en la interacción con las microvellosidades intestinales sólo participa el flagelo (Gonzalez et al. 1999).

En este proceso de adhesión intervienen varias moléculas tanto por parte del parásito como por parte del vector. En los triatominos se ha visto la expresión de diferentes glicosaminoglicanos en su tracto digestivo, que interaccionan con el epimastigote (Oliveira et al. 2012). Por su parte *T. cruzi*, además de moléculas como lectinas y carbohidratos que van a mediar la unión receptor-ligando (Gonzalez et al. 2013; Oliveira et al. 2012), necesita del reconocimiento y las interacciones que se producen a través de glicoinositolfofolípidos (GIPLs), los cuales se encuentran abundantemente en la superficie de los epimastigotes (Azambuja, Ratcliffe, Garcia 2005; Garcia et al. 2007; Gonzalez et al. 2013).

Por otro lado, durante la metacicloénesis se produce un cambio drástico de la expresión génica con una larga cantidad de proteínas sintetizadas de manera

exclusiva durante este proceso, encontrándose diferencias de hasta el 50% de *spots* entre el perfil electroforético de proteínas en 2D de epimastigotes y tripomastigotes metacíclicos (Parodi-Talice et al. 2007) y con hasta 1114 proteínas (574 proteínas hipotéticas) expresadas de manera diferencial (522 sobreexpresadas y 592 con una expresión menor) (de Godoy et al. 2012). Así, se ha visto como durante metacicloogénesis, el descenso de proteínas ribosomales va acompañado de un aumento en la producción de enzimas y proteínas relacionadas con la defensa antioxidante, transialidasas, AMP cíclico y adenil ciclasas (Atwood et al. 2005; de Godoy et al. 2012; Parodi-Talice et al. 2007). Aunque en la actualidad no existen evidencias que vinculen de una forma directa la familia multigénica MASP con el desarrollo de la metacicloogénesis, estas proteínas se han encontrado expresadas en distintas cepas del parásito (De Pablos 2010; De Pablos and Osuna 2012b). Además, Godoy *et al.* (de Godoy et al. 2012) detectaron la sobreexpresión de la glicoproteína de adhesión al flagelo (*flagellum-adhesion glycoprotein*) en el transcurso de la transformación. Así la metacicloogénesis, tal y como se ha comentado anteriormente, se muestra como un suceso dependiente de la adhesión de los epimastigotes a las paredes del recto (Kollien, Schmidt, Schaub 1998). Incluso experimentos *in vitro* revelan que es necesaria la adhesión de los epimastigotes a las paredes del frasco de cultivo para que éstos se diferencien en la formas tripomastigotes metacíclicas (Bonaldo et al. 1988). Con todo ello las proteínas de membrana con motivo GPI (como las proteínas MASP), van a mediar la adhesión de los epimastigotes al lumen del estómago y del intestino del triatomino, siendo moléculas clave para el establecimiento de la infección y para el desarrollo de la metacicloogénesis.

De esta manera, ya en el intestino de los triatóminos y con el fin de establecer la infección y llevar a cabo la metacicloogénesis, el repertorio clonal de MASP en los epimastigotes interactuaría con los receptores del insecto vector para adherirse a la superficie de las células del endotelio intestinal (Buscaglia et al. 2006; Garcia et al. 2007; Mortara et al. 1992).

Pese a todo, sería necesarios ensayos *in vitro* para demostrar si la capacidad de adhesión y la capacidad de inducir metacicloogénesis están relacionadas con la presencia de MASP en la superficie de los epimastigotes.

En el marco de la ausencia de un control transcripcional que caracteriza a los kinetoplástidos, y ante un ambiente tan variable como es en el que se mueven las formas epimastigotes de *T. cruzi*, donde dependiendo de la zona en la que se encuentren dentro del tracto digestivo del triatomino van a variar los diferentes aspectos bioquímicos y fisiológicos que se han descrito anteriormente, la existencia en la población de conjuntos de células fenotípicamente heterogéneos, puede ser clave para la supervivencia del parásito (Rovira-Graells et al. 2012a; Seco-Hidalgo, Osuna, Pablos 2015; Ubeda et al. 2014). Así, los resultados obtenidos en los análisis desarrollados durante este trabajo nos animan a sugerir que el alto grado de heterogeneidad clonal de la familia MASP podría permitir a los epimastigotes de *T. cruzi* establecer una rápida respuesta ante un ambiente tan cambiante (Figura 5.3). De esta manera, el amplio repertorio de proteínas MASP, expresados de una manera clonal, ofrecería a *T. cruzi* protección contra la acción de las enzimas digestivas del insecto vector. También le permitiría al parásito evadirse de la respuesta inmune del triatomino, así como podría intervenir en la adhesión a las microvellosidades intestinales, algo fundamental para el proceso de metacicloogénesis.











## **VI. CONCLUSIONES/CONCLUSIONS**



Los resultados obtenidos en esta Tesis Doctoral nos han llevado a las siguientes conclusiones:

1. La técnica PCR anidada de tres pasos, basada en las regiones conservadas de la familia MASP, se ha mostrado como un método eficaz para el estudio de la familia MASP, tal y como demuestra el alto nivel de especificidad así como el número de amplicones obtenido tanto en la librería de expresión como en la electroforesis capilar, ambos realizados a partir de muestras de ADNc de las fases tripomastigote y epimastigote de *T. cruzi*.
2. Mediante "*hierarchical cluster analysis*" se pudieron comparar las secuencias obtenidas en la librería de expresión MASP y sus ortólogos en diferentes cepas de *T. cruzi*, demostrándose la existencia de dos velocidades evolutivas para los miembros (genes y pseudogenes) de esta librería. Los genes MASP se correlacionaron de acuerdo a su origen filogenético, mientras que dicha correlación se perdió cuando se compararon los pseudogenes de esta familia.
3. A pesar de la tendencia general hacia una expresión de la familia MASP altamente variable, las técnicas "*hierarchical cluster analysis*", western blot, electroforesis capilar así como la comparación entre diferentes librerías de expresión MASP, han permitido detectar varios miembros del transcriptoma MASP (incluidos pseudogenes) conservados entre clones, cepas e incluso entre estadíos de *T. cruzi*. Estos resultados sugieren la acción de una presión selectiva positiva para mantener dichos transcritos y sus productos protéicos.
4. El nivel de expresión de los transcritos MASP detectados mediante northern blot por la sonda *masp19*, fue aproximadamente 3 veces superior en la fase tripomastigote respecto a la de la fase epimastigote. Sin embargo, el tiempo de vida del ARN MASP fue el mismo, resultado que confirmaría el alto grado de conservación de las secuencias *cis*-reguladoras dentro las regiones 3' UTR de los transcritos MASP.
5. El estudio de la variabilidad clonal de la fase epimastigote, llevado a cabo sobre 8 líneas isogénicas de la cepa Pan4, ha permitido caracterizar por primera vez la heterogeneidad transcripcional en *T. cruzi* mediante el estudio

de la expresión de la familia MASP (tanto a nivel de ARN como de proteínas), así como analizar la variabilidad fenotípica en aspectos tales como crecimiento, metacicloogénesis espontánea y secreción de diferentes catabolitos. La variabilidad clonal demostrada en este trabajo, puede suponer una importante ventaja adaptativa para *T. cruzi*, permitiendo una rápida respuesta de las formas epimastigotes del parásito ante las presiones y cambios ambientales que existen en el tracto digestivo del triatomino.

The results obtained in this thesis have led us to the following conclusions:

1. The 3-step nested PCR, based on the conserved MASP regions, has been shown as an efficient method to the study of the MASP family, such as demonstrate the high specificity and the number of amplicons obtained by the expression library and by the capillary electrophoresis, both processes realized from cDNA samples of trypomastigote and epimastigote stages of *T. cruzi*.
2. By hierarchical cluster analysis was possible to compare the sequence of the transcripts belonging to the MASP expression library and their orthologs in different *T. cruzi* strains, demonstrating the existence of two different evolutionary speeds for the members (genes and pseudogenes) of this library. The MASP genes showed a correlation with their phylogenetic origin, while the MASP pseudogenes showed a loss of this correlation.
3. Despite a general trend expressing a great variability of the MASP family, the hierarchical cluster analysis, western blot, capillary electrophoresis, and sequence comparison among different MASP cDNA libraries, have allowed for the detection of some members of the MASP transcriptome (including pseudogenes) which are conserved among clones, strains and even stages of *T. cruzi*. This results suggest the action of a positive selection for keeping these transcripts and their protein products.
4. The MASP level expression of the MASP transcripts detected by northern blot using the *masp19* probe, was approximately 3-fold higher in trypomastigotes than in epimastigote forms. However, the lifespan of the detected MASP RNAs was the same in the both stages of *T. cruzi*, which is in accordance with the extremely high



degree of conservation of the cis-sequences within the 3'UTR regions of MASP transcripts.

5. The study of the clonal variability in the epimastigote stage, conducted in 8 isogenic cell lines of the Pan4 strain, has allowed for the first characterization of the transcriptional heterogeneity in *T. cruzi* through the study of the MASP family expression (at RNA and protein levels), as well as analyze the phenotypical variation in aspects such as growth, spontaneous metacyclogenesis and secretion of different catabolites. The clonal variability demonstrated in this work might provide an adaptative advantage for *T. cruzi*, allowing a rapid response of the epimastigote stage of the parasite to environmental pressures or changes that exist in digestive tract of the triatomine.





