



Escuela de Bio  
y Nanotecnologías  
EByN\_UNSAM

# Análisis molecular y funcional de las secuencias asociadas a telómeros de *Toxoplasma gondii*

Tesis de Doctorado en Biología Molecular y Biotecnología  
Laboratorio de Parasitología Molecular

**Susana Marisol Contreras**  
Licenciada en Genética

**Director: Dr. Sergio Oscar Angel**



Universidad  
Nacional  
de San Martín

I N T E C H

Instituto Tecnológico de Chascomús

Chascomús, Provincia de Buenos Aires, Argentina

**-2022-**

### ***Trabajo final de Tesis Doctoral***

Tesis para optar por el título de Doctora en Biología Molecular y Biotecnología de la Escuela de Bio y NanoTecnologías de la Universidad Nacional de San Martín.

Este trabajo de investigación fue financiado por el Consejo Nacional de Investigaciones Científicas y Técnicas, CONICET; y el Instituto Nacional de Salud de los Estados Unidos, NIH.

El desarrollo de esta tesis doctoral fue realizado con dirección del Dr. Sergio Oscar Angel en el Laboratorio de Parasitología Molecular del Instituto Tecnológico de Chascomús.

Chascomús, Buenos Aires, Argentina.

Atentamente.

Lic. Susana Marisol Contreras



Dr. Sergio O. Angel



**Dedicatoria**

*A mi abuela Claudina*

*Cuando lo fantástico se cuela en lo real nos muestra la otra cara de la realidad.*

*Lxs mejores alquimistas son lxs que se animan a mezclar, combinar y fundir lo real y lo surreal.*

*Son lxs que sueñan despiertxs. Son lxs que crean lo increíble.*

## **Agradecimientos**

*A las entidades que financiaron este proyecto, CONICET y NIH, como así también a la Universidad Nacional de San Martín.*

*A la Comisión de Doctorado y las juradas.*

*En especial a Sergio por hacer honor a su apellido. Por confiar, por creer y por crear.*

*Por la posibilidad.*

*Al INTECH por abrirme sus puertas y dar alojamiento; por hacerme parte de su comunidad.*

*A Alexandra Elbakyan por su página Sci-Hub gracias a la cual el acceso a la información científica para dar discusión y fundamento a esta tesis fue factible.*

*A las bases de datos que procesan y disponen la información.*

*A los compañeros del lab. Gracias por la disponibilidad, la ayuda y los momentos de felicidad. En especial a Silvina Solange, por ser la "Gurú" y por enseñarme que siempre hay que soñar.*

*A los compañeros de la UB6 por estar. A María Corvi por su buena voluntad.*

*A la Dra. Josefina Ocampo por la gran ayuda.*

*Al señor Eduardo por instalar todos los programas bien intencionados que permitieron los datos analizar.*

*A mis guías que de alguna u otra manera estuvieron presentes para enseñarme a tener paciencia y templanza.*

*A todos y cada uno de las personas que acompañaron esta vuelta de consciencia en el pensar, en el sentir y el accionar.*

*A mi familia por recordarme la simplicidad en el amar. Por ser y estar.*

*A les amigues, a l@s de acá, a l@s de allá. A l@s que están cerca y a l@s de más allá... A l@s de ahora, a los de antes, a todes los que me regalaron su cariño y familiaridad. A cada ser que compartió el tan preciado tiempo. Gracias por invertir en abrazos y caricias al alma. Les amo, Marieli, Silvi, Ane, Romi, Fran, Vale, Eli, Andre, Paulita, Nadia, Pablo, Nati C., Dami, Nati O., Santi, Maxi, Sabri, Meli, Luis, Agus, Mapi, Nico, Gabi, Javi, Mati, Chris, Ima, Lucre, Sole, Ema, Juampi. Gracias.*

*A los asaditos y a lxs compañeres de la resi.*

*A Graciela, por las enseñanzas a través de la práctica del Yoga. A Sole, por introducirme en los saberes ancestrales astrológicos. A Fede por regalarme el libro.*

*A las vecindades, Beatriz, Fernando, Wally. A Marcela Ferraro por compartir su casa y a Ana y Eze por habitarla juntas.*

*A Blanca Rosa y a la Abuela Chuli por el hogar y el amparo.*

*A India por las pautas de Improvisación, a Gaviota por proveer el sagrado alimento sanador.*

*Al círculo de canto por recibirme, por su medicina por y para la expansión.*

*A Jilda. Por enseñarme el amor incondicional.*

*A la fuerza creadora, por permitirme aprender cada día como ser humana.*

*¡Gracias, gracias, gracias!*

## **Resumen**

Las regiones subteloméricas o subtelómeros (STs) podrían tener un rol en la regulación de genes específicos de respuesta ambiental y patogenicidad; en la estructuración de la cromatina, la replicación del ADN y/o división celular. En este trabajo definimos los STs de todos los extremos cromosómicos del parásito de importancia para la salud humana y animal, *Toxoplasma gondii*. Se definieron en función de una densidad génica baja, enriquecimiento de marcas de histona de heterocromatina (H2A.X y H3.3) y la presencia de ADN satelital. La distribución de otras marcas epigénéticas: H3K4me3, H4K31me; y las variantes H2A.X/H2B.Z nos permitió detectar posibles genes activos, ambiguos o silenciados, como también inferir posibles genes ARN largos no codificantes (lncRNA). Generamos un mapa de genes STs, algunos de ellos genes de copia única funcionales anotados y otros hipotéticos. También identificamos a las familias multigénicas de proteínas integrales de membrana FamB y FamC. Los genes FamC mostraron expresarse, al menos uno con posible localización apical y al menos 2 proteínas FamC resultaron levemente antigénicas con sueros de ratones e individuos infectados. En otras especies, el mantenimiento y silenciamiento de los STs están ligados a la actividad de enzimas deacetilasas del tipo sirtuinas como Sir2. Para indagar sobre la modulación de la cromatina en los STs, analizamos los efectos de 2 fármacos, un activador (Resveratrol –RSV-) y un inhibidor de sirtuinas (sirtinol). Se encontró que RSV y sirtinol tendrían un efecto inhibitorio sobre la replicación de *T. gondii*, que RSV reduciría los niveles de acetilación de H3 y H4K16 e incrementaría la fosforilación de H2A.X, una marca de daño de ADN. El tratamiento con sirtinol no mostró un efecto claro sobre las marcas de histonas. Nuestros hallazgos sugieren un posible vínculo entre el activador y el daño o los procesos de reparación del ADN. El análisis de localización subcelular de Sir2a y Sir2b de *T. gondii* mostró que serían citoplasmática y nuclear respectivamente. A futuro se analizará si están asociadas a los STs y si modulan la heterocromatina en *T. gondii*.

**Palabras claves:** *Toxoplasma gondii*, telómeros, subtelómeros, cromatina, familias multigénicas, tratamiento, histona deacetilasa.

## **Abstract**

*Subtelomeres or subtelomeric regions (STs) would be involved in the regulation of specific genes of environmental response and pathogenicity. They would also participate in structuring chromatin, DNA replication, and cell division mechanisms. Here, we define STs at all chromosomal ends of Toxoplasma gondii, a parasite of great importance for human and animal health. We define STs based on low gene density, enrichment of specific heterochromatic marks (H2A.X and H3.3), and the presence of satellite DNA. The distribution of other epigenetic marks: H3K4me3, H4K31me; and the variants H2A.X/H2B.Z allowed us to detect possible active, ambiguous, or silenced genes, as well as, to infer putative long non-coding RNA (lncRNA). We generated a map of STs genes, some of them annotated functional single-copy genes and others hypothetical. We also identified the multigene families of integral membrane proteins FamB and FamC. The FamC genes showed expression, one with a possible apical location, and at least 2 FamC proteins were slightly antigenic with sera from mice and infected individuals. The maintenance and silencing of STs in other species are linked to the activity of sirtuin-type deacetylase enzymes such as Sir2. To investigate the modulation of chromatin in STs, we analyzed the effects of 2 drugs, an activator (Resveratrol -RSV-) and a sirtuin inhibitor (sirtinol). The RSV and sirtinol showed an inhibitory effect on T. gondii replication. The RSV reduced H3 and H4K16 acetylation levels and increased phosphorylation of H2A.X (a marker of DNA damage). Sirtinol treatment effect on histone marks was not evident. Our findings suggest a possible link between the activator and DNA damage or repair processes. Subcellular localization analysis of Sir2a and Sir2b from T. gondii showed that they are cytoplasmic and nuclear, respectively. Future experiments will show if they are associated with STs and if they modulate heterochromatin in T. gondii.*

**Keywords:** *Toxoplasma gondii, telomeres, subtelomeres, chromatin, multigenic families, treatment, histone deacetylase.*

## **Publicaciones**

Los resultados de esta tesis han sido presentados en:

### *Revistas científicas y capítulo de libro:*

✓ Contreras SM, Zambrano Siri RT, Rivera EM, Cristaldi C., Kamenetzky L., Kim K., Clemente M., Ocampo J., Vanagas L., Angel SO. Architecture, chromatin and gene organization of *Toxoplasma gondii* subtelomeres. Under review. Epigenomes. 2022, 14, x. <https://doi.org/10.3390/xxxxx>.

✓ Contreras SM, Ganuza A, Corvi MM, Angel SO. Resveratrol induces H3 and H4K16 deacetylation and H2A.X phosphorylation in *Toxoplasma gondii*. BMC Research Notes 14:19. 2021 January 7. DOI: 10.1186/s13104-020-05416-4.

✓ Laura Vanagas; Susana Marisol Contreras; Sergio O. Angel. Apicomplexan and histone variants: What's new? Chromatin and Epigenetics. InTeChOpen. 2019 April 25. DOI: 10.5772/intechopen.81409.

### *Congresos y reuniones científicas:*

✓ Contreras SM, Ocampo J, Rivera EM, Cristaldi C, Ganuza A, Kamenetzky L, Clemente M, Vanagas L, Angel SO. Características únicas de los subtelómeros de *Toxoplasma gondii*. XI Congreso argentino de la Sociedad Argentina de Protozoología. Marzo 2022. Mendoza, Argentina.

✓ Susana M. Contreras, Elías M. Rivera, Laura Kamenetzky, Josefina Ocampo, Kami Kim, Laura Vanagas, Sergio O. Angel. *Characterization of subtelomeric regions and associated genes of Toxoplasma gondii*. XXXII Annual Molecular Parasitology Meeting, MPM. 5-9 Octubre, 2021. Woods Holes, Massachusetts, USA (presentación *on line*).

✓ Contreras SM, Angel SO. *The Telomeric Associated Sequences of Toxoplasma gondii-TgTAS*. MBL summer course Biology of Parasitism: Modern Approaches. June 17-August 6, 2016. Woods Holes, Massachusetts, USA.

✓ Contreras SM, Campo V, Vanagas L, Angel SO. *Effect of the activation of histone deacetylation by SIR2 on the lytic cycle of Toxoplasma gondii*. XXVII Reunion anual de Protozoología. Noviembre 2015. Buenos Aires, Argentina.

✓ Contreras SM, Dalmaso MC, Angel SO. *Screening of putative Toxoplasma gondii Telomeric Associated Sequences-Like (TgTASL) binding proteins*. X Congreso de la Sociedad Argentina de Parasitología. Noviembre 2014. Mar del Plata, Buenos Aires Argentina.

✓ Contreras SM, Dalmaso MC, Angel SO. *Identificación de las proteínas asociadas a las regiones subteloméricas de Toxoplasma gondii*. XV Reunión Anual de la Sociedad Argentina de Biología. Chascomús, Buenos Aires Argentina. Diciembre 2013.

## **Abreviaturas y palabras tomadas del inglés**

<b>aa</b>	aminoácidos
<b>ADN</b>	Ácido Desoxiribo Nucleico
<b>ADNc</b>	Ácido Desoxiribo Nucleico copia
<b>ADNg</b>	Ácido Desoxiribo Nucleico genómico
<b>APR</b>	<i>Del inglés: Apical Polar Ring</i>
<b>ApiAp2</b>	Apicomplexa Apetala 2, factor de transcripción de plantas
<b>ARN</b>	Ácido Ribo Nucleico
<b>ARNm</b>	Ácido Ribo Nucleico mensajero
<b>ATG</b>	Sitio de inicio de la transcripción codificante de la Metionina 1
<b>ATM</b>	<i>Del inglés: Ataxia Telangiectasia Mutated</i>
<b>BSA</b>	<i>Del inglés: Bovine Serum Albumin</i>
<b>CDK</b>	<i>Del inglés: Cyclin-Dependent Kinases</i>
<b>CenH3</b>	<i>Del inglés: Centromeric H3</i>
<b>CONICET</b>	Consejo Nacional de Investigaciones Científicas y Técnicas
<b>ChiP</b>	<i>Del inglés: Chromatin Immuno-Precipitation</i>
<b>ChiP-ChiP</b>	<i>Del inglés: ChiP DNA microarrays or ChiP on ChiP</i>
<b>ChiP-qPCR</b>	<i>Del inglés: ChiP-quantitative PCR</i>
<b>ChiP-Seq</b>	<i>Del inglés: ChiP-Sequencing</i>
<b>CRISPR/Cas9</b>	<i>Del inglés: Clustered Regularly Interspaced Palindromic Repeats/Caspase 9 system</i>
<b>DAPI</b>	4,6-diamidino 2-fenil-indol
<b>DDR</b>	<i>Del inglés: DNA Damage Response</i>
<b>DEA</b>	<i>Del inglés: Differential Expression Analysis</i>
<b>DMEM</b>	<i>Del inglés: Dulbecco's Modified Eagle's Medium, Gibco</i>
<b>DMSO</b>	Di-Metil-Sulfóxido
<b>DSB</b>	<i>Del inglés: Double-Strand Break</i>
<b>ELISA</b>	<i>Del inglés: Enzyme-Linked Immuno Sorbent Assay</i> o Ensayo por inmunoabsorción ligado a enzimas.
<b>FACS</b>	<i>Del inglés: Fluorescence-Activated Cell Sorting</i>
<b>famb</b>	familia multigénica b
<b>famc</b>	familia multigénica c
<b>FT</b>	Factor de Transcripción
<b>GO</b>	<i>Del inglés: Gene Ontology</i>

<b>h</b>	<i>humanx, de humanxs</i>
<b>H</b>	Histona
<b>HDAC</b>	Histona Deacetilasa
<b>Hh</b>	<i>Hammondia hammondi, H. hammondi</i>
<b>hhb</b>	<i>Del inglés: H. hammondi b family</i>
<b>hhc</b>	<i>Del inglés: H. hammondi c family</i>
<b>HMT</b>	Histona Metil Transferasa
<b>hp</b>	gen hipotético
<b>HP1</b>	<i>Del inglés: Heterochromatin Protein 1</i>
<b>HRP</b>	<i>Del inglés: Horse Radish Peroxidase</i>
<b>HRR</b>	<i>Del inglés: Homologous Recombinational Repair</i>
<b>hs</b>	horas
<b>H2A.X</b>	Histona 2A variante X
<b>H2A.Z</b>	Histona 2A variante Z
<b>H2Ba</b>	Histona 2B canónica
<b>H2B.Z</b>	Histona 2B variante Z, H2Bv
<b>Hv</b>	Histonas variantes
<b>H3.3</b>	Histona 3 variante 3
<b>H3K4me1/3</b>	Histona 3 mono/tri-metilada en la lisina (K) 4
<b>H3K9/14ac</b>	Histona 3 acetilada en la K 9 o 14
<b>H3K9me3</b>	Histona 3 tri-metilada en la K 9
<b>H4K20me3</b>	Histona 4 tri-metilada en la K 20
<b>H4K31ac</b>	Histona 4 acetilada en K 31
<b>H4K31me1</b>	Histona 4 mono-metilada en la K 31
<b>HFF</b>	<i>Del inglés: Human Foreskin Fibroblast, células fibroblásticas humanas</i>
<b>HRP</b>	<i>Del inglés</i>
<b>hSIR</b>	<i>Del inglés: human SIR2</i>
<b>Hst1p-Hst4p</b>	<i>Del inglés: Homologue of SIR two 1 to 4 protein</i>
<b>hTERT</b>	Células HFF con Telomerasa mutada
<b>HU</b>	Hidroxiurea
<b>hxpprt</b>	gen codificante de la enzima Hipoxantina-Xantina-Guanina-Fosforibosil-Transferasa
<b>IC<sub>50</sub></b>	<i>Del inglés: Inhibitory Concentration 50% of maximum</i>
<b>ID</b>	<i>Del inglés: Identification number, o número de identificación en base de datos</i>

<b>IFI</b>	Inmuno Fluorescencia Indirecta
<b>IMC</b>	<i>Del inglés: Inner Membrane Complex</i>
<b>IPTG</b>	Isopropil-β-D-1-tiogalactopiranosido
<b>ISP</b>	<i>Del inglés: IMC Sub-compartment Proteins</i>
<b>Kpb</b>	Kilopares de bases
<b>KDa</b>	Kilo Daltons
<b>lncRNA</b>	<i>Del inglés: long non-coding RNA</i>
<b>LOPIT</b>	<i>Del inglés: Localization of Organelle Proteins by Isotype Targeting</i>
<b>Mlp</b>	<i>Del inglés: Myosin like protein</i>
<b>M</b>	Molar
<b>mM</b>	milimolar
<b>ml</b>	mililitros
<b>mm</b>	milímetros
<b>MORN1</b>	<i>Del inglés: Membrane Occupation and Recognition Nexus protein 1</i>
<b>n</b>	número haploide
<b>Nc</b>	<i>Neospora caninum, N. caninum</i>
<b>ncb</b>	<i>Del inglés: N. caninum b family</i>
<b>ncc</b>	<i>Del inglés: N. caninum c family</i>
<b>NCBI</b>	<i>Del inglés: National Center for Biotechnology Information</i>
<b>NHEJ</b>	<i>Del inglés: Non-Homologous End Joining</i>
<b>NIH</b>	<i>Del inglés: National Institute of Health</i>
<b>NUCDISC</b>	<i>Del inglés: discrimination of nuclear localization signals</i>
<b>ORC1</b>	<i>Del inglés: Origin Recognition Complex 1</i>
<b>OTU</b>	<i>Del inglés: Ovarian Tumor</i>
<b>pb</b>	pares de bases
<b>PBS</b>	<i>Del inglés: Phosphate Buffer Saline</i>
<b>PCR</b>	<i>Del inglés: Polimerase Chain Reaction</i>
<b>PDVF</b>	<i>Del inglés: Polyvinylidene fluoride or polyvinylidene difluoride</i>
<b>PEV</b>	<i>Del inglés: Position Effect Variegation</i>
<b>Pf</b>	<i>Plasmodium falciparum, P. falciparum</i>
<b>PI</b>	Punto Isoeléctrico
<b>PIR</b>	<i>Del inglés: Plasmodium Interspersed Repeat superfamily</i>
<b>PLP</b>	<i>Del inglés: Piridoxal-5'-phosphate</i>

<b>POT1</b>	<i>Del inglés: Protection Of Telomeres 1</i>
<b>PTM</b>	<i>Del inglés: Post Traslational Modifications</i>
<b>qPCR</b>	<i>Del inglés: quantitative real time PCR</i>
<b>RAP1</b>	<i>Del inglés: Repressor Activator Protein</i>
<b>RC</b>	Reversa Complementaria, ADN
<b>RE</b>	Retículo Endoplasmático
<b>RFU</b>	<i>Del inglés: Relative Fluorescence Unit</i>
<b>RH_RFP</b>	<i>Del inglés: RH parasites with Red Fluorescent Protein</i>
<b>Rif1</b>	<i>Del inglés: RAP1-interacting factor 1</i>
<b>RISC</b>	<i>Del inglés: RNA-Induced Silencing Complexes</i>
<b>RNA</b>	<i>Del inglés: Ribo Nucleic Acid</i>
<b>rSIR2A</b>	Proteína recombinante de TgSIR2A
<b>rSIR2B</b>	Proteína recombinante de TgSIR2B
<b>RSV</b>	Resveratrol, (3,5,4'-trihidroxiestilbeno)
<b>RT-PCR</b>	<i>Del inglés: reverse transcription PCR</i>
<b>SAG1</b>	<i>Del inglés: Surface Antigen-1</i>
<b>SDS-PAGE</b>	<i>Del inglés: Sodium Dodecyl Sulfate-Poly Acrylamide Gel Electrophoresis</i>
<b>SFA</b>	<i>Del inglés: Striated Fiber Assembly</i>
<b>SIR2</b>	<i>Del inglés: Silent Information Regulator's 2, Sirtuin 2, Sirtuina</i>
<b>SIRT</b>	<i>Del inglés: SIR Two Inhibitor Naftol, Sirtinol</i>
<b>SMC1</b>	<i>Del inglés: Structural Maintenance of Chromosomes 1</i>
<b>SNC</b>	Sistema Nervioso Central
<b>SRS</b>	<i>Del inglés: SAG1-related sequences</i>
<b>ST</b>	Subtelómero, subtelomérico/a
<b>subTERRA</b>	lncRNA subtelomérico
<b>TAS</b>	<i>Del inglés: Telomeric Associated Sequences</i>
<b>TBS</b>	<i>Del inglés: Tris Buffer Saline</i>
<b>Tg</b>	<i>Toxoplasma gondii, T. gondii</i>
<b>tgb</b>	<i>Del inglés: T. gondii b family</i>
<b>tgc</b>	<i>Del inglés: T. gondii c family</i>
<b>TARE</b>	<i>Del inglés: Telomeric Associated Repetitive Elements</i>
<b>Taz1</b>	<i>Del inglés: Telomere-associated in Schizosaccharomyces pombe</i>
<b>TERRA</b>	<i>Del inglés: Telomeric Repeats containing lncRNAs</i>

<b>TERT</b>	<i>Del inglés: Telomerase Retro Transcriptase</i>
<b>TgIRE</b>	<i>Del inglés: Toxoplasma Interspersed Repetitive Element</i>
<b>TLA</b>	<i>Del inglés: Toxoplasma Lysate Antigen</i>
<b>TM</b>	Trans-Membrana
<b>TPE</b>	<i>Del inglés: Telomeric Possition Effect</i>
<b>TPP1</b>	<i>Del inglés: TIN2-interacting protein 1</i>
<b>TRF1/2</b>	<i>Del inglés: Telomeric Repeat binding Factors 1/2</i>
<b>Tsf</b>	<i>Del inglés: Toxoplasma specific family</i>
<b>TS</b>	Trans-Sialidasa
<b>TSS</b>	<i>Del inglés: Transcription Start Site</i>
<b>3'UTR</b>	<i>Del inglés: Three prime Untraslated Region</i>
<b>5'UTR</b>	<i>Del inglés: Five prime Untraslated Region</i>
<b>VP</b>	Vacuola Parasitófora
<b>VSG</b>	<i>Del inglés: Variable Surface Glicoprotein</i>
<b>WB</b>	<i>Del inglés: Western Blot, inmunoblot o electrotransferencia</i>
<b>Δ</b>	Delección de un gen, cepa mutante deleccionada
<b>μg</b>	Microgramos
<b>μm</b>	Micrometros
<b>μM</b>	Micromolar

## Índice

Análisis molecular y funcional de las secuencias asociadas a telómeros de <i>Toxoplasma gondii</i> .....	i
Trabajo final de Tesis Doctoral .....	ii
Dedicatoria .....	iii
Agradecimientos .....	v
Resumen .....	vii
Abstract .....	viii
Publicaciones .....	ix
Abreviaturas y palabras tomadas del inglés .....	xi
Índice .....	xvi
<b>INTRODUCCIÓN   TOXOPLASMA GONDII Y TOXOPLASMOSIS</b> .....	1
1. Generalidades.....	2
1.1 <i>Toxoplasma gondii</i> : un oportunista como modelo de estudio .....	2
1.2 La célula: el zoíto.....	5
1.3 La replicación celular .....	7
1.4 Ciclo celular y transcripción .....	11
1.5 Organización del genoma.....	13
1.6 Heterocromatina en <i>T. gondii</i> .....	15
<b>CAPÍTULO I   REGIONES SUBTELOMÉRICAS DE <i>T. GONDII</i></b> .....	19
2. Regiones asociadas al telómero .....	20

2.1	Secuencias asociadas a telómeros en <i>T. gondii</i> .....	21
2.2	Organización nucleosomal de la heterocromatina subtelomérica .....	24
3.	Hipótesis del capítulo I.....	28
4.	Objetivos del capítulo I.....	29
4.1	Objetivos específicos del capítulo I.....	29
5.	Resultados del capítulo I.....	30
5.1	Regiones teloméricas de <i>T. gondii</i> .....	30
5.2	Definición de regiones subteloméricas de <i>T. gondii</i> : TgSTs .....	31
5.2.1	Cromosoma Ia.....	34
5.2.2	Cromosoma Ib.....	36
5.2.3	Cromosoma II.....	38
5.2.4	Cromosoma III.....	40
5.2.5	Cromosoma IV .....	42
5.2.6	Cromosoma V .....	45
5.2.7	Cromosoma VI .....	47
5.2.8	Cromosoma VIIa .....	49
5.2.9	Cromosoma VIIb-VIII.....	51
5.2.10	Cromosoma IX .....	52
5.2.11	Cromosoma X .....	55
5.2.12	Cromosoma XI .....	57
5.2.13	Cromosoma XII .....	59
5.3	ADN satelital y subtelómeros.....	63

5.4 Análisis de la cromatina en las regiones TgSTs .....	64
5.4.1 Posicionamiento de histonas variantes .....	64
5.4.2 Posicionamiento de PTMs específicas .....	67
5.4.3 Detección de putativos lncRNAs STs .....	71
6. Discusión del capítulo I .....	76
7. Conclusiones del capítulo I .....	83
<b>CAPÍTULO II   GENES SUBTELOMÉRICOS DE <i>T. GONDII</i></b> .....	<b>84</b>
8. Familias de genes subteloméricos .....	85
9. Hipótesis del capítulo II .....	86
10. Objetivo del capítulo II .....	87
10.1. Objetivos específicos del capítulo II .....	87
11. Resultados del capítulo II .....	88
11.1 Descripción de los genes subteloméricos .....	88
11.1.1 Genes de copia única .....	88
11.1.1.1 Genes con función anotada .....	88
11.1.1.2 Genes hipotéticos .....	98
11.1.2 Familias multigénicas .....	103
11.1.2.1 Genes <i>famb</i> .....	105
11.1.2.2 Genes <i>famc</i> .....	107
11.2 Las proteínas FamB y FamC y la virulencia .....	114
11.3. Caracterización de la familia FamC en <i>T. gondii</i> : TgCs .....	119
11.3.1 Caracterización de la expresión génica <i>tgc</i> .....	121

11.3.1.1	Expresión en condiciones normales de crecimiento .....	121
11.3.1.2	Expresión génica en condiciones de estrés ambiental.....	123
11.3.2	Caracterización de la expresión proteica TgC .....	127
11.3.2.1	Producción de proteínas recombinantes .....	127
11.3.2.2	Producción de anticuerpos específicos.....	129
11.3.2.3	Inmunodetección en condiciones normales de crecimiento .....	130
11.3.2.4	Inmunolocalización subcelular .....	131
11.3.2.5	Sobreexpresión de TgC6 .....	132
11.3.2.6	Ensayos de inmunogenicidad .....	134
12.	Discusión del capítulo II .....	137
13.	Conclusiones del capítulo II .....	147
<b>CAPÍTULO III</b>	<b>  PROTEÍNAS INTERACTORAS DE SUBTELÓMEROS.....</b>	<b>148</b>
14.	Proteínas asociadas a telómeros y subtelómeros.....	149
15.	Control epigenético de la cromatina subtelomérica.....	153
16.	Drogas anti- deacetil transferasas .....	156
17.	Hipótesis del capítulo III .....	158
18.	Objetivo del capítulo III.....	159
18.1	Objetivos específicos del capítulo III.....	159
19.	Resultados del capítulo III.....	160
19.1	Búsqueda <i>in silico</i> de proteínas ortólogas descritas en otras especies .....	160
19.2	Efecto del activador e inhibidor de Sirtuinas en la replicación de <i>T. gondii</i> .....	162
19.2.1	Efecto del RVS sobre la replicación .....	162

19.2.1 Efecto de SIRT en la replicación .....	162
19.3 Efectos de activador e inhibidor de Sirtuinas sobre PTMs de <i>T. gondii</i> .....	163
19.3.1 Efecto de RSV sobre las acetilaciones en H3 y H4 .....	163
19.3.2 Efecto de SIRT sobre las acetilaciones en H3 y H4 .....	165
19.4 Efectos del activador e inhibidor de Sirtuinas sobre el ciclo celular .....	167
19.5 Efecto de RSV sobre el daño al ADN .....	170
19.6 Localización genómica y subcelular de TgSIR2A y TgSIR2B .....	173
19.6.1 Análisis del rol biológico .....	173
19.6.2 Análisis <i>in silico</i> de Sirtuinas .....	173
19.6.3 Producción de proteínas recombinantes, generación de anticuerpos específicos e inmunodetección .....	178
19.6.3.1 TgSIR2A recombinante (rSIR2A) .....	178
19.6.3.2 TgSIR2B recombinante (rTgSIR2B) .....	180
20. Discusión del capítulo III .....	183
21. Conclusiones del capítulo III .....	190
22. Conclusiones finales .....	191
23. Perspectivas futuras .....	194
<b>CAPÍTULO IV   METODOLOGÍA</b> .....	195
24. Análisis de la cromatina .....	196
25. Definición de LncRNAs .....	196
26. Análisis bioinformático .....	197
27. Análisis evolutivo .....	198

28. Cultivo de parásitos y reactivos .....	198
29. Análisis de la expresión de genes por RT-PCR y qPCR.....	202
30. Obtención de proteínas recombinantes .....	202
30.1 Clonado de proteínas .....	202
30.1.1. Proteínas TgCs .....	202
30.1.1.1 TgC3 .....	202
30.1.1.2 TgC5 .....	203
30.1.1.3 TgC6 .....	203
30.1.1.4 TgC8 .....	203
30.1.2 Sirtuinas .....	204
30.1.2.1 TgSIR2A .....	204
30.1.2.2 TgSIR2B .....	207
30.2. Expresión en bacterias .....	208
30.3. Purificación de proteínas recombinantes .....	208
31. Generación de anticuerpos policlonales .....	209
32. Sobreexpresión de proteínas en <i>Toxoplasma</i> .....	210
33. Transfección de parásitos .....	211
34. Ensayos de especificidad por WB.....	212
35. Ensayos de IFI .....	212
36. Titulación de los anticuerpos.....	213
37. Ensayo de inmunogenicidad .....	214
38. Ensayos de citotoxicidad y crecimiento con drogas .....	214

38.1 Ensayo de toxicidad celular .....	214
38.2 Ensayos de replicación intracelular .....	215
39. Medición del contenido de ADN por citometría de flujo.....	215
Bibliografía.....	217

## INTRODUCCIÓN | *TOXOPLASMA GONDII* Y TOXOPLASMOSIS

## 1. Generalidades

### 1.1 *Toxoplasma gondii*: un oportunista como modelo de estudio

*Toxoplasma gondii* (*T. gondii*) es un parásito oportunista intracelular obligado el cual es agente causal de la toxoplasmosis. Actualmente se estima que más de un tercio de la población mundial se encuentra infectada crónicamente, aunque la mayoría resulta asintomática (Tenter, Heckeroth, & Weiss 2000).

*T. gondii* presenta un ciclo de vida complejo (Dubey, Lindsay, & Speer 1998, Sibley et al. 2009), ya que puede reproducirse sexual y asexualmente (**Fig. 1**). Su ciclo de reproducción sexual se lleva a cabo en los felinos (hospedadores definitivos) mientras que su ciclo de reproducción asexual puede ocurrir en cualquier organismo de sangre caliente (hospedadores intermediarios), infectando a su vez, cualquier tipo de célula nucleada (Laliberte & Carruthers, 2008). La reproducción sexual tiene como resultado la liberación al ambiente de ooquistes (que contienen a los esporozoítos, formas unicelulares infectivas) presentes en las heces de los felinos (**Fig. 1 y Fig.2A**). Estos pueden ser ingeridos por el hospedador intermediario, como los humanos, en el consumo de verduras crudas mal lavadas o agua contaminada, reiniciando el ciclo asexual del parásito y constituyendo la primera vía de infección horizontal (Tenter, Heckeroth, & Weiss 2000). Cuando los hospedadores intermediarios ingieren los ooquistes, los esporozoítos se liberan del mismo en el intestino e infectan las células epiteliales. Subsecuentemente estos esporozoítos se desarrollarán en taquizoítos (la forma celular asexual de replicación activa) (**Fig. 1 y Fig.2A**), los cuales comenzarán a invadir rápidamente a las células hospedadoras. Este estadio de la infección se conoce como fase aguda, en la cual el parásito lleva a cabo varias rondas de invasión y replicación dentro de las células hospedadoras durante aproximadamente 14 días. En esta fase, los taquizoítos lisan y destruyen las células que los contienen para colonizar nuevas células vecinas, así como también infectar células dendríticas y macrófagos que colaboran en transportarlos a otras partes del cuerpo, accediendo a todos los tejidos del hospedador (Sibley et al. 2009). Esta fase de la infección puede presentar algunos síntomas febriles causados por la

destrucción de los tejidos por parte de los taquizoítos y termina cuando el sistema inmune logra controlar la replicación de los parásitos impulsándolos a diferenciarse al estadio de bradizoíto (Kim & Weiss 2008) (**Fig. 1 y Fig.2A**).

Los bradizoítos constituyen el estadio persistente de *Toxoplasma*, y llevan a la infección a un estadio más avanzado denominado fase crónica. Estos son de replicación lenta y se mantienen durante toda la vida del hospedador alojados preferentemente en el SNC y en el músculo esquelético. La persistencia de los bradizoítos de manera latente se ve favorecida debido a que las células del sistema inmune tienen muy poco acceso al SNC, resultando, al mismo tiempo, resistentes a las terapias actuales de tratamiento (Kim & Weiss 2008).

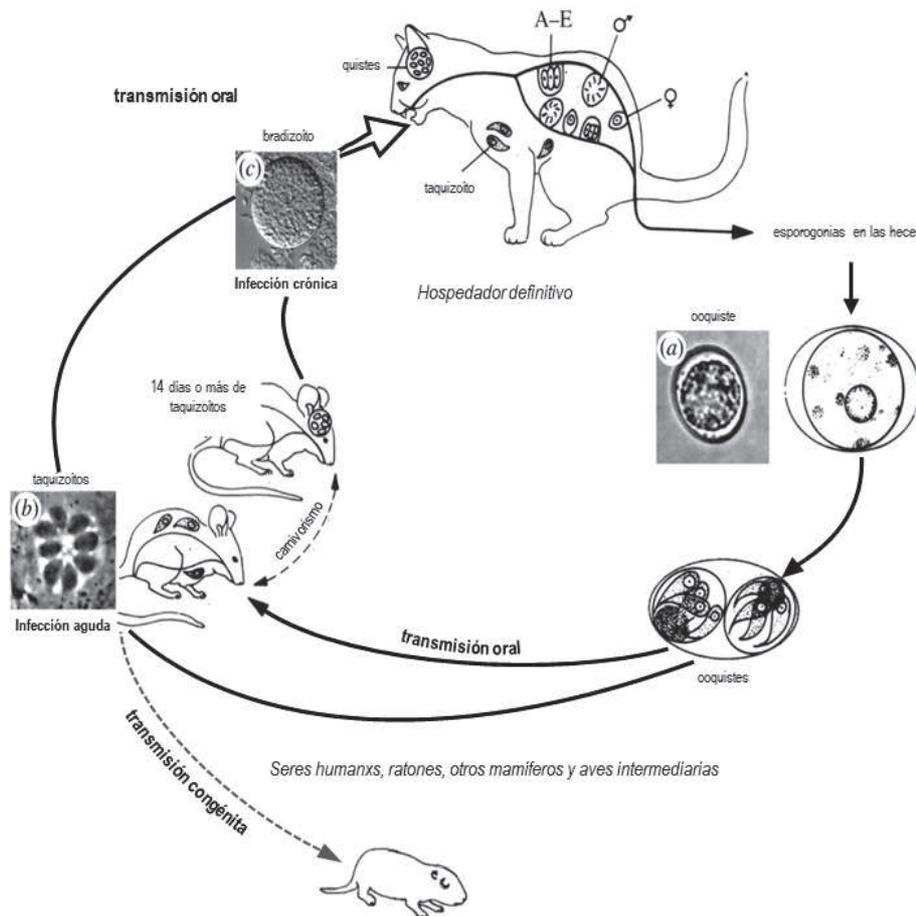
El consumo de tejidos infectados con bradizoítos (carne cruda o mal cocida) constituye la segunda vía de transmisión horizontal (Tenter, Heckeroth, & Weiss 2000). Dentro del hospedador, además, los bradizoítos pueden diferenciarse nuevamente en taquizoítos si dejara de existir la presión de la respuesta inmune, lo cual constituye la mayor causa de mortalidad y morbilidad asociada a toxoplasmosis en pacientes inmunocomprometidos asociados a HIV, transplantados, etc (Kim & Weiss 2008), en este último caso debido a que este parásito también puede transmitirse vía transfusiones de sangre y trasplante de órganos (Tenter, Heckeroth, & Weiss 2000).

La forma replicativa y activa del parásito puede causar encefalitis y enfermedades neurológicas, pero también puede afectar el corazón, el hígado, el oído interno y los ojos causando coriorretinitis (Weiss & Dubey 2009). El estadio latente, por su presencia en el SNC, se lo ha relacionado con afecciones como el cáncer de cerebro, trastornos de hiperactividad y déficit de atención, trastorno obsesivo-compulsivo y esquizofrenia (Brynska, Tomaszewicz-Libudzic, & Wolanczyk 2001; Yolken, Dickerson, & Fuller 2009; Miman et al. 2010; Vittecoq et al. 2012).

Finalmente, *Toxoplasma* puede transmitirse verticalmente durante la gestación estableciendo una infección congénita (Tenter, Heckeroth, & Weiss 2000) (**Fig. 1**). En este caso, los taquizoítos son capaces de atravesar la barrera placentaria e infectar al feto en desarrollo, causar

malformaciones e inclusive la muerte del embrión (Carlier et al. 2012). En el caso del humano, la mujer embarazada solo puede transmitírselo al feto si se infecta por primera vez durante el embarazo, ya que el establecimiento de una infección crónica está ligada a la ausencia de taquizoítos, pero también a la imposibilidad de re-infectarse (Kravetz & Federman 2005).

**Fig. 1. Ciclo de vida de *T. gondii*.** Los ciclos sexuales y asexuales del ciclo de vida son potencialmente independientes. La replicación sexual solo ocurre en el intestino de los felinos, lo que resulta en el desprendimiento de (a) ooquistes diploides que experimentan esporogonia (meiosis) en el ambiente para producir esporozoítos (haploides). A-E se refieren a las etapas de desarrollo dentro de los enterocitos. Los ooquistes son extremadamente resistentes en el ambiente y dan lugar a infecciones en una variedad de hospedadores intermediarios. La fase asexual puede ciclar teóricamente entre hospedadores intermediarios *ad infinitum*. Durante la infección aguda, los taquizoítos (b) proliferan rápidamente (endodiogenia) y se diseminan dentro del hospedador. Los quistes quiescentes en los tejidos los cuales contienen (c) bradizoítos dan lugar a una infección a largo plazo (crónica) y se transmiten asexualmente después de la ingestión oral por una variedad de hospedadores. Adaptado de (Sibley et al. 2009).



La toxoplasmosis se ha convertido en una enfermedad parasitaria desatendida y si bien existen algunos tratamientos que pueden ser aplicados durante la fase temprana de la infección, estos suelen tener efectos colaterales adversos y/o ser poco tolerados. Los tratamientos actuales de elección emplean pirimetamina y sulfadiazina, las cuales actúan en el estadio de taquizoíto inhibiendo el metabolismo del folato (Tenter, Heckeroth, & Weiss 2000). Por otra parte, existen cepas resistentes a estas drogas y el desarrollo de vacunas todavía no ha sido efectivo en humanos. Es por ello que es necesario indagar en posibles moléculas asociadas a mecanismos claves para el crecimiento y desarrollo del parásito dando lugar a la búsqueda de nuevos blancos terapéuticos que propicien tratamientos más eficientes contra la infección.

Por último, estudiar la toxoplasmosis a nivel molecular a través de los mecanismos que regulan la expresión génica y la replicación de los taquizoítos en modelos *in vitro*, constituye un aporte para el conocimiento, no solo para dilucidar la maquinaria de replicación y expresión génica en un modelo de fácil manipulación como lo es *Toxoplasma*, sino también a nivel de todo un Phylum. De hecho, *T. gondii* se propone como un buen modelo de estudio para todo el Phylum Apicomplexa al cual pertenece.

## **1.2 La célula: el zoíto**

El Phylum Apicomplexa comprende, además de *T. gondii*, otras especies de parásitos también de importancia para la salud humana y animal como *Plasmodium*, *Cryptosporidium* y *Babesia* (Tenter, Heckeroth, & Weiss 2000). A nivel celular, la mayoría de ellos presentan una organela particular denominada apicoplasto derivada de su antecesor, un alga flagelada, y un complejo apical también asociado a proteínas del mismo origen (Levine et al. 1980; Janouskovec et al. 2010; Sato 2011)(**Fig.2**). El complejo apical participa en la invasión, y la formación de parásitos hijos dentro de células hospedadoras (Morrisette 2015; Blader et al. 2015) (**Fig.2**).

Entre los Apicomplejos, *Toxoplasma* pertenece al grupo de los Coccidia, los cuales son formadores de quistes (Tenter & Johnson 1997). Junto a los Ciliados y Dinoflagelados, los Apicomplexa

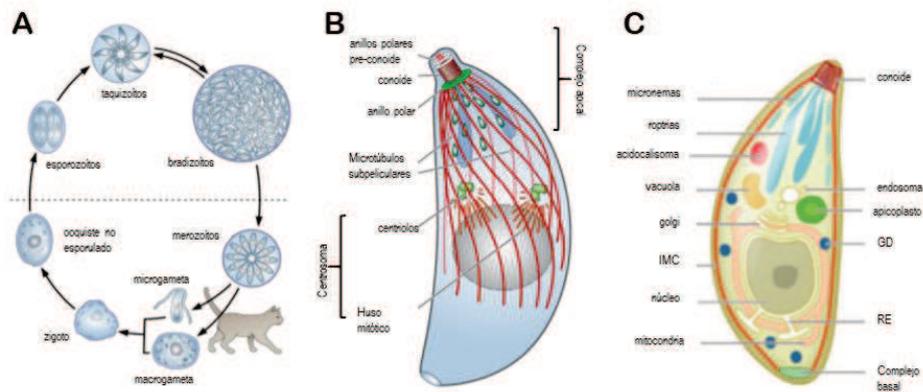
constituyen el Superphyllum Alveolata. Todos ellos comparten la presencia de un sistema de membranas alveolares periféricas internas que en los Apicomplexa se conoce como IMC (Blader et al. 2015) (**Fig.2**).

*T. gondii* tiene un ciclo de vida durante el cual se desarrolla en distintas formas que ocupan nichos específicos. La morfología del citoesqueleto de microtúbulos se conserva en gran medida en cuatro etapas celulares o de "zoíto" que comparten las propiedades de invasión y de replicación asexual: los taquizoítos, los esporozoítos, los merozoítos y los bradizoítos (**Fig.2A**). Este grupo de zoítos se caracteriza por la presencia de cinco estructuras distintas que contienen tubulina y componen el citoesqueleto: los microtúbulos subpeliculares, los husos mitóticos, los centriolos, el conoide y los microtúbulos intraconoide (Morrissette 2015) (**Fig.2B**).

El zoíto es una célula polarizada, en donde el citoesqueleto presenta estructuras especializadas en el extremo anterior y posterior conocidas como complejo apical y complejo basal respectivamente (Striepen et al. 2007) (**Fig.2C**). Este eje también polariza la motilidad celular, conduciendo al parásito dentro de las células hospedadoras a través de su ápice. En el extremo apical, el IMC está conectado al conoide (Hu, Roos, & Murray 2002), que protruye durante la invasión y la motilidad en un mecanismo dependiente de  $Ca^{+2}$ . El IMC también está asociado a los 22 microtúbulos subpeliculares que emanan del extremo apical hasta 2/3 de la longitud del parásito. Estos a su vez están unidos por el anillo apical que coordina la formación de los microtúbulos (Morrissette & Sibley 2002). Cabe destacar que el extremo apical es también el sitio para la secreción de organelas específicas de invasión como las roptrias y los micronemas que se encuentran empaquetadas en la parte anterior de la célula. El conoide es una estructura con forma de canasta en el ápice de la célula que está construido por láminas de tubulina curvadas en forma de cono hueco, es importante para la invasión de la célula hospedadora y se asocia con proteínas que se encuentran típicamente en el aparato flagelar (Hu, Roos, & Murray 2002; Morrissette 2015). El conoide junto a los microtúbulos subpeliculares (los cuales le dan forma y polaridad al zoíto, uniéndose además a los centriolos) y el anillo polar apical: APR, se asocian

en el extremo anterior de los parásitos formando el complejo apical (De Leon et al. 2013; Katris et al. 2014) (**Fig.2C**).

**Fig.2. Morfología general de *Toxoplasma*.** **A.** Estadios de zoíto en *Toxoplasma*. **B.** Organización del citoesqueleto de microtúbulos durante el ciclo de vida del zoíto. Adaptado de (Morrissette 2015). **C.** Morfología de las estructuras subcelulares. GD, granulo denso; RE, retículo endoplasmático; IMC, complejo de membranas internas. Adaptado de Blader et al. 2015 y Morrissette 2015.



### 1.3 La replicación celular

El resto de la célula contiene el núcleo y las herramientas para crecer y desarrollarse una vez adentrada en la célula hospedadora, incluida una única mitocondria, un plastidio (apicoplasto) - ambos con material genético propio- y el aparato de Golgi (Striepen et al. 2007). En el extremo opuesto, el IMC termina en un complejo basal; el cual es una estructura multimérica compuesta por la proteína nexo de reconocimiento y ocupación de la membrana MORN1 y un set de proteínas de IMC específicas (Gubbels et al. 2006; Lorestani et al. 2010), como así también la copa posterior, compuesta por la proteína Centrina 2 (Mann & Beckers 2001; Hu 2008). Este complejo basal funciona como un anillo contráctil para completar la división celular (Blader et al. 2015) (**Fig.2C**). Por otro lado, un centrosoma no canónico (centriolos paralelos sin la matriz pericentriolar) se asocian a un bolsillo de membrana especializado que alberga los polos del huso mitótico y el cinetocoro, conocido como centrocono, conectando así con el núcleo (Morrissette 2015; Francia, Dubremetz, & Morrissette 2016) (**Fig.2B**).

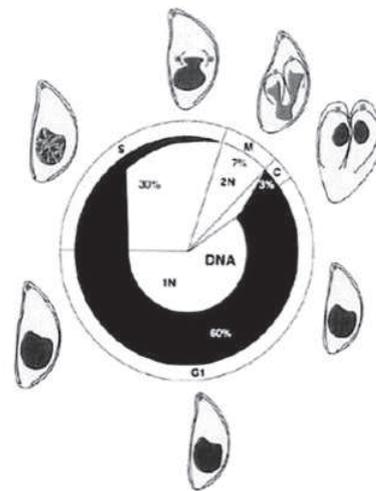
Los zoítos presentan dos estadios biológicos muy distintos cuando están dentro de la célula hospedadora (intracelulares) o fuera de ella (extracelulares). Los zoitos extracelulares no se dividen, son muy móviles, extruyen el conoide y secretan el contenido de los micronemas, estos tienen un único objetivo, censar, encontrar e invadir una nueva célula hospedadora. En cambio, los zoítos intracelulares se dividen activamente, no son motiles, no extruyen el conoide y no liberan el contenido de los micronemas (Striepen et al. 2007). El intercambio entre estos dos estadios es acompañado por cambios en los niveles de expresión génica (Lescault et al. 2010; Gaji et al. 2011), en la disponibilidad de ARN mensajero, ARNm, para la traducción (formación de gránulos de ARNm en taquizoítos extracelulares) (Lirussi & Matrajt 2011), y la translocación de enzimas glicolíticas desde el citoplasma a la película (microtúbulos subpeliculares e IMC) (Pomel, Luk, & Beckers 2008).

Uno de los pasos claves en la biología y desarrollo de este parásito es la replicación de los taquizoítos intracelulares. Una vez que los parásitos ingresan a la célula hospedadora comienzan a dividirse dentro de una estructura de membrana que se desarrolla alrededor de los mismos, denominada vacuola parasitófora (VP). Estos se replican por un proceso conocido como endodiogenia, en donde, la división nuclear ocurre por mitosis cerrada, es decir sin pérdida de la envoltura nuclear, dando cada taquizoíto origen a dos células hijas en el interior del mismo con la participación del IMC en el proceso (Gubbels, White, & Szatanek 2008). La mitosis ocurre en presencia de la envoltura nuclear y con una condensación de la cromatina poco apreciable (Francia et al. 2020). Si bien la división nuclear y la formación de células están coordinadas en el tiempo, son procesos separados físicamente dentro de la célula madre (Striepen et al. 2007). En los Apicomplexa, las células hijas no se generan por fisión, sino que se forman de nuevo en el citosol de la célula madre (Francia et al. 2020). En la mitosis de *T. gondii* todas las etapas de la mitosis eucariota canónica son observadas (es decir, metafase, anafase y telofase) (Francia et al. 2020).

Cada ciclo celular consta de una fase G1 (con un contenido de ADN haploide, 1N), seguida de una fase S (con un contenido de ADN diploide; 1,8N), y finalmente una mitosis que llega al contenido

1N previo al inicio de la citocinesis de las células hijas (Radke et al. 2001). La población con un contenido de ADN 2N es muy baja, indicando que no habría una fase G2. Esta característica del ciclo celular de los taquizoítos cambia frente a la presencia de un estrés, donde, los ciclos de división del parásito se enlentecen y el desarrollo de bradizoítos se hace evidente. En estos parásitos en proceso de diferenciación aparece una nueva subpoblación con un contenido de ADN nuclear de 1,8-2N, indicando la presencia de una fase S-tardía/G2 en los parásitos en transición, aunque el contenido de ADN final de los bradizoítos sería 1N en fase G0; sugiriendo que los parásitos con fase S tardía/G2 constituirían un estadio intermedio en la vía de desarrollo de los quistes tisulares (Radke et al. 2003) (Fig. 3).

**Fig. 3. Esquema del ciclo celular de un taquizoito intracelular.** Fases del ciclo celular: G1 (crecimiento, aproximadamente 5 h), S (síntesis, aproximadamente 2 h 30' minutos), mientras que la M (mitosis) y C (citocinesis) serían bastante rápidas en comparación, tardando 1 hora y 20' minutos aproximadamente. Se muestra además la morfología del parásito en cada fase. Se observa el porcentaje de células en cada fase y el contenido de ADN: 1N en G1; 1,8N en S y 2N en M (Adaptado de Radke et al. 2001).



Los mecanismos moleculares que subyacen a la progresión del ciclo celular están basados al igual que el resto de los eucariotas en un mecanismo dependiente de ciclinas y quinasas dependientes de ciclinas, CDK; aunque las CDK encontradas en *Toxoplasma* pertenecen a un tipo de quinasas diferentes de las CDK halladas en otros organismos. En total, se han descrito un total de 7 ciclinas atípicas y 10 proteínas con dominios relacionados a quinasas que regularían estos procesos celulares (Alvarez & Suvorova 2017).

Actualmente se sabe que *Toxoplasma* presenta un genoma haploide compuesto por 13 cromosomas (n=13) (Bunnik et al. 2019; Berná et al. 2021) que suman aproximadamente 62.4

megabases de ADN. Como se mencionó, la mitosis que se da es cerrada, es decir, que la segregación de los cromosomas duplicados ocurre dentro de los límites de la envoltura nuclear (Sheffield & Melton 1968), la cual nunca desaparece (Gubbels, White, & Szatanek 2008). El extremo del huso mitótico está incrustado en la envoltura nuclear en la estructura conocida como centrocono, que parece ser un poro nuclear modificado o especializado (*Fig. 4*). El centrocono, pero no los microtúbulos del huso mitótico, se mantiene durante todo el ciclo celular. Como en otros eucariotas, los cromosomas se unen a los microtúbulos del huso a través del complejo de proteína del cinetocoro ensamblado en el centrómero (Farrell & Gubbels 2014). Los centrómeros de todos los cromosomas se agrupan siempre, y se anclan al centrocono (*Fig. 4*) (Brooks et al. 2011), que, en la interfase, es independiente de la presencia de microtúbulos (Farrell & Gubbels 2014). La razón de este agrupamiento de los cromosomas por sus centrómeros en forma persistente no se conoce bien, pero podría ser relevante para mantener la integridad del genoma, “contar” los cromosomas y facilitar su segregación durante la citocinesis. El agrupamiento de los centrómeros no estaría mediado por los microtúbulos del huso mitótico sino por un componente de unión a la cromatina homólogo al factor eucariota de mantenimiento estructural de los cromosomas 1, SMC1. TgSMC1 junto a la histona variante centromérica de H3, CenH3, interactuarían con TgExportin1, un componente soluble del complejo de poros nucleares, sugiriendo que la envoltura nuclear, y, en particular, el complejo de poros nucleares podría desempeñar un papel en el posicionamiento de los centrómeros en *T. gondii* (Francia et al. 2020).

*Toxoplasma* comienza a ensamblar sus células hijas internamente en el citoplasma. El ciclo de división comienza con la duplicación del aparato de Golgi seguido de la duplicación del centrosoma (Hartmann et al. 2006). Dicha organela coordina la división nuclear y permanece asociada al centrocono (Francia & Striepen 2014). El andamiaje para el ensamblado del contorno de las células hijas nuevas lo proporcionan los microtúbulos subpelliculares y el IMC (Beck et al. 2010). Los microtúbulos emergen desde el anillo apical y se extienden a lo largo del eje longitudinal de la célula hasta contactar con el centrosoma. Por otra parte, el centrosoma se ancla al citoesqueleto en formación y a la base del conoide

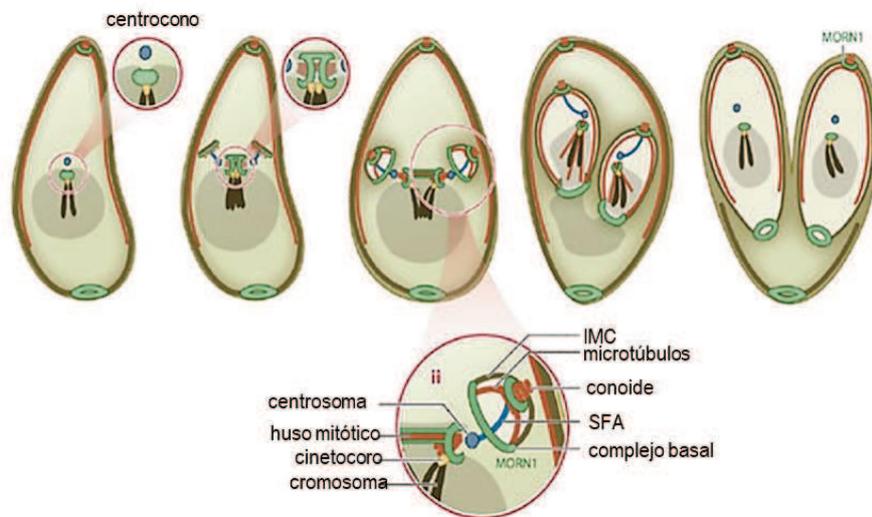
mediante la formación de una fibra de ensamblaje estriada compuesta por proteínas de origen flagelar denominada SFA (Chen & Gubbels 2013; Francia et al. 2012). Los siguientes pasos implican la incorporación de proteínas del citoesqueleto como MORN1 y las proteínas ISP. MORN1 se asocia con las vesículas alveolares tempranas y luego transiciona entre los complejos apical y basal (Gubbels et al. 2008; Hu 2008), mientras que ISP2 es esencial para el desarrollo temprano del IMC (Beck et al. 2010). A partir de aquí, el citoesqueleto crece desde el extremo apical mediante la deposición de varias proteínas IMC y el crecimiento de microtúbulos subpeliculares en forma de un cono ensanchado. Finalmente, los microtúbulos subpeliculares dejan de crecer cuando alcanzan 2/3 de la longitud máxima de la célula hija, coincidiendo con la finalización de la citocinesis, que también concluye la división del retículo endoplásmico (**Fig. 4**) (Chen & Gubbels 2013).

Particularmente, se ha visto que el núcleo presenta el mismo aspecto a lo largo del ciclo celular aunque este parecería no ser homogéneo, dado que la envoltura nuclear y los cromosomas tendrían una re-localización y/o rotación dinámica dentro del mismo durante la división, según lo observado en la localización de la epicromatina, la cual es un epítipo conformacional formado por ADN y las histonas H2A y H2B localizadas solo en la superficie exterior de la cromatina (Olins et al. 2010; Vanagas et al. 2013).

#### **1.4 Ciclo celular y transcripción**

La progresión de *Toxoplasma* durante su ciclo celular está caracterizada por dos subtranscriptomas bien definidos, uno correspondiente a G1 (donde se activan genes de biosíntesis y vías metabólicas) y otro que se corresponde a la fase S/M (donde se activan genes correspondientes a la replicación del ADN, el citoesqueleto y las organelas de invasión). De hecho, un tercio de los genes se expresan de manera bimodal a través del ciclo, controlados al parecer por una única familia de factores de transcripción (FTs) con dominios de unión a ADN de plantas denominados, ApiAp2 (Behnke et al. 2010; Radke et al. 2013).

**Fig. 4. Proceso de división celular.** (Arriba). La mitosis y la división celular están coordinadas por el centrosoma. A lo largo del ciclo celular, los cromosomas se agrupan por los centrómeros y se anclan al huso mitótico por el cinetocoro (centrocono). El ensamblaje de las células hijas nuevas comienza antes de la finalización de la mitosis y se controla mediante el ensamblaje del citoesqueleto cortical en una dirección apical-basal. El núcleo resulta anclado en los andamios de las células hijas a través del centrosoma y la fibra SFA. La constricción del complejo basal impulsa el estrechamiento de la célula hija hacia el extremo basal y estimula la separación de los parásitos hijos. (Abajo) Ampliación de la interacción entre el núcleo, el centrosoma y el citoesqueleto de la célula hija. Tomado de Blader et al. 2015.



Estos FTs son proteínas relacionadas a la familia de proteínas de unión al ADN reguladoras de la transcripción, *Ap2*, descubiertas en primer lugar en plantas. Las proteínas *Ap2* de plantas presentan un solo dominio *Ap2* si están asociadas a la respuesta ambiental a estrés (Ohme-Takagi & Shinshi 1995) o dos dominios *Ap2* en tándem si están asociadas a la regulación de la expresión génica durante el desarrollo (Jofuku et al. 1994); estos FTs no estarían presentes en animales (Painter, Campbell, & Llinás 2011).

*T. gondii* presenta 68 FT AP2s en su genoma (Altschul et al. 2010), de los cuales 44 fueron detectados por *microarrays* en taquizoítos, y 24 de ellos estarían regulados durante el ciclo celular, muchos serían constitutivos y un grupo pequeño sería específico de bradizoíto (Behnke et al. 2010). A la actualidad pocos FTs AP2s de *Toxoplasma* han sido caracterizados; sin embargo, hay trabajos que

han avanzado en elucidar el rol de estos en la activación génica estadio específica, como es el caso del control de la expresión de genes de bradizoíto y formación de quistes por parte de TgAP2XI-4 (Walker et al. 2013a); en la supresión de genes durante el desarrollo del taquizoíto, como en el caso de TgAP2XI-9, el cual es continuamente degradado actuando como represor de la formación de bradizoítos, censando tal vez cuando las condiciones sean las apropiadas para la interconversión (Radke et al. 2013) y finalmente en la regulación de genes específicos de virulencia durante la fase S y M del ciclo celular por parte de TgAP2XI-5 (Walker et al. 2013b).

### **1.5 Organización del genoma**

En los organismos eucariotas, el ADN genómico se envuelve alrededor de las histonas para conformar la cromatina. Los cromosomas son el resultado de la compactación extrema de la cromatina, es decir, del ADN asociado a proteínas. La cromatina puede presentarse en estados de menor o mayor compactación. El primer estado, más laxo, conocido como eucromatina, característica de regiones con abundancia de genes, se asocia a procesos de transcripción génica, replicación temprana del ADN y reparación del daño. Mientras que el estado más condensado que acompañaría los mecanismos de silenciamiento génico y de mantenimiento de la integridad cromosómica, se conoce como heterocromatina. En esta tesis nos referiremos como eucromatina a regiones con alta densidad de genes, similar al estudio de Brown et al. (2010) (Brown, Murray, & Verstrepen 2010). Las regiones de heterocromatina, a su vez pueden tener la característica de ser constitutivas o facultativas (Wang, Jia, & Jia 2016). A menudo las regiones heterocromáticas facultativas están asociadas a genes regulados por el desarrollo y su nivel de compactación cambiaría en respuesta a señales del mismo desarrollo, de estrés y/o ambientales (Trojer & Reinberg 2007). En contraste, la heterocromatina constitutiva se ensambla preferentemente en regiones pobres en genes y enriquecidas en elementos repetitivos, como el ADN satélite y los transposones, manteniendo niveles altos de compactación (Grewal & Jia 2007). A diferencia de los micro (<10 pares de bases -pb-) y mini (10-100 pb) satélites, que están presentes en la eucromatina, el ADN satelital es característico de la heterocromatina, y por lo tanto de las regiones

subteloméricas. Las familias de ADN satélite están organizadas en secuencias cortas (de 10 a unos miles de pb) altamente repetitivas en grupos en tándem mayores a 1000 pb (Garrido-Ramos 2017). Significativamente, estudios recientes demostraron que la heterocromatina constitutiva también estaría regulada dinámicamente y respondería a los estímulos. Si bien estos cambios podrían potencialmente ayudar a los organismos a adaptarse a nuevos entornos, en ciertos casos, como en el ser humano, estos cambios podrían causar ciertas enfermedades (Wang, Jia, & Jia 2016).

En general, la heterocromatina asociada a procesos biológicos como la división celular y/o el manteniendo integral, es un tipo de cromatina constitutiva, como lo son los centrómeros y telómeros. Mientras que la heterocromatina asociada a la regulación del silenciamiento génico es del tipo facultativa, es decir, que puede revertir rápidamente su conformación ante un estímulo. La heterocromatina constitutiva presenta además un conjunto de propiedades distintivas que la diferencia del resto del genoma: presenta replicación tardía durante la fase S, un nivel de recombinación meiótica muy reducida, inactivación en mosaico de la expresión de genes cuando se encuentran físicamente asociados, fenómeno denominado efecto de posición variegada, PEV, enriquecimiento de secuencias satélites altamente repetitivas, y remanentes de elementos transponibles (Dimitri et al. 2009).

Actualmente se sabe que, la eucromatina se encuentra asociada a histonas con marcas postraduccionales (PTMs) que inducen un estado hiperacetilado, mientras que la heterocromatina estaría asociada a histonas hipoacetiladas (Wang, Jia, & Jia 2016). Por otra parte, en la eucromatina es característico encontrar metilaciones sobre histonas del tipo H3K4me1/2/3. Se sabe además que las PTMs de histonas asociadas a la eucromatina ocupan estructuras diferentes, por ejemplo, la marca H3K4me3 está asociada a promotores, mientras que la marca H3K36me3 está asociada a las regiones codificantes (Martin & Zhang 2005; Bannister & Kouzarides 2011). Por otro lado, la heterocromatina constitutiva se encuentra generalmente en las regiones centroméricas, subteloméricas y teloméricas asociada a la marca H3K9me3 y a metilaciones en las citosinas del ADN, mientras que la heterocromatina facultativa estaría asociada a las marcas H3K27me2/3 (Beisel & Paro 2011).

En general, las secuencias subteloméricas asociadas a los telómeros han sido asociadas a un rol facultativo de la cromatina ligado a mecanismos epigenéticos de regulación de la expresión de genes. Dichos genes se encuentran físicamente comprendidos dentro de estas regiones y quedarían influenciados por el efecto de posición variegada o también conocido como “Efecto de Posición Telomérica: TPE”. El TPE es un efecto de silenciamiento de la actividad génica debido el estado de represión de la heterocromatina circundante. Este mecanismo epigenético es llevado a cabo por proteínas que introducen, reconocen, leen e implementan un estado represivo sobre la expresión génica bajo condiciones ambientales normales. En la mayoría de los casos, estos genes se corresponden con genes de respuesta a estrés ambiental y son especie específicos (Ottaviani, Gilson, & Magdinier 2008; Hernandez-Rivas et al. 2010). Por otra parte, se sabe que, en algunas especies, los genes subteloméricos reprimidos se activan en respuesta a condiciones de estrés ambiental o nutricional, promoviendo el crecimiento y la supervivencia de los organismos (Scherf, Figueiredo, & Freitas-junior 2001; Harari et al. 2013).

### **1.6 Heterocromatina en *T. gondii***

El centrómero se corresponde con un locus genómico donde las fibras del huso mitótico se anclan a los cromosomas para formar el cinetocoro (**Fig.2 y Fig. 4**), estructura requerida para la correcta segregación de los cromosomas durante la división celular (Verdaasdonk & Bloom 2011). Las secuencias de ADN centroméricas no son conservadas entre especies y difieren incluso entre cromosomas del mismo organismo. Sin embargo, la cromatina centromérica de todos los organismos estudiados hasta la fecha contienen la variante CenH3 (Verdaasdonk & Bloom 2011; Henikoff & Dalal 2005). Los centrómeros de *T. gondii* co-localizan con CenH3 en el centrocono (**Fig. 4**) en un extremo de la envoltura nuclear (Brooks et al. 2011) y con la proteína heterocromática de cromodominio HP1 (TgChromo1) (Gissot et al. 2012). Brooks y col., (2011), demostraron que TgCenH3 se localiza en los dominios centroméricos y es detectada como un único punto durante todo el ciclo celular cerca de la membrana nuclear (Francia et al. 2020). Esta se encuentra unida al centrocono; estableciéndolo como el

compartimiento subnuclear que serviría como organizador de la localización cromosómica durante el ciclo celular, anclando los cromosomas al mismo.

Por su parte, los telómeros son secuencias en los extremos cromosómicos que se encargan de mantener el estado del extremo final de las hebras de ADN. En la mayoría de los organismos son secuencias repetitivas que se asocian a proteínas formando estructuras específicas denominadas telosomas (Ottaviani, Gilson, & Magdinier 2008). En *T. gondii*, las secuencias teloméricas están conformadas por secuencias repetitivas del tipo TTTAGGG (Dalmasso et al. 2014) al igual que en plantas (Garrido-Ramos 2017) y si bien aún no se han descrito proteínas estructurales asociadas a las mismas, si se ha realizado una búsqueda y revisión en bases de datos de posibles proteínas homólogas descritas en otras especies (Vanagas, Contreras, & Angel 2019). A su vez, al igual que otras especies de levaduras y parásitos como *Plasmodium falciparum*, *T. gondii* posee regiones subteloméricas asociadas a los telómeros (TgTAS) en la región subsiguiente a dichos extremos cromosómicos (Dalmasso et al. 2014).