## **VII. BIBLIOGRAFIA**



- Abad-Franch F, Santos WS, Schofield CJ. 2010. Research needs for chagas disease prevention. *Acta Trop* .
- Abranches E, Bekman E, Henrique D. 2013. Generation and characterization of a novel mouse embryonic stem cell line with a dynamic reporter of nanog expression. *PLoS One* 8(3):e59928.
- Acosta Rodriguez EV, Zuniga EI, Montes CL, Merino MC, Bermejo DA, Amezcua Vesely MC, Motran CC, Gruppi A. 2007. *Trypanosoma cruzi* infection beats the B-cell compartment favouring parasite establishment: Can we strike first? *Scand J Immunol* 66(2-3):137-42.
- Adad SJ, Andrade DC, Lopes ER, Chapadeiro E. 1991. Pathological anatomy of chagasic megaesophagus. *Rev Inst Med Trop Sao Paulo* 33(6):443-50.
- Afonso AM, Ebell MH, Tarleton RL. 2012. A systematic review of high quality diagnostic tests for chagas disease. *PLoS Negl Trop Dis* 6(11):e1881.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215(3):403-10.
- Amambua-Ngwa A, Tetteh KK, Manske M, Gomez-Escobar N, Stewart LB, Deerpake ME, Cheeseman IH, Newbold CI, Holder AA, Knuepfer E, et al. 2012. Population genomic scan for candidate signatures of balancing selection to guide antigen characterization in malaria parasites. *PLoS Genet* 8(11):e1002992.
- Andrade SG. 1974. Caracterização de cepas do *Trypanosoma cruzi* isoladas no recôncavo bahiano. *Rev Pat Trop* 3:65-121.
- Andrade LO and Andrews NW. 2004. Lysosomal fusion is essential for the retention of *Trypanosoma cruzi* inside host cells. *J Exp Med* 200(9):1135-43.
- Andrade LO, Machado CR, Chiari E, Pena SD, Macedo AM. 1999. Differential tissue distribution of diverse clones of *Trypanosoma cruzi* in infected mice. *Mol Biochem Parasitol* 100(2):163-72.
- Andrade V, Brodskyn C, Andrade SG. 1983. Correlation between isoenzyme patterns and biological behaviour of different strains of *Trypanosoma cruzi*. *Trans R Soc Trop Med Hyg* 77(6):796-9.
- Atwood JA,3rd, Minning T, Ludolf F, Nuccio A, Weatherly DB, Alvarez-Manilla G, Tarleton R, Orlando R. 2006. Glycoproteomics of *Trypanosoma cruzi* trypomastigotes using subcellular fractionation, lectin affinity, and stable isotope labeling. *J Proteome Res* 5(12):3376-84.
- Atwood JA,3rd, Weatherly DB, Minning TA, Bundy B, Cavola C, Opperdoes FR, Orlando R, Tarleton RL. 2005. The *Trypanosoma cruzi* proteome. *Science* 309(5733):473-6.
- Auld JR, Agrawal AA, Relyea RA. 2010. Re-evaluating the costs and limits of adaptive phenotypic plasticity. *Proc Biol Sci* 277(1681):503-11.

- Avendaño A, Calderón-Arguedas O, Valerio I, Chinchilla M. 2006. Heterogeneity in the in vitro metacyclogenesis of clones derived from a costa rican strain of *Trypanosoma cruzi* (Kinetoplastida: Trypanosomatidae). Parasitol Latinoam 61:12-6.
- Azambuja P, Garcia ES, Ratcliffe NA. 2005. Gut microbiota and parasite transmission by insect vectors. Trends Parasitol 21(12):568-72.
- Azambuja P, Ratcliffe NA, Garcia ES. 2005. Towards an understanding of the interactions of *Trypanosoma cruzi* and *Trypanosoma rangeli* within the reduviid insect host *Rhodnius prolixus*. An Acad Bras Cienc 77(3):397-404.
- Azambuja P, Feder D, Garcia ES. 2004. Isolation of *Serratia marcescens* in the midgut of *Rhodnius prolixus*: Impact on the establishment of the parasite *Trypanosoma cruzi* in the vector. Exp Parasitol 107(1-2):89-96.
- Baida RC, Santos MR, Carmo MS, Yoshida N, Ferreira D, Ferreira AT, El Sayed NM, Andersson B, da Silveira JF. 2006. Molecular characterization of serine-, alanine-, and proline-rich proteins of *Trypanosoma cruzi* and their possible role in host cell infection. Infect Immun 74(3):1537-46.
- Bailey TL and Elkan C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol 2:28-36.
- Balakirev, E.S. and Ayala, F.J. 2003. Pseudogenes: Are they "junk" or functional DNA? . 123 p.
- Balazsi G, van Oudenaarden A, Collins JJ. 2011. Cellular decision making and biological noise: From microbes to mammals. Cell 144(6):910-25.
- Barbosa MP, Carmo AA, Rocha MO, Ribeiro AL. 2015. Ventricular arrhythmias in chagas disease. Rev Soc Bras Med Trop 48(1):4-10.
- Bartholomeu DC, Cerqueira GC, Leao AC, daRocha WD, Pais FS, Macedo C, Djikeng A, Teixeira SM, El-Sayed NM. 2009. Genomic organization and expression profile of the mucin-associated surface protein (masp) family of the human pathogen *Trypanosoma cruzi*. Nucleic Acids Res 37(10):3407-17.
- Bayer-Santos E, Aguilar-Bonavides C, Rodrigues SP, Cordero EM, Marques AF, Varela-Ramirez A, Choi H, Yoshida N, da Silveira JF, Almeida IC. 2013. Proteomic analysis of *Trypanosoma cruzi* secretome: Characterization of two populations of extracellular vesicles and soluble proteins. J Proteome Res 12(2):883-97.
- Betran E, Wang W, Jin L, Long M. 2002. Evolution of the phosphoglycerate mutase processed gene in human and chimpanzee revealing the origin of a new primate gene. Mol Biol Evol 19(5):654-63.
- Bitar M, Boroni M, Macedo AM, Machado CR, Franco GR. 2013. The spliced leader trans-splicing mechanism in different organisms: Molecular details and possible biological roles. Front Genet 4:199.

- Bonaldo MC, Souto-Padron T, de Souza W, Goldenberg S. 1988. Cell-substrate adhesion during *Trypanosoma cruzi* differentiation. *J Cell Biol* 106(4):1349-58.
- Borges EC, Machado EM, Garcia ES, Azambuja P. 2006. *Trypanosoma cruzi*: Effects of infection on cathepsin D activity in the midgut of *Rhodnius prolixus*. *Exp Parasitol* 112(2):130-3.
- Brener Z and Gazzinelli RT. 1997. Immunological control of *Trypanosoma cruzi* infection and pathogenesis of chagas' disease. *Int Arch Allergy Immunol* 114(2):103-10.
- Bringaud F, Berriman M, Hertz-Fowler C. 2009. Trypanosomatid genomes contain several subfamilies of ingi-related retroposons. *Eukaryot Cell* 8(10):1532-42.
- Bringaud F, Bartholomeu DC, Blandin G, Delcher A, Baltz T, El-Sayed NM, Ghedin E. 2006. The *Trypanosoma cruzi* L1Tc and NARTc non-LTR retrotransposons show relative site specificity for insertion. *Mol Biol Evol* 23(2):411-20.
- Bringaud F, Biteau N, Zuiderwijk E, Berriman M, El-Sayed NM, Ghedin E, Melville SE, Hall N, Baltz T. 2004. The ingi and RIME non-LTR retrotransposons are not randomly distributed in the genome of *Trypanosoma brucei*. *Mol Biol Evol* 21(3):520-8.
- Briones MR, Souto RP, Stolf BS, Zingales B. 1999. The evolution of two *Trypanosoma cruzi* subgroups inferred from rRNA genes can be correlated with the interchange of american mammalian faunas in the cenozoic and has implications to pathogenicity and host specificity. *Mol Biochem Parasitol* 104(2):219-32.
- Brisse S, Dujardin JC, Tibayrenc M. 2000. Identification of six *Trypanosoma cruzi* lineages by sequence-characterised amplified region markers. *Mol Biochem Parasitol* 111(1):95-105.
- Bristow J, Gitelman SE, Tee MK, Staels B, Miller WL. 1993. Abundant adrenal-specific transcription of the human P450c21A "pseudogene". *J Biol Chem* 268(17):12919-24.
- Brown D and Waneck GL. 1992. Glycosyl-phosphatidylinositol-anchored membrane proteins. *J Am Soc Nephrol* 3(4):895-906.
- Burleigh BA and Woolsey AM. 2002. Cell signalling and *Trypanosoma cruzi* invasion. *Cell Microbiol* 4(11):701-11.
- Buscaglia CA and Di Noia JM. 2003. *Trypanosoma cruzi* clonal diversity and the epidemiology of chagas' disease. *Microbes Infect* 5(5):419-27.
- Buscaglia CA, Campo VA, Frasch AC, Di Noia JM. 2006. *Trypanosoma cruzi* surface mucins: Host-dependent coat diversity. *Nat Rev Microbiol* 4(3):229-36.
- Caballero ZC, Sousa OE, Marques WP, Saez-Alquezar A, Umezawa ES. 2007. Evaluation of serological tests to identify *Trypanosoma cruzi* infection in