Durante millones de años los parásitos Apicomplexa han co-evolucionado con sus huéspedes y han desarrollado grandes familias multigénicas específicas de especies involucradas en las interacciones huésped-parásito (Reid et al. 2012). Estas familias de genes son importantes para la supervivencia de los mismos, la patogénesis, la virulencia y la evasión de la respuesta inmune. Todas las especies de *Plasmodium*, por ejemplo, contienen genes de virulencia que pertenecen a la superfamilia de repeticiones intercaladas de *Plasmodium*: PIR. *P. falciparum* y *P. knowlesi* han desarrollado familias únicas de genes que orquestan mecanismos de variación antigénica, conocidos como genes *var* y *SICAvar*, respectivamente presentes en regiones subteloméricas (Galinski et al. 2018; Deitsch and Dzikowski 2017). Durante el proceso de la variación antigénica, los parásitos pueden escapar de la respuesta inmune del huésped al cambiar qué miembros de estas grandes familias de antígenos se expresan en los parásitos. La capacidad de estos parásitos para cambiar su perfil antigénico se correlaciona con su alta virulencia y persistencia ante las respuestas inmunes adaptativas

del hospedador (Bunnik et al. 2019). En *T. gondii* también existen varias familias multigénicas que codifican para proteínas de secreción o membrana involucradas en adhesión celular e invasión (Ej. SRS, Rop, Mic, Gra) y las familias A, B, C, D y E (Lorenzi et al. 2016), de las cuales solo FamC se ubicó en regiones asociadas a los telómeros (Dalmasso et al. 2014).

Por otra parte, la heterocromatina subtelomérica también podría tener un rol en la compartimentalización, dinámica nuclear y proceso de división celular. Es importante mencionar que las distintas configuraciones conformacionales del estado de la cromatina no están libradas al azar dentro núcleo; sino que están asociadas a regiones nucleares específicas. Este posicionamiento de las regiones cromosómicas en el núcleo de las células eucariotas presenta una íntima relación asociada a la expresión génica. Tal es el caso de *P. falciparum*, en donde el agrupamiento de una familia específica de genes de virulencia en las regiones TAS se correlaciona con su silenciamiento coordinado influenciando toda la organización del genoma. De hecho, el mecanismo de variación antigénica está también asociado a un efecto en el plegamiento de los cromosomas, indicando una conexión entre la organización del genoma, la expresión génica y una mayor virulencia de estos patógenos (Bunnik et al. 2019). Las especies virulentas de *P. falciparum* y *P. knowlesi* mostraron agrupamiento de centrómeros, telómeros y genes de virulencia. Mientras que, el genoma de *T. gondii* solo mostró agrupamiento de centrómeros, sin asociación, hasta ahora, de agrupamientos de telómeros y familias de genes de virulencia (Bunnik et al. 2019). Cabe destacar que en *T. gondii* no se ha observado mecanismos de variación antigénica, al igual que otras especies de *Plasmodium* menos virulentas, aunque si posee familias de genes específicos que pueden estar involucradas en el escape de la respuesta inmune (Lekutis et al. 2001).

Bunnik y col., (2019) indican que los organismos que no experimentan variación antigénica podrían tener menos requisitos con respecto a la estructuración 3D de su genoma en el núcleo. Sin embargo, Gissot y col., (2012) encontraron que la proteína TgChromo1-la cual reconoce la tri-metilación en la (K) (lisina) 9 de la histona H3 (H3K9me3)- está presente solo en regiones heterocromáticas y co-

localizaría en la periferia nuclear con regiones subteloméricas y peri-centroméricas dando la pauta de que los subtelómeros podrían tener un rol en la compartimentalización del genoma de este parásito.

## CAPÍTULO I | REGIONES SUBTELOMÉRICAS DE *T. GONDII*

## 2. Regiones asociadas al telómero

Los subtelómeros (STs), regiones próximas o contiguas a los telómeros, son regiones de transición entre el telómero y la secuencia de ADN específica del cromosoma. Los STs son difíciles de definir, y su longitud varía desde los 20 Kpb en algunas cepas de levadura hasta 100 Kpb en varios eucariotas superiores (Mefford & Trask 2002; Brown, Murray, & Verstrepen 2010). El mayor problema para el ensamblaje de los STs es que a menudo poseen una alta densidad de secuencias repetidas en tándem de ADN satelital (Britten & Kohne 1968; Eichler 2001). Por lo tanto, los STs permanecen relativamente poco estudiados. Los STs a menudo se ven como loci variables que contienen familias de genes de evolución rápida involucradas en procesos adaptativos (Brown, Murray, & Verstrepen 2010). Además, algunos genomas de plantas como el maíz, *Arabidopsis* o el arroz contienen bloques de heterocromatina, conocidos como "protuberancias", que contienen ADN satélite muy repetido (McClintock 1929; Fransz et al. 2000; Cheng et al. 2001).

Las regiones subteloméricas han ganado reconocimiento de su importancia al poder ligarlas con diferentes funciones biológicas (Kwapisz & Morillon 2020). Están involucradas en (1) el mantenimiento de los telómeros y la regulación de la longitud de los telómeros, (2) la segregación cromosómica adecuada, (3) el reconocimiento y emparejamiento de los cromosomas durante la meiosis (4) la senescencia replicativa, (5) el posicionamiento nuclear de los telómeros, (6) el control de la propagación de la heterocromatina telomérica, (7) el polimorfismo fenotípico y plasticidad del genoma y, finalmente, en (8) la regulación de la transcripción de TERRA y subTERRA (Kwapisz et al. 2015), familias de ARN largos no codificantes, lncRNA de telómeros y subtelómeros respectivamente (Kwapisz & Morillon 2020).

En general, los subtelómeros se componen de dos regiones: una región próxima a los telómeros, también llamada TAS (por secuencias asociadas a los telómeros), asociada a elementos repetitivos y una región distal a los telómeros asociada a genes subteloméricos (Kwapisz & Morillon 2020). En general, son regiones pobres en genes, cuya expresión está bajo el efecto de posicionamiento

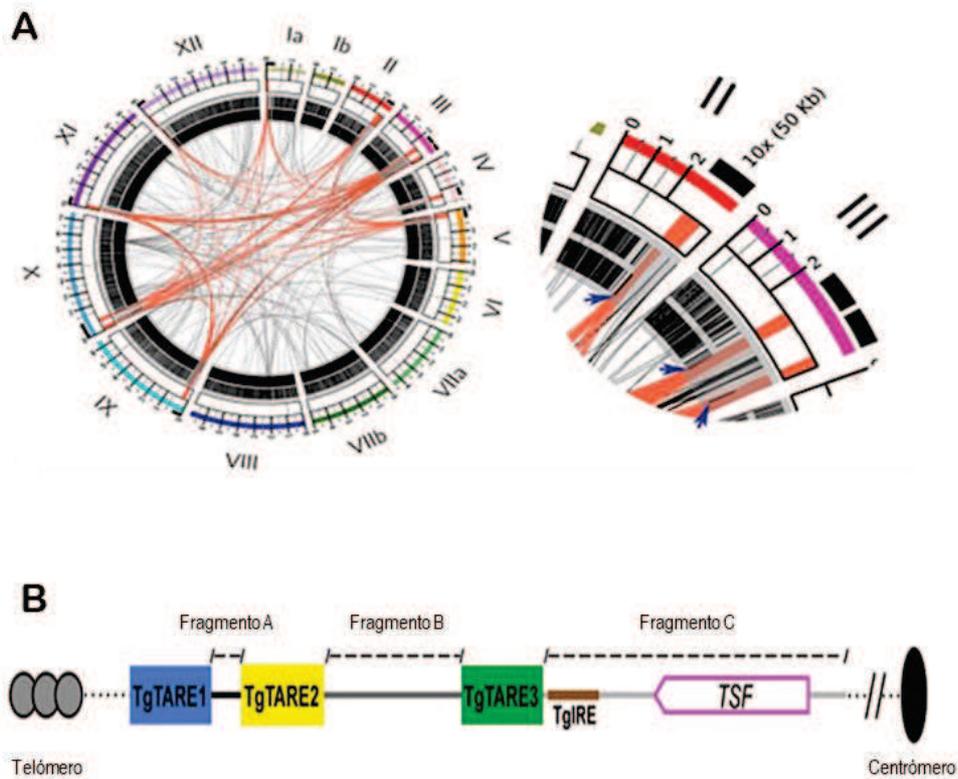
TPE arriba mencionado, el cual va disminuyendo su intensidad a medida que se aleja del telómero (Kwapisz & Morillon 2020). Los análisis transcripcionales han proporcionado evidencia con respecto a la actividad transcripcional de los telómeros y subtelómeros sugiriendo que la producción de lncRNAs sentido y antisentido es su característica común, pero no la única (Kwapisz & Morillon, 2020).

## **2.1 Secuencias asociadas a telómeros en *T. gondii***

Si bien *Toxoplasma* no posee mecanismos de variación antigénica si posee una estructuración del ADN adyacente a los telómeros, que fue inicialmente denominada como secuencias asociadas a telómeros de *T. gondii* (TgTAS). Esta estructuración se encontró en asociación a una familia específica de genes de *Toxoplasma*, denominada *Tsf* (Dalmasso et al. 2014) o FamC (Lorenzi et al. 2016).

En el trabajo de Dalmasso y col., (Dalmasso et al. 2014) las regiones TAS fueron identificadas en base a un estudio de genómica comparada entre todos los cromosomas, ya que las regiones subteloméricas o TAS presentan altos niveles de conservación inter-cromosómica (John et al. 1985). De esa manera se detectó la presencia de una región sinténica de entre 25-30 Kpb en uno de los extremos de 9 de los 14 cromosomas de *Toxoplasma*. En ese trabajo se encontró que los cromosomas Ia, II, IV, V, IX, XI y XII presentaban una única región TAS asociada, mientras que los cromosomas III y X contenían dos regiones TAS cada uno, habiéndose definido solo 9 regiones TAS de las 26 posibles (**Fig.5A**). Se pudo observar que las TgTAS estaban organizadas desde el telómero hacia el centrómero en tres regiones conteniendo elementos repetitivos en tándem denominados TARE 1, 2 y 3. Las unidades monoméricas de los TARE presentaron tamaños desde los ~350 pb a ~680 pb para TgTARE1, ~ 250–300 pb para TgTARE2, y de 133 a ~250–300 pb para TgTARE3, en donde todos se encontraron separados por tres fragmentos de ADN no repetitivo diferentes entre sí, denominados fragmentos A, B y C (**Fig.5B**). Cabe mencionar que TgTARE1 se correspondió con el elemento repetitivo de ADN satelital de la familia sat350/sat680 descrito previamente para *Toxoplasma* (Matrajt et al. 1999).

**Fig.5. Estructura de TgTAS.** **A.** Comparación de todos contra todos los cromosomas del genoma de la cepa ME49 (panel izquierdo). Los cromosomas se presentan como parte de un círculo, cada uno en un color diferente. La identidad de secuencia se representa con cintas, y las correspondientes a TgTAS se resaltan en rojo. Los cromosomas del genoma que comparten genes anotados (barras negras sobre los dos círculos concéntricos de fondo gris) además de TgTAS, están representados con cintas grises. Una inspección más cercana de las regiones TgTAS (panel derecho) muestra que éstas casi no poseen genes anotados con excepción de un solo gen perteneciente a la familia multigénica *tsf*, localizado en un extremo de la región (flechas azules). Se encontró además que cada TAS estaría asociado a un miembro distinto de la familia. **B.** Composición de la estructura interna de un TgTAS. Adaptado de Dalmaso et al. 2014.



Otro elemento de ADN repetitivo encontrado en asociación a TgTAS fue TgIRE (Echeverría et al. 2000; Dalmaso et al. 2006), ubicado contiguo al TgTARE3 dentro el fragmento C (**Fig.5B**).

Las regiones TgTAS se encontraron flanqueadas hacia el extremo centromérico (dentro del fragmento C) por la familia de genes *Tsf*, cuya función hasta el momento no ha sido descrita. Estos genes sólo se habían detectado en el genoma de *T. gondii* y de ahí su nomenclatura inicial descrita por Dalmaso y col., (2014). Luego, esta familia génica fue detectada en el genoma de *Hammondia hammondi* (*H. hammondi*), pero no en el genoma de la especie más relacionada *Neospora caninum* (*N.*

*caninum*) (Lorenzi et al. 2016). Por otra parte, a través de un análisis de CHIP-qPCR se demostró que las regiones TgTAS analizadas estaban enriquecidas en los marcadores de heterocromatina H4K20me1 y H2A.X, ambos asociados a promotores inactivos y/o silenciados (Dalmaso et al. 2014; Bogado et al. 2014).

Además de la familia C, *T. gondii* presenta otras familias multigénicas de genes que codifican para proteínas de función desconocida y con potencial de localizarse en la membrana, denominadas FamA, FamB, la mencionada familia FamC, FamD y FamE (Lorenzi et al. 2016).

La proteína TgChromo1 indicada anteriormente se encontró principalmente asociada a las regiones peri-centroméricas (Gissot et al. 2012), pero también estaría asociada a las repeticiones presentes al final del cromosoma Ia y IX, regiones del tipo subteloméricas descritas por Dalmaso y col., (2014). Curiosamente, la sonda del cromosoma Ia utilizada por Gissot y col., (2012); corresponde a TgTARE1, lo que sugiere que los elementos TgTAS podrían ubicarse en la periferia nuclear. En un conjunto separado de experimentos de CHIP-qPCR se detectó además un enriquecimiento de H4K20me1 y H3K9me1 asociadas con el elemento repetitivo de ADN satélite, sat350, el cual se encuentra presente en el TgTARE1 (Braun et al. 2010).

En un trabajo previo realizado por Sautel y col., (2007) se demostró que la histona Metiltransferasa SET8 mediaría la mono- y tri-metilación de las lisinas en la posición 20 de H4 en *Toxoplasma* y que tanto SET8 como H4K20me1 estarían reguladas durante el ciclo celular. En este trabajo, los experimentos de ChiP-Chip demostraron que tanto SET8 como H4K20me1 y H4K20me3 estarían asociadas a las regiones repetitivas teloméricas y al elemento repetitivo sat350 al igual que H3K9me3, pero en menor medida (Sautel et al. 2007).

Finalmente, Dalmaso y col., (2014) demostraron que los nucleosomas asociados a las regiones TgTAS, II, V y XII (regiones del fragmento A, cercanas al TgTARE1) estaban enriquecidas en ambos marcadores, H4K20me1 y H4K20me3 (Dalmaso et al. 2014), como así también encontraron

enriquecimiento de H4K20me1 en otras regiones repetitivas dentro del fragmento A de TgTAS río abajo del TgTARE1 (TgIRE100 y TgIRE600) y del fragmento C (TgIRE) (Dalmaso et al. 2009).

## **2.2 Organización nucleosomal de la heterocromatina subtelomérica**

La heterocromatina ha desarrollado formas específicas para compartimentalizar, silenciar, y reparar el ADN. El delicado equilibrio entre el mantenimiento epigenético de la heterocromatina y los procesos celulares como la mitosis, la reparación y replicación del ADN, revelan una gran dinámica y plasticidad de estos dominios; pero, que pueden ser perturbados por múltiples estímulos, con consecuencias de gran alcance para la integridad del genoma. De hecho, la disfunción de la heterocromatina provoca un “alboroto” genético tal, que induce una reparación aberrante del ADN, errores de segregación en los cromosomas, activación de transposones, y/o, estrés en la replicación, lo cual, en organismos superiores, está íntimamente implicado con el envejecimiento y la tumorigénesis (Janssen, Colmenares, & Karpen 2018).

En general, *T. gondii* presenta una composición nucleosomal conservada de las histonas canónicas H2A, H2B, H3 y H4. Estas histonas también presentan modificaciones o PTMs conservadas (Vanagas, Contreras, & Angel 2019). Las modificaciones PTMs de las histonas ocurren principalmente en las regiones N-terminales de las mismas, las cuales son flexibles y protruyen por fuera del nucleosoma. Estas incluyen acetilación, metilación, fosforilación, ubiquitinización, citrulinación, sumoilación y ADP-ribosilación (Fischle, Wang, & Allis 2003). La adición o eliminación de estos grupos específicos en las histonas puede alterar la asociación histona-ADN y servir como señal para ser reconocida por complejos modificadores o remodeladores de la cromatina que finalmente alteran la actividad transcripcional, replicación y reparación del ADN. La interconexión entre estas diferentes modificaciones da como resultado mecanismos dependientes del contexto a través de los cuales la expresión génica y las demás funciones genómicas estarían reguladas. Si bien es probable que la acetilación de las lisinas neutralice las cargas positivas del ADN y la fosforilación al agregar cargas negativas provoquen la descondensación de la cromatina, el código de marcas postraduccionales de las

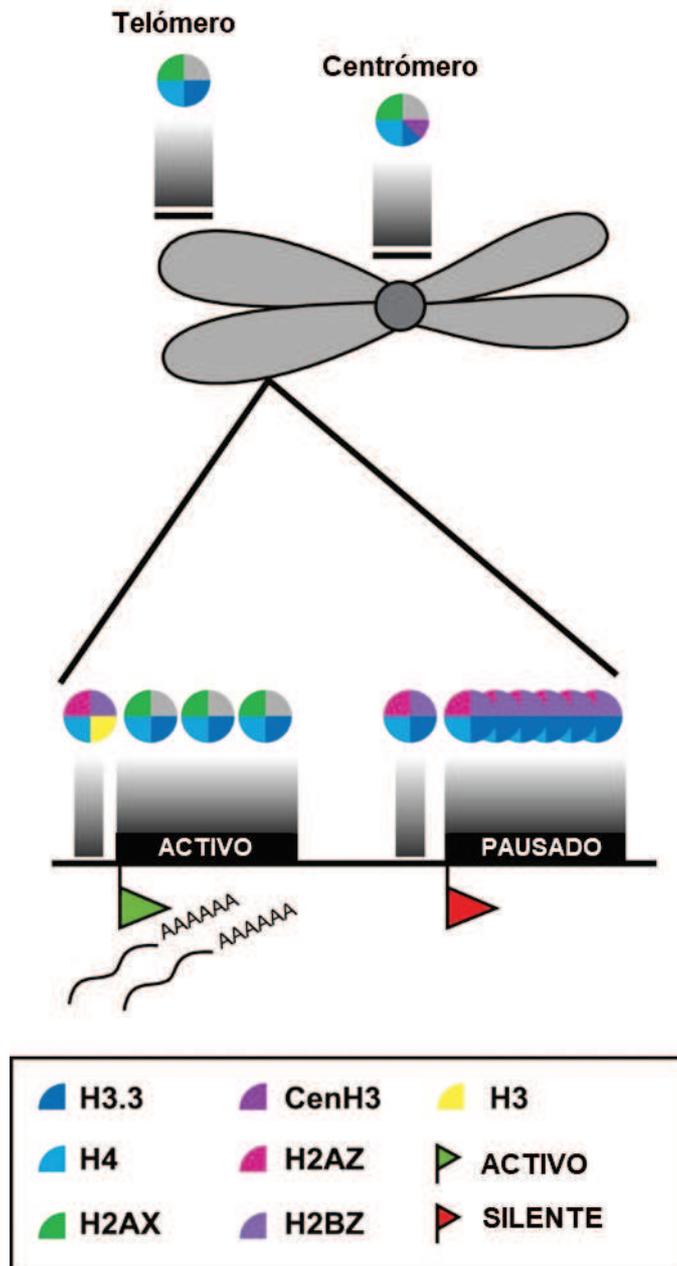
histonas es un mecanismo mucho más complejo (Fischle, Wang, & Allis 2003). Se acepta que las histonas hiperacetiladas se encuentran mayormente asociadas a regiones genómicas activas; mientras que la deacetilación principalmente resulta en represión y silenciamiento como ya se mencionó (Wang, Jia, & Jia 2016).

Uno de los descubrimientos más sorprendentes en los parásitos Apicomplexa fue la presencia de una nueva variante H2B, anteriormente denominada H2Bv (Dalmasso et al. 2006; Dalmasso et al. 2009; Dalmasso, Sullivan, & Angel 2011), pero reclasificada como H2B.Z (Talbert et al. 2012). *Toxoplasma* cuenta además con variantes de H2A, H2A.X que dimeriza con H2Ba canónica y una variante H2A.Z, la cual dimeriza con H2B.Z (Dalmasso et al. 2009; Bogado et al. 2014). Este hecho también se observó en *P. falciparum* (Petter et al. 2013), aunque este parásito carece de la histona variante H2A.X (Miao et al. 2006). *Toxoplasma* posee también variantes de H3: la variante centromérica ya mencionada, CenH3 y la variante H3.3 (Vanagas, Contreras, & Angel 2019). En un análisis genómico global reciente propuesto por Nardelli y col., (2022) detectaron que H3.3 y H2A.X localizan en telómeros, mientras que H3.3, H2A.X y CenH3 están enriquecidas en centrómeros. Las histonas H2A.Z y H2B.Z colocalizan con la marca de activación transcripcional H3K4me3 en regiones promotoras que rodean la región libre de nucleosomas aguas arriba del sitio de inicio de la transcripción, TSS incluyendo parte del 5'UTR. Las histonas H2A.Z/H2B.Z también se localizan en los cuerpos génicos de genes que están silenciados pero preparados para la activación, incluidos los genes específicos de la etapa de bradizoíto. La mayoría de los loci ocupados por H2A.X y H2A.Z/H2B.Z no se superponen, lo cual concuerda con que las variantes de histonas delimitan regiones funcionales específicas de la cromatina (Nardelli et al. 2022). En *P. falciparum*, se propuso que los nucleosomas que contienen H2A.Z demarcan las regiones intergénicas/reguladoras del genoma, que sirven como andamiaje para el reclutamiento de enzimas modificadoras de histonas (Bártfai et al. 2010) y que las modificaciones H3K9ac -acetilación en la (K) 9 de H3- y H3K4me3 -metilación en la (K) 4 de H3- se encuentran posicionadas preferentemente en o junto a los nucleosomas doble variante que contienen H2A.Z/H2B.Z (Salcedo-Amaya et al. 2009). Por

otra parte, se encontró que los genes subteloméricos silentes no estarían asociados a estos nucleosomas doble variante, pero si lo harían cuando estos se activan durante el proceso infectivo (Petter et al. 2011). Por lo tanto, en este contexto, la regulación epigenética se ha propuesto como un mecanismo clave en los procesos de regulación génica, replicación y reparación del ADN en organismos Apicomplexa. En la **Fig. 6** se propone una representación de la ocupación nucleosomal en la cromatina de *T. gondii* y un modelo de regulación génica a través del posicionamiento de las mismas propuesto por Nardelli y col., (2022).

La cromatina de las regiones subteloméricas, así como la regulación de los genes presentes en la misma, pueden poseer mecanismos diferentes a los observados para los genes localizados en el genoma propiamente dicho. La histona H2A.Z es una barrera para la dispersión de la heterocromatina (Meneghini, Wu, & Madhani 2003; Bártfai et al. 2010; Petter et al. 2011); y en este sentido, la localización y rol de las histonas H2A.Z y H2B.Z en las regiones subteloméricas de *T. gondii* permanece sin estudiarse, y, por lo tanto, se considera que el análisis de sus posicionamientos es un paso fundamental

**Fig. 6. Modelo de regulación génica específica asociada a variantes de histonas.** Se muestra un modelo de la ocupación nucleosomal en la cromatina de *T. gondii*. Las diferentes regiones de la cromatina tienen configuraciones distintas que cambian diferencialmente la arquitectura de la cromatina de acuerdo con los procesos relacionados al ADN. Los nucleosomas en las regiones típicas de heterocromatina, como los centrómeros, contienen la histona centromérica específica CenH3, junto con H2A.X. El enriquecimiento de H3.3 también es evidente en esas regiones, lo que sugiere un heterodímero atípico con CenH3 y H3.3 o nucleosomas alternativos que contienen CenH3 y H3.3. H2A.X probablemente forma un tetrámero con una H2B desconocida, probablemente H2Ba. La H2Ba canónica estaría ubicada junto a H2A.X en sitios opuestos al dímero H2A.Z/H2B.Z. H3.3 y H2A.X también están enriquecidos en los extremos de los cromosomas. Los sitios de inicio de la transcripción (TSS) de los genes activos se enriquecen con el dímero variante H2A.Z/H2B.Z. En el CDS de estos genes, H2A.Z/H2B.Z se sustituyen por H2A.X (con H2Ba), y la histona canónica H3 por H3.3. Los genes silenciados transcripcionalmente tienen una composición similar con H2A.Z, H2B.Z y H3.3 en el TSS. Pero esta configuración se extiende por todo el CDS de estos genes, especialmente en genes que son específicos para bradizoítos. Estas regiones no contienen H3K4me3, un punto de referencia para la transcripción activa, lo que sugiere que estos genes podrían estar equilibrados. H4 no tiene variantes, por lo que debería estar presente en todo el genoma. Tomado de Nardelli et al. 2022.



para ampliar el conocimiento de las regiones subteloméricas y su dinámica. Como así también aportar al entendimiento de los mecanismos adaptativos de *T. gondii*.

### 3. Hipótesis del capítulo I

Dado que los primeros estudios definieron regiones subteloméricas TgTAS para la mayoría de los cromosomas, pero no para todos, y únicamente asociadas a un solo extremo; sumada a la actualización constante del proyecto genoma de *T. gondii* (toxodb.org) es que se propone un re análisis de dichas regiones.

Una hipótesis de este capítulo se basa en que se disponen de nuevas herramientas para definir las regiones subteloméricas de los cromosomas de *T. gondii* aún no descritas o para redefinir las actualmente anotadas. Dada la particular composición de histonas presente en la cromatina de las regiones subteloméricas y la disponibilidad de nuevos estudios genómicos masivos de posicionamiento de histonas, consideramos que la localización de las mismas pueden ser una herramienta que contribuya para delimitar las regiones subteloméricas existentes e identificar nuevas.

Otra hipótesis de este capítulo es que las regiones subteloméricas están compuestas por ADN satelital y genes. Los STs son pobres en genes, pero aun así existen genes que tendrían patrones de expresión particular, generalmente condicionados a situaciones de estrés o con roles asociados a patogenicidad, como los genes antígeno variable de *Plasmodium* y *Trypanosoma*.

#### 4. Objetivos del capítulo I

Definir regiones subteloméricas de *T. gondii*, TgSTs.

Describir su composición génica e inferir la posible implicancia en la biología celular de este parásito.

##### 4.1 Objetivos específicos del capítulo I

- ✓ Recopilar datos de las regiones teloméricas y asociadas a telómeros descritas para *Toxoplasma* según versión actualizada del genoma en ToxoDB y bibliografía.
- ✓ Identificar y delimitar las regiones TgSTs en base a la densidad de genes subteloméricos, enriquecimiento de H2A.X, H3.3, y la presencia de ADN satelital.
- ✓ Analizar la organización de la cromatina en las regiones ST de *T. gondii*.

## 5. Resultados del capítulo I

### 5.1 Regiones teloméricas de *T. gondii*

En primer lugar, se realizó una recopilación de datos de las regiones heterocromáticas descritas para *Toxoplasma*. Las secuencias asociadas a telómeros (TgTAS) fueron descritas por primera vez por Dalmaso y col., (2014). En dicho estudio, solo se describieron 9 regiones TAS para uno de los dos extremos cromosómicos, derecho o izquierdo, asociadas a una única familia de genes *Tsf*. Los telómeros en *T. gondii* están compuestos por repeticiones simples del tipo TTTAGGG, los cuales solo se habían encontrado en algunos extremos cromosómicos (Dalmaso et al. 2014) y sus longitudes varían entre cepas (Xia et al. 2021). En este análisis se realizó una revisión genómica detallada de todos los extremos cromosómicos (izquierdo y derecho por convención según programa IGV, <https://software.broadinstitute.org/software/igv/>) corroborando la presencia de secuencias repetitivas en tándem TTTAGGG como factor determinante de la región telomérica (**Tabla 1**).

**Tabla 1. Regiones teloméricas en *T. gondii*.**

Entre paréntesis se muestra el lado del brazo cromosómico en el cual se encontraron (L): del inglés, *Left*, lado izquierdo; (R): del inglés, *Right*, lado derecho. El guion indica que no se encontraron secuencias correspondientes en la última versión revisada (versión 51).

Cromosoma	Telómero (TTTAGGG)	Cadena sentido	Tamaño (pb)
TGME49_chrla: 1-1,859,933	(R)	(+)	797
TGME49_chrlb: 1-1,955,354	(R)	(+)	966
TGME49_chrlI: 1-2,347,032	(R)	(+)	281
TGME49_chrIII: 1-2,532,871	(L)	(-)	766
	(R)	(+)	1107
TGME49_chrIV: 1-2,686,605	(L)	(-)	764
TGME49_chrV: 1-3,331,915	(L)	(-)	47
TGME49_chrVI: 1-3,656,745	-	-	-
TGME49_chrVIIa: 1-4,541,629	(R)	(+)	1078
TGME49_chrVIIb: 1-5,069,724	(L)	(-)	828
TGME49_chrVIII: 1-6,970,285	(R)	(+)	126
TGME49_chrIX: 1-6,327,655	(R)	(+)	769
TGME49_chrX: 1-7486,190	-	-	-
TGME49_chrXI: 1-6,623,461	(R)	(+)	1148
TGME49_chrXII: 1-7,094,428	(R)	(+)	1070

En este análisis se pudo observar la presencia de repeticiones TTTAGGG en los extremos del lado izquierdo solamente en los cromosomas III, IV, V y VIIb. Mientras que los cromosomas Ia, Ib, II, III, VIIa, VIII, IX, XI, y XII mostraron repeticiones TTTAGGG en los extremos derecho de los mismos (**Tabla 1**). En el caso del cromosoma II, las secuencias TTTAGGG se encontraron a una distancia ~1.4

Kpb río arriba del extremo del brazo derecho. Para los cromosomas VI y X no se encontraron repeticiones teloméricas en tándem (*Tabla 1*).

## 5.2 Definición de regiones subteloméricas de *T. gondii*: TgSTs

Las regiones subteloméricas son difíciles de definir debido a su variabilidad en tamaño entre cromosomas, aunque según el principio de equi-localidad, las regiones heterocromáticas tienden a ocupar las mismas posiciones en los distintos cromosomas (John et al. 1985). Son regiones pobres en genes y ricas en ADN satelital que evolucionan rápido (Kwapisz & Morillon 2020). A menudo comprenden grupos de genes asociados a una ruta metabólica o función biológica (Teunissen & Steensma 1995; Hasin et al. 2008; Brown, Murray, & Verstrepen 2010; Li et al. 2021) y están bajo el efecto TPE, el cual mantiene silenciada la cromatina, incluyendo los genes presentes en la misma (Gottschling et al. 1990). A su vez, la extraordinaria inestabilidad de los subtelómeros permite una rápida adaptación a nuevos nichos, resultando en una alta tasa de duplicación de genes, seguida de una divergencia funcional de los alelos (Brown, Murray, & Verstrepen 2010). Eso implica que también es común detectar familias multigénicas como en *Plasmodium* y *Trypanosoma* (Omelianczyk et al. 2020; Ramirez 2019).

En este trabajo, las regiones TAS han sido revisadas y redefinidas como subteloméricas: STs, sobre la base de que están asociadas a los extremos de los cromosomas, son pobres en genes y ricas en ADN satelital, (secuencias de ADN repetido en tándem de más de 1000 pb asociadas a la heterocromatina).

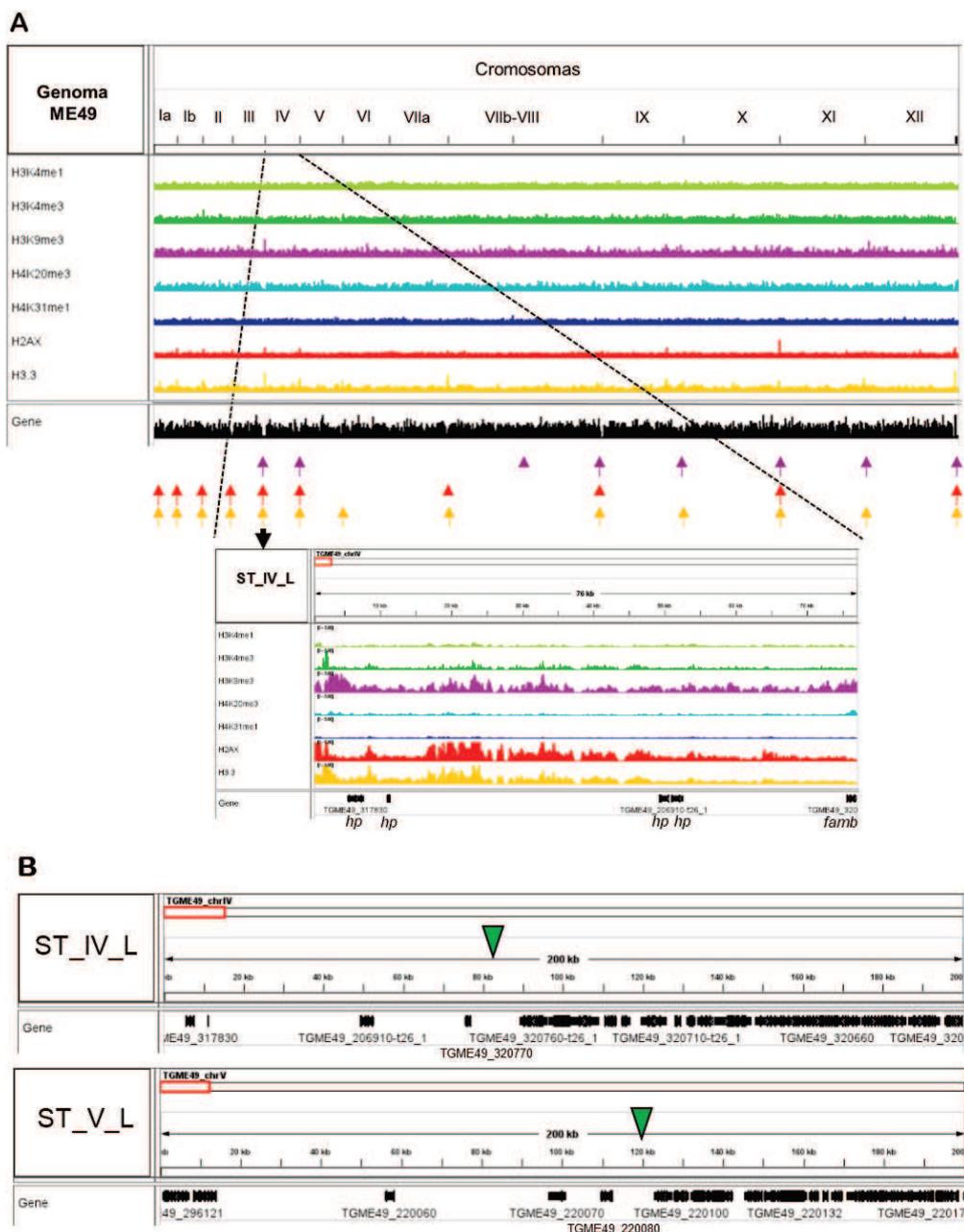
Dado que en *T. gondii* no se realizó un análisis masivo de marcas de heterocromatina como, por ejemplo, HP1; empleamos como marca de heterocromatina el enriquecimiento de la histona H2A.X en las regiones subteloméricas (Bogado et al. 2014). Por otra parte, según lo observado en el análisis genómico global de Nardelli y col., (2022), H2A.X de *T. gondii* está enriquecida en regiones heterocromáticas centrómeras junto a CenH3, teloméricas y secuencias asociadas a telómeros junto a H3.3. Se encuentra depletada, además, en los sitios de inicio de la transcripción (TSS) de regiones genómicas activas, y mantiene un enriquecimiento parcial a lo largo del resto del cromosoma mostrando

importantes picos de enriquecimiento en los extremos de algunos de ellos (Nardelli et al. 2022). Este mismo set de datos de ChiP-Seq se utilizó para abordar el análisis del enriquecimiento de H2A.X en los extremos cromosómicos (**Fig. 7A**). En esta instancia se incorporó al análisis la distribución y enriquecimiento de las marcas de heterocromatina H3K9me3 y H4K20me3, como así también las marcas propuestas por Sindikubwabo y col., (2017), H4K31me1 y H3K4me1 (que acompaña el enriquecimiento de H4K31me1); las cuales ocuparían regiones TSS/promotoras junto a H2A.X y H3.3 en genes silentes (**Fig. 7A**).