- humans and determine cross-reactivity with *Trypanosoma rangeli* and *Leishmania* spp. Clin Vaccine Immunol 14(8):1045-9.
- Caballero ZC, Costa-Martins AG, Ferreira RC, P Alves JM, Serrano MG, Camargo EP, Buck GA, Minoprio P, G Teixeira MM. 2015. Phylogenetic and syntenic data support a single horizontal transference to a *Trypanosoma* ancestor of a prokaryotic proline racemase implicated in parasite evasion from host defences. Parasit Vectors 8:222,015-0829-y.
- Caler EV, Vaena de Avalos S, Haynes PA, Andrews NW, Burleigh BA. 1998. Oligopeptidase B-dependent signaling mediates host cell invasion by *Trypanosoma cruzi*. EMBO J 17(17):4975-86.
- Campo VA, Buscaglia CA, Di Noia JM, Frasch AC. 2006. Immunocharacterization of the mucin-type proteins from the intracellular stage of *Trypanosoma cruzi*. Microbes Infect 8(2):401-9.
- Canepa GE, Mesias AC, Yu H, Chen X, Buscaglia CA. 2012. Structural features affecting trafficking, processing, and secretion of *Trypanosoma cruzi* mucins. J Biol Chem 287(31):26365-76.
- Carvalho-Moreira CJ, Spata MC, Coura JR, Garcia ES, Azambuja P, Gonzalez MS, Mello CB. 2003. In vivo and in vitro metacyclogenesis tests of two strains of *Trypanosoma cruzi* in the triatomine vectors triatoma pseudomaculata and *Rhodnius neglectus*: Short/long-term and comparative study. Exp Parasitol 103(3-4):102-11.
- Casadesus J and Low DA. 2013. Programmed heterogeneity: Epigenetic mechanisms in bacteria. J Biol Chem 288(20):13929-35.
- Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, Robertson M, Vrana J, Jones K, Grotewold L, Smith A. 2007. Nanog safeguards pluripotency and mediates germline development. Nature 450(7173):1230-4.
- Cortez M, Atayde V, Yoshida N. 2006. Host cell invasion mediated by *Trypanosoma cruzi* surface molecule gp82 is associated with F-actin disassembly and is inhibited by enteroinvasive escherichia coli. Microbes Infect 8(6):1502-12.
- Cosentino RO and Aguero F. 2012. A simple strain typing assay for *Trypanosoma cruzi*: Discrimination of major evolutionary lineages from a single amplification product. PLoS Negl Trop Dis 6(7):e1777.
- Coura JR. 2013. Chagas disease: Control, elimination and eradication. is it possible? Mem Inst Oswaldo Cruz 108(8):962-7.
- Coura JR. 2007. Chagas disease: What is known and what is needed--a background article. Mem Inst Oswaldo Cruz 102 Suppl 1:113-22.
- Coura JR, Junqueira AC, Fernandes O, Valente SA, Miles MA. 2002. Emerging chagas disease in amazonian brazil. Trends Parasitol 18(4):171-6.
- Cross GA, Kim HS, Wickstead B. 2014. Capturing the variant surface glycoprotein repertoire (the VSGnome) of *Trypanosoma brucei* lister 427. Mol Biochem Parasitol 195(1):59-73.

- Cummings KL and Tarleton RL. 2003. Rapid quantitation of *Trypanosoma cruzi* in host tissue by real-time PCR. *Mol Biochem Parasitol* 129(1):53-9.
- Cura CI, Mejia-Jaramillo AM, Duffy T, Burgos JM, Rodriguero M, Cardinal MV, Kjos S, Gurgel-Goncalves R, Blanchet D, De Pablos LM, et al. 2010. *Trypanosoma cruzi* I genotypes in different geographical regions and transmission cycles based on a microsatellite motif of the intergenic spacer of spliced-leader genes. *Int J Parasitol* 40(14):1599-607.
- Cura CI, Duffy T, Lucero RH, Bisio M, Peneau J, Jimenez-Coello M, Calabuig E, Gimenez MJ, Valencia Ayala E, Kjos SA, et al. 2015. Multiplex real-time PCR assay using TaqMan probes for the identification of *Trypanosoma cruzi* DTUs in biological and clinical samples. *PLoS Negl Trop Dis* 9(5):e0003765.
- da Silveira JF, Umezawa ES, Luquetti AO. 2001. Chagas disease: Recombinant *Trypanosoma cruzi* antigens for serological diagnosis. *Trends Parasitol* 17(6):286-91.
- Davis WR, Gabbara S, Hupe D, Peliska JA. 1998. Actinomycin D inhibition of DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase and nucleocapsid protein. *Biochemistry* 37(40):14213-21.
- de Freitas JM, Augusto-Pinto L, Pimenta JR, Bastos-Rodrigues L, Goncalves VF, Teixeira SM, Chiari E, Junqueira AC, Fernandes O, Macedo AM, et al. 2006. Ancestral genomes, sex, and the population structure of *Trypanosoma cruzi*. *PLoS Pathog* 2(3):e24.
- De Gaudenzi J, Noé G, Campo VA, Frasch AC, Cassola A. 2011. Gene expression regulation in trypanosomatids. *Essays in Biochemistry* 51:31-46.
- de Godoy LM, Marchini FK, Pavoni DP, Rampazzo Rde C, Probst CM, Goldenberg S, Krieger MA. 2012. Quantitative proteomics of *Trypanosoma cruzi* during metacyclogenesis. *Proteomics* 12(17):2694-703.
- de Lana M, da Silveira Pinto A, Barnabe C, Quesney V, Noel S, Tibayrenc M. 1998. *Trypanosoma cruzi*: Compared vectorial transmissibility of three major clonal genotypes by triatoma infestans. *Exp Parasitol* 90(1):20-5.
- De Pablos LM. 2010. Análisis global de la familia multigénica MASP (mucin associated surface proteins) de *Trypanosoma cruzi*. .
- De Pablos LM and Osuna A. 2012a. Multigene families in *Trypanosoma cruzi* and their role in infectivity. *Infect Immun* 80(7):2258-64.
- De Pablos LM and Osuna A. 2012b. Conserved regions as markers of different patterns of expression and distribution of the mucin-associated surface proteins of *Trypanosoma cruzi*. *Infect Immun* 80(1):169-74.
- De Pablos LM, Gonzalez GG, Solano Parada J, Seco Hidalgo V, Diaz Lozano IM, Gomez Samblas MM, Cruz Bustos T, Osuna A. 2011. Differential expression and characterization of a member of the mucin-associated surface protein family secreted by *Trypanosoma cruzi*. *Infect Immun* 79(10):3993-4001.

- De Souza W. 2002. From the cell biology to the development of new chemotherapeutic approaches against trypanosomatids: Dreams and reality. *Kinetoplastid Biol Dis* 1(1):3.
- D'Errico I, Gadaleta G, Saccone C. 2004. Pseudogenes in metazoa: Origin and features. *Brief Funct Genomic Proteomic* 3(2):157-67.
- Dewitt TJ, Sih A, Wilson DS. 1998. Costs and limits of phenotypic plasticity. *Trends Ecol Evol* 13(2):77-81.
- Di Girolamo C, Martelli G, Ciannamè A, Vocale C, Fini M, Stefanini A, Landini MP, Viale P, Verucchi G. 2015. Chagas disease in a non-endemic country: A multidisciplinary research, Bologna, Italy. *J Immigr Minor Health*.
- Di Noia JM, Pollevick GD, Xavier MT, Previato JO, Mendoca-Previato L, Sanchez DO, Frasch AC. 1996. High diversity in mucin genes and mucin molecules in *Trypanosoma cruzi*. *J Biol Chem* 271(50):32078-83.
- Dias JC, Silveira AC, Schofield CJ. 2002. The impact of chagas disease control in Latin America: A review. *Mem Inst Oswaldo Cruz* 97(5):603-12.
- Ding W, Lin L, Chen B, Dai J. 2006. L1 elements, processed pseudogenes and retrogenes in mammalian genomes. *IUBMB Life* 58(12):677-85.
- Docampo R, Jimenez V, Lander N, Li ZH, Niyogi S. 2013. New insights into roles of acidocalcisomes and contractile vacuole complex in osmoregulation in protists. *Int Rev Cell Mol Biol* 305:69-113.
- Donelson JE and Zeng W. 1990. A comparison of trans-RNA splicing in trypanosomes and nematodes. *Parasitol Today* 6(10):327-34.
- dos Santos SL, Freitas LM, Lobo FP, Rodrigues-Luiz GF, Mendes TA, Oliveira AC, Andrade LO, Chiari E, Gazzinelli RT, Teixeira SM, et al. 2012. The MASP family of *Trypanosoma cruzi*: Changes in gene expression and antigenic profile during the acute phase of experimental infection. *PLoS Negl Trop Dis* 6(8):e1779.
- El-Sayed NM, Myler PJ, Blandin G, Berriman M, Crabtree J, Aggarwal G, Caler E, Renault H, Worthey EA, Hertz-Fowler C, et al. 2005a. Comparative genomics of trypanosomatid parasitic protozoa. *Science* 309(5733):404-9.
- El-Sayed NM, Myler PJ, Bartholomeu DC, Nilsson D, Aggarwal G, Tran AN, Ghedin E, Worthey EA, Delcher AL, Blandin G, et al. 2005b. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* 309(5733):409-15.
- Epting CL, Coates BM, Engman DM. 2010. Molecular mechanisms of host cell invasion by *Trypanosoma cruzi*. *Exp Parasitol* 126(3):283-91.
- Espino CI, Gomez T, Gonzalez G, do Santos MF, Solano J, Sousa O, Moreno N, Windsor D, Ying A, Vilchez S, et al. 2009. Detection of Wolbachia bacteria in multiple organs and feces of the triatomine insect *Rhodnius palllescens* (Hemiptera, Reduviidae). *Appl Environ Microbiol* 75(2):547-50.