Por otro lado, los genes *Tsffamc* fueron propuestos como un límite de las regiones STs, según lo descrito para TgTAS anteriormente (Dalmaso et al. 2014). En el análisis genómico local de las regiones TgTAS y datos de ChiP-Seq identificamos otra familia de genes que estarían dentro de las regiones STs, denominada *famb*, la cual será descrita en el capítulo II. En esta sección asumimos por lo tanto que los genes de las familias *famb* y *famc/Tsf* están asociados a los STs de los cromosomas de *T. gondii*, empleándose en muchos casos como marcador del límite propuesto. Por otra parte, se contempló si las regiones TSS de los genes presentes en las regiones STs reanalizadas tenían marcas de cromatina activa (H3K4me3/H2A.Z/H2B.Z), silente (H2A.X) o ambigua (H2A.Z/H2B.Z) según lo reportado por Nardelli y col., (2022). En la eucromatina de *T. gondii* los genes poseen poca distancia entre sí, estando a una separación de solapamiento de 7000 pb, aunque la mayoría está a menos de 2000 pb (datos no mostrados). Para confirmar la región ST propuesta, nos extendimos 100 Kpb más allá del límite propuesto analizando que este tuviera densidad de genes compatible con eucromatina (**Fig. 7B**).

El análisis de enriquecimiento de H2A.X, otras histonas variantes (Hv), PTMs de histonas, y la presencia de genes se realizó mediante el programa IGV. El enriquecimiento de regiones de ADN satelital se realizó mediante el estudio de gráficas de puntos, Dot plot, alcanzando los 100 Kpb desde los extremos excepto en los casos donde el ST definido fue superior a ese largo. En esos casos se llegó hasta el límite propuesto.

**Fig. 7: Definición de las regiones subtelerómicas de *T. gondii* (TgSTs).** **A.** Posicionamiento cromosómico global de Hv y PTMs sobre el genoma de la cepa ME49 de *T. gondii*. Se muestran las localizaciones de las histonas y sus marcas en picos de colores a lo largo de cada cromosoma: H3K4me1 (verde claro), H3K4me3 (verde oscuro), H3K9me3 (violeta), H4K20me3 (celeste), H4K31me1 (azul), H2A.X (rojo), H3.3 (amarillo); en donde el enriquecimiento coincide con la acumulación de las mismas (flechas hacia arriba). Las barras inferiores de color negro se corresponden con la presencia de genes a lo largo de los cromosomas. Se muestra la versión 53 del genoma ME49. Abajo se muestra una ampliación del brazo izquierdo del cromosoma IV, ST\_IV\_L junto a los ID de genes en la región. **B.** Densidad génica en una ventana de hasta 200 Kbps, incluidas dos regiones ST. La flecha verde puntiaguda muestra la región límite ST, ST\_IV\_L: *tgb16* (TGME49\_320770). ST\_V\_L: *tgc8* (TGME49\_220080).



En este análisis se proponen nuevas regiones STs, una en el extremo del brazo izquierdo (ST\_Cromosoma\_L) y otra en el extremo del brazo derecho (ST\_Cromosoma\_R). Según lo revisado por Kwapisz & Morillon (2020) las regiones subteloméricas están compuestas de dos regiones: una proximal o asociada al telómero (TAS) y una distal del telómero. El inicio de cada región TgST estaría definido desde el telómero o extremo hacia el centrómero y el límite interno lo marcaría la presencia de algún gen desde donde comienza la eucromatina. En todos los casos se confirmó la región propuesta en base a la presencia de ADN satelital.

A continuación, se describen en detalle las características de cada región ST en cada extremo cromosómico en particular.

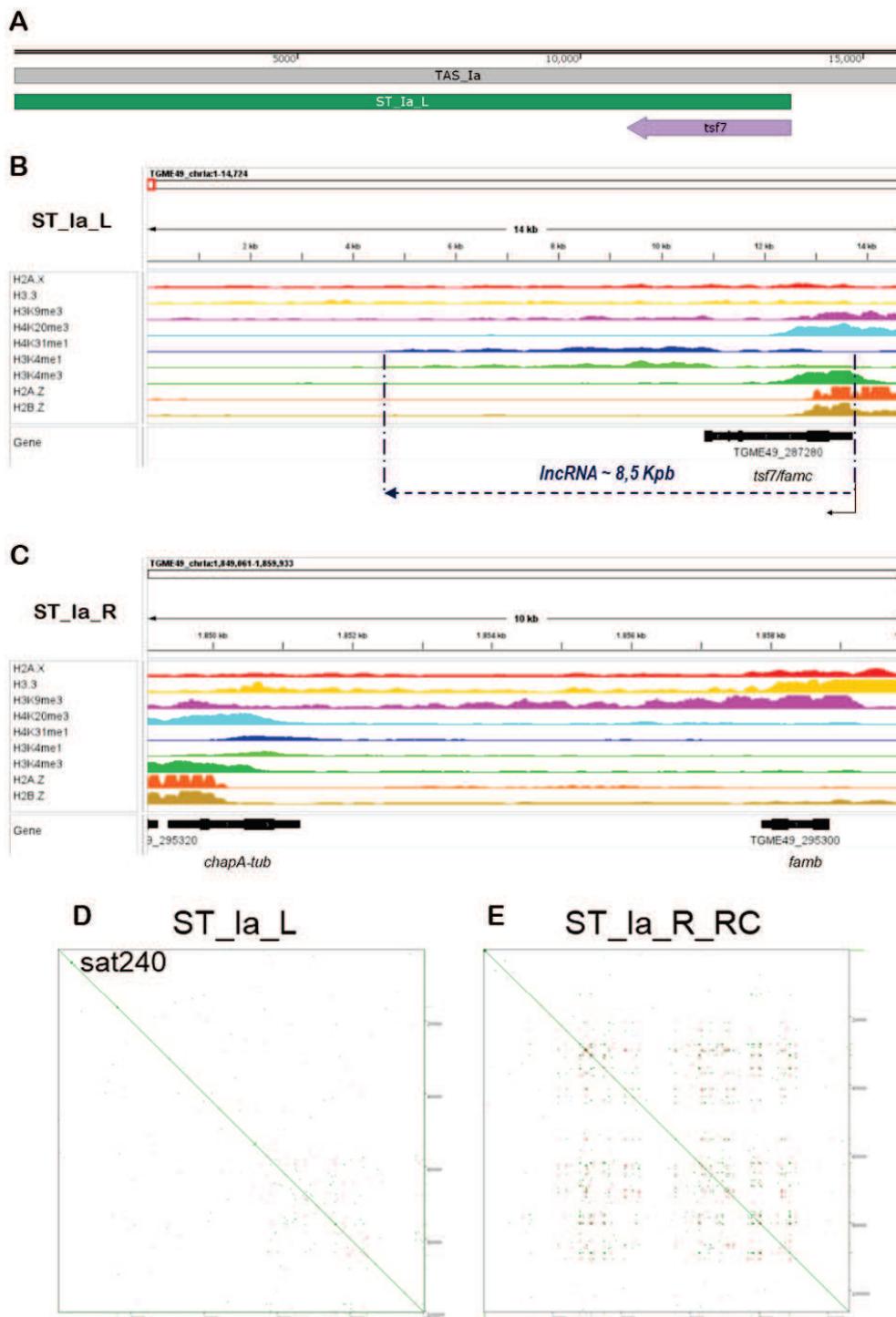
### **5.2.1 Cromosoma la**

La región ST\_la de ~15 Kpb definida para este cromosoma se hallaba descrita tomando el brazo derecho como inicio del mismo (Dalmasso et al. 2014), pero en este análisis lo hemos invertido por cuestiones de estructuración del genoma en el programa IGV, quedando el inicio (1) en el extremo izquierdo del cromosoma. Esta región fue asociada previamente a la presencia del gen *Tsf7/famc* (TGME49\_287280) en el límite interno de la misma (Dalmasso et al. 2014). En este análisis se redefinió (**Fig. 8A**).

*Brazo izquierdo:* tal como se visualiza en la **Fig. 8B**, esta región no presenta un enriquecimiento de H2A.X marcado. Basados en la presencia del gen *Tsf7* (TGME49\_287280), la región ST\_la\_L tendría una longitud de ~13,7 Kpb, abarcando hasta la región 5'UTR de dicho gen, cuya región TSS presenta marcas de activación (**Fig. 8B**).

*Brazo derecho:* se encontró una región de ~10 Kpb enriquecida en H2A.X en el extremo derecho delimitada por el gen codificante para la chaperona de tubulina, *chapA-tub* (TGME49\_295310) cuya región TSS se encuentra activa (**Fig. 8C**). Dicha región incluye además un miembro de la familia *famb* en estado silente (TGME49\_295300) (**Fig. 8C**).

**Fig. 8. Mapa cromosómico de la localización genómica de ST\_Ia\_L y ST\_Ia\_R. A.** TAS\_Ia (15,7 Kpb) en color gris (Dalmasso et al. 2014). En verde, ST\_Ia\_L. **B.** Enriquecimiento de Hv y PTMs en la región genómica ST\_Ia\_L. H2A.Z (naranja), H2B.Z (marrón). **C.** Enriquecimiento de Hv y PTMs en la región genómica ST\_Ia\_R. IncRNA putativo TSS<sup>+</sup> (ver sección 5.4.3). **D.** Mapeo de ADN satelital repetitivo en ST\_Ia\_L. **E.** Mapeo de ADN satelital repetitivo en ST\_Ia\_R. RC: Reversa Complementaria. Líneas verdes por fuera de la diagonal indican similitud directa, mientras que las líneas rojas indican similitud con la cadena RC.



El análisis de Dot plot de ST\_la\_L mostró la presencia de ADN repetitivo en tándem compatible con ADN satelital sat240 dentro los primeros 10 Kpb (*Fig. 8D*), mientras que ST\_la\_R (*Fig. 8E*) no mostró la presencia de ADN satelital dentro de los 100 Kpb abarcados en el extremo derecho cromosómico, aunque se observaron pequeños fragmentos de ADN repetido disperso (fragmentos de 20-30 pb) en forma directa (cadena sentido) e inversa (cadena reversa complementaria) a lo largo de los 100 Kpb de ST\_la\_R (*Fig. 8E*).

En conclusión, la región ST\_la\_L, tendría al menos un tamaño de ~13,7 Kpb, mientras que para el extremo en el brazo derecho se describe una nueva región ST\_la\_R de ~10 Kpb (*Tabla 2*).

### **5.2.2 Cromosoma Ib**

Este cromosoma no presentaba regiones TAS descritas por Dalmaso y col., (2014), por lo tanto, se analizaron ambos extremos cromosómicos en búsqueda de algún marcador que permitiera delimitar al menos parte de la región subtelomérica.

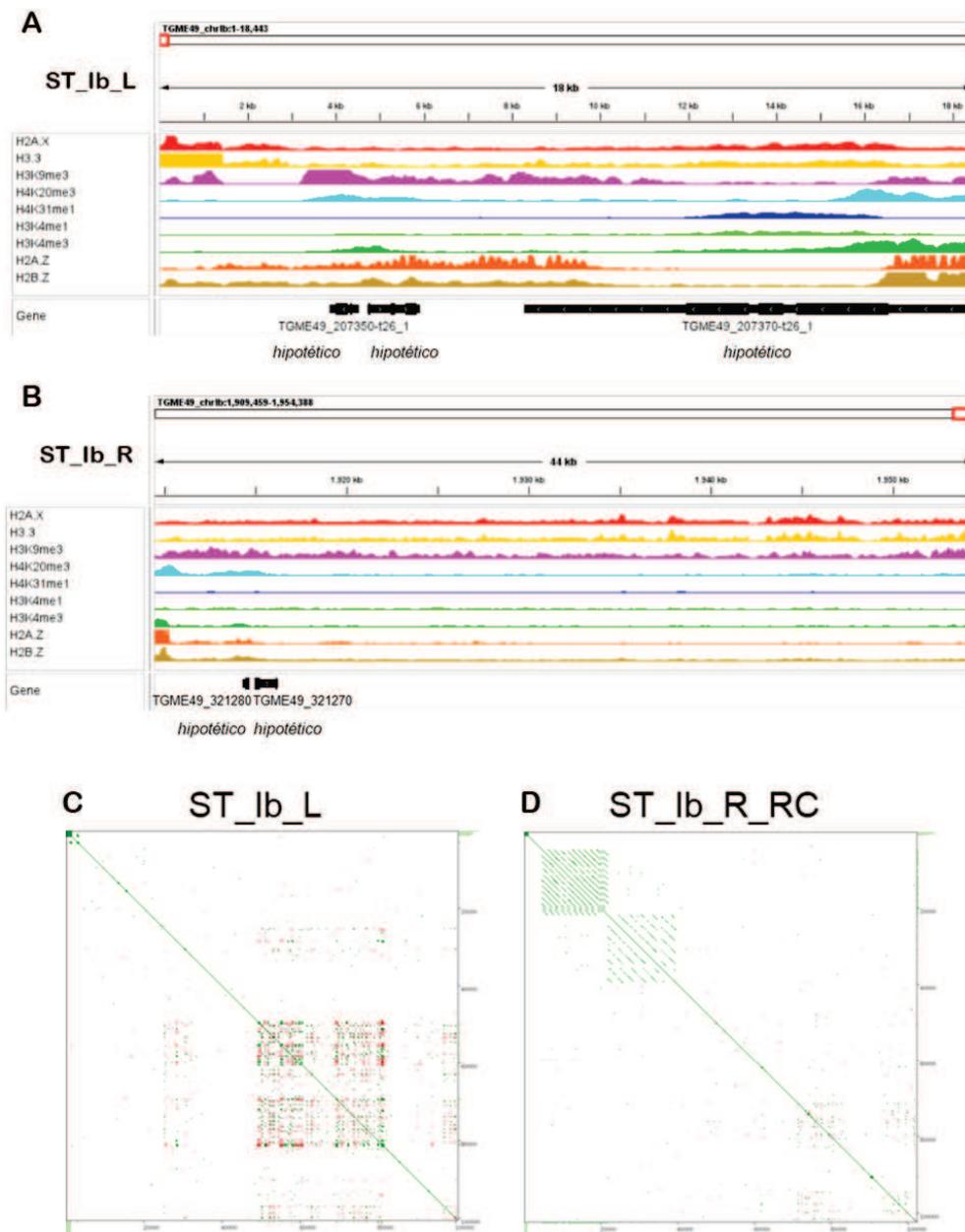
*Brazo izquierdo:* en este análisis se define ST\_lb\_L, como una región de ~16,5 Kpb enriquecida en H2A.X hasta la región TSS/5'UTR del gen hipotético TGME49\_207370, la cual presenta marcas activas marcando el límite de la región. Dentro de la región definida se incluyen dos genes hipotéticos TGME49\_207350 y TGME49\_207360, ambos con marcas de cromatina ambigua (H2A.Z/H2B.Z y bajos niveles de H3K4me3) en su región TSS (*Fig. 9A*).

*Brazo derecho:* en el extremo opuesto encontramos la región ST\_lb\_R de ~40 Kpb, parcialmente enriquecida en picos aislados de H2A.X, libre de genes hasta la presencia de los genes hipotéticos TGME49\_321270 (silente) y TGME49\_321280 (ambiguo) (*Fig. 9B*).

El análisis de Dot plot para ST\_lb\_L no mostró regiones de ADN satelital estricta, aunque se observó una región rica en pequeños fragmentos de ADN repetido disperso en forma directa e inversa entre los 50 – 80 Kpb (*Fig. 9C*). Asimismo, se observó para ST\_lb\_R la presencia de dos regiones ricas en ADN repetitivo en tándem, compatible con ADN satelital, dentro de los primeros 40 Kpb (*Fig. 9D*).

Por lo tanto, se describen nuevas regiones ST para el cromosoma Ib, ST\_Ib\_L y ST\_Ib\_R, de ~16,5 y ~40 Kpb respectivamente (*Tabla 2*).

**Fig. 9. Mapa cromosómico de la localización genómica de ST\_Ib\_L y ST\_Ib\_R. A.** Enriquecimiento de Hv y PTMs en la región genómica ST\_Ib\_L. **B.** Enriquecimiento de Hv y PTMs en la región genómica ST\_Ib\_R. **C.** Mapeo de ADN satelital repetitivo en ST\_Ib\_L. **D.** Mapeo de ADN satelital repetitivo en ST\_Ib\_R.



### 5.2.3 Cromosoma II

La región TAS descrita previamente para este cromosoma se encontró en el extremo del brazo derecho del mismo, TAS\_II (Dalmasso et al. 2014). Dalmasso y col., la definieron como una región de 43 Kpb asociada a la presencia de un miembro *famc*, *Tsf11*, (TGME49\_298060) y localizada internamente en la región subtelomérica, sin incluir el extremo cromosómico telomérico (**Fig. 10A**).

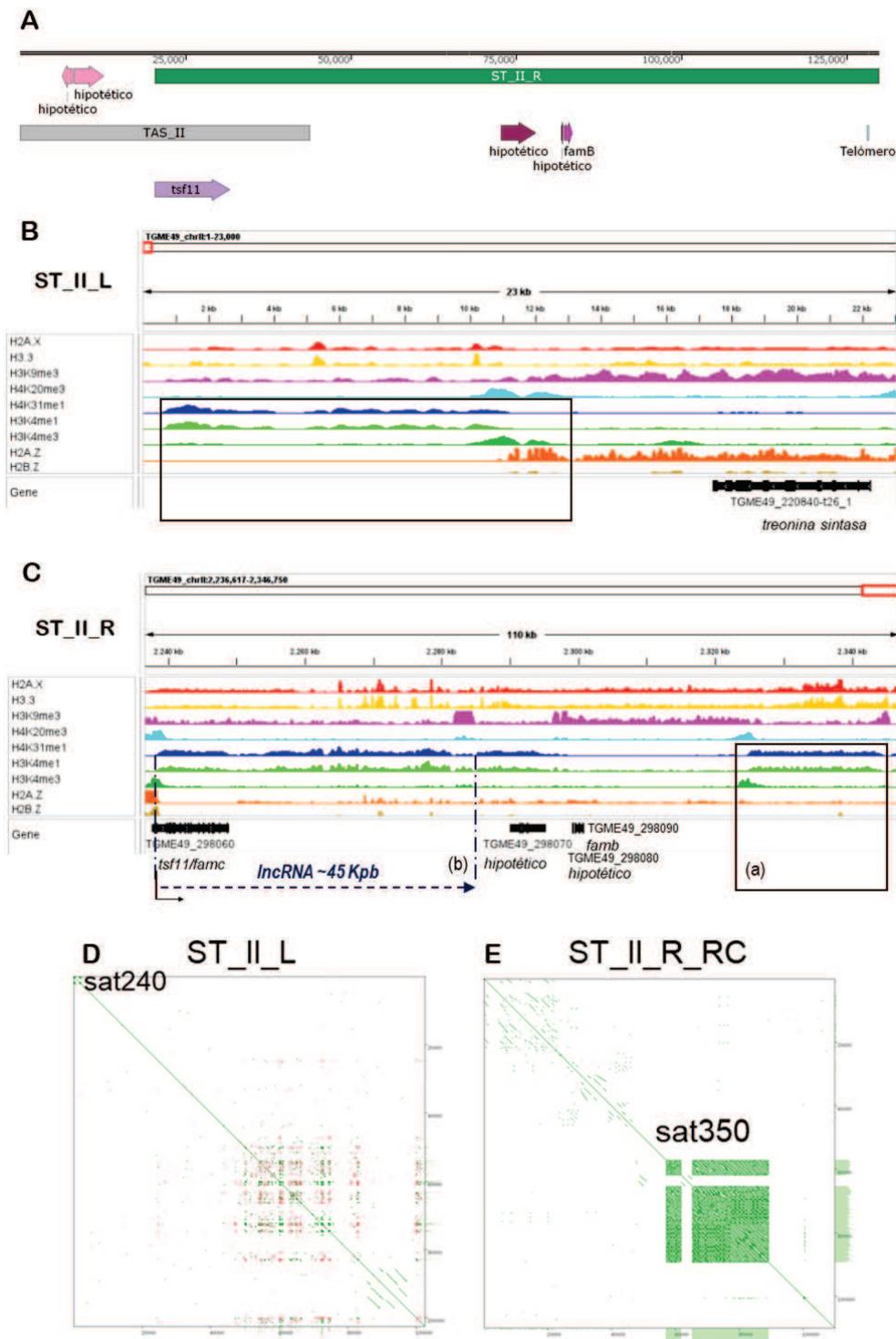
*Brazo izquierdo:* se propone una región de ~22,2 Kpb designada aquí ST\_II\_L, parcialmente enriquecida en H2A.X y desprovista de genes abarcando una región donde además comienzan los nucleosomas enriquecidos en H2A.Z y H2B.Z (con picos de H3K4me3) procedida por el gen codificante para una treonina sintasa: TGME49\_220840 con TSS en estado ambiguo, el cual marcaría el límite interno de la región ST\_II\_L (**Fig. 10B**).

*Brazo derecho:* la región ST\_II\_R abarcaría ~109,4 Kpb, estando parcialmente enriquecida en H2A.X, definiéndose su límite por la presencia del gen de la familia *famc*: *Tsf11* (TGME49\_298060) cuya región TSS se encontraría activa. Dentro de ella se encontraron varios genes, el más cercano al telómero, el gen *famb* (TGME49\_298090), y contiguo a este los genes hipotéticos TGME49\_298080 y TGME49\_298070, todos ellos con marca de cromatina silente en sus regiones TSS (**Fig. 10A y C**).

El análisis por Dot plot de ST\_II\_L mostró una región dentro de los primeros 10 Kpb compatible con ADN satelital sat240 y también se encontró la presencia de regiones cortas directa e inversamente repetidas a partir de los 40 Kpb (**Fig. 10D**). A su vez, ST\_II\_R, presenta una región rica (~60–90 Kpb) en elementos de ADN repetitivos en tándem, compatible con ADN satelital del tipo sat350 (**Fig. 10E**).

En este caso, se plantea entonces una nueva región subtelomérica para el brazo izquierdo del cromosoma II, ST\_II\_L de ~22,2 Kpb y una región ST\_II\_R en el brazo derecho que tendría ~109,4 Kpb (**Tabla 2**).

**Fig. 10. Mapa cromosómico de la localización genómica de ST\_II\_L y ST\_II\_R. A.** Región del extremo derecho del cromosoma II. TAS\_II (Dalmasso et al. 2014) en color gris. Los genes en color rosa son los genes hipotéticos que se encontraron dentro de dicha región pero que quedaron por fuera de este análisis (TGME49\_298050 y TGME49\_298040). ST\_II\_R descrita en este estudio (en verde oscuro). **B.** Enriquecimiento de Hv y PTMs en la región genómica ST\_II\_L. Recuadro negro: lncRNA compatible con TSS\* (ver sección 5.4.3). **C.** Enriquecimiento de Hv y PTMs en la región genómica ST\_II\_R. (a) lncRNA compatible con TSS\*, (b) lncRNA putativo TSS+ (ver sección 5.4.3). **D.** Mapeo de ADN satelital repetitivo en ST\_II\_L. **E.** Mapeo de ADN satelital repetitivo en ST\_II\_R. Las rayas verdes continuas corresponden a regiones repetitivas en tándem de ADN satelital.



### 5.2.4 Cromosoma III

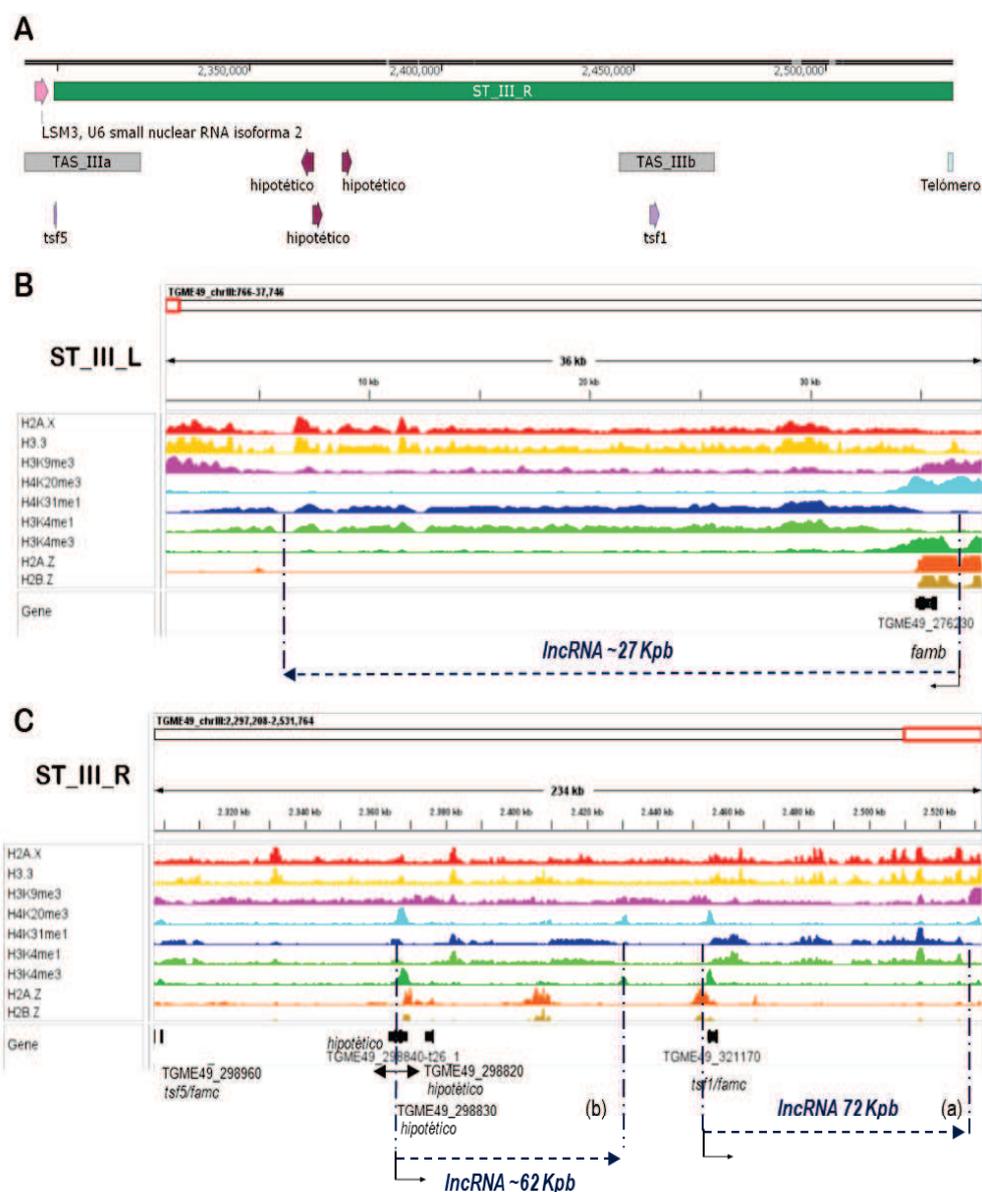
Previamente se encontró que había dos regiones TAS en el extremo derecho de este cromosoma, una primera región cercana al telómero de ~24 Kpb denominada TAS\_IIIb y río arriba una segunda de ~30 Kpb denominada TAS\_IIIa (Dalmaso et al. 2014). Ambas regiones se encontraron asociadas a genes de la familia *famc*: *Tsf1* (TGME49\_321170) y *Tsf5* (TGME49\_298960) respectivamente (**Fig. 11A**).

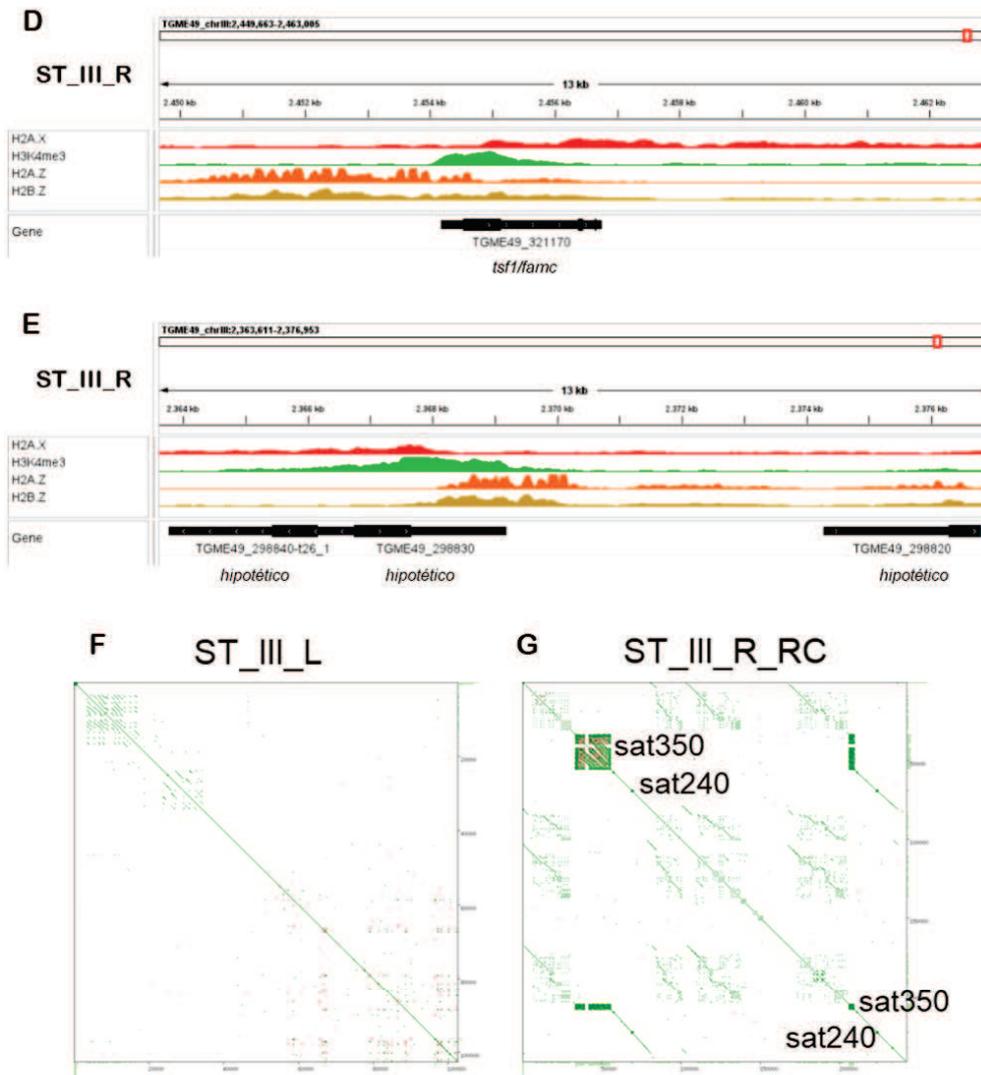
*Brazo izquierdo*: la región ST\_III\_L abarcaría ~35 Kpb, con enriquecimiento parcial de H2A.X, la cual estaría desprovista de genes hasta el límite interno propuesto por el gen de la familia *famb* (TGME49\_276230) (**Fig. 11B**). La región TSS de TGME49\_276230 estaría asociada a marcas de transcripción activa.

*Brazo derecho*: la región ST\_III\_R se tomó desde la región TSS ambigua del gen de la familia *famc/Tsf5* (TGME49\_298960) hasta el telómero quedando una región de ~232,4 Kpb que abarcaría, además, tres genes hipotéticos: TGME49\_298820, TGME49\_298830 y TGME49\_298840; y el gen *Tsf1* de la familia *famc* (TGME49\_321170) (**Fig. 11C**). ST\_III\_R presenta un enriquecimiento de H2A.X notorio de ~85 Kpb, desde el extremo hasta el gen *Tsf1*, a partir del cual el enriquecimiento de H2A.X se muestra como picos aislados (**Fig. 11C**). El gen *Tsf1* mostraría marcas de cromatina activa en su región TSS (**Fig. 11D**), mientras que los genes hipotéticos, excepto TGME49\_298840, mostrarían marcas de cromatina silente en sus regiones TSS (**Fig. 11E**). Es de destacar que el gen hipotético TGME49\_298830 presentaría marcas de histonas activas en la región 3', aunque también podrían ser marcas de la región TSS del gen TGME49\_298840, el cual solapa su región 5'/TSS, con la región 3' del gen TGME49\_298830 (**Fig. 11E**).