- Fallahi-Sichani M, Honarnejad S, Heiser LM, Gray JW, Sorger PK. 2013. Metrics other than potency reveal systematic variation in responses to cancer drugs. *Nat Chem Biol* 9(11):708-14.
- Ferguson MA. 1999. The structure, biosynthesis and functions of glycosylphosphatidylinositol anchors, and the contributions of trypanosome research. *J Cell Sci* 112 ( Pt 17)(Pt 17):2799-809.
- Ferguson MA, Brimacombe JS, Cottaz S, Field RA, Guther LS, Homans SW, McConville MJ, Mehlert A, Milne KG, Ralton JE. 1994. Glycosylphosphatidylinositol molecules of the parasite and the host. *Parasitology* 108 Suppl:S45-54.
- Fernandez-Moya SM and Estevez AM. 2010. Posttranscriptional control and the role of RNA-binding proteins in gene regulation in trypanosomatid protozoan parasites. *Wiley Interdiscip Rev RNA* 1(1):34-46.
- Ferreira D, Cortez M, Atayde VD, Yoshida N. 2006. Actin cytoskeleton-dependent and -independent host cell invasion by *Trypanosoma cruzi* is mediated by distinct parasite surface molecules. *Infect Immun* 74(10):5522-8.
- Field MC, Medina-Acosta E, Cross GA. 1991. Characterization of a glycosylphosphatidylinositol membrane protein anchor precursor in *Leishmania mexicana*. *Mol Biochem Parasitol* 48(2):227-9.
- Forsman A. 2014. Rethinking phenotypic plasticity and its consequences for individuals, populations and species. *Heredity* (Edinb) .
- Franzen O, Ochaya S, Sherwood E, Lewis MD, Llewellyn MS, Miles MA, Andersson B. 2011. Shotgun sequencing analysis of *Trypanosoma cruzi* I sylvio X10/1 and comparison with T. cruzi VI CL Brener. *PLoS Negl Trop Dis* 5(3):e984.
- Franzen O, Talavera-Lopez C, Ochaya S, Butler CE, Messenger LA, Lewis MD, Llewellyn MS, Marinkelle CJ, Tyler KM, Miles MA, et al. 2012. Comparative genomic analysis of human infective *Trypanosoma cruzi* lineages with the bat-restricted subspecies *T. cruzi* marinkellei. *BMC Genomics* 13:531,2164-13-531.
- Freitas LM, dos Santos SL, Rodrigues-Luiz GF, Mendes TA, Rodrigues TS, Gazzinelli RT, Teixeira SM, Fujiwara RT, Bartholomeu DC. 2011. Genomic analyses, gene expression and antigenic profile of the transsialidase superfamily of *Trypanosoma cruzi* reveal an undetected level of complexity. *PLoS One* 6(10):e25914.
- Fujita M and Jigami Y. 2008. Lipid remodeling of GPI-anchored proteins and its function. *Biochim Biophys Acta* 1780(3):410-20.
- Fujiwara H. 2015. Site-specific non-LTR retrotransposons. *Microbiol Spectr* 3(2):MDNA3,0001-2014.
- Garcia ES and Azambuja P. 1991. Development and interactions of *Trypanosoma cruzi* within the insect vector. *Parasitol Today* 7(9):240-4.

- Garcia ES and Dvorak JA. 1982. Growth and development of two *Trypanosoma cruzi* clones in the arthropod dipetalogaster maximus. *Am J Trop Med Hyg* 31(2):259-62.
- Garcia ES and Gilliam FC. 1980. *Trypanosoma cruzi* development is independent of protein digestion in the gut of *Rhodnius prolixus*. *J Parasitol* 66(6):1052-3.
- Garcia ES, Ratcliffe NA, Whitten MM, Gonzalez MS, Azambuja P. 2007. Exploring the role of insect host factors in the dynamics of *Trypanosoma cruzi*-*Rhodnius prolixus* interactions. *J Insect Physiol* 53(1):11-21.
- Gazos-Lopes F, Mesquita RD, Silva-Cardoso L, Senna R, Silveira AB, Jablonka W, Cudishevitch CO, Carneiro AB, Machado EA, Lima LG, et al. 2012. Glycoinositolphospholipids from trypanosomatids subvert nitric oxide production in *Rhodnius prolixus* salivary glands. *PLoS One* 7(10):e47285.
- Ghalambor CK, McKay JK, Carroll SP, Reznick DN. 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *British Ecological Society* .
- Gonzalez MS, Nogueira NF, Mello CB, De Souza W, Schaub GA, Azambuja P, Garcia ES. 1999. Influence of brain and azadirachtin on *Trypanosoma cruzi* development in the vector, *Rhodnius prolixus*. *Exp Parasitol* 92(2):100-8.
- Gonzalez MS, Souza MS, Garcia ES, Nogueira NF, Mello CB, Canepa GE, Bertotti S, Durante IM, Azambuja P, Buscaglia CA. 2013. *Trypanosoma cruzi* TcSMUG L-surface mucins promote development and infectivity in the triatomine vector *Rhodnius prolixus*. *PLoS Negl Trop Dis* 7(11):e2552.
- Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R. 2010. A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* 38(Web Server issue):W695-9.
- Guedes PM, Silva GK, Gutierrez FR, Silva JS. 2011. Current status of chagas disease chemotherapy. *Expert Rev Anti Infect Ther* 9(5):609-20.
- Guo X, Zhang Z, Gerstein MB, Zheng D. 2009. Small RNAs originated from pseudogenes: Cis- or trans-acting? *PLoS Comput Biol* 5(7):e1000449.
- Hall JP, Wang H, Barry JD. 2013. Mosaic VSGs and the scale of *Trypanosoma brucei* antigenic variation. *PLoS Pathog* 9(7):e1003502.
- Harlow, E., Lane, D. 1988. *Antibodies: A laboratory manual* .
- Hartel AJ, Glogger M, Guigas G, Jones NG, Fenz SF, Weiss M, Engstler M. 2015. The molecular size of the extra-membrane domain influences the diffusion of the GPI-anchored VSG on the trypanosome plasma membrane. *Sci Rep* 5:10394.
- Hayashi K, Lopes SM, Tang F, Surani MA. 2008. Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell* 3(4):391-401.

- Herman JJ, Spencer HG, Donohue K, Sultan SE. 2014. How stable 'should' epigenetic modifications be? insights from adaptive plasticity and bet hedging. *Evolution* 68(3):632-43.
- Higgins MK, Tkachenko O, Brown A, Reed J, Raper J, Carrington M. 2013. Structure of the trypanosome haptoglobin-hemoglobin receptor and implications for nutrient uptake and innate immunity. *Proc Natl Acad Sci U S A* 110(5):1905-10.
- Horn D. 2008. Codon usage suggests that translational selection has a major impact on protein expression in trypanosomatids. *BMC Genomics* 9:2,2164-9-2.
- Hotez PJ, Bottazzi ME, Franco-Paredes C, Ault SK, Periago MR. 2008. The neglected tropical diseases of latin america and the caribbean: A review of disease burden and distribution and a roadmap for control and elimination. *PLoS Negl Trop Dis* 2(9):e300.
- Jackson AP, Allison HC, Barry JD, Field MC, Hertz-Fowler C, Berriman M. 2013. A cell-surface phylome for african trypanosomes. *PLoS Negl Trop Dis* 7(3):e2121.
- Jackson AP, Berry A, Aslett M, Allison HC, Burton P, Vavrova-Anderson J, Brown R, Browne H, Corton N, Hauser H, et al. 2012. Antigenic diversity is generated by distinct evolutionary mechanisms in african trypanosome species. *Proc Natl Acad Sci U S A* 109(9):3416-21.
- Jackson Y, Varcher Herrera M, Gascon J. 2014. Economic crisis and increased immigrant mobility: New challenges in managing chagas disease in europe. *Bull World Health Organ* 92(10):771-2.
- Jager AV, Muia RP, Campetella O. 2008. Stage-specific expression of *Trypanosoma cruzi* trans-sialidase involves highly conserved 3' untranslated regions. *FEMS Microbiol Lett* 283(2):182-8.
- Kafsack BF, Rovira-Graells N, Clark TG, Bancells C, Crowley VM, Campino SG, Williams AE, Drought LG, Kwiatkowski DP, Baker DA, et al. 2014. A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature* 507(7491):248-52.
- Kalyana-Sundaram S, Kumar-Sinha C, Shankar S, Robinson DR, Wu YM, Cao X, Asangani IA, Kothari V, Prensner JR, Lonigro RJ, et al. 2012. Expressed pseudogenes in the transcriptional landscape of human cancers. *Cell* 149(7):1622-34.
- Kissinger JC. 2006. A tale of three genomes: The kinetoplastids have arrived. *Trends Parasitol* 22(6):240-3.
- Kochin BF, Bull JJ, Antia R. 2010. Parasite evolution and life history theory. *PLoS Biol* 8(10):e1000524.
- Kollien AH and Schaub GA. 2000. The development of *Trypanosoma cruzi* in triatominae. *Parasitol Today* 16(9):381-7.

- Kollien AH, Schmidt J, Schaub GA. 1998. Modes of association of *Trypanosoma cruzi* with the intestinal tract of the vector *Triatoma infestans*. *Acta Trop* 70(2):127-41.
- Korneev SA, Park JH, O'Shea M. 1999a. Neuronal expression of neural nitric oxide synthase (nNOS) protein is suppressed by an antisense RNA transcribed from an *NOS* pseudogene. *J Neurosci* 19(18):7711-20.
- Korneev SA, Park JH, O'Shea M. 1999b. Neuronal expression of neural nitric oxide synthase (nNOS) protein is suppressed by an antisense RNA transcribed from an *NOS* pseudogene. *J Neurosci* 19(18):7711-20.
- Kuhn RE. 1994. Macrophages in experimental chagas' disease. *Immunol Ser* 60:495-502.
- Kumar RM, Cahan P, Shalek AK, Satija R, DaleyKeyser AJ, Li H, Zhang J, Pardee K, Gennert D, Trombetta JJ, et al. 2014. Deconstructing transcriptional heterogeneity in pluripotent stem cells. *Nature* 516(7529):56-61.
- Laffitte MC, Genois MM, Mukherjee A, Legare D, Masson JY, Ouellette M. 2014. Formation of linear amplicons with inverted duplications in *Leishmania* requires the MRE11 nuclease. *PLoS Genet* 10(12):e1004805.
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. 2007. Clustal W and clustal X version 2.0. *Bioinformatics* 23(21):2947-8.
- Lauria-Pires L and Teixeira AR. 1996. Virulence and pathogenicity associated with diversity of *Trypanosoma cruzi* stocks and clones derived from chagas' disease patients. *Am J Trop Med Hyg* 55(3):304-10.
- LeBowitz JH, Smith HQ, Rusche L, Beverley SM. 1993. Coupling of poly(A) site selection and trans-splicing in *Leishmania*. *Genes Dev* 7(6):996-1007.
- Lent, H., Wygodzinsky, P. 1979. Revision of triatominae (hemiptera, reduvidae), and their significance as vectors of chagas disease. .
- Levy SF, Ziv N, Siegal ML. 2012. Bet hedging in yeast by heterogeneous, age-correlated expression of a stress protectant. *PLoS Biol* 10(5):e1001325.
- Ley V, Robbins ES, Nussenzweig V, Andrews NW. 1990. The exit of *Trypanosoma cruzi* from the phagosome is inhibited by raising the pH of acidic compartments. *J Exp Med* 171(2):401-13.
- Li B. 2015. DNA double-strand breaks and telomeres play important roles in *Trypanosoma brucei* antigenic variation. *Eukaryot Cell* 14(3):196-205.
- Li W, Yang W, Wang XJ. 2013. Pseudogenes: Pseudo or real functional elements? *J Genet Genomics* 40(4):171-7.
- Lima FM, Souza RT, Santori FR, Santos MF, Cortez DR, Barros RM, Cano MI, Valadares HM, Macedo AM, Mortara RA, et al. 2013. Interclonal variations in the molecular karyotype of *Trypanosoma cruzi*: Chromosome

- rearrangements in a single cell-derived clone of the G strain. PLoS One 8(5):e63738.
- Lima L, Espinosa-Alvarez O, Ortiz PA, Trejo-Varon JA, Carranza JC, Pinto CM, Serrano MG, Buck GA, Camargo EP, Teixeira MM. 2015. Genetic diversity of *Trypanosoma cruzi* in bats, and multilocus phylogenetic and phylogeographical analyses supporting tcbat as an independent DTU (discrete typing unit). Acta Trop 151:166-77.
- Lima VS, Mangia RH, Carreira JC, Marchevsky RS, Jansen AM. 1999. *Trypanosoma cruzi*: Correlations of biological aspects of the life cycle in mice and triatomines. Mem Inst Oswaldo Cruz 94(3):397-402.
- Lopez L, Morales G, Ursic R, Wolff M, Lowenberger C. 2003. Isolation and characterization of a novel insect defensin from *Rhodnius prolixus*, a vector of chagas disease. Insect Biochem Mol Biol 33(4):439-47.
- Macedo AM, Oliveira RP, Pena SD. 2002. Schematic representation of the life cycle of the flagellate protozoan *Trypanosoma cruzi*. Expert Reviews in Molecular Medicine .
- Macedo AM, Machado CR, Oliveira RP, Pena SD. 2004. *Trypanosoma cruzi*: Genetic structure of populations and relevance of genetic variability to the pathogenesis of chagas disease. Mem Inst Oswaldo Cruz 99(1):1-12.
- Maganto-Garcia E, Punzon C, Terhorst C, Fresno M. 2008. Rab5 activation by toll-like receptor 2 is required for *Trypanosoma cruzi* internalization and replication in macrophages. Traffic 9(8):1299-315.
- Manne-Goehler J, Reich MR, Wirtz VJ. 2015. Access to care for chagas disease in the united states: A health systems analysis. Am J Trop Med Hyg 93(1):108-13.
- Mantel PY and Marti M. 2014. The role of extracellular vesicles in *Plasmodium* and other protozoan parasites. Cell Microbiol 16(3):344-54.
- Marcili A, Lima L, Cavazzana M, Junqueira AC, Veludo HH, Maia Da Silva F, Campaner M, Paiva F, Nunes VL, Teixeira MM. 2009. A new genotype of *Trypanosoma cruzi* associated with bats evidenced by phylogenetic analyses using SSU rDNA, cytochrome b and histone H2B genes and genotyping based on ITS1 rDNA. Parasitology 136(6):641-55.
- Marcon GE, Andrade PD, de Albuquerque DM, Wanderley Jda S, de Almeida EA, Guariento ME, Costa SC. 2002. Use of a nested polymerase chain reaction (N-PCR) to detect *Trypanosoma cruzi* in blood samples from chronic chagasic patients and patients with doubtful serologies. Diagn Microbiol Infect Dis 43(1):39-43.
- Mazet M, Morand P, Biran M, Bouyssou G, Courtois P, Daulouede S, Millerioux Y, Franconi JM, Vincendeau P, Moreau P, et al. 2013. Revisiting the central metabolism of the bloodstream forms of *Trypanosoma brucei*: Production of acetate in the mitochondrion is essential for parasite viability. PLoS Negl Trop Dis 7(12):e2587.



- Mester B, Elguero J, Claramunt RM, Castanys S, Mascaro ML, Osuna A, Vilaplana MJ, Molina P. 1987. Activity against *Trypanosoma cruzi* of new analogues of nifurtimox. Arch Pharm (Weinheim) 320(2):115-20.
- Miles MA, Feliciangeli MD, de Arias AR. 2003. American trypanosomiasis (chagas' disease) and the role of molecular epidemiology in guiding control strategies. BMJ 326(7404):1444-8.
- Miles MA, Lanham SM, de Souza AA, Povoá M. 1980. Further enzymic characters of *Trypanosoma cruzi* and their evaluation for strain identification. Trans R Soc Trop Med Hyg 74(2):221-37.
- Miles MA, Toye PJ, Oswald SC, Godfrey DG. 1977. The identification by isoenzyme patterns of two distinct strain-groups of *Trypanosoma cruzi*, circulating independently in a rural area of Brazil. Trans R Soc Trop Med Hyg 71(3):217-25.
- Miles MA, Souza A, Povoá M, Shaw JJ, Lainson R, Toye PJ. 1978. Isozymic heterogeneity of *Trypanosoma cruzi* in the first autochthonous patients with chagas' disease in Amazonian Brazil. Nature 272(5656):819-21.
- Ming M, Chuenkova M, Ortega-Barria E, Pereira ME. 1993. Mediation of *Trypanosoma cruzi* invasion by sialic acid on the host cell and trans-sialidase on the trypanosome. Mol Biochem Parasitol 59(2):243-52.
- Minning TA, Weatherly DB, Flibotte S, Tarleton RL. 2011. Widespread, focal copy number variations (CNV) and whole chromosome aneuploidies in *Trypanosoma cruzi* strains revealed by array comparative genomic hybridization. BMC Genomics 12:139,2164-12-139.
- Moncayo A and Silveira AC. 2009. Current epidemiological trends for chagas disease in Latin America and future challenges in epidemiology, surveillance and health policy. Mem Inst Oswaldo Cruz 104 Suppl 1:17-30.
- Montefusco MC, Merlo K, Bryan CD, Surks HK, Reis SE, Mendelsohn ME, Huggins GS. 2010. Little ROCK is a ROCK1 pseudogene expressed in human smooth muscle cells. BMC Genet 11:22.
- Moraes Barros RR, Marini MM, Antonio CR, Cortez DR, Miyake AM, Lima FM, Ruiz JC, Bartholomeu DC, Chiurillo MA, Ramirez JL, et al. 2012. Anatomy and evolution of telomeric and subtelomeric regions in the human protozoan parasite *Trypanosoma cruzi*. BMC Genomics 13:229,2164-13-229.
- Moreno SN and Docampo R. 2009. The role of acidocalcisomes in parasitic protists. J Eukaryot Microbiol 56(3):208-13.
- Morillo CA, Marin-Neto JA, Avezum A, Sosa-Estani S, Rassi A, Jr, Rosas F, Villena E, Quiroz R, Bonilla R, Britto C, et al. 2015. Randomized trial of benznidazole for chronic chagas' cardiomyopathy. N Engl J Med .
- Morrow CA and Fraser JA. 2013. Ploidy variation as an adaptive mechanism in human pathogenic fungi. Semin Cell Dev Biol 24(4):339-46.