En el caso de ST\_III\_L se detectaron 2 regiones levemente enriquecidas en ADN repetido en tándem dentro de los 40 Kpb (**Fig. 11F**). Mientras que el análisis de Dot plot para ST\_III\_R mostró la presencia de dos regiones ricas en elementos de ADN repetitivos entre 30 y 50 Kpb (sat350 y sat240) y

**Fig. 11. Mapa cromosómico de la localización genómica de ST\_III\_L y ST\_III\_R. A.** Región en el extremo derecho del cromosoma III. TAS\_IIIa y TAS\_IIIb (Dalmasso et al. 2014) en color gris. El gen en color rosa corresponde a un gen que se encontró dentro de TAS\_IIIa pero que quedó por fuera de este análisis. ST\_III\_R (en verde oscuro). **B.** Enriquecimiento de Hv y PTMs en la región genómica ST\_III\_L. lncRNA putativo TSS+ (ver sección 5.4.3). **C.** Enriquecimiento de Hv y PTMs en la región genómica ST\_II\_R. (a), (b) lncRNAs compatibles con TSS+ (ver sección 5.4.3). **C.** Enriquecimiento de Hv y PTMs en la región genómica ST\_III\_R. lncRNA (ver sección 5.4.3). Las flechas negras indican la posición y dirección de dos genes: TGME49\_298840, hacia la izquierda (-) y TGME49\_298830 hacia la derecha (+). **D.** Ampliación de región enriquecida en marcas activas en *tsf1*. **E.** Ampliación de la región de genes hipotéticos. **F.** Mapeo de ADN satelital repetitivo en ST\_III\_L. **G.** Mapeo de ADN satelital repetitivo en ST\_III\_R.





otras dos cajas de repetitivos más pequeña (sat240 y sat350) localizada después de los 200 Kpb, todas compatibles con ADN satelital (*Fig. 11G*). Por lo tanto, en este cromosoma se encontró una nueva región ST en el brazo izquierdo de ~35 Kpb, ST\_III\_L, y una secuencia ST\_III\_R, de 232,4 Kpb (*Tabla 2*).

### 5.2.5 Cromosoma IV

El caso de este cromosoma es particular, la región TAS\_ IV de 21 Kpb descrita por Dalmaso y col., (2014) correspondiente al brazo derecho no se condijo con la localización el extremo cromosómico ni tampoco con el miembro de la familia *Tsf* previamente asociado, *Tsf12* (TGME49\_300990). Dado que esta posición no mantenía la equilocalidad establecida para las regiones

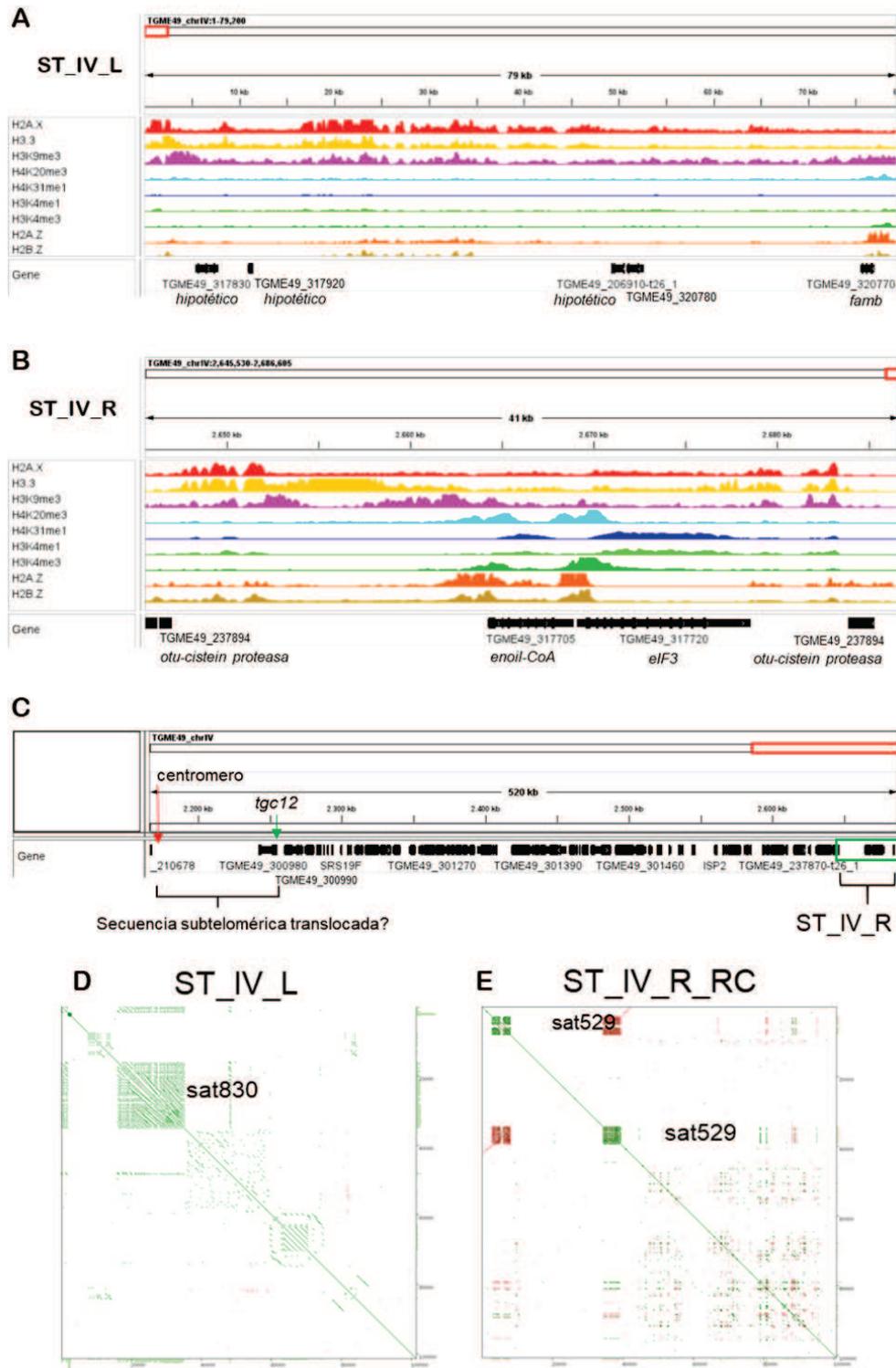
subteloméricas decidimos desestimarla y tomar como región telomérica el extremo derecho del cromosoma IV.

*Brazo izquierdo:* se identificó una región de ~77 Kpb enriquecida en H2A.X, abarcando desde el extremo del cromosoma al gen de la familia *famb* (TGME49\_320770) cuya región TSS se encuentra en estado ambiguo. Dicho gen establecería el límite de ST\_IV\_L en el lado izquierdo. Por otra parte, se observó dentro de ST\_IV\_L, la presencia de los genes hipotéticos TGME49\_206910, TGME49\_317820, TGME49\_317830 y TGME49\_320780, todos ellos con marca de cromatina silente en sus regiones TSS (*Fig. 12A*).

*Brazo derecho:* se propone una región ST\_IV\_R de ~41 Kb con enriquecimiento y picos de H2A.X desde la región TSS (ambigua) del gen TGME49\_237894 (Cisteína proteasa), hasta el extremo del cromosoma (*Fig. 12B*). Aguas arriba de este gen se observó una región extensa (analizamos solo 100 Kpb siguientes) de genes cuyo grado de separación, es compatible con una región de eucromatina. Luego detectamos a *Tsf12* (TGME49\_300990) a 438 Kpb del extremo derecho. Sin embargo, desde su posición río abajo, hacia el telómero, existe una alta densidad de genes, compatible con eucromatina. Cabe resaltar que río arriba de *Tsf12* se encuentra un gen contiguo hipotético, TGME49\_300980 y luego 82 Kpb libre de genes. Un análisis detallado de la región detectó que en esos 82 Kpb estaría el centrómero (2,166,000-2,177,000), por lo tanto, habría una región libre de genes entre *Tsf12* y el centrómero de 71,7 Kpb (*Fig. 12C*). Nos preguntamos si antiguamente ese fragmento, o parte de él, de 71,7 Kpb, incluido *Tsf12*, estarían formando parte del subtelómero y luego podría haberse translocado en forma invertida en la nueva posición cercana al centrómero.

Cabe mencionar que también se observó la presencia de otros genes dentro de la región ST\_IV\_R: TGME49\_317705 y TGME49\_317720 con TSS activos y finalmente TGME49\_237900 cercano al telómero con TSS silente (*Fig. 12B*).

**Fig. 12. Mapa cromosómico de la localización genómica de ST\_IV\_L y ST\_IV\_R. A.** Enriquecimiento de Hv y PTMs a lo largo de la región genómica ST\_IV\_L. **B.** Enriquecimiento de Hv y PTMs en la región genómica ST\_IV\_R. **C.** Un supuesto evento de translocación del ST a la posición del centrómero. ST\_IV\_R se describió *de novo* (cuadrado verde). **D.** Mapeo de ADN satelital repetitivo en ST\_IV\_L. **E.** Mapeo de ADN satelital repetitivo en ST\_IV\_R.



El análisis de Dot plot de ST\_IV\_L mostró una región enriquecida en secuencias de ADN repetitivas en tándem, compatible con ADN satelital, sat830, que abarca desde los 15 hasta los 30 Kpb, y otra, menos enriquecida en secuencias de ADN repetidas en tándem, entre los 60 y 70 Kpb (**Fig. 12C**). En el caso de ST\_IV\_R se encontraron dos regiones enriquecidas en secuencias de ADN repetitivas en tándem, compatible con ADN satelital, sat529, invertidas entre ellas, una dentro de los primeros 10 Kpb y la otra caja entre 30-40 Kpb (**Fig. 12D**).

Se presenta entonces una nueva región ST para el brazo izquierdo del cromosoma IV, ST\_IV\_L de ~77 Kpb; y, por otra parte, una región ST para el brazo derecho, ST\_IV\_R, de ~41 Kpb (**Tabla 2**).

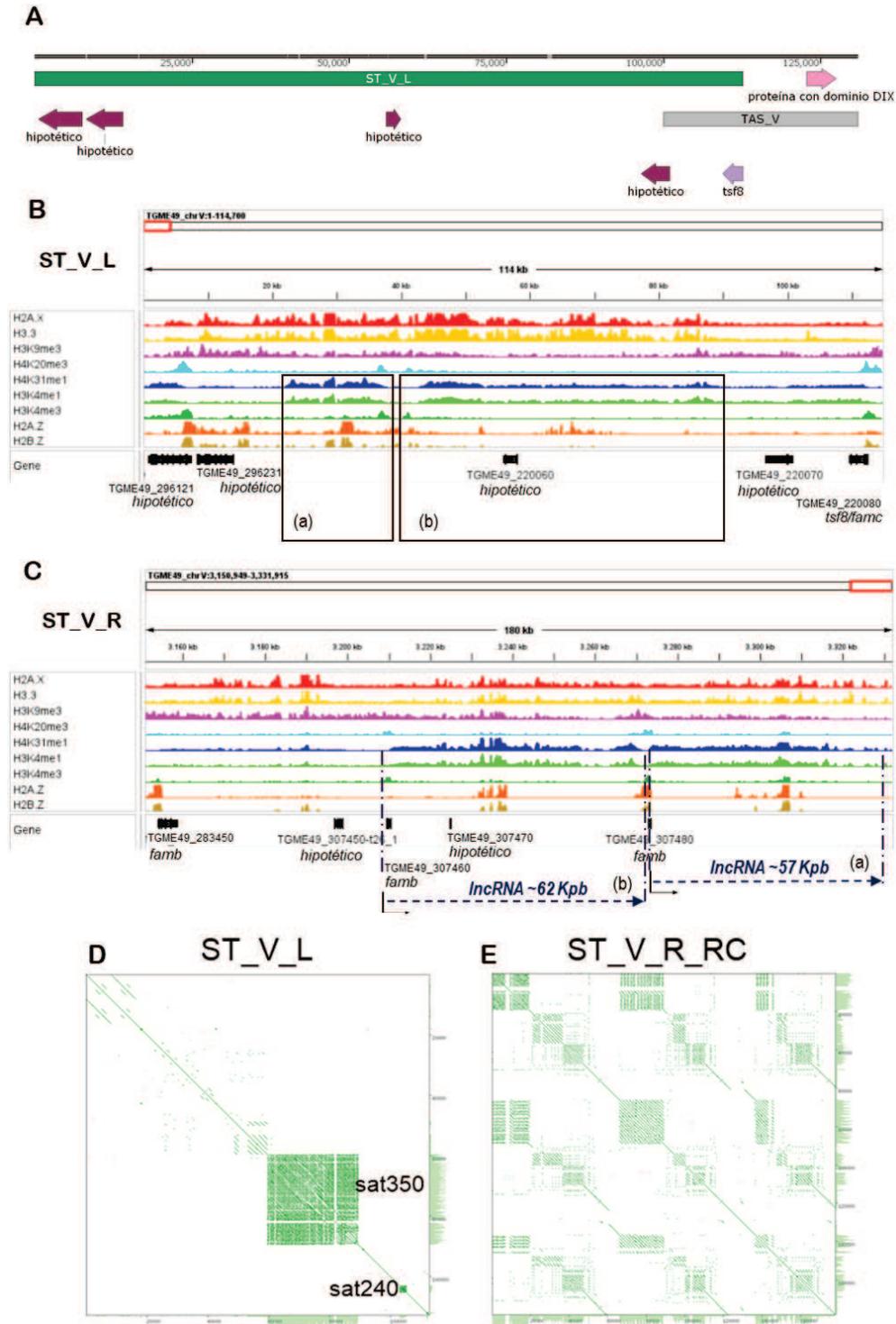
### **5.2.6 Cromosoma V**

La región TAS originalmente descrita para este cromosoma, TAS\_V de 30 Kpb (Dalmasso et al. 2014) se correspondía con una región interna del extremo izquierdo de este cromosoma que no llegaba al extremo cromosómico y dentro de la cual se observó la presencia de un gen codificante para una proteína con dominio DIX (TGME49\_220090) no descrito previamente, además de la presencia del gen *Tsf8* (TGME49\_220080) correspondiente a la familia *famc* (**Fig. 13A**).

*Brazo izquierdo:* se identificó una región enriquecida en la histona H2A.X hasta la posición cercana a los 90 Kpb. No obstante, la región de baja densidad génica se dispuso hasta al gen *Tsf8/famc* (TGME49\_220080), el cual mostró un estado de cromatina ambiguo en su TSS, tomándose como límite interno. Esto generó una región ST\_V\_L de 112,6 Kpb (**Fig. 13B**). Esta región incluye la presencia de varios genes hipotéticos: TGME49\_220070, TGME49\_220060, TGME49\_296231 y TGME49\_296121. Los genes TGME49\_220070 y TGME49\_220060 presentarían una cromatina silente en sus regiones TSS mientras que los otros dos genes hipotéticos tendrían marcas en sus TSS de cromatina activa (**Fig. 13B**).

*Brazo derecho:* Al analizar el extremo derecho se encontró un enriquecimiento de H2A.X hasta el TSS del gen *famb* (TGME49\_283450) en aparente estado ambiguo, el cual se tomó como límite

**Fig. 13. Mapa cromosómico de la localización genómica de ST\_V\_L y ST\_V\_R. A.** Región en el extremo derecho del cromosoma V. TAS\_V (Dalmasso et al. 2014) en color gris. ST\_V\_R (en verde oscuro). **B.** Enriquecimiento de Hv y PTMs en la región genómica ST\_V\_L. (a), (b) lncRNAs compatibles con TSS\* (ver sección 5.4.3). **C.** Enriquecimiento de Hv y PTMs en la región genómica ST\_V\_R. (a), (b) lncRNAs compatibles con TSS\* (ver sección 5.4.3). **D.** Mapeo de ADN satelital repetitivo en ST\_V\_L. **E.** Mapeo de ADN satelital repetitivo en ST\_V\_R. C. Enriquecimiento de Hv y PTMs en la región genómica ST\_II\_R.



de la región ST\_V\_R. De esa manera, la región ST\_V\_R tendría una extensión de 177,9 Kpb, y la presencia de tres genes de la familia *famb* (TGME49\_307460 y TGME49\_307480, además del mencionado arriba), y los genes hipotéticos TGME49\_307450 y TGME49\_307470 (**Fig. 13C**). Los genes hipotéticos y el *famb* TGME49\_307460 tendrían marcas de cromatina silente en sus regiones TSS. El gen *famb* TGME49\_307480 tendría marca de cromatina ambigua a activa en su TSS.

El análisis de Dot plot en ST\_V\_L mostró una región de ADN repetido en tándem, sat350, compatible con ADN satelital, entre 60 y 90 Kpb, y otra más pequeña, sat240, río abajo de los 100 Kpb (**Fig. 13D**). Mientras que para ST\_V\_R se muestra la presencia de ADN repetido en tándem, compatible con ADN satelital, en varias posiciones a lo largo del ST, incluyendo tres repeticiones de regiones extensas de ADN dentro del mismo (**Fig. 13E**).

Es así que se re-definieron los parámetros de la región ST para el cromosoma V. La región ST en el extremo izquierdo, ST\_V\_L, tendría una longitud de 112,6 Kpb; mientras que ST en el extremo derecho, ST\_V\_R, tendría una longitud de 177,9 Kpb (**Tabla 2**).

### **5.2.7 Cromosoma VI**

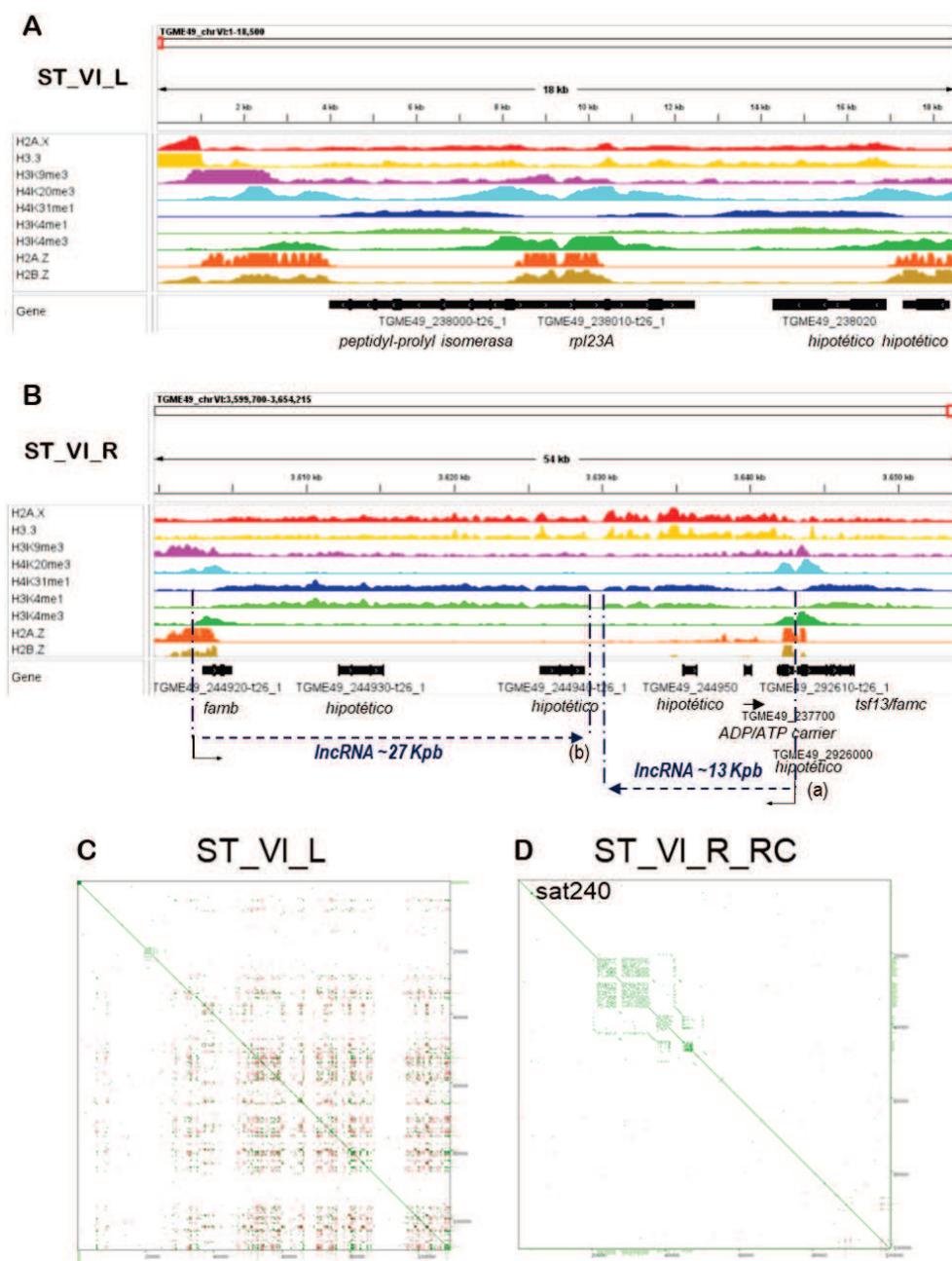
En el caso de este cromosoma, Dalmasso y col., (2014), no describieron regiones TAS asociadas a alguno de los extremos cromosómicos; por lo tanto, se analizaron ambos extremos.

*Brazo izquierdo:* se encontró una región enriquecida en H2A.X de ~2 Kpb sin genes, y luego una región corta hasta el gen codificante *peptidil-prolil isomerasa*, *ppi* (TGME49\_238000) cuya región TSS activa demarcaría el fin de ST\_VI\_L, de 8,1 Kpb (**Fig. 14A**).

*Brazo derecho:* En el extremo opuesto, para ST\_VI\_R, se observó la presencia de una región enriquecida en H2A.X de ~25 Kpb que llega hasta el gen hipotético TGME49\_244940. La región con baja densidad génica se extendió mas allá, delimitándose a esta región ST en función de la posición del gen de la familia *famb* (TGME49\_244920) aguas arriba (**Fig. 14B**) cuya región TSS se observó en estado activo. De esa manera, la región abarca ~53,7 Kpb y contiene un gen de la familia *famc* correspondiente a un gen *Tsf* no descrito previamente, *Tsf13* (TGME49\_292610) (**Fig. 14B**), con marca de cromatina activa en su TSS. La región ST\_VI\_R también comprende los genes hipotéticos

TGME49\_292600, TGME49\_244950, TGME49\_244940 y TGME49\_244930; y un gen codificante para una proteína ADP/ATP *carrier* (TGME49\_237700), todos ellos con marcas de cromatina silente en sus regiones TSS (Fig. 14B).

**Fig. 14. Mapa cromosómico de la localización genómica de ST\_VI\_L y ST\_VI\_R. A.** Enriquecimiento de Hv y PTMs en la región genómica ST\_VI\_L. Se muestran los genes río arriba de la región. **B.** Enriquecimiento de Hv y PTMs en la región genómica ST\_VI\_R. (a),(b) lncRNAs putativos TSS<sup>+</sup> (ver sección 5.4.3). **C.** Mapeo de ADN satelital repetitivo en ST\_VI\_L. **D.** Mapeo de ADN satelital repetitivo en ST\_VI\_R.



El análisis por Dot plot para ST\_VI\_L no mostró regiones ricas en ADN satelital, excepto una pequeña caja de ADN en tándem a ~50 Kpb y numerosos fragmentos pequeños de ADN (20-60 nucleótidos) repetido disperso en forma directa e inversa desde los 30 Kpb hasta cerca de los 90 Kpb (**Fig. 14C**), región compatible con eucromatina. El análisis de Dot plot para ST\_VI\_R mostró la presencia de regiones de elementos repetitivos en tándem, dentro de los 20-50 Kpb donde una de ellas sería compatible con ADN satelital sat240 (**Fig. 14D**).

Por lo tanto, para el cromosoma VI se describe una región ST\_VI\_L en el extremo izquierdo de al menos 8,1 Kpb; y una región ST\_VI\_R de 53,7 Kpb en el brazo derecho (**Tabla 2**).

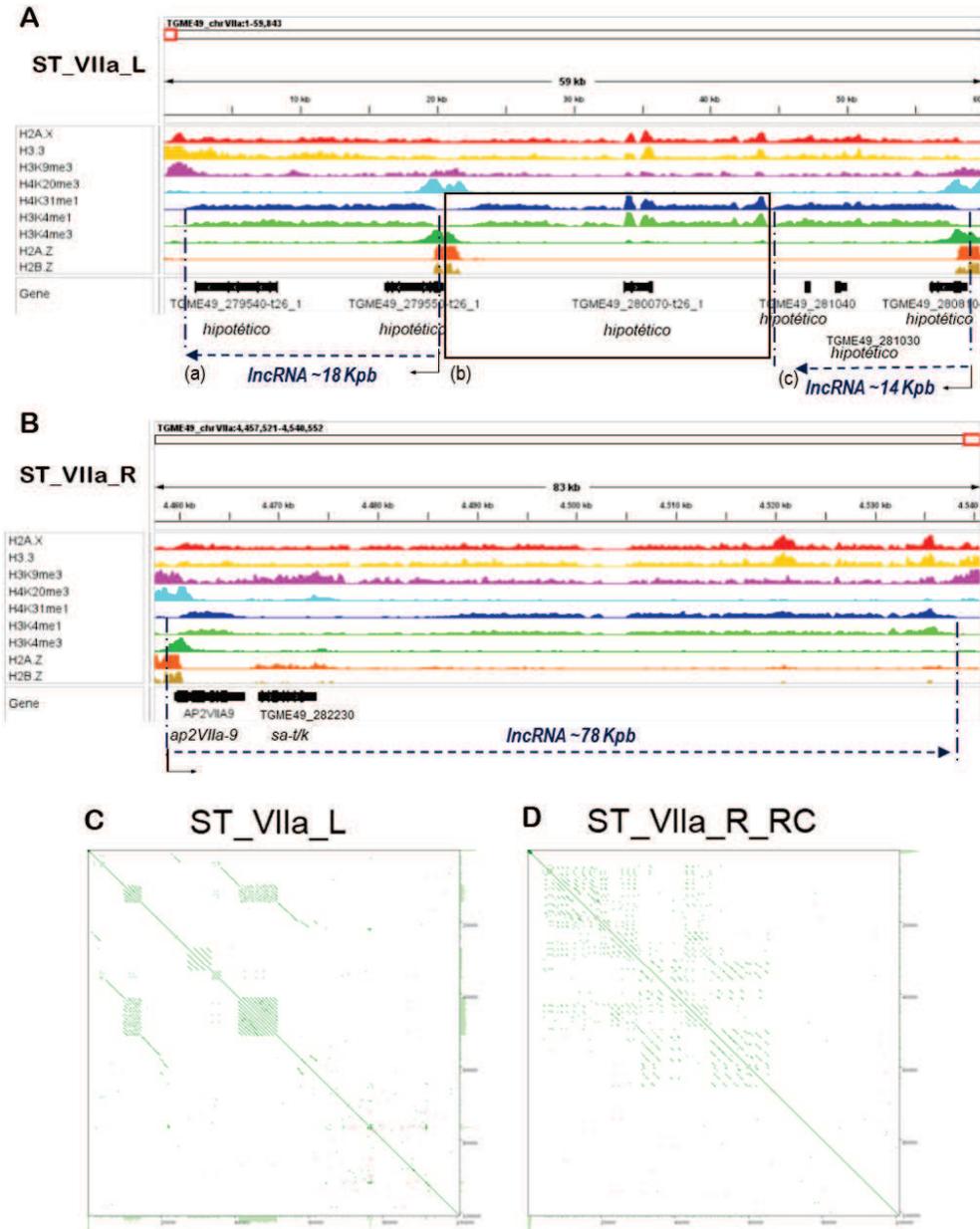
### **5.2.8 Cromosoma VIIa**

Según Dalmasso y col., (2014), este cromosoma no presentaba regiones TAS asociadas, por lo tanto, al igual que para los demás cromosomas, se procedió a buscar enriquecimiento de H2A.X en los extremos y genes asociados.

*Brazo izquierdo:* en su extremo izquierdo se encontró una región parcialmente enriquecida en H2A.X de ~58,4 Kpb, la cual se denominó ST\_VIIa\_L. No se encontraron genes de las familias *famb* o *famc*, pero si se observaron genes hipotéticos: TGME49\_279540, TGME49\_279550, TGME49\_280070, TGME49\_281040, TGME49\_281030 y TGME49\_280810. De estos genes hipotéticos, solo dos presentaban marcas de cromatina activa en sus regiones TSS (TGME49\_279550 y TGME49\_280810) el resto tendrían TSS en estado silente. Se consideró hasta el gen TGME49\_280810 (con TSS activo), como límite de la región ST\_VIIa\_L, desde donde comienza una densidad génica compatible con eucromatina (**Fig. 15A**).

*Brazo derecho:* Mientras que, el extremo derecho del cromosoma se observó una región de ~81 Kpb enriquecida en varios picos de H2A.X. Basados en la densidad génica, se propuso como el límite interno de ST\_VIIa\_R el gen *ap2VIIa-9* (TGME49\_282220), con marca de TSS activo, quedando incluido el gen *sulfato adeniltransferasa/adenilsulfato kinasa* (TGME49\_282230) (*sa-t/k*), el cual muestra una marca de cromatina ambigua en su región TSS (H2A.Z/H2B.Z) (**Fig. 15B**).

**Fig. 15. Mapa cromosómico de la localización genómica de ST\_VIIa\_L y ST\_VIIa\_R. A.** Enriquecimiento de Hv y PTMs en la región genómica ST\_VIIa\_L. (a),(c) IncRNAs putativos *TSS*<sup>+</sup>; (b) IncRNA compatible con *TSS*<sup>+</sup> (ver sección 5.4.3). **B.** Enriquecimiento de Hv y PTMs en la región genómica ST\_VIIa\_R. IncRNA putativo *TSS*<sup>+</sup> (ver sección 5.4.3) **C.** Mapeo de ADN satelital repetitivo en ST\_VIIa\_L. **D.** Mapeo de ADN satelital repetitivo en ST\_VIIa\_R.



El análisis de Dot plot para ST\_VIIa\_L mostró la presencia de una región de ADN repetido en tándem, compatible con ADN satelital, entre 40-50 Kpb, que también estaría presente en menor grado de extensión en la posición de 10 Kpb, reforzando la región ST\_VIIa\_L propuesta (Fig. 15C). Mientras

que el análisis de Dot plot para ST\_VIIa\_R mostró la presencia de fragmentos de ADN repetidos en forma directa a lo largo de los primeros 60 Kpb de la región ST propuesta (**Fig. 15D**). Por lo tanto, se proponen aquí dos regiones ST nuevas asociadas al cromosoma VIIa, ST\_VIIa\_L de ~58,4 Kpb en el extremo izquierdo, y la región ST\_VIIa\_R de ~81 Kpb en el extremo derecho (**Tabla 2**).

### **5.2.9 Cromosoma VIIb-VIII**

Estos cromosomas estarían fusionados según un trabajo de Bunnik y col., (2019); por lo tanto, solo se analizó el extremo izquierdo del cromosoma VIIb (VIIb\_L) y el extremo derecho del cromosoma VIII (VIII\_R).

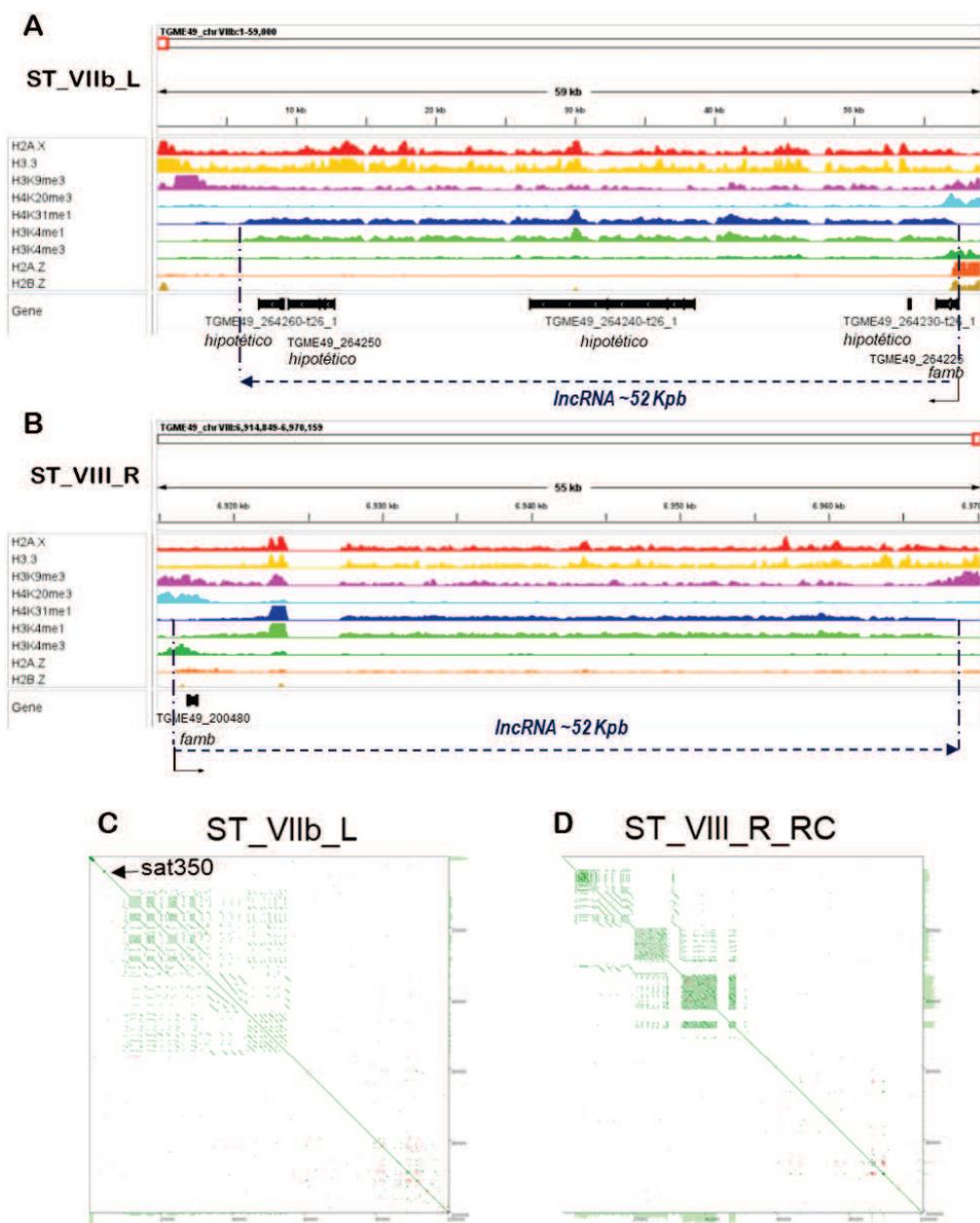
*Brazo izquierdo:* En el caso de este extremo se encontró una región parcialmente enriquecida en H2A.X de ~56,7 Kpb, hasta el gen miembro de la familia *famb* en estado activo (TGME49\_264225) el cual marcaría el límite interno de una nueva región ST\_VIIb\_L. En esta región se observó, además, la presencia de cuatro genes hipotéticos: TGME49\_264260, TGME49\_264250, TGME49\_264240 y TGME49\_264230, todos ellos con marcas de cromatina silente en sus regiones TSS (**Fig.16A**).

*Brazo derecho:* Asimismo, en el extremo derecho del cromosoma VIII se encontró una región enriquecida en H2A.X de ~53,1 Kpb, ST\_VIII\_R, libre de genes, a la cual se la delimitó por un miembro de la familia *famb* (TGME49\_200480) el cual solo presentó marca de H3K4me3, compatible con un estado activo (**Fig.16B**).

El análisis de Dot plot de ST\_VIIb\_L mostró un enriquecimiento de elementos repetidos en tándem, compatibles con ADN satelital, sat350, entre los 10–50 Kpb aproximadamente (**Fig.16C**), mientras que el análisis para ST\_VIII\_R mostró un enriquecimiento de elementos repetidos en tándem, compatibles con ADN satelital, a lo largo de los primeros 50 Kpb (**Fig.16D**).

Por lo tanto, se describen aquí dos nuevas regiones ST para el cromosoma VIIb-VIII, ST\_VIIb\_L de ~56,7 Kpb, y derecho, ST\_VIII\_R de ~53,1 Kpb (**Tabla 2**).

**Fig.16.** Mapa cromosómico de la localización genómica de ST\_VIIb\_L y ST\_VIII\_R. **A.** Enriquecimiento de Hv y PTMs en la región genómica ST\_VIIb\_L. lncRNA putativo TSS<sup>+</sup> (ver sección 5.4.3). **B.** Enriquecimiento de Hv y PTMs en la región genómica ST\_VIII\_R. lncRNA putativo TSS<sup>+</sup> (ver sección 5.4.3). **D.** Mapeo de ADN satelital repetitivo en ST\_VIIb\_L. **E.** Mapeo de ADN satelital repetitivo en ST\_VIII\_R.



### 5.2.10 Cromosoma IX

La región asociada al telómero descrita para este cromosoma, TAS\_IX (Dalmasso et al. 2014) se correspondía con una región de 20 Kpb asociada al gen *famc Tsf4* (TGME49\_200130) en el

brazo izquierdo del cromosoma. Cuando se analizó la localización genómica se encontró una región enriquecida en H2A.X, pero sin el gen descrito por Dalmasso y col., (2014) (**Fig. 17A**). Cabe señalar que, en este caso, la localización genómica de *Tsf4* estaría asociada a un cromosoma distinto (cromosoma XII) según la versión 53 del genoma de referencia (TGME49\_chrXII: 7,033,492-7,035,777).

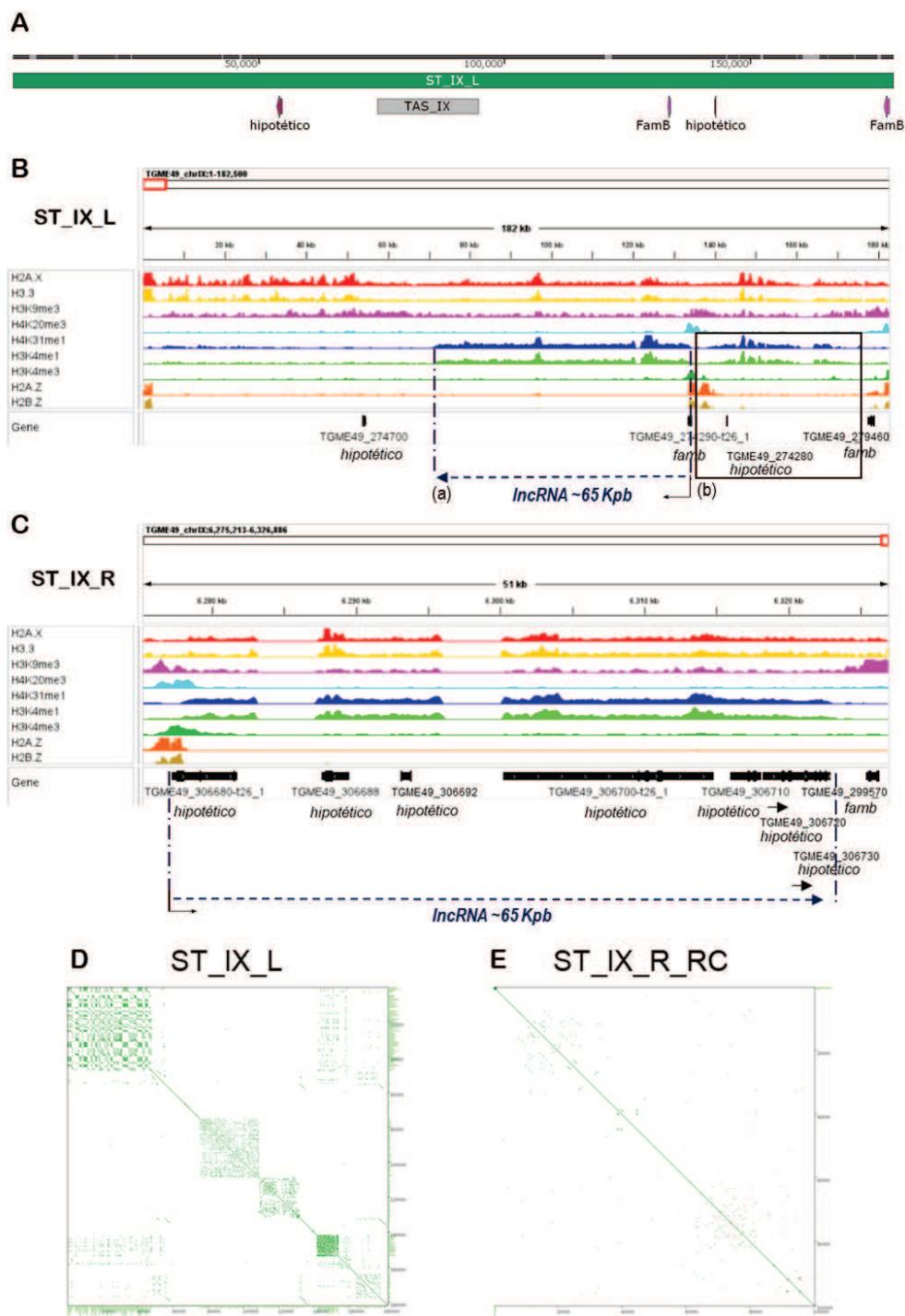
*Brazo izquierdo:* Debido a lo expuesto más arriba, se amplió el rango de esta región en dirección al centrómero y se encontró la presencia de dos genes de la familia *famb* (TGME49\_274290, TGME49\_279460), el primero con marca de cromatina activa en su región TSS, y el último en estado ambiguo, a partir de la cual comenzaba la eucromatina. El gen TGME49\_279460 fue propuesto para delimitar internamente la región (**Fig. 17B**), Se estableció así una nueva región la cual se denominó, ST\_IX\_L de ~179,1 Kpb. ST\_IX\_L estaría enriquecida en H2A.X en los primeros 50 Kpb, aunque presentaría enriquecimientos parciales hasta el gen *famb* TGME49\_279460. Contaría además con la presencia de dos genes hipotéticos: TGME49\_274700 y TGME49\_274280 dentro de la misma, ambos con marca de cromatina silente en sus regiones TSS (**Fig. 17B**).

*Brazo derecho:* Cuando se analizó el extremo derecho, el gen TGME49\_306680 (activo) fue propuesto como posible límite de ST\_IX\_R estableciendo una región de ~49,6 Kpb (**Fig. 17C**). Muy cerca del extremo, a ~2 Kpb se identificó el gen *famb* (TGME49\_299570) con TSS con marcas de cromatina silente. Al seguir analizando la región río arriba de *famb*, se encontró una región rica en genes hipotéticos: TGME49\_306730, TGME49\_306720, TGME49\_306710, TGME49\_306700, TGME49\_306692 y TGME49\_306688; todos ellos también con marcas de cromatina silente. La región ST\_IX\_R sugerida está parcialmente enriquecida en H2A.X (**Fig. 17C**).

La región ST\_IX\_L propuesta presentaría a lo largo de ella numerosas regiones ricas en ADN repetitivo en tándem, compatibles con ADN satelital (**Fig. 17D**); mientras que la región ST\_IX\_R no mostraría regiones ricas en ADN repetitivo (**Fig. 17E**).

En conclusión, se re-definió la región en el brazo izquierdo, ST\_IX\_L, como una región del cromosoma IX enriquecida en H2A.X de ~179,1 Kpb; mientras que en el extremo contrario se encontró y definió una región ST\_IX\_R de ~49,6 Kpb (**Tabla 2**).

**Fig. 17. Mapa cromosómico de la localización genómica de ST\_IX\_L y ST\_IX\_R. A.** Región en el extremo izquierdo del cromosoma IX. TAS\_IX (Dalmaso et al. 2014) en color gris. ST\_IX\_L (en verde oscuro). **B.** Enriquecimiento de Hv y PTMs a lo largo de la región genómica ST\_IX\_L. (a) IncRNA putativo TSS<sup>+</sup>; (b) IncRNA compatible con TSS<sup>\*</sup> (ver sección 5.4.3). **C.** Enriquecimiento de Hv y PTMs en la región genómica ST\_IX\_R. **D.** IncRNA putativo TSS<sup>+</sup> (ver sección 5.4.3). Mapeo de ADN satelital repetitivo en ST\_IX\_L. **E.** Mapeo de ADN satelital repetitivo en ST\_IX\_R.



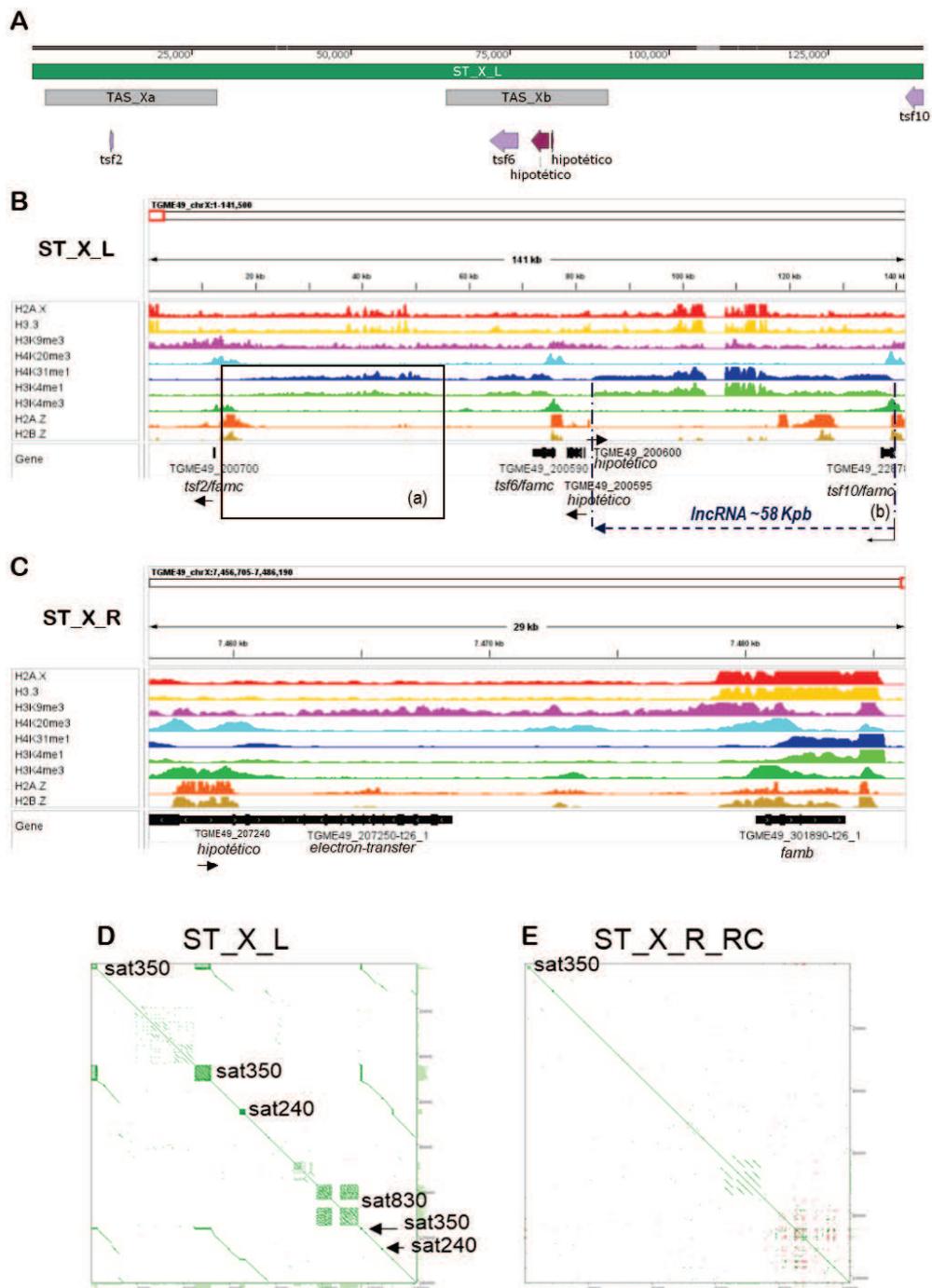
### 5.2.11 Cromosoma X

En este caso se describieron previamente dos regiones TAS: TAS\_Xa y TAS\_Xb en el extremo izquierdo del cromosoma, las cuales se correspondían con regiones de 25 y 26 Kpb respectiva y originalmente (Dalmasso et al. 2014). Al igual que en otros extremos cromosómicos las regiones TAS\_Xa y TAS\_Xb mostraron asociación a un miembro de la familia *famc*, *Tsf2* (TGME49\_200700) y *Tsf6* (TGME49\_200590) cada una respectivamente (**Fig. 18A**).

*Brazo izquierdo:* Debido a lo mencionado anteriormente, se redefinió la región ST en el cromosoma X. Un aspecto novedoso de esta región en el brazo izquierdo es que se encontró aguas arriba de *Tsf6* la presencia de un tercer gen de la familia *famc*, *Tsf10* (TGME49\_228780) descrito previamente en asociación al cromosoma XII cuya región TSS estaría en asociación a nucleosomas activos. A partir de *Tsf10* comenzaría la eucromatina, por lo tanto, se describe a este gen como límite de la región ST en el extremo izquierdo (ST\_X\_L). De esta manera la región ST\_X\_L abarcaría toda la zona con picos de enriquecimiento de H2A.X desde el inicio del telómero y como límite interno la región TSS del gen *Tsf10*, alcanzando una región de ~138,9 Kpb (**Fig. 18B**). Dentro del ST propuesto se observa la presencia de los otros genes de la familia *famc* como *Tsf2* (TGME49\_200700) y *Tsf6* (TGME49\_200590) con regiones TSS con marcas de cromatina ambigua (*Tsf2*) y activa (*Tsf6*) y dos genes hipotéticos (TGME49\_200595 y TGME49\_200600) con regiones TSS tipo silentes (**Fig. 18B**).

*Brazo derecho:* Al analizar extremo derecho del cromosoma X (ST\_X\_R) se encontró una región enriquecida en H2A.X, que incluye un miembro de la familia *famb* (TGME49\_301890) con marcas de cromatina activa. También se observó una región parcialmente rica en H2A.X hasta el gen hipotético TGME49\_207240, donde comienza la eucromatina. Dicha región incluye además un gen que codifica para la flavoproteína de transferencia de electrones: TGME49\_207250, el cual presenta marca de cromatina silenciada en su región TSS (**Fig. 18C**). Mientras que el gen TGME49\_207240 posee una región TSS activa y se propone como límite interno de la región ST\_X\_R, la cual tendría una longitud de ~27,5 Kpb y (**Fig. 18C**).

**Fig. 18. Mapa cromosómico de la localización genómica de ST\_X\_L y ST\_X\_R. A.** Región en el extremo izquierdo del cromosoma X. TAS\_Xa y TAS\_Xb (Dalmaso et al. 2014) en color gris. ST\_X\_L (en verde oscuro). **B.** Enriquecimiento de Hv y PTMs en la región genómica ST\_X\_L. (a) lncRNA compatible con TSS\*, (b) lncRNA putativo con TSS\* (ver sección 5.4.3). **C.** Enriquecimiento de Hv y PTMs en la región genómica ST\_X\_R. **D.** Mapeo de ADN satelital repetitivo en ST\_X\_L. **E.** Mapeo de ADN satelital repetitivo en ST\_X\_R.



La región ST\_X\_L propuesta presentaría además numerosas regiones ricas en ADN repetitivo en tándem, compatibles con ADN satelital, principalmente en la posición 40-50 Kpb (sat350, sat240), 60 Kpb y entre 90-120 Kpb (sat830, sat350 y sat240) (**Fig. 18D**). Mientras que, la región ST\_X\_R propuesta mostró una región enriquecida en ADN repetitivo compatible con sat350 en sus primeros 10 Kpb. También existiría una región con presencia de ADN repetitivo en tándem entre los 60–80 Kpb, formado por al menos tres copias de un fragmento cercano a los 2000 pb (**Fig. 18E**).

Por lo tanto, se redefinió una nueva región ST para el brazo izquierdo del cromosoma X, ST\_X\_L de ~139,8 Kpb; y una nueva región ST en el extremo cromosómico derecho, ST\_X\_R, de ~27,5 Kpb (**Tabla 2**).

### **5.2.12 Cromosoma XI**

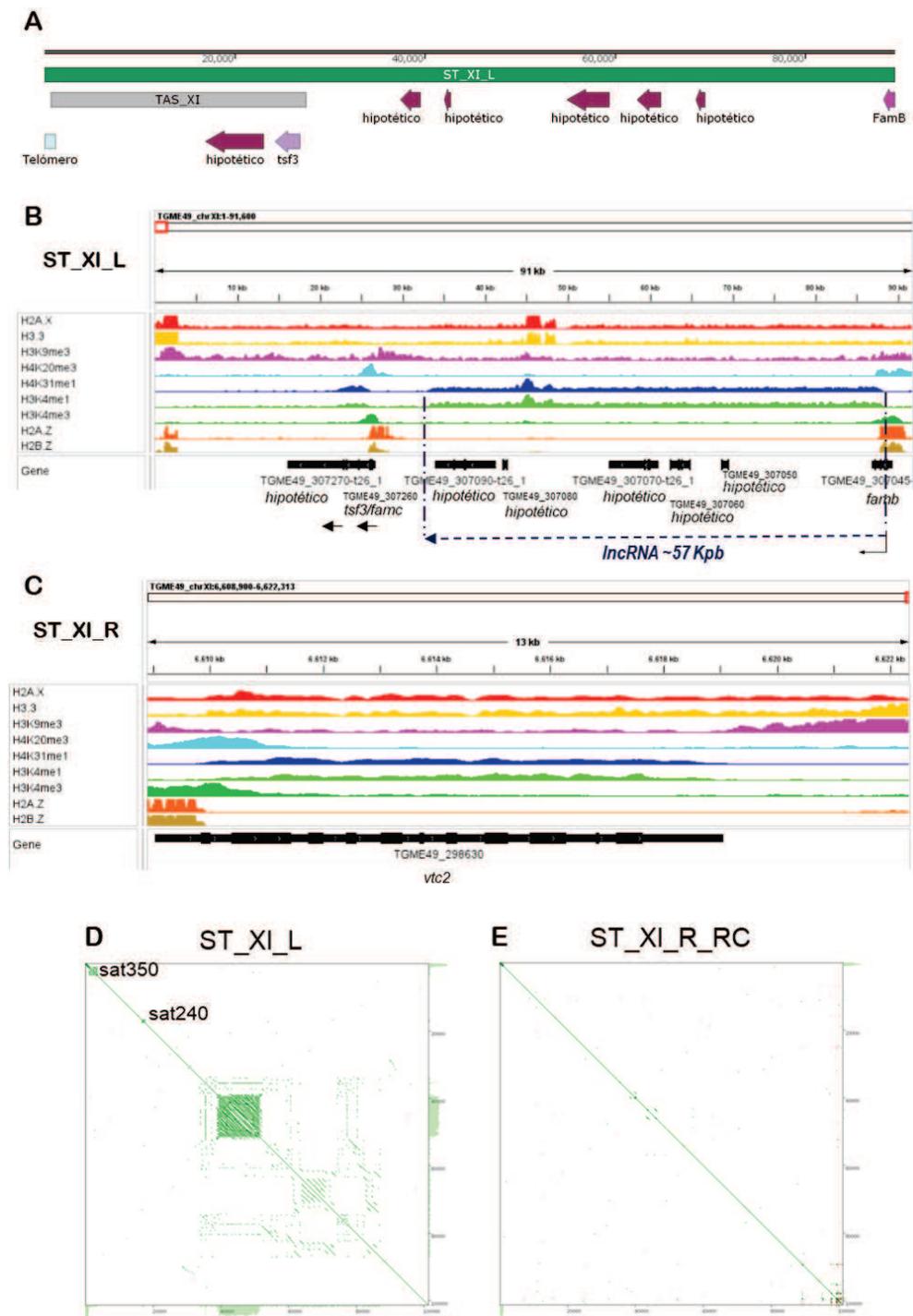
Para este cromosoma se había descrito una región TAS de 26 Kpb asociada al extremo izquierdo del mismo, TAS\_XI, y al gen *Tsf3* (TGME49\_307260) de la familia *famc* (Dalmasso et al. 2014) (**Fig. 19A**).

*Brazo izquierdo:* Por lo tanto, debido a lo observado para esta región se procedió a redefinir la zona como en los casos anteriores. ST\_XI\_L abarcaría entonces desde el telómero hasta el TSS activo de *famb*: TGME49\_307045, con un largo de ~88,2 Kpb. Esta región presentaría una región con picos de enriquecimiento de H2A.X y la presencia del gen de la familia *famc/Tsf3* (TGME49\_307260) y genes hipotéticos: TGME49\_307270, TGME49\_307090, TGME49\_307080, TGME49\_307070, TGME49\_307060, TGME49\_307050 (**Fig. 19B**). Los genes TGME49\_307050-307060-307070 tienen distancias de 3790 y 1713 pb entre sí, pero TGME49\_307050 está separado del gen TGME49\_307045 por 17391 pb, y el gen TGME49\_307070 está separado de TGME49\_307080 por 12329 pb. A excepción de los genes de la *famb* y *famc*, el resto, de los genes hipotéticos, presentan marcas de cromatina silente en sus regiones TSS (**Fig. 19B**).

*Brazo derecho:* En el extremo opuesto se observó una región parcialmente enriquecida en H2A.X de ~13,2 Kpb, hasta la región TSS del gen *vtc2* (TGME49\_298630), la cual presenta marcas de

cromatina activa (*Fig. 19C*). La región TSS de este gen fue propuesta como límite interno de la región ST\_XI\_R (*Fig. 19C*).

**Fig. 19. Mapa cromosómico de la localización genómica de ST\_XI\_L y ST\_XI\_R. A.** Región en el extremo izquierdo del cromosoma X. TAS\_XI (Dalmasso et al. 2014) en color gris. ST\_X\_L (en verde oscuro). **B.** Enriquecimiento de Hv y PTMs en la región genómica ST\_XI\_L. IncRNA putativo TSS<sup>+</sup> (ver sección 5.4.3). **C.** Enriquecimiento de Hv y PTMs en la región genómica ST\_XI\_R. **D.** Mapeo de ADN satelital repetitivo en ST\_XI\_L. **E.** Mapeo de ADN satelital repetitivo en ST\_XI\_R.



La región ST\_XI\_L propuesta presentaría una región rica en ADN repetitivo en tándem, compatible con ADN satelital, sat350 y sat240 dentro de los primeros 20 Kpb (*Fig. 19D*). Mientras que los 100 Kpb desde el extremo derecho del cromosoma XI que comprenden a ST\_XI\_R no presentan regiones ricas en ADN repetitivo en tándem (*Fig. 19E*).

Por lo tanto, para el cromosoma XI se re-definió la región subtelomérica ST\_XI\_L en el brazo izquierdo de ~88,2 Kpb, mientras que en el extremo derecho para ST\_XI\_R se propuso una región de ~13,2 Kpb (*Tabla 2*).

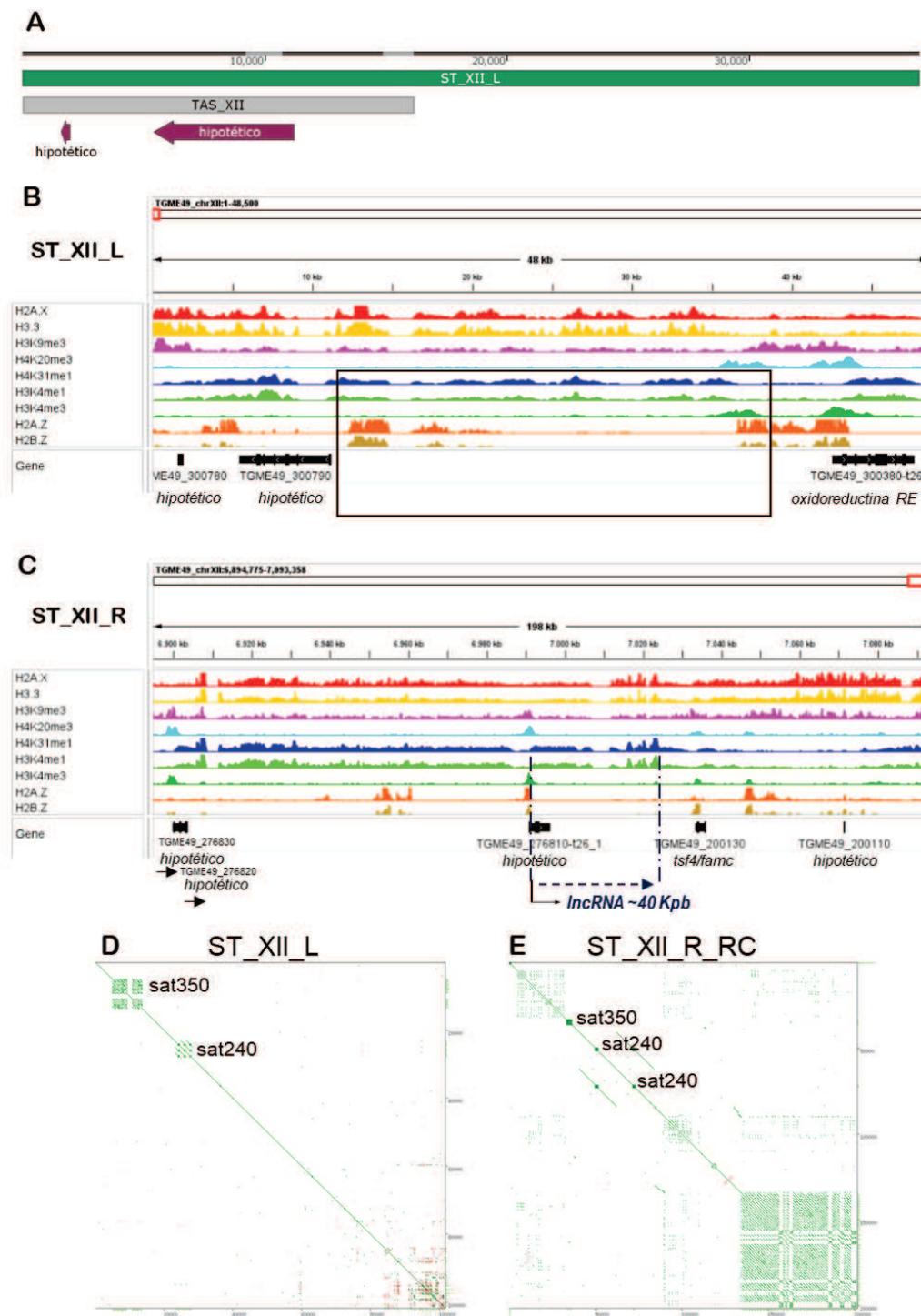
### **5.2.13 Cromosoma XII**

La localización genómica de TAS\_XII de 16 Kpb descrita previamente (Dalmaso et al. 2014) para este cromosoma en el extremo izquierdo y asociada al gen de la familia *famc Tsf10* (TGME49\_228780) no se condijo con lo observado en este análisis (*Fig. 20A*). Cabe mencionar que la localización del gen *Tsf10* descrito para este TAS, no se encuentra en el cromosoma XII como se planteaba, sino que se encuentra en el cromosoma X (TGME49\_chrX:137206-139810), formando parte de la región ST\_X\_L.

*Brazo izquierdo:* se observó una región enriquecida en la histona H2A.X hasta la región TSS activa del gen codificante para una oxidoreductina de retículo endoplasmático: TGME49\_300380 abarcando una región de ~47,6 Kpb (*Fig. 20B*). Esta región ST\_XII\_L propuesta incluiría la presencia de dos genes hipotéticos: TGME49\_300780 y TGME49\_300790, ambos casos con marca de cromatina ambigua río arriba de sus regiones TSS, y localizados dentro de los 11 Kpb del extremo, generando una región libre de genes de cerca de 36 Kpb (*Fig. 20B*).

*Brazo derecho:* Cuando se analizó el extremo opuesto se encontró una primera región enriquecida en H2A.X de ~61 Kpb hasta llegar a la región TSS silente a ambigua del gen de la familia *famc Tsf4* (TGME49\_200130) (previamente descrito para el TAS\_IX) (*Fig. 20C*). Luego se observó hacia el telómero aguas abajo de *Tsf4* la presencia de un gen hipotético (TGME49\_200110) también con marca de cromatina silente a ambigua (*Fig. 20C*). También se observaron picos de enriquecimiento de H2A.X en dirección al centrómero abarcando dos genes hipotéticos más, TGME49\_276810 (con TSS ambiguo)

**Fig. 20. Mapa cromosómico de la localización genómica de ST\_XII\_L y ST\_XII\_R. A.** Región en el extremo izquierdo del cromosoma XII. TAS\_XII (Dalmasso et al. 2014) en color gris. ST\_XII\_L (en verde oscuro). **B.** Enriquecimiento de Hv y PTMs en la región genómica ST\_XII\_L. LncRNA compatible con TSS<sup>+</sup> (ver sección 5.4.3). **C.** Enriquecimiento de Hv y PTMs en la región genómica ST\_XII\_R. IncRNA putativo con TSS<sup>+</sup> (ver sección 5.4.3). **D.** Mapeo de ADN satelital repetitivo en ST\_XII\_L. **E.** Mapeo de ADN satelital repetitivo en ST\_XII\_R. Dot plot de 200 Kpb.