- Mortara RA, da Silva S, Araguth MF, Blanco SA, Yoshida N. 1992. Polymorphism of the 35- and 50-kilodalton surface glycoconjugates of *Trypanosoma cruzi* metacyclic trypomastigotes. *Infect Immun* 60(11):4673-8.
- Mortara RA, Procopio DO, Barros HC, Verbisck NV, Andreoli WK, Silva RB, da Silva S. 1999. Features of host cell invasion by different infective forms of *Trypanosoma cruzi*. *Mem Inst Oswaldo Cruz* 94 Suppl 1:135-7.
- Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, Cook K, Stepansky A, Levy D, Esposito D, et al. 2011. Tumour evolution inferred by single-cell sequencing. *Nature* 472(7341):90-4.
- Navin NE. 2015. The first five years of single-cell cancer genomics and beyond. *Genome Res* 25(10):1499-507.
- Navin NE. 2014. Cancer genomics: One cell at a time. *Genome Biol* 15(8):452,014-0452-9.
- Nguyen T, Sunahara R, Marchese A, Van Tol HH, Seeman P, O'Dowd BF. 1991. Transcription of a human dopamine D5 pseudogene. *Biochem Biophys Res Commun* 181(1):16-21.
- Niang M, Yan Yam X, Preiser PR. 2009. The *Plasmodium falciparum* STEVOR multigene family mediates antigenic variation of the infected erythrocyte. *PLoS Pathog* 5(2):e1000307.
- GeneDoc: a tool for editing and annotating multiple sequence alignments [Internet]; c1997. Available from: [www.psc.edu/biomed/genedoc](http://www.psc.edu/biomed/genedoc) .
- Noireau F, Diosque P, Jansen AM. 2009. *Trypanosoma cruzi*: Adaptation to its vectors and its hosts. *Vet Res* 40(2):26.
- Oelemann WM, Teixeira MD, Verissimo Da Costa GC, Borges-Pereira J, De Castro JA, Coura JR, Peralta JM. 1998. Evaluation of three commercial enzyme-linked immunosorbent assays for diagnosis of chagas' disease. *J Clin Microbiol* 36(9):2423-7.
- Oliveira FO,Jr, Alves CR, Souza-Silva F, Calvet CM, Cortes LM, Gonzalez MS, Toma L, Boucas RI, Nader HB, Pereira MC. 2012. *Trypanosoma cruzi* heparin-binding proteins mediate the adherence of epimastigotes to the midgut epithelial cells of *Rhodnius prolixus*. *Parasitology* 139(6):735-43.
- Opperdoes FR. 1990. The glycosome of trypanosomes and *Leishmania*. *Biochem Soc Trans* 18(5):729-31.
- La enfermedad de Chagas (tripanosomiasis americana). [Internet] [cited 2015 Junio/17]. Available from: <http://www.who.int/mediacentre/factsheets/fs340/es/> .
- Osuna Carrillo A, Jimenez Ortiz A, Lozano Maldonado J. 1979. Medios de cultivo para la obtención de formas metacíclicas de *Trypanosoma cruzi* . *Revista Ibérica De Parasitología* 39:129-33.

- Osuna A, Castanys S, Rodriguez-Cabezas MN, Gamarro F. 1990. *Trypanosoma cruzi*: Calcium ion movement during internalization in host HeLa cells. *Int J Parasitol* 20(5):673-6.
- Osuna A, Gamarro F, Castanys S, Ruiz-Perez LM. 1986. Inhibition of lysosomal fusion by *Trypanosoma cruzi* in peritoneal macrophages. *Int J Parasitol* 16(6):629-32.
- Osuna A, Jimenez-Ortiz A, Mascaro C, Alonso C. 1983. *Trypanosoma cruzi*: Arrested division of amastigote forms in enucleated HeLa cells. *J Parasitol* 69(3):629-31.
- Osuna A, Ortega G, Gamarro F, Castanys S, Mascaro MC. 1984a. Some factors affecting the in vitro invasion of HeLa cells by *Trypanosoma cruzi*. *Int J Parasitol* 14(3):253-7.
- Osuna A, Ortega G, Gamarro F, Castanys S, Mascaro MC. 1984b. Some factors affecting the in vitro invasion of HeLa cells by *Trypanosoma cruzi*. *Int J Parasitol* 14(3):253-7.
- Ouellette M and Papadopoulou B. 2009. Coordinated gene expression by post-transcriptional regulons in african trypanosomes. *J Biol* 8(11):100.
- Palacios-Pru E, Carrasco H, Scorza C, Espinoza R. 1989. Ultrastructural characteristics of different stages of human chagasic myocarditis. *Am J Trop Med Hyg* 41(1):29-40.
- Parodi-Talice A, Monteiro-Goes V, Arrambide N, Avila AR, Duran R, Correa A, Dallagiovanna B, Cayota A, Krieger M, Goldenberg S, et al. 2007. Proteomic analysis of metacyclic trypomastigotes undergoing *Trypanosoma cruzi* metacyclogenesis. *J Mass Spectrom* 42(11):1422-32.
- Pays E, Lips S, Nolan D, Vanhamme L, Perez-Morga D. 2001. The VSG expression sites of *Trypanosoma brucei*: Multipurpose tools for the adaptation of the parasite to mammalian hosts. *Mol Biochem Parasitol* 114(1):1-16.
- Pelkmans L. 2012. Cell biology. using cell-to-cell variability--a new era in molecular biology. *Science* 336(6080):425-6.
- Pena SD, Machado CR, Macedo AM. 2009. *Trypanosoma cruzi*: Ancestral genomes and population structure. *Mem Inst Oswaldo Cruz* 104 Suppl 1:108-14.
- Perlowagora-Szumlewicz A and Moreira CJ. 1994. In vivo differentiation of *Trypanosoma cruzi*--1. experimental evidence of the influence of vector species on metacyclogenesis. *Mem Inst Oswaldo Cruz* 89(4):603-18.
- Petherick A. 2010. Chagas disease in the chaco. *Nature* 465(7301):S18-20.
- Pink RC, Wicks K, Caley DP, Punch EK, Jacobs L, Carter DR. 2011. Pseudogenes: Pseudo-functional or key regulators in health and disease? *RNA* 17(5):792-8.

- Pinto A, Pett S, Jackson Y. 2014. Identifying chagas disease in australia: An emerging challenge for general practitioners. *Aust Fam Physician* 43(7):440-2.
- Podlaha O and Zhang J. 2010. Pseudogenes and their evolution. *Encyclopedia of Life Sciences* .
- Poliseno L. 2012. Pseudogenes: Newly discovered players in human cancer. *Sci Signal* 5(242):re5.
- Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP. 2010. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* 465(7301):1033-8.
- Pollitt LC, MacGregor P, Matthews K, Reece SE. 2011. Malaria and trypanosome transmission: Different parasites, same rules? *Trends Parasitol* 27(5):197-203.
- Postan M, McDaniel JP, Dvorak JA. 1987. Comparative studies of the infection of lewis rats with four *Trypanosoma cruzi* clones. *Trans R Soc Trop Med Hyg* 81(3):415-9.
- Prata A. 2001. Clinical and epidemiological aspects of chagas disease. *Lancet Infect Dis* 1(2):92-100.
- Previato JO, Wait R, Jones C, DosReis GA, Todeschini AR, Heise N, Previato LM. 2004. Glycoinositolphospholipid from *Trypanosoma cruzi*: Structure, biosynthesis and immunobiology. *Adv Parasitol* 56:1-41.
- Qvarnstrom Y, Schijman AG, Veron V, Aznar C, Steurer F, da Silva AJ. 2012. Sensitive and specific detection of *Trypanosoma cruzi* DNA in clinical specimens using a multi-target real-time PCR approach. *PLoS Negl Trop Dis* 6(7):e1689.
- R: A language and environment for statistical computing [Internet]; c2014 [cited 2015]. Available from: <http://www.R-project.org/> .
- Ramirez JC, Cura CI, da Cruz Moreira O, Lages-Silva E, Juiz N, Velazquez E, Ramirez JD, Alberti A, Pavia P, Flores-Chavez MD, et al. 2015. Analytical validation of quantitative real-time PCR methods for quantification of *Trypanosoma cruzi* DNA in blood samples from chagas disease patients. *J Mol Diagn* 17(5):605-15.
- Rassi A,Jr, Rassi A, Marin-Neto JA. 2010. Chagas disease. *Lancet* 375(9723):1388-402.
- Rassi A,Jr, Rassi A, Little WC. 2000. Chagas' heart disease. *Clin Cardiol* 23(12):883-9.
- Recker M, Buckee CO, Serazin A, Kyes S, Pinches R, Christodoulou Z, Springer AL, Gupta S, Newbold CI. 2011. Antigenic variation in *Plasmodium falciparum* malaria involves a highly structured switching pattern. *PLoS Pathog* 7(3):e1001306.

- Reis-Cunha JL, Rodrigues-Luiz GF, Valdivia HO, Baptista RP, Mendes TA, de Morais GL, Guedes R, Macedo AM, Bern C, Gilman RH, et al. 2015. Chromosomal copy number variation reveals differential levels of genomic plasticity in distinct *Trypanosoma cruzi* strains. BMC Genomics 16:499,015-1680-4.
- Requena-Mendez A, Aldasoro E, de Lazzari E, Sicuri E, Brown M, Moore DA, Gascon J, Munoz J. 2015. Prevalence of chagas disease in latin-american migrants living in europe: A systematic review and meta-analysis. PLoS Negl Trop Dis 9(2):e0003540.
- Riviere L, Moreau P, Allmann S, Hahn M, Biran M, Plazolles N, Franconi JM, Boshart M, Bringaud F. 2009. Acetate produced in the mitochondrion is the essential precursor for lipid biosynthesis in procyclic trypanosomes. Proc Natl Acad Sci U S A 106(31):12694-9.
- Rocha GM, Brandao BA, Mortara RA, Attias M, de Souza W, Carvalho TM. 2006. The flagellar attachment zone of *Trypanosoma cruzi* epimastigote forms. J Struct Biol 154(1):89-99.
- Rodriguez A, Rioult MG, Ora A, Andrews NW. 1995. A trypanosome-soluble factor induces IP3 formation, intracellular Ca<sup>2+</sup> mobilization and microfilament rearrangement in host cells. J Cell Biol 129(5):1263-73.
- Rodriguez A, Samoff E, Rioult MG, Chung A, Andrews NW. 1996. Host cell invasion by trypanosomes requires lysosomes and microtubule/kinesin-mediated transport. J Cell Biol 134(2):349-62.
- Ropert C, Ferreira LR, Campos MA, Procopio DO, Travassos LR, Ferguson MA, Reis LF, Teixeira MM, Almeida IC, Gazzinelli RT. 2002. Macrophage signaling by glycosylphosphatidylinositol-anchored mucin-like glycoproteins derived from *Trypanosoma cruzi* trypomastigotes. Microbes Infect 4(9):1015-25.
- Rovira-Graells N, Gupta AP, Planet E, Crowley VM, Mok S, Ribas de Pouplana L, Preiser PR, Bozdech Z, Cortes A. 2012a. Transcriptional variation in the malaria parasite *Plasmodium falciparum*. Genome Res 22(5):925-38.
- Rovira-Graells N, Gupta AP, Planet E, Crowley VM, Mok S, Ribas de Pouplana L, Preiser PR, Bozdech Z, Cortes A. 2012b. Transcriptional variation in the malaria parasite *Plasmodium falciparum*. Genome Res 22(5):925-38.
- Ruiz RC, Favoreto S, Jr, Dorta ML, Oshiro ME, Ferreira AT, Manque PM, Yoshida N. 1998. Infectivity of *Trypanosoma cruzi* strains is associated with differential expression of surface glycoproteins with differential Ca<sup>2+</sup> signalling activity. Biochem J 330 ( Pt 1)(Pt 1):505-11.
- Salles NA, Sabino EC, Cliquet MG, Eluf-Neto J, Mayer A, Almeida-Neto C, Mendonca MC, Dorliach-Llacer P, Chamone DF, Saez-Alquezar A. 1996. Risk of exposure to chagas' disease among seroreactive brazilian blood donors. Transfusion 36(11-12):969-73.
- Salmena L. 2014. Pseudogene redux with new biological significance. Methods Mol Biol 1167:3-13.

- Salomone OA, Basquiera AL, Sembaj A, Aguerri AM, Reyes ME, Omelianuk M, Fernandez RA, Enders J, Palma A, Barral JM, et al. 2003. *Trypanosoma cruzi* in persons without serologic evidence of disease, argentina. *Emerg Infect Dis* 9(12):1558-62.
- Sanchez-Moreno M, Fernandez-Becerra MC, Castilla-Calvente JJ, Osuna A. 1995. Metabolic studies by <sup>1</sup>H NMR of different forms of *Trypanosoma cruzi* as obtained by 'in vitro' culture. *FEMS Microbiol Lett* 133(1-2):119-25.
- Sant'Anna C, Nakayasu ES, Pereira MG, Lourenco D, de Souza W, Almeida IC, Cunha-E-Silva NL. 2009. Subcellular proteomics of *Trypanosoma cruzi* reservosomes. *Proteomics* 9(7):1782-94.
- Scharfstein J, Schmitz V, Morandi V, Capella MM, Lima AP, Morrot A, Juliano L, Muller-Esterl W. 2000. Host cell invasion by *Trypanosoma cruzi* is potentiated by activation of bradykinin B(2) receptors. *J Exp Med* 192(9):1289-300.
- Schenkman S, Robbins ES, Nussenzweig V. 1991. Attachment of *Trypanosoma cruzi* to mammalian cells requires parasite energy, and invasion can be independent of the target cell cytoskeleton. *Infect Immun* 59(2):645-54.
- Schijman AG, Bisio M, Orellana L, Sued M, Duffy T, Mejia Jaramillo AM, Cura C, Auter F, Veron V, Qvarnstrom Y, et al. 2011. International study to evaluate PCR methods for detection of *Trypanosoma cruzi* DNA in blood samples from chagas disease patients. *PLoS Negl Trop Dis* 5(1):e931.
- Schmunis G. 2013. Status of and cost of chagas disease worldwide. *Lancet Infect Dis* 13(4):283-4.
- Schmunis GA. 2007. Epidemiology of chagas disease in non-endemic countries: The role of international migration. *Mem Inst Oswaldo Cruz* 102 Suppl 1:75-85.
- Schmunis GA and Yadon ZE. 2009. Chagas disease: A latin american health problem becoming a world health problem. *Acta Trop* .
- Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH image to ImageJ: 25 years of image analysis. *Nat Methods* 9(7):671-5.
- Schneider P, Bell AS, Sim DG, O'Donnell AJ, Blanford S, Paaijmans KP, Read AF, Reece SE. 2012. Virulence, drug sensitivity and transmission success in the rodent malaria, *Plasmodium chabaudi*. *Proc Biol Sci* 279(1747):4677-85.
- Seco-Hidalgo V, Osuna A, Pablos LM. 2015. To bet or not to bet: Deciphering cell to cell variation in protozoan infections. *Trends Parasitol* .
- Seco-Hidalgo V, De Pablos LM, Osuna A. 2015a. Transcriptional and phenotypical heterogeneity of *Trypanosoma cruzi* cell populations. *Open Biol* 5(12):10.1098/rsob.150190.

- Seco-Hidalgo V, De Pablos LM, Osuna A. 2015b. Heterogeneity of *Trypanosoma cruzi* cell populations: Clonal variation in the expression of the MASP multigene family. Genbank .
- Siegel TN, Hekstra DR, Wang X, Dewell S, Cross GA. 2010. Genome-wide analysis of mRNA abundance in two life-cycle stages of *Trypanosoma brucei* and identification of splicing and polyadenylation sites. *Nucleic Acids Res* 38(15):4946-57.
- Silva DC, Silva RC, Ferreira RC, Briones MR. 2013. Examining marginal sequence similarities between bacterial type III secretion system components and *Trypanosoma cruzi* surface proteins: Horizontal gene transfer or convergent evolution? *Front Genet* 4:143.
- Simons AM. 2011. Modes of response to environmental change and the elusive empirical evidence for bet hedging. *Proc Biol Sci* 278(1712):1601-9.
- Smircich P, Eastman G, Bispo S, Duhagon MA, Guerra-Slompo EP, Garat B, Goldenberg S, Munroe DJ, Dallagiovanna B, Holetz F, et al. 2015. Ribosome profiling reveals translation control as a key mechanism generating differential gene expression in *Trypanosoma cruzi*. *BMC Genomics* 16:443,015-1563-8.
- Soares TS, Buarque DS, Queiroz BR, Gomes CM, Braz GR, Araujo RN, Pereira MH, Guarneri AA, Tanaka AS. 2015. A kazal-type inhibitor is modulated by *Trypanosoma cruzi* to control microbiota inside the anterior midgut of *Rhodnius prolixus*. *Biochimie* 112:41-8.
- Souto RP and Zingales B. 1993. Sensitive detection and strain classification of *Trypanosoma cruzi* by amplification of a ribosomal RNA sequence. *Mol Biochem Parasitol* 62(1):45-52.
- Souto RP, Fernandes O, Macedo AM, Campbell DA, Zingales B. 1996. DNA markers define two major phylogenetic lineages of *Trypanosoma cruzi*. *Mol Biochem Parasitol* 83(2):141-52.
- Souza RT, Santos MR, Lima FM, El-Sayed NM, Myler PJ, Ruiz JC, da Silveira JF. 2007. New *Trypanosoma cruzi* repeated element that shows site specificity for insertion. *Eukaryot Cell* 6(7):1228-38.
- Souza RT, Lima FM, Barros RM, Cortez DR, Santos MF, Cordero EM, Ruiz JC, Goldenberg S, Teixeira MM, da Silveira JF. 2011. Genome size, karyotype polymorphism and chromosomal evolution in *Trypanosoma cruzi*. *PLoS One* 6(8):e23042.
- Souza W. 2009. Structural organization of *Trypanosoma cruzi*. *Mem Inst Oswaldo Cruz* 104 Suppl 1:89-100.
- Souza W. 2008. Electron microscopy of trypanosomes-a historical view. *Mem Inst Oswaldo Cruz* 103(4):313-25.
- Stanne TM and Rudenko G. 2010. Active VSG expression sites in *Trypanosoma brucei* are depleted of nucleosomes. *Eukaryot Cell* 9(1):136-47.

- Stuart K, Brun R, Croft S, Fairlamb A, Gurtler RE, McKerrow J, Reed S, Tarleton R. 2008. Kinetoplastids: Related protozoan pathogens, different diseases. *J Clin Invest* 118(4):1301-10.
- Stuart KD, Schnauffer A, Ernst NL, Panigrahi AK. 2005. Complex management: RNA editing in trypanosomes. *Trends Biochem Sci* 30(2):97-105.
- Sturm NR and Campbell DA. 2010. Alternative lifestyles: The population structure of *Trypanosoma cruzi*. *Acta Trop* 115(1-2):35-43.
- Sullivan DT, Starmer WT, Curtiss SW, Menotti-Raymond M, Yum J. 1994. Unusual molecular evolution of an adh pseudogene in drosophila. *Mol Biol Evol* 11(3):443-58.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30(12):2725-9.
- Tardieux I, Nathanson MH, Andrews NW. 1994. Role in host cell invasion of *Trypanosoma cruzi*-induced cytosolic-free Ca<sup>2+</sup> transients. *J Exp Med* 179(3):1017-22.
- Tardieux I, Webster P, Ravesloot J, Boron W, Lunn JA, Heuser JE, Andrews NW. 1992. Lysosome recruitment and fusion are early events required for trypanosome invasion of mammalian cells. *Cell* 71(7):1117-30.
- Tarleton RL. 2005. New approaches in vaccine development for parasitic infections. *Cell Microbiol* 7(10):1379-86.
- Tarleton RL, Schulz CL, Grogl M, Kuhn RE. 1984. Diagnosis of chagas' disease in humans using a biotin-3H-avidin radioimmunoassay. *Am J Trop Med Hyg* 33(1):34-40.
- Tarleton RL, Reithinger R, Urbina JA, Kitron U, Gurtler RE. 2007. The challenges of chagas disease-- grim outlook or glimmer of hope. *PLoS Med* 4(12):e332.
- Teive HA and Arruda WO. 2009. Did darwin have syncope caused by chagas' disease? *Arq Neuropsiquiatr* 67(2A):371.
- The World Health Organization. 2002a. Control of chagas disease. . Report nr 905. 1 p. .
- The World Health Organization. 2002b. Control of chagas disease. . Report nr 811. 38 p. .
- The World Health Organization. [Http://Www.who.int/tdr/diseases/chagas/direction.htm](http://www.who.int/tdr/diseases/chagas/direction.htm). .
- The World Health Organization report 2004. Changing history. 2004. World health organization. .
- Tibayrenc M. 1998. Genetic epidemiology of parasitic protozoa and other infectious agents: The need for an integrated approach. *Int J Parasitol* 28(1):85-104.