TgSTs	Localización genómica	Tamaño ~Kpb	Gen Interno	ID	5'-3'	TSS
ST_la_L	TGME49_chrla:1-13,724	13,7	<i>tgc7</i>	_287280	(-)	activo
ST_la_R	TGME49_chrla:1,849,361-1,859,933	10	<i>tgb</i>	_295300	(+)	silente
			<i>chapA-tub</i>	_295310	(+)	activo
ST_lb_L	TGME49_chrlb:1-16,551	16,5	<i>tgh</i>	_207350	(-)	ambiguo
			<i>tgh</i>	_207360	(+)	ambiguo
			<i>tgh</i>	_207370	(-)	activo
ST_lb_R	TGME49_chrlb:1,909,459-1,954,388	40	<i>tgh</i>	_321270	(+)	silente
			<i>tgh</i>	_321280	(+)	ambiguo
ST_II_L	TGME49_chrII:1-22,252	22,2	<i>ts</i>	_220840	(-)	ambiguo
ST_II_R	TGME49_chrII:2,237,617-2,346,750	109,4	<i>tgb</i>	_298090	(+)	silente
			<i>tgh</i>	_298080	(+)	silente
			<i>tgh</i>	_298070	(+)	silente
			<i>tgc11</i>	_298060	(+)	activo
ST_III_L	TGME49_chrIII:766-35,746	35	<i>tgb</i>	_276230	(-)	activo
ST_III_R	TGME49_chrIII:2,299,208-2,531,764	232,4	<i>tgc1</i>	_321170	(+)	activo
			<i>tgh</i>	_298820	(+)	silente
			<i>tgh</i>	_298830	(+)	silente
			<i>tgh</i>	_298840	(-)	activo
			<i>tgc5</i>	_298960	(+)	ambiguo
ST_IV_L	TGME49_chrIV:764-77,001	77	<i>tgh</i>	_317830	(-)	silente
			<i>tgh</i>	_317820	(+)	silente
			<i>tgh</i>	_206910	(-)	silente
			<i>tgh</i>	_320780	(-)	silente
			<i>tgb</i>	_320770	(-)	ambiguo
ST_IV_R	TGME49_chrIV:2,645,530-2,686,605	41	<i>otu-cp</i>	_237900	(-)	silente
			<i>eIF3</i>	_317720	(+)	activo
			<i>enoil-CoA</i>	_317705	(+)	activo
			<i>otu-cp</i>	_237894	(+)	ambiguo
ST_V_L	TGME49_chrV:47-112,661	112,6	<i>tgh</i>	_296121	(-)	activo
			<i>tgh</i>	_296231	(-)	activo
			<i>tgh</i>	_220060	(+)	silente
			<i>tgh</i>	_220070	(-)	silente
			<i>tgc8</i>	_220080	(-)	ambiguo
ST_V_R	TGME49_chrV:3,153,949-3,331,915	177,9	<i>tgb</i>	_307480	(+)	ambiguo
			<i>tgh</i>	_307470	(+)	silente
			<i>tgb</i>	_307460	(+)	silente
			<i>tgh</i>	_307450	(+)	silente
			<i>tgb</i>	_283450	(+)	ambiguo
ST_VI_L	TGME49_chrVI:1-8,500	8,5	<i>ppi</i>	_238000	(-)	activo
ST_VI_R	TGME49_chrVI:3,602,980-3,654,215	53,7	<i>tgc13</i>	_292610	(+)	activo
			<i>tgh</i>	_292600	(+)	silente
			<i>adp/atp</i>	_237700	(+)	silente
			<i>tgh</i>	_244950	(+)	silente
			<i>tgh</i>	_244940	(+)	silente
			<i>tgh</i>	_244930	(+)	silente
			<i>tgb</i>	_244920	(+)	activo
ST_VIIa_L	TGME49_chrVIIa:1-58,843	58,4	<i>tgh</i>	_279540	(-)	silente
			<i>tgh</i>	_279550	(-)	activo
			<i>tgh</i>	_280070	(+)	silente
			<i>tgh</i>	_281040	(-)	silente
			<i>tgh</i>	_281030	(+)	silente
			<i>tgh</i>	_280810	(-)	activo

ST_VIIa_R	TGME49_chrVIIa:4,459,521-4,540,552	81	<i>sa-t/k</i>	_282230	(+)	ambiguo
			<i>ap2VIIa-9</i>	_282220	(+)	activo
ST_VIIb_L	TGME49_chrVIIb:828-57,559	56,7	<i>tgh</i>	_264260	(-)	silente
			<i>tgh</i>	_264250	(-)	silente
			<i>tgh</i>	_264240	(-)	silente
			<i>tgh</i>	_264230	(-)	silente
			<i>tgb</i>	_264225	(-)	activo
ST_VIII_R	TGME49_chrVIII:6,916,849-6,970,159	53,1	<i>tgb</i>	_200480	(+)	activo
ST_IX_L	TGME49_chrIX:1-179,116	179,1	<i>tgh</i>	_274700	(-)	silente
			<i>tgb</i>	_274290	(-)	activo
			<i>tgh</i>	_274280	(+)	silente
			<i>tgb</i>	_279460	(-)	ambiguo
ST_IX_R	TGME49_chrIX:6,277,213-6,326,886	49,6	<i>tgb</i>	_299570	(+)	silente
			<i>tgh</i>	_306730	(+)	silente
			<i>tgh</i>	_306720	(+)	silente
			<i>tgh</i>	_306710	(+)	silente
			<i>tgh</i>	_306700	(+)	silente
			<i>tgh</i>	_306692	(+)	silente
			<i>tgh</i>	_306688	(-)	silente
			<i>tgh</i>	_306680	(+)	activo
ST_X_L	TGME49_chrX:1-139,810	139,8	<i>tgc2</i>	_200700	(-)	ambiguo
			<i>tgc6</i>	_200590	(-)	activo
			<i>tgh</i>	_200595	(-)	silente
			<i>tgh</i>	_200600	(+)	silente
			<i>tgc10</i>	_228780	(-)	activo
ST_X_R	TGME49_chrX:7,458,705-7,486,190	27,5	<i>tgb</i>	_301890	(+)	activo
			<i>electron-t</i>	_207250	(-)	silente
			<i>tgh</i>	_207240	(+)	activo
ST_XI_L	TGME49_chrXI:1-89,389	88,2	<i>tgh</i>	_307270	(-)	silente
			<i>tgc3</i>	_307260	(-)	activo
			<i>tgh</i>	_307090	(-)	silente
			<i>tgh</i>	_307080	(-)	silente
			<i>tgh</i>	_307070	(-)	silente
			<i>tgh</i>	_307060	(-)	silente
			<i>tgh</i>	_307050	(-)	silente
			<i>tgb</i>	_307045	(-)	activo
ST_XI_R	TGME49_chrXI:6,609,042-6,622,313	13,2	<i>vtc2</i>	_298630	(+)	activo
ST_XII_L	TGME49_chrXII:1-47,600	47,6	<i>tgh</i>	_300780	(-)	ambiguo
			<i>tgh</i>	_300790	(-)	ambiguo
			<i>oxired-RE</i>	_300380	(+)	activo
ST_XII_R	TGME49_chrXII:6,901,775-7,093,358	191,5	<i>tgh</i>	_200110	(+)	ambiguo
			<i>tgc4</i>	_200130	(+)	ambiguo
			<i>tgh</i>	_276810	(+)	ambiguo
			<i>tgh</i>	_276820	(+)	silente

**Tabla 2. Regiones TgSTs definidas en este estudio.** Se describen regiones subteloméricas, STs, para el brazo izquierdo: ST\_Cromosoma\_L, y para el brazo derecho, ST\_Cromosoma\_R. Los genes asociados a cada región se describen en orden de aparición desde el telómero hacia el límite interno subtelomérico. (L): del inglés, *Left*, lado izquierdo; (R): del inglés, *Right*, lado derecho. Se tomó como lado derecho el final de la secuencia genómica según programa IGV. ID: TGME49; *tgc*: genes *famc*, *tgb*: genes *famb*, *tgh*: genes hipotéticos. 5'-3': (+) cadena de ADN sentido, (-) antisentido.

y TGME49\_276820 (con TSS silente), el cual marcaría el límite interno de ST\_XII\_R estableciendo una región de ~191,5 Kpb. Entre el gen 276820 y el gen 276810 hay 87,3 Kpb libre de genes, y entre *Tsf4* y el gen 276810 hay 42,5 Kpb libre de genes, denotando la característica de regiones libres de genes de los STs, diferente a lo que se observa en regiones de eucromatina. La región ST\_XII\_L propuesta presentaría regiones ricas en ADN repetitivo en tándem, compatibles con ADN satelital, sat350 y sat240, entre los 5-15Kpb (**Fig. 20D**). Por su parte, la región ST\_XII\_R propuesta presentaría además numerosas regiones ricas en ADN repetitivo en tándem, algunas compatibles con ADN satelital (sat350 y sat240), entre los 40 Kpb y 70 Kpb y finalmente una región de ADN repetitivo satelital entre los 130-200 Kpb (**Fig. 20E**). También se observa la presencia de fragmentos de ADN repetido en forma directa en la región comprendida por los 60 a 80 Kpb (**Fig. 20E**).

Por lo tanto, para el cromosoma XII se re-definió la región subtelomérica ST\_XII\_L en el brazo izquierdo de ~47,6 Kpb, mientras que en el extremo derecho se propuso una región ST\_XII\_R de ~191,5 Kpb (**Tabla 2**).

### **5.3 ADN satelital y subtelómeros**

Como se mencionó, se pudo detectar mediante Dot plot la presencia y composición de los elementos de ADN repetitivo de las regiones STs de *T. gondii*. De esta forma fue posible revelar regiones de ADN repetitivo en tándem, compatibles con el ADN satélite, en casi todos los STs. Los diferentes elementos repetitivos no presentarían una alta identidad entre ellos (datos no mostrados). Esto no es inesperado ya que el elemento satélite de *T. gondii* más reconocido, sat350, presente en los ST\_II\_R, III\_R, V\_L, VIIb\_L, X\_L, XI\_L, XII\_L y XII\_R, presentó según un estudio del laboratorio niveles de identidad intra-monoméricos entre 58 y 95% (Clemente et al. 2004).

Curiosamente, este nuevo análisis no concuerda con la organización anterior de TARE 1 a TARE 3 observada por Dalmaso y col., (2014). De hecho, casi no hay relación de identidad entre los ADN satelitales detectados en cada subtelómero, exceptos los descritos en las **Fig. 8 a 20**: sat350, n=

12; sat240, n=12; sat830, n=2 y sat529, el cual se detectó uno solo, pero había sido caracterizado en su localización centromérica (Xia et al. 2021).

En general, el elemento repetitivo más consistente como ADN satélite corresponde a la familia sat350/sat680. Mientras que los ADN satélites son bastante homogéneos y se observan diferencias a nivel de secuencia entre las especies, el ADN sat350 consta de unidades o monómeros con alta variabilidad (Clemente et al. 2004). Otra pregunta es el origen de los otros satélites de ADN que no tienen relación con sat350. De estos, solo sat240 mostró alguna propagación. Sat830 y el resto parecen ser satélites de ADN únicos, aunque habría pequeñas regiones de identidad entre todos ellos. La presencia de grandes regiones de ADN satélite en los STs, mayormente diferentes y la presencia de solo pocas familias de ADN satélite, cuyos monómeros presentan un bajo porcentaje de identidad entre ellos, sugieren que el ADN satelital en estas regiones podría contribuir a la plasticidad y evolución de los subtelómeros.

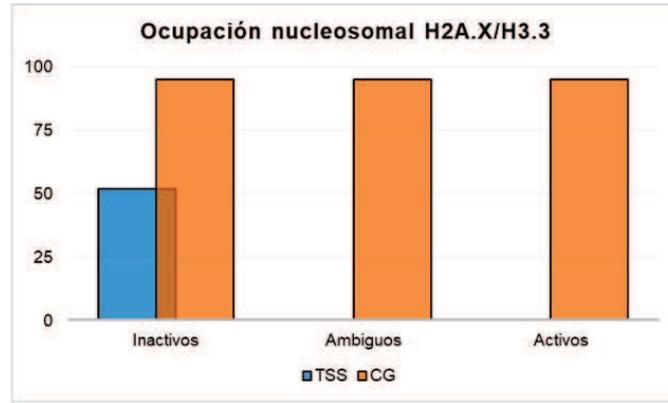
## **5.4 Análisis de la cromatina en las regiones TgSTs**

### **5.4.1 Posicionamiento de histonas variantes**

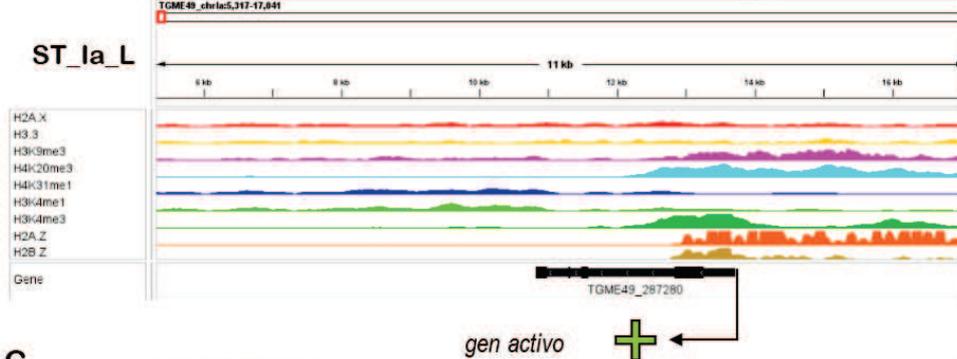
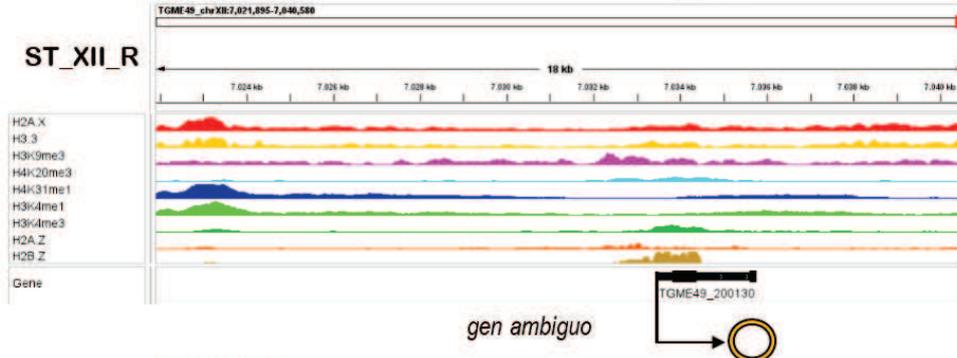
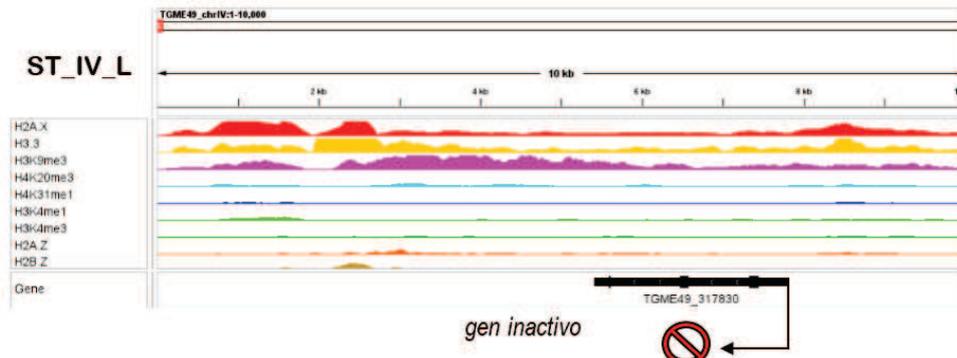
Como se describió en la sección 5.2, el enriquecimiento de la variante de histonas H2A.X es un componente fundamental en la definición de las regiones TgSTs. Específicamente, se evaluó la distribución nucleosomal de H2A.X y H3.3 sobre las regiones TSS y cuerpos génicos (CGen) de los 96 genes que integrarían las regiones STs (*Fig. 21A*). Las histonas variantes H2A.X y H3.3 se encontraron asociadas a casi todos los CGen de los genes STs independientemente de su estado: activo, ambiguo o silente (*Fig. 21B*) con excepción del gen TGME49\_237900, donde se encontraron posicionadas sobre el 3'UTR (*Fig. 21C*). Sin embargo, 52 genes STs mostraron enriquecimiento de H2A.X y H3.3 en sus

**Fig. 21. Ocupación nucleosomal en genes STs. A.** Distribución de H2A.X y H3.3 en función de la cromatina asociada a la actividad génica. TSS: barra azul, enriquecimiento de H2A.X/H3.3 sobre región TSS. CG: barras naranjas, enriquecimiento de H2A.X/H3.3 sobre el CGen. **B.** Se muestra el enriquecimiento de las Hv junto a PTMs específicas en genes STs y su relación con los estados de expresión génica. **C.** Gen STs con estado de cromatina ST indefinida (TGME49\_237900).

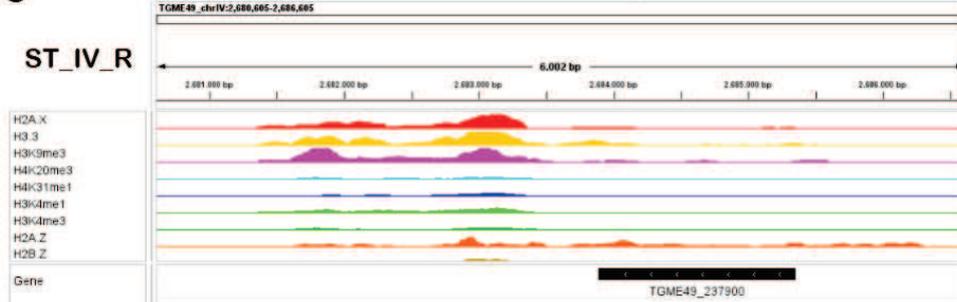
**A**



**B**



**C**



regiones TSS además del enriquecimiento de H2A.X y H3.3 en el CGen (**Fig. 21A y B**); los cuales se corresponderían con un estado de cromatina silente (**Tabla 2**). Estas observaciones establecen una relación inversa con lo observado por Nardelli y col., (2022) cuando H2A.X ocupa los CGen en genes activos como por ejemplo genes de taquizoito: *ldh1*, pero también ocupa los CGen en el caso de genes estadio específico de bradizoito como ejemplo: *eno2* (Nardelli et al. 2022).

Durante el análisis, también se observó el patrón de distribución de las histonas variantes H2A.Z y H2B.Z. Se encontró que se posicionan junto a la marca H3K4me3 en 26 regiones TSS de los genes STs (**Fig. 21B, Tabla 2**). En las regiones STs, el doble dímero de histonas variantes H2A.Z/H2B.Z marca regiones TSS de genes activos junto a H3K4me3 como lo describieron Nardelli y col., (2022). Sin embargo, a diferencia de lo observado para genes silentes en el resto del genoma (Nardelli et al. 2022), los CGen de dichos genes STs no se encontraron asociados al doble dímero (**Fig. 21B**). Mientras que en 17 genes STs, se encontró enriquecimiento únicamente de H2A.Z y/o H2B.Z en la región TSS, definiéndolos como genes ambiguos (**Fig. 21B**).

El análisis de las 3 variantes de la histona H2A y la marca H3K4me3 en las regiones río arriba del extremo 5' UTR de los genes, compatible con la región TSS, nos permitió respaldar la selección de algunas regiones STs ricas en genes, según si los genes mostraban cromatina en un estado silente, ambigua o activa a saber: genes STs ambiguos (17%) (donde los genes poseen TSS únicamente con H2A.Z/H2B.Z) y genes STs silentes (55%) (cuyo TSS está asociado a H2A.X y H3.3) (**Tabla 2**).

También se pudo observar la presencia de picos de H2A.Z y/o H2B.Z a veces acompañado por H3K4me3, en regiones libres de genes como las observadas en ST\_III\_R (posición 2,400,500-2,410,000, H2A.Z/H2B.Z), ST\_V\_L (posición 30,000-33,000, H2A.Z/H2B.Z/H3K4me3), ST\_V\_R (posición 3,304,000-3,307,000, H2A.Z/H2B.Z/H3K4me3), ST\_X\_L (posición 118,000-129,000, H2A.Z/H2B.Z/H3K4me3), ST\_X\_R (posición 7,472,000-7,474,300, H2A.Z/H2B.Z) y ST\_XII\_R (posición 7,045,500-7,050,000, H2A.Z/H2B.Z/H3K4me3) (**Fig. 22**).

#### 5.4.2 Posicionamiento de PTMs específicas

Sorprendentemente, también pudimos observar la presencia de picos de H3K4me3 sin acompañamiento de H2A.Z/H2B.Z cerca de los extremos de los cromosomas (**Fig. 23**), con excepción de los extremos de ST\_Ia\_L, ST\_II\_L, ST\_II\_R, ST\_IV\_R, ST\_V\_L, ST\_V\_R, ST\_VI\_R, ST\_X\_R. De los 18 STs que mostraron estos picos, 11 coincidieron con la presencia de la repetición telomérica TTTAGGG en el extremo del cromosoma, exceptuando solo a 3 (ver **Tabla 1**); lo que sugiere que, la presencia de H3K4me3 en los nucleosomas de los extremos de las regiones STs estaría estrechamente asociada al funcionamiento de los telómeros de *T. gondii* en cuanto a la transcripción de lncRNA TERRA.

De manera global se vio que H3K4me3 tiene una ocupancia relativamente baja (27%) en genes STs (26) en comparación con las marcas de metilación asociadas a la heterocromatina: H3K9me3, H3K4me1 y H4K31me1 que están presentes en un ~70% de los genes y en un 45% en el caso de H4K20me3 (**Fig.24A**).

La marca H3K4me3 se encontró exclusivamente en regiones TSS y no en los CGen (**Fig.24B**). Dada su asociación como marca de gen activo, cuando estuvo sola o en presencia de H2A.Z/H2B.Z se la empleó para definir cromatina activa. El 82% de los genes activos STs (22) presentó en promedio correspondencia entre el enriquecimiento de H3K4me3 y datos de transcriptómica (**Tabla 3**, **Tabla 5**, **Tabla 7**, **Tabla 8**), aunque solo el 34,6% de los genes (9) coincidieron con valores promedios de transcripción del 70% o más.

Las metilaciones sobre H3K9me3 y H4K20me3 se encontraron mayormente asociadas a las regiones TSS de los genes STs independientemente del estado de los mismos, aunque H4K20me3 se asociaría mayoritariamente con genes STs activos. Dichas marcas también se observaron distribuidas sobre el CGen en algunos pocos casos (**Fig.24B**).

Por otra parte, las marcas H3K4me1/H4K31me1 se encontraron preferencialmente en las regiones TSS de genes con cromatina silente (**Fig.24B**). De manera notable, una gran parte de los genes STs con TSS asociados a cromatina silente no presentan asociación a H3K9me3 ni H4K20me3, sino que estarían únicamente asociados a H2A.X/H3.3/H3K4me3/H4K31me1 tanto en la región TSS como

en el CGen (*Fig.24C y D*). Se ha podido corroborar, por lo tanto, que tal como lo vieron Sindikubwabo y col., (2017); H4K31me1 y H3K4me1 tendrían distribución cromosómica a lo largo del CGen en genes activos y ambiguos (*Fig.24B*). Sin embargo, también ocuparían regiones TSS junto a H2A.X y H3.3 en la mayoría de los genes STs silentes, pudiendo tal vez establecerse como marcas de inactivación específica en las regiones heterocromáticas subteloméricas de *T. gondii*.

**Fig. 22.** Mapeo de histonas variantes H2A.Z/H2B.Z y H3K4me3 en regiones STs sin genes.

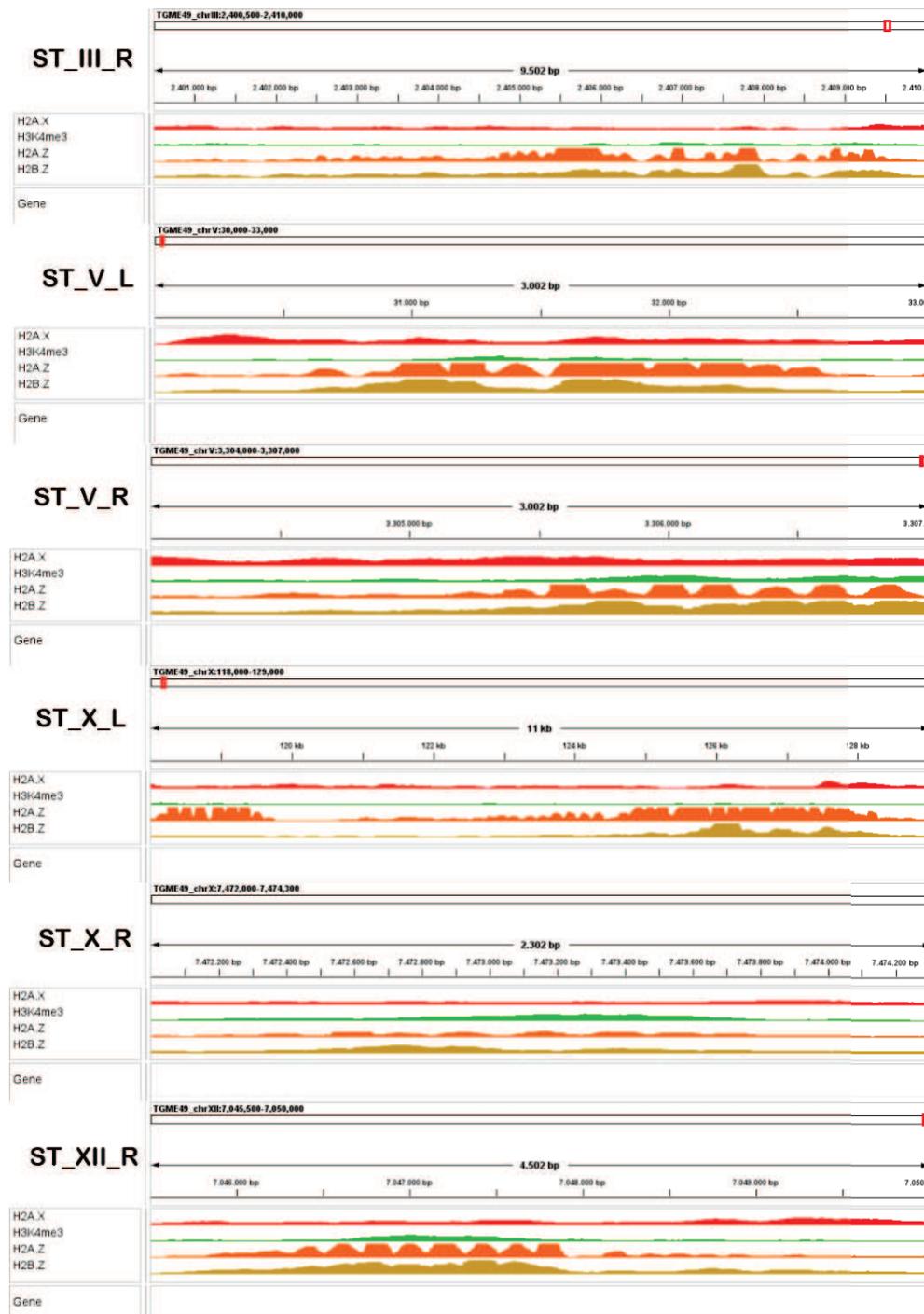
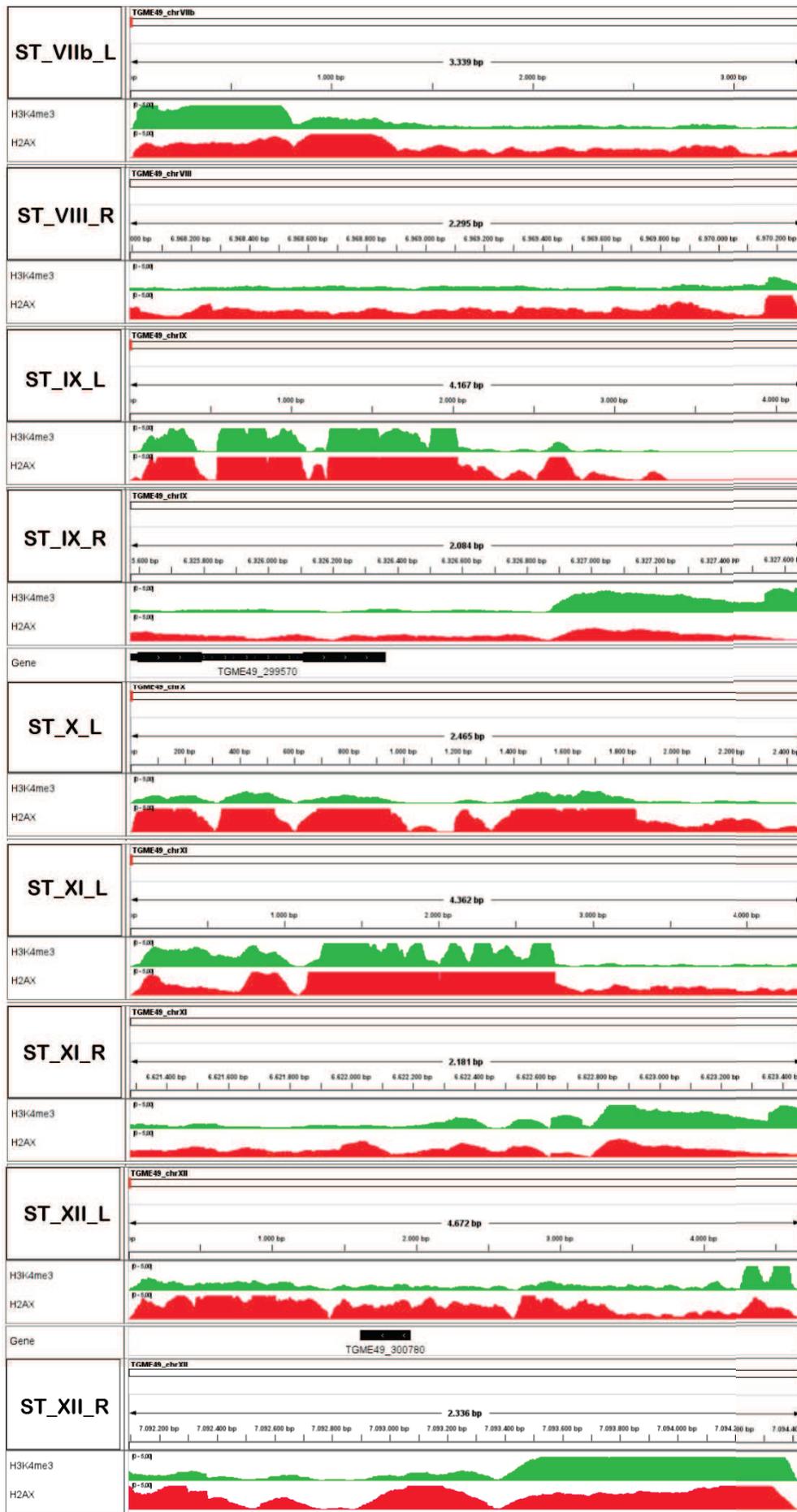
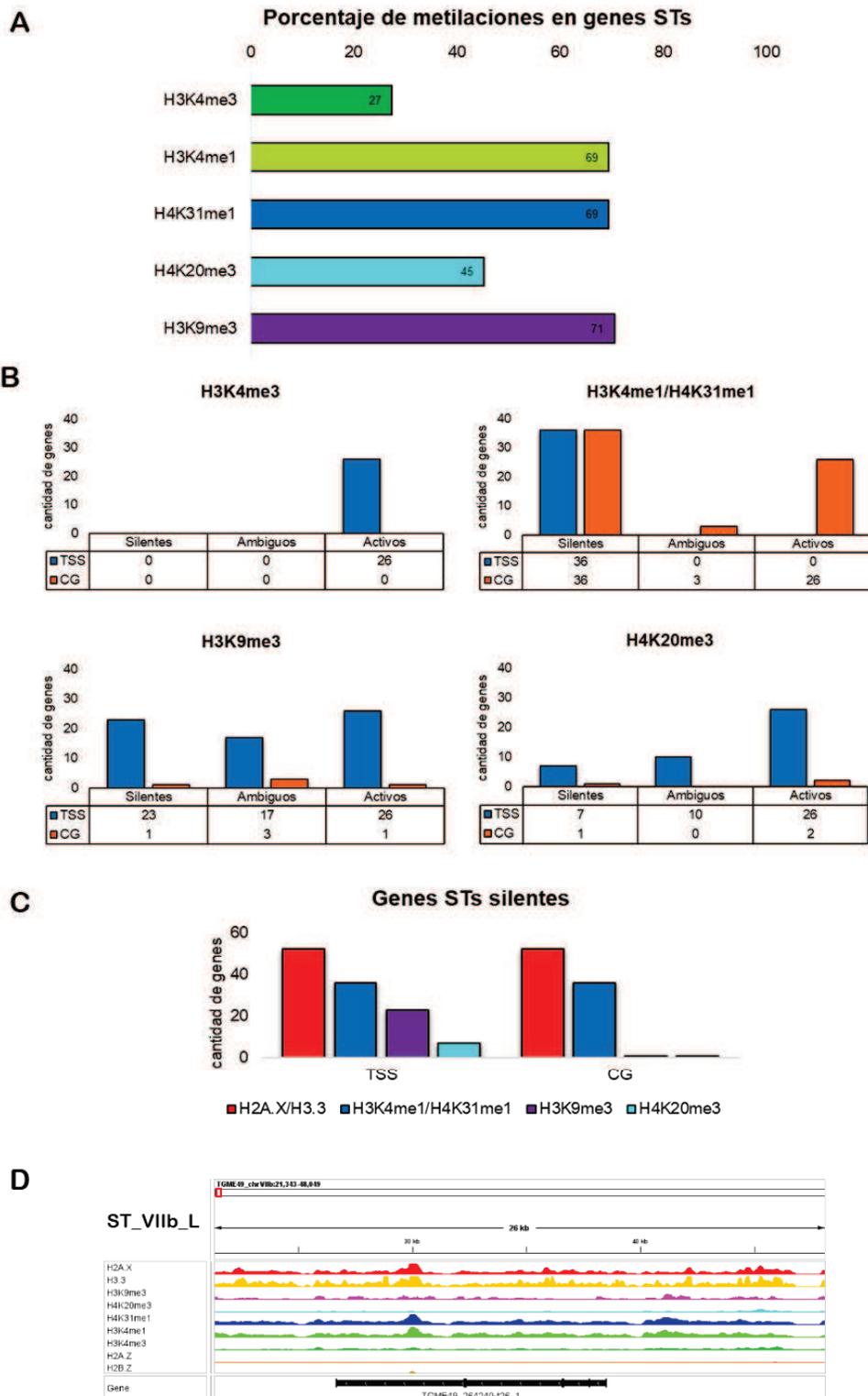


Fig. 23. Mapeo de H3K4me3 en regiones teloméricas.





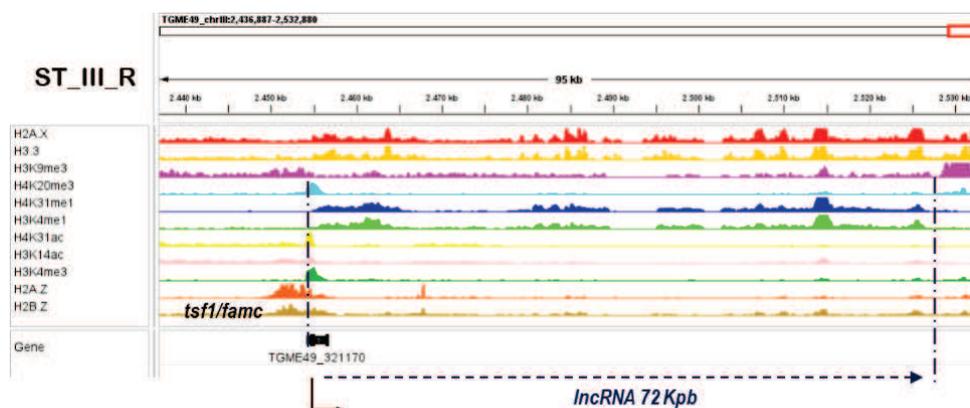
**Fig.24. Análisis de metilaciones específicas en los genes STs.** **A.** Porcentaje de metilaciones asociadas a los genes STs. El cálculo se realizó sobre el total de genes (95) que presentaban asociación o no a cada una de las marcas. **B.** Posicionamiento genómico de las metilaciones en relación a la estructura de los genes STs. **C.** Asociación de ocupación nucleosómica de Hv y metilaciones específicas sobre genes STs silentes. **D.** Ejemplo de posicionamiento de Hv y PTMs sobre TSS y CG de gen ST silente (TGME49\_264240).