- Tibayrenc M, Kjellberg F, Ayala FJ. 1990. A clonal theory of parasitic protozoa: The population structures of *entamoeba*, *giardia*, *Leishmania*, *naegleria*, *Plasmodium*, *trichomonas*, and *Trypanosoma* and their medical and taxonomical consequences. *Proc Natl Acad Sci U S A* 87(7):2414-8.
- Tibayrenc M, Ward P, Moya A, Ayala FJ. 1986. Natural populations of *Trypanosoma cruzi*, the agent of chagas disease, have a complex multiclonal structure. *Proc Natl Acad Sci U S A* 83(1):115-9.
- Tilmon KJ. 2008. Specialization, speciation, and radiation: The evolutionary biology of herbivorous insects. University of California Press .
- Todorov AG, Einicker-Lamas M, de Castro SL, Oliveira MM, Guilherme A. 2000. Activation of host cell phosphatidylinositol 3-kinases by *Trypanosoma cruzi* infection. *J Biol Chem* 275(41):32182-6.
- Tomasini N and Diosque P. 2015. Evolution of *Trypanosoma cruzi*: Clarifying hybridisations, mitochondrial introgressions and phylogenetic relationships between major lineages. *Mem Inst Oswaldo Cruz* 110(3):403-13.
- Tomlinson S, Vandekerckhove F, Frevert U, Nussenzweig V. 1995. The induction of *Trypanosoma cruzi* trypomastigote to amastigote transformation by low pH. *Parasitology* 110 ( Pt 5)(Pt 5):547-54.
- Torres-Padilla ME and Chambers I. 2014. Transcription factor heterogeneity in pluripotent stem cells: A stochastic advantage. *Development* 141(11):2173-81.
- Towbin H, Staehelin T, Gordon J. 1992. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. 1979. *Biotechnology* 24:145-9.
- Toyooka Y, Shimosato D, Murakami K, Takahashi K, Niwa H. 2008. Identification and characterization of subpopulations in undifferentiated ES cell culture. *Development* 135(5):909-18.
- Tyler KM, Luxton GW, Applewhite DA, Murphy SC, Engman DM. 2005. Responsive microtubule dynamics promote cell invasion by *Trypanosoma cruzi*. *Cell Microbiol* 7(11):1579-91.
- Ubeda JM, Raymond F, Mukherjee A, Plourde M, Gingras H, Roy G, Lapointe A, Leprohon P, Papadopoulou B, Corbeil J, et al. 2014. Genome-wide stochastic adaptive DNA amplification at direct and inverted DNA repeats in the parasite *Leishmania*. *PLoS Biol* 12(5):e1001868.
- Umezawa ES, Nascimento MS, Kesper N,Jr, Coura JR, Borges-Pereira J, Junqueira AC, Camargo ME. 1996. Immunoblot assay using excreted-secreted antigens of *Trypanosoma cruzi* in serodiagnosis of congenital, acute, and chronic chagas' disease. *J Clin Microbiol* 34(9):2143-7.
- Urbina JA. 2009. Specific chemotherapy of chagas disease: Relevance, current limitations and new approaches. *Acta Trop* .
- Ursic-Bedoya RJ and Lowenberger CA. 2007. *Rhodnius prolixus*: Identification of immune-related genes up-regulated in response to pathogens and

- parasites using suppressive subtractive hybridization. *Dev Comp Immunol* 31(2):109-20.
- Vihinen M. 2014. Contribution of pseudogenes to sequence diversity. *Methods Mol Biol* 1167:15-24.
- Villalta F, Smith CM, Ruiz-Ruano A, Lima MF. 2001. A ligand that *Trypanosoma cruzi* uses to bind to mammalian cells to initiate infection. *FEBS Lett* 505(3):383-8.
- Viney M and Reece SE. 2013. Adaptive noise. *Proc Biol Sci* 280(1767):20131104.
- Wang Y, Waters J, Leung ML, Unruh A, Roh W, Shi X, Chen K, Scheet P, Vattathil S, Liang H, et al. 2014. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* 512(7513):155-60.
- Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25(9):1189-91.
- Weatherly DB, Boehlke C, Tarleton RL. 2009. Chromosome level assembly of the hybrid *Trypanosoma cruzi* genome. *BMC Genomics* 10:255,2164-10-255.
- Wen YZ, Zheng LL, Qu LH, Ayala FJ, Lun ZR. 2012. Pseudogenes are not pseudo any more. *RNA Biol* 9(1):27-32.
- Wen YZ, Zheng LL, Liao JY, Wang MH, Wei Y, Guo XM, Qu LH, Ayala FJ, Lun ZR. 2011. Pseudogene-derived small interference RNAs regulate gene expression in african *Trypanosoma brucei*. *Proc Natl Acad Sci U S A* 108(20):8345-50.
- Westenberger SJ, Barnabe C, Campbell DA, Sturm NR. 2005. Two hybridization events define the population structure of *Trypanosoma cruzi*. *Genetics* 171(2):527-43.
- Wilkinson SR, Taylor MC, Horn D, Kelly JM, Cheeseman I. 2008. A mechanism for cross-resistance to nifurtimox and benznidazole in trypanosomes. *Proc Natl Acad Sci U S A* 105(13):5022-7.
- Wolfe AJ. 2005. The acetate switch. *Microbiol Mol Biol Rev* 69(1):12-50.
- Woodruff, A.W. 1965. Darwin's health in relation to his voyage to south america. *Br Med J* 1(5437):745-50.
- Woolsey AM and Burleigh BA. 2004. Host cell actin polymerization is required for cellular retention of *Trypanosoma cruzi* and early association with endosomal/lysosomal compartments. *Cell Microbiol* 6(9):829-38.
- Woolsey AM, Sunwoo L, Petersen CA, Brachmann SM, Cantley LC, Burleigh BA. 2003. Novel PI 3-kinase-dependent mechanisms of trypanosome invasion and vacuole maturation. *J Cell Sci* 116(Pt 17):3611-22.

- World Health Organization. 2015. Chagas disease in latin america: An epidemiological update based on 2010 estimates In: Weekly epidemiological record Geneva: . 33 p.
- Chagas disease: control and elimination. [Internet]; c2010 [cited 2015 . Available from: [http://apps.who.int/gb/ebwha/pdf\\_files/WHA63/A63\\_17-en.pdf](http://apps.who.int/gb/ebwha/pdf_files/WHA63/A63_17-en.pdf) .
- Xia X, Owen MS, Lee RE, Gaudet S. 2014. Cell-to-cell variability in cell death: Can systems biology help us make sense of it all? *Cell Death Dis* 5:e1261.
- Xong HV, Vanhamme L, Chamekh M, Chimfwembe CE, Van Den Abbeele J, Pays A, Van Meirvenne N, Hamers R, De Baetselier P, Pays E. 1998. A VSG expression site-associated gene confers resistance to human serum in *Trypanosoma rhodesiense*. *Cell* 95(6):839-46.
- Yadon ZE and Schmunis GA. 2009. Congenital chagas disease: Estimating the potential risk in the united states. *Am J Trop Med Hyg* 81(6):927-33.
- Yano Y, Saito R, Yoshida N, Yoshiki A, Wynshaw-Boris A, Tomita M, Hirotsune S. 2004. A new role for expressed pseudogenes as ncRNA: Regulation of mRNA stability of its homologous coding gene. *J Mol Med (Berl)* 82(7):414-22.
- Zalzman M, Falco G, Sharova LV, Nishiyama A, Thomas M, Lee SL, Stagg CA, Hoang HG, Yang HT, Indig FE, et al. 2010. Zscan4 regulates telomere elongation and genomic stability in ES cells. *Nature* 464(7290):858-63.
- Zarate-Blades CR, Blades N, Nascimento MS, da Silveira JF, Umezawa ES. 2007. Diagnostic performance of tests based on *Trypanosoma cruzi* excreted-secreted antigens in an endemic area for chagas' disease in bolivia. *Diagn Microbiol Infect Dis* 57(2):229-32.
- Zhang J, Pontoppidan B, Xue J, Rask L, Meijer J. 2002. The third myrosinase gene TGG3 in arabidopsis thaliana is a pseudogene specifically expressed in stamen and petal. *Physiol Plant* 115(1):25-34.
- Zheng D, Zhang Z, Harrison PM, Karro J, Carriero N, Gerstein M. 2005. Integrated pseudogene annotation for human chromosome 22: Evidence for transcription. *J Mol Biol* 349(1):27-45.
- Zingales B, Miles MA, Campbell DA, Tibayrenc M, Macedo AM, Teixeira MM, Schijman AG, Llewellyn MS, Lages-Silva E, Machado CR, et al. 2012. The revised *Trypanosoma cruzi* subspecific nomenclature: Rationale, epidemiological relevance and research applications. *Infect Genet Evol* 12(2):240-53.
- Zingales B, Andrade SG, Briones MR, Campbell DA, Chiari E, Fernandes O, Guhl F, Lages-Silva E, Macedo AM, Machado CR, et al. 2009. A new consensus for *Trypanosoma cruzi* intraspecific nomenclature: Second revision meeting recommends TcI to TcVI. *Mem Inst Oswaldo Cruz* 104(7):1051-4.

Zou M, Baitei EY, Alzahrani AS, Al-Mohanna F, Farid NR, Meyer B, Shi Y. 2009. Oncogenic activation of MAP kinase by *BRAF* pseudogene in thyroid tumors. *Neoplasia* 11(1):57-65.