### 5.4.3 Detección de putativos lncRNAs STs

De manera extraordinaria, durante el análisis de las metilaciones se han observado enriquecimientos de las marcas H4K31me1/H3K4me1 en regiones subteloméricas no génicas en los extremos cromosómicos. Este enriquecimiento en regiones no génicas fue previamente observado por Sindikubwabo y col., (2017) en el extremo del cromosoma III y en otras regiones cromosómicas internas argumentado que se correspondían con regiones codificantes de ARN largos no codificantes (lncRNA) no caracterizados de 20 a 70 Kpb (Sindikubwabo et al. 2017). En este sentido, cuando se analizó la posición genómica del lncRNA de 72 Kpb propuesta por Sindikubwabo y col., (2017) (TGME49\_chrIII:2,436,887-2,532,880) se encontró que estaba inmerso dentro del ST\_III\_R y que el inicio de esta región en el extremo interno correspondía con la región TSS del gen *Tsf1/famc* TGME49\_321170 (Fig. 25).

**Fig. 25. ARN no codificantes (lncRNA) en TgSTs.** Se muestra la posición genómica del lncRNA de 72 Kpb descrito por Sindikudwabo et al. 2017, correspondiente con el extremo derecho del cromosoma III junto a las Hv y las PTMs correspondientes. Se muestra, además, el posicionamiento genómico de las marcas de cromatina activa: H4K31ac y H3K14ac.



Por lo tanto, teniendo en cuenta esta conformación del enriquecimiento de H3K4me1/H4K31me1 por fuera de los cuerpos génicos (CGen) y la presencia de una región TSS con marcas activas (H4K31ac/H3K14ac/H3K4me3) (Sindikubwabo et al. 2017) y H2A.Z/H2B.Z (Nardelli et al. 2022) cercana a alguno de los extremos, se revisaron todas las regiones STs definidas en búsqueda de

posibles lncRNAs (*Fig. 8 a Fig. 20*). Mediante este análisis se detectaron regiones compatibles con putativos lncRNA subteloméricos, subTERRA, según lo planteado por Sindikubwabo y col., (2017) (*Fig. 25*). Estas regiones lncRNA presentaron: a) regiones TSS coincidentes con regiones TSS de genes activos en su mayoría en 15 STs (TSS<sup>+</sup>) (*Fig. 8B; Fig. 10C; Fig. 11B,C; Fig. 13C; Fig. 14B; Fig. 15A,B; Fig. 16A,B; Fig. 17B,C; Fig. 18B; Fig. 19B; Fig. 20C*) y b) regiones TSS compatibles con regiones activas pero no coincidentes con regiones TSS de genes en 6 STs (compatibles con TSS<sup>\*</sup>) (*Fig. 11C; Fig. 13B; Fig. 15A; Fig. 17A; Fig. 18B; Fig. 20B*) (*Tabla 3*). Mientras que 8 STs no mostraron marcas de H3K4me1/H4K31me1 extensas (>200bp) fuera del CGen, es decir, no mostraron regiones no génicas con enriquecimiento solo de H3K4me1/H4K31me1 (*Fig. 8C; Fig. 9A,B; Fig. 12A,B; Fig. 14A; Fig. 18C; Fig. 19C*). Los STs presentarían putativos lncRNAs asociados a regiones TSS activas en: ST\_Ia\_L, ~8,5 Kpb (*Fig. 8B*); ST\_II\_R(b), ~45 Kpb (*Fig. 10C*); ST\_III\_L, ~27 Kpb (*Fig. 11B*); ST\_III\_R(a), 72 Kpb; ST\_III\_R(b), ~62 Kpb (*Fig. 11C*); ST\_V\_R(a), ~57 Kpb; ST\_V\_R(b), ~62 Kpb (*Fig. 13C*); ST\_VI\_R(a), ~13Kpb; ST\_VI\_R(b), ~27 Kpb (*Fig. 14B*); ST\_VIIa\_L(a), ~18 Kpb; ST\_VIIa\_L(c), ~14 Kpb (*Fig. 15A*); ST\_VIIa\_R, ~78 Kpb (*Fig. 15B*); ST\_VIIb\_L, ~52 Kpb (*Fig. 16A*); ST\_VIII\_R, ~52 Kpb (*Fig. 16B*); ST\_IX\_L(a), ~48 Kpb (*Fig. 17B*); ST\_IX\_R, ~48 Kpb (*Fig. 17C*); ST\_X\_L(b), ~58 Kpb (*Fig. 18B*); ST\_XI\_L, ~57 Kpb (*Fig. 19B*) y ST\_XII\_R, ~40 Kpb (*Fig. 20C*). Por otra parte, en algunos STs la región compatible con lncRNA se superpone con la presencia de otros genes dentro de la región sin TSS con cromatina activa como por ejemplo el gen TGME49\_244930 dentro de ST\_VI\_R entre otros (ver *Tabla 3*). Esto abre la pregunta de si estos genes serían productos de los ARN no codificantes largos.

Otro aspecto fundamental es el posicionamiento de las PTMs fuera de los contextos de estructura génica. Tal es el caso de H3K4me3 en los extremos teloméricos y de H4K31me1/H3K4me1 en regiones no codificantes dentro de los STs. El posicionamiento de H3K4me3 indicaría la asociación de esta marca al funcionamiento y expresión en los telómeros como ARN no codificantes TERRA; mientras que el posicionamiento de H4K31me1/H3K4me1 a partir de regiones TSS de genes subteloméricos activos, en su mayoría de las familias b y c, indicaría la posibilidad de expresión de ARN largos no codificantes en los STs de *T. gondii*, subTERRA.

La mayoría de los genes que se encontraron delimitando la región enriquecida en H3K4me1/H4K31me1 pertenecen a familias multigénicas b y c, aunque también se observaron genes hipotéticos como ya se mencionó.

Por lo tanto, el análisis de la ocupación nucleosomal de las variantes de histonas y de las PTMs específicas nos abren el panorama acerca de la constitución de la cromatina en las regiones STs de *T. gondii*.

Regiones IncRNA	~Kpb	gen	ID	Transcrip.	condición
<i>TSS<sup>+</sup></i>					
ST_la_L	8,5	_287280	<i>tgc7</i>	53.0/36.9	activo
ST_II_R(b)	45	_298060	<i>tgc11</i>	14,5/32.6	activo
ST_III_L	27	_276230	<i>tgb5</i>	69.9/85.4	activo
ST_III_R (a)	72	_321170	<i>tgc1</i>	44.7/61.2	activo
ST_III_R (b) <sup>1</sup>	62	_298830	<i>tgh33</i>	70.0/53.1	activo
ST_V_R (a) <sup>4</sup>	57	_307480	<i>tgb14</i>	NDA	(?)
ST_V_R (b) <sup>5</sup>	62	_307460	<i>tgb13</i>	NDA	(?)
ST_VI_R (a) <sup>6</sup>	13	_292600	<i>tgh28</i>	84.6/85.8	activo
ST_VI_R (b) <sup>7</sup>	27	_244920	<i>tgb2</i>	NDA	(?)
ST_VIIa_L (a) <sup>8</sup>	18	_279550	<i>tgh23</i>	3.01/3.03	activo
ST_VIIa_L (c) <sup>10</sup>	14	_280810	<i>tgh25</i>	5.5/15.7	activo
ST_VIIa_R <sup>11</sup>	78	_282220	<i>ap2VIIa-9</i>	32.8/19.4	activo
ST_VIIb_L <sup>12</sup>	52	_264225	<i>tgb3</i>	NDA	(?)
ST_VIII_R	52	_200480	<i>tgb1</i>	NDA	(?)
ST_IX_L (a)	65	_274290	<i>tgb4</i>	91.8/87.7	activo
ST_IX_R <sup>14</sup>	48	_306680	<i>tgh38</i>	31.8/37.9	activo
ST_X_L (b) <sup>15</sup>	58	_228780	<i>tgc10</i>	15.2/23.9	activo
ST_XI_L	57	_307045	<i>tgb12</i>	63.7/37.9	activo
ST_XII_R	40	_276810	<i>tgh20</i>	NDA	(?)
<i>compatibles con TSS*</i>					
ST_II_L	10,5	(-)			(?)
ST_II_R(a)	25	(-)			(?)
ST_V_L (a) <sup>2</sup>	13	(-)			(?)
ST_V_L (b) <sup>3</sup>	48	(-)			(?)
ST_VIIa_L (b) <sup>9</sup>	25	(-)			(?)
ST_IX_L (b) <sup>13</sup>	35	(-)			(?)
ST_X_L (a)	44	(-)			(?)
ST_XII_L	23	(-)			(?)

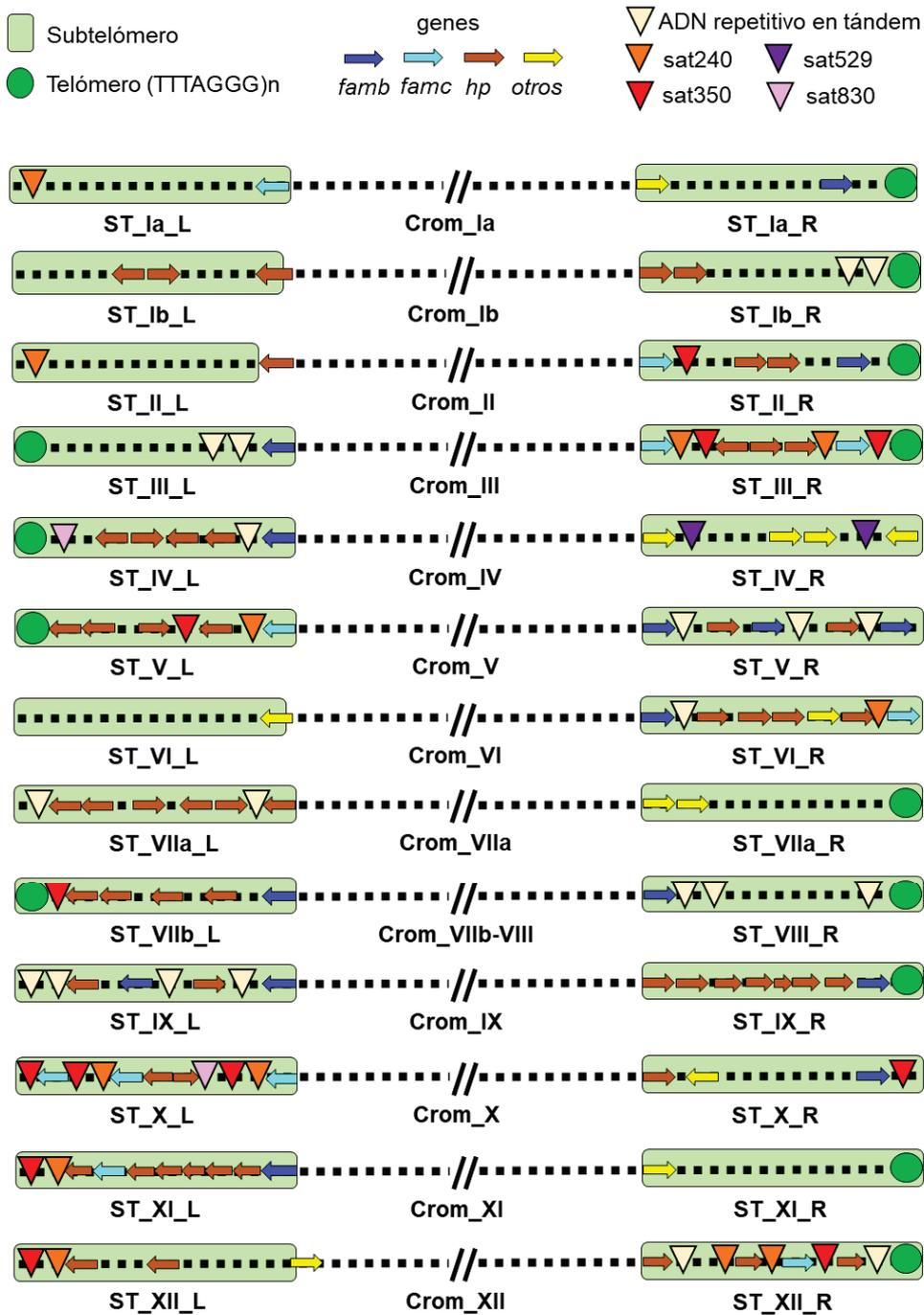
**Tabla 3. Putativos lncRNAs en subtelómeros de *T. gondii*.** TSS\*: regiones compatibles con TSS de genes STs. (a), (b), (c), etc; hacen referencia a más de una región lncRNA putativa, siendo la primera (a) las más cercana al telómero. TSS\*: putativos lncRNA definidos a partir de una región con marca compatible de TSS (H3K4me3/H2A.Z/H2B.Z) seguido de una región amplia con marca H3K4me1/H4K31me1 sin genes STs asociados. <sup>1</sup>La región putativa del lncRNA abarca una región interna la cual posee picos de H4K31ac/H2A.Z/H2B.Z aguas abajo del gen interno TGME49\_298820 ubicado dentro de la región sin TSS activo. <sup>2</sup>La región putativa del lncRNA abarca una región interna la cual posee picos de las PTMs: H4K31ac/H3K4me3/H2A.Z/H2B.Z. <sup>3</sup>La región putativa del lncRNA abarca una región interna la cual posee picos de H2A.Z/H2B.Z aguas abajo del gen interno TGME49\_220060 sin TSS activo. <sup>4</sup>La región putativa del lncRNA abarca una región interna la cual posee picos de las PTMs: H4K31ac/H3K4me3/H2A.Z/H2B.Z/H3K9me3/H4K20me3, sin estructura génica asociada. <sup>5</sup>La región putativa del lncRNA abarca una región interna la cual posee picos de las PTMs: H4K31ac/H3K4me3/H2A.Z/H2B.Z/H3K9me3/H4K20me3, aguas abajo del gen interno TGME49\_307470 sin TSS activo. <sup>6</sup>La región putativa incluye dos posibles genes internos: \_237700 y \_244950 sin TSS activos. <sup>7</sup>La región putativa incluye dos posibles genes internos: \_244930 y \_244940 sin TSS activos. <sup>8</sup>La región putativa incluye gen interno: \_279540 sin TSS activo. <sup>9</sup>La región putativa incluye gen interno: \_280070 sin TSS activo. <sup>10</sup>La región putativa incluye dos posibles genes internos: \_281040 y \_281030 sin TSS activos. <sup>11</sup>La región putativa incluye gen interno: \_282230 sin TSS activo. <sup>12</sup>La región putativa incluye genes internos: \_264230, \_264240, \_264250, \_264260 sin TSS activos. <sup>13</sup>La región putativa incluye gen interno: \_284280 sin TSS activo. <sup>14</sup>La región putativa incluye genes internos: \_306688, \_306692, \_306700, \_306710 sin TSS activos. <sup>15</sup>La región putativa del lncRNA abarca una región interna la cual posee picos de las PTMs: H2A.Z/H2B.Z, sin estructura génica asociada aguas abajo del gen *tgc10*.

## 6. Discusión del capítulo I

En este trabajo se encontraron 17 regiones STs nuevas, siete de ellas asociadas a los extremos en los brazos izquierdos cromosómicos: ST\_Ib\_L, ST\_II\_L, ST\_III\_L, ST\_IV\_L, ST\_VI\_L, ST\_VIIa\_L, ST\_VIIb\_L; diez asociadas a los extremos en los brazos derechos cromosómicos: ST\_Ia\_R, ST\_Ib\_R, ST\_V\_R, ST\_VI\_R, ST\_VIIa\_R, ST\_VIII\_R, ST\_IX\_R, ST\_X\_R, ST\_XI\_R, ST\_XII\_R; y una de ellas (ST\_VI\_R) asociada a un gen *Tsf famc* no descrito previamente (TGME49\_292610). También se corrigieron y re-definieron las nueve regiones TAS descritas previamente, ahora STs, seis en los extremos izquierdos: ST\_Ia\_L, ST\_V\_L, ST\_IX\_L, ST\_X\_L, ST\_XI\_L y ST\_XII\_L; y tres en los extremos derechos cromosómicos: ST\_II\_R, ST\_III\_R, y ST\_IV\_R (*Tabla 2*). En la *Fig. 26* se muestra para cada cromosoma el panorama global de la estructura de las nuevas regiones subtelméricas de *T. gondii* planteadas por este análisis, TgSTs.

A diferencia del trabajo de Dalmasso y col., (2014); el cual baso su búsqueda de regiones TAS mediante un análisis de genómica comparada fundamentada en que las regiones subtelméricas presentan un alto grado de conservación inter-cromosómica, aquí nos basamos en diferentes parámetros: el enriquecimiento evidente de H2A.X, la existencia de regiones extensas libres de genes, la presencia de ADN satelital, y en el caso de que hubieran genes, se reforzó con el hecho de que estos estuvieran en estado silente o ambiguo. Al momento, solo se habían detectado 9 de 26 posibles STs (Dalmasso et al. 2014). En otro trabajo, se basaron en que los STs son regiones con baja densidad génica, realizando un estudio estadístico desde los extremos comparando las distribuciones uniformes de genes contra la observada, con la consigna de que las regiones con distribuciones uniformes (asociadas a eucromatina) no son subtelméricas. En ese estudio, realizado en levaduras, definieron a los STs como una región de 33 Kpb (Brown, Murray, & Verstrepen 2010). No obstante, los STs pueden tener largos diferentes aún entre los cromosomas de la misma especie. El tamaño de las regiones subtelméricas en *Trypanosoma cruzi* varía de 5 a 182 Kpb y la más pequeña de estas regiones podría haber sido generada por una ruptura cromosómica reciente y un evento de reparación de los telómeros.

**Fig. 26. Mapa cromosómico putativo de las regiones STs de *T. gondii*.** A la izquierda se muestra el extremo izquierdo del cromosoma y a la derecha el extremo derecho de los cromosomas. Las regiones enriquecidas en H2A.X se presentan en verde claro. La presencia de repeticiones TTTAGGG en los telómeros se representan en círculos verde oscuro según la versión 53 del genoma de referencia ME49. La presencia de genes *famb* se representan en azul, de genes *famc* en turquesa, de genes hipotéticos (*hp*) en marrón y de otros genes en amarillo. Los triángulos corresponden a ADN satelital.



Las regiones subteloméricas en este parásito están ligadas a los genes *ts*, que codifican para proteínas de superficie con actividad de transialidasa, los cuales son genes que juegan un rol fundamental en la adaptación de *T. cruzi* a los distintos nichos (Moraes Barros et al. 2012). En ese sentido, es de alta relevancia conocer el conjunto de genes asociados a las regiones subteloméricas de *T. gondii* y sus posibles roles. En concordancia con lo observado en *T. cruzi*, se ha observado aquí que el tamaño de los TgSTs inferidos varía entre 13,2 a 232,4 Kpb abarcando en sus territorios uno, a varios genes, dándonos la pauta de que serían regiones dinámicas en las que probablemente estén asociados mecanismos de rearrreglos cromosómicos durante la recombinación genética.

Como se ha mostrado, el análisis genómico a través de la composición nucleosomal permitió tener un nuevo panorama acerca de la composición estructural cromosómica de los telómeros y más específicamente de las regiones STs de *T. gondii* en la cepa de referencia ME49. Se pudo observar que a diferencia de lo planteado por Dalmasso y col., (2014) las regiones STs estarían asociadas a más de una familia de genes específicos, tal como se analizará en detalle en el capítulo siguiente, y a variantes de ADN satelital. Estas cuestiones se hacen notar a medida que los curados de los genomas van avanzando y van apareciendo nuevas regiones que antes no habían sido correctamente ensambladas. Estas regiones nuevas muchas veces se corresponden con regiones repetitivas de los genomas que terminan resultando en huecos durante el ensamblaje, y que en ocasiones pueden conducir a predicciones incorrectas de la estructura cromosómica (Xia et al. 2021). Este hecho quedó demostrado en el trabajo comparativo realizado por Xia y col., (2021) y Berná y col., (2021). Dichos grupos demostraron que los cromosomas VIIb y VIII están fusionados, que no en todas las regiones de los extremos se detectaron secuencias repetitivas teloméricas, y que las regiones repetitivas asociadas a las regiones STs como, por ejemplo: TARE1, en algunos casos, serían más largas que las descritas por las actuales bases de datos. En ese sentido, en este trabajo, la presencia de ADN repetitivo en tándem, compatible con ADN satelital, fue observado en la mayoría de los STs redefinidos, pero sin seguir estrictamente el orden TARE1-3, tal como había sido descrito previamente (Dalmasso et al. 2014), sino mas bien se observó la presencia de diferentes regiones de ADN satelital,

sin relación entre ellas, salvo algunas familias (sat240, sat350); y con diferentes longitudes y dispersión. En general, el elemento repetitivo más consistente como ADN satelital corresponde a la familia sat350/sat680. De manera interesante, las unidades o monómeros de sat350 han mostrado alta variabilidad entre sí, sugiriendo incluso la posibilidad de la existencia de elementos de ADN repetitivo de alto orden (Clemente et al. 2004). Sat350 también mostró variabilidad en número entre cepas y también en número de copias en la progenie F1 respecto a sus parentales (Xia et al. 2021). Por lo pronto, la gran diversidad y plasticidad en número y secuencia de estos ADN satelitales en los STs de *T. gondii* sugieren una tasa de evolución y dinamismo importantes para ese dominio cromosómico. Su significado biológico es aún desconocido e intrigante. En este sentido, hay un trabajo que ha observado que las regiones repetitivas asociadas a las secuencias TARE de *P. falciparum* están asociadas a la replicación del ADN (Broadbent et al. 2011). Por lo tanto, resulta imprescindible ampliar el estudio y caracterización de estas secuencias repetitivas en los extremos teloméricos-subteloméricos para entender los mecanismos adaptativos de *T. gondii*.

En *Toxoplasma* se describieron ARN no codificantes pequeños codificados por el elemento repetitivo sat350 ubicado dentro del TgTARE1 los que a su vez estaban enriquecidos en marcadores H4K20me1, H4K20me3, H3K9me1 y H3K9me3 (Sautel et al. 2007; Braun et al. 2010). Estos ARN pequeños fueron clasificados como ARN de interferencia pequeños (ARNi) y asociados al complejo de silenciamiento inducido por ARN o RISC junto con la proteína TgAGO (Argonauta 1) descrita por primera vez en *Toxoplasma* (Braun et al. 2010) por lo que sería interesante detectar en el futuro los posibles blancos de ARNi provenientes de estos elementos repetitivos como así también estudiar sus patrones de expresión.

El enriquecimiento de H2A.X a lo largo de toda la región subtelomérica observada a través del análisis de los datos de Chip-Seq respalda el enriquecimiento de H2A.X observado por ChIP-qPCR por Dalmasso y col., (2014), en las regiones correspondientes al Fragmento A de las regiones TAS (Fig.5). De hecho, el estudio deja claro que principalmente los picos de H2A.X se detectaron

mayoritariamente hacia los extremos y luego van decayendo hacia la región interna de los brazos de los cromosomas. Eso podría estar indicando un rol importante de H2A.X en el efecto de silenciamiento telomérico o TPE, asociado a su vez con que este efecto se va tornando más débil a medida que se aleja de los telómeros (Kwapisz & Morillon 2020).

Los picos de H2A.Z/H2B.Z en regiones libre de genes es más intrigante. En aquellos casos donde además de detectarse H2A.Z/H2B.Z se detectó H3K4me3, se podría inferir la presencia de algún promotor de algún gen no identificado o de un promotor asociado a la expresión de miRNA (Mei et al. 2014) o de lncRNA en asociación a la presencia de otras marcas específicas como H3K4me1 y H4K31me1 (Sindikubwabo et al. 2017), que de hecho se ha observado. Por lo que, el enriquecimiento de H2A.Z/H2B.Z en regiones libre de genes apoyaría el hecho de la transcripción de lncRNAs subTERRA. Por otro lado, H2A.Z, en combinación con H3K4me1 y H3K27ac, se posiciona en *enhancers* activos (Jin et al. 2009; Heintzman et al. 2009), sugiriendo que la presencia de picos de H2A.Z/H2B.Z en regiones libres de genes podría corresponderse a *enhancers*. Finalmente, recientemente se observó que la histona H2A.Z se posiciona en los orígenes de replicación tempranos, facilitando la unión de H4K20me2 a dichos sitios, la cual es relevante para el reconocimiento de las proteínas ORC1 a dicho sitio y comenzar de esa manera la replicación del ADN (Long et al. 2019). En los casos de regiones STs, se trataría de orígenes de replicación tardíos (ST\_III\_R y ST\_X\_R). Sin embargo, podría ser una posibilidad a estudiar en el futuro.

Otra observación importante es el hecho de la presencia de H3K4me3 en los telómeros y en algunos casos también, enriquecimiento de H2A.Z y H2B.Z. En mamíferos, los telómeros consisten en repeticiones de TTAGGG organizadas en nucleosomas y asociadas con un complejo de seis proteínas conocido como Shelterinas, las cuales preservan la estructura y protegen los extremos de los cromosomas de la respuesta al daño del ADN (de Lange 2005). Estudios recientes han encontrado que los telómeros de mamíferos se transcriben en ARN telomérico que contiene repeticiones UUAGGG (TERRA) a partir de regiones teloméricas. La unión de TERRA en los telómeros pareciera estar implicada

principalmente en los mecanismos *cis* de la organización y el mantenimiento de la cromatina telomérica. Se demostró además que la enzima HMT de la histona 3 en la lisina 4 (H3K4me3) (con dominio SET) y el regulador transcripcional MLL se asocian con los telómeros y contribuyen a su metilación y transcripción de una manera dependiente de la longitud del telómero, influyendo en el estado general de la cromatina telomérica (Caslini et al. 2009). Un aspecto interesante es que los lncRNA TERRA no se expresan en todos los cromosomas, sino en algunos (uno o dos) como fue observado en mamíferos (Montero et al. 2016). Queda abierta la pregunta de si la presencia de H3K4me3 en el borde de los telómeros de varios cromosomas de *T. gondii* estarían implicados en la transcripción de varios ARN TERRA o también solo se expresan unos pocos de ellos.

El análisis de las regiones TSS de los genes STs ha permitido establecer distintos estados para la cromatina subtelomérica en la cual habría: genes silentes, ambiguos, activos.

Por otra parte, se ha descrito la asociación de las metilaciones específicas de heterocromatina como H3K9me3 y H4K20me3 que si bien se encontraron asociadas a promotores silentes también estarían enriquecidas en promotores activos y ambiguos. La marca de histona H3K9me3 en regiones TSS activas, junto a H3K4me3, podría deberse a una posible reacción cruzada del anticuerpo con H3K4me3 (K. Kim, comunicación personal). En forma similar podría ocurrir con la marca H4K20me3 en regiones TSS activas. Por lo tanto, se requieren de mayores estudios para confirmar la localización de esas histonas en los TSS activos.

Ahora bien, también se describieron otras marcas específicas en asociación a las regiones STs, H3K4me1 y H4K31me1, las cuales fueron planteadas como nuevos marcadores de regiones silenciadas que acompañan el enriquecimiento de H2A.X y H3.3 tanto en el TSS como en el cuerpo CGen de la mayoría de los genes STs silentes. Por otra parte, la presencia de estas marcas en regiones no génicas con algún extremo en cercanías a regiones TSS activas (presencia de H3K3me/H2A.Z/H2B.Z) podría estar indicando la presencia de lncRNA en los STs de *T. gondii* aún no descritos los cuales funcionarían en *cis* cerca del sitio de su producción (p.ej., podrían funcionar en la

homeostasis de los telómeros) o en *trans* para alterar la forma de la cromatina y la expresión génica en loci distantes, como se informó en otros organismos modelo (Azzalin & Lingner 2015) (Sindikubwabo et al. 2017). De hecho, Sindikubwabo y col., (2017) detectaron enriquecimiento de estas marcas en regiones TARE de *P. falciparum* por lo que se sugiere seguir avanzando de manera experimental en la comprobación de la expresión de estos lncRNAs putativos subTERRA de *T. gondii*.

## 7. Conclusiones del capítulo I

Basados en la densidad de genes se ha podido establecer las regiones STs de todos los cromosomas de *T. gondii*, las cuales presentan como característica general un enriquecimiento notorio de H2A.X y H3.3 hacia sus extremos con el telómero; la identificación de numerosos genes únicos y dos familias multigénicas, cuyas regiones TSS presentan diferentes composiciones de cromatina según estuvieran activas, ambiguas o silentes. La presencia de picos aislados de H3K4me3 fuera de los genes podría estar implicada en la expresión de lncRNAs TERRA. Mientras que la presencia de H3K4me1 y H4K31me1 en regiones no génicas podría indicar la presencia de lncRNA en los STs.

La identificación de numerosas regiones de ADN satélite en la mayoría de los STs, sin relación evidente entre ellos en cuanto a origen evolutivo, estaría indicando que los STs de *T. gondii* podrían presentar una dinámica de evolución rápida. Finalmente, el hallazgo de numerosos genes, entre ellos, las familias multigénicas, permite establecer que las regiones TgSTs podrían tener roles en aspectos de la biología del parásito que pueden estar relacionados a resistencia al estrés, virulencia, patogenicidad, etc. Todos estos aspectos podrían ser de gran relevancia para comprender la evolución y adaptabilidad de *T. gondii*.

## CAPÍTULO II | GENES SUBTELOMÉRICOS DE *T. GONDII*

## 8. Familias de genes subteloméricos

Si bien los STs son conocidos como regiones pobres en genes, en muchos organismos se observó la presencia de algunos grupos de genes cuyas funciones pueden ser relevantes para los mismos. Es así que en *Trypanosoma sp* y *Plasmodium sp*, se encuentran una serie de familias multigénicas que codifican para proteínas de membrana cuya función está asociada a la evasión de la respuesta inmune y a la patogenicidad (Omelianczyk et al. 2020; Ramirez 2019). En levaduras se ha observado el enriquecimiento subtelomérico de genes relacionados con la respuesta al estrés y al metabolismo que son esenciales en condiciones de crecimiento no óptimas (Ottaviani, Gilson, & Magdinier 2008). En especies de hongos patógenos como *Candida albicans* y *Candida glabrata* existen familias de genes subteloméricos (genes ALS y genes EPA, respectivamente), que codifican para proteínas responsables de la adherencia a los tejidos de mamíferos (Cormack, Ghori, & Falkow 1999; de Las Peñas et al. 2003; Hoyer 2001). En *Saccharomyces cerevisiae* los genes subteloméricos *flo* codifican proteínas que confieren adherencia al agar, superficies sólidas y a otras células de levadura (Guo et al. 2000; Reynolds & Fink 2001). En estos ejemplos, la proteína codificada por cada miembro de las familias mencionadas es capaz de alterar las propiedades de la superficie celular logrando generar variación fenotípica en respuesta al ambiente del hospedador (Halme et al. 2004).

En *T. gondii* se había asociado a las secuencias TgTAS (aquí TgSTs) a la familia multigénica FamC específica de la especie, denominada *Tsf* (aquí, *tgc*) (Dalmasso et al. 2014; Lorenzi et al. 2016). En el capítulo anterior encontramos que en las regiones STs definidas existiría otra familia multigénica (FamB, aquí *tgb*) y genes de copia única tanto con identidad asociada como hipotéticos. En este capítulo describiremos los genes presentes en las regiones STs propuestas.

## 9. Hipótesis del capítulo II

La identificación, descripción y análisis *in silico* de los genes subteloméricos de *T. gondii* es esencial para detectar grupos de genes asociados por alguna condición metabólica, localización subcelular y/o familias específicas de especie, y/o habilitar un conocimiento que facilite futuros estudios de genes de interés.

## 10. Objetivo del capítulo II

Indagar acerca de la estructura génica y proteica de las familias de genes subteloméricos y la posible función en la biología del parásito.

### 10.1. Objetivos específicos del capítulo II

- ✓ Describir *in silico* los genes asociados a TgST, analizar la expresión génica, las estructuras proteicas correspondientes e indagar acerca de su origen filogenético y sintenia.
- ✓ Analizar las proteínas FamB y FamC como posibles factores de virulencia.
- ✓ Caracterizar la expresión génica de familia FamC en *T. gondii*.
- ✓ Caracterizar la expresión proteica de familia FamC en *T. gondii*.

## 11. Resultados del capítulo II

### 11.1 Descripción de los genes subteloméricos

Las regiones subteloméricas de *T. gondii* (TgSTs) se encontraron asociadas a un total de 96 genes, donde solo el 11% se correspondió con la familia multigénica *Tsf* (*famc/tgc*) descrita por Dalmasso et al. 2014 (**Fig.27A**). Se pudo observar que un 16% se correspondió con la familia multigénica *famb/tgb* y un 14% con genes de copia única con alguna función anotada. Un gran porcentaje de genes (59%) se correspondió a genes únicos hipotéticos sin aparente relación evolutiva entre ellos. Sin embargo, hemos visto que la contribución de los genes delimitando las regiones STs sería bastante proporcionada entre los genes multigénicos y de copia única, aunque los genes *tgb* parecerían ocupar prioritariamente los límites propuestos (**Fig.27B**). Mediante un primer análisis del estado de expresión de estos genes con datos de transcriptómica recabados de la base de datos ToxoDB se pudo ver que muchos de ellos se encontrarían en estado activo de expresión (**Fig.27C**). La mayoría de los genes únicos con función anotada (FA) se transcribirían y un 60% de los genes *tgc* también lo haría. Mientras que los genes *tgb* y *hp* presentarían una mayor proporción de genes en estado inactivo de transcripción (**Fig.27C**).

A continuación, se detallan las características de los genes encontrados en asociación a las regiones TgSTs junto a la descripción de su composición, función, estructura y expresión según la base de datos ToxoDB.

#### 11.1.1 Genes de copia única

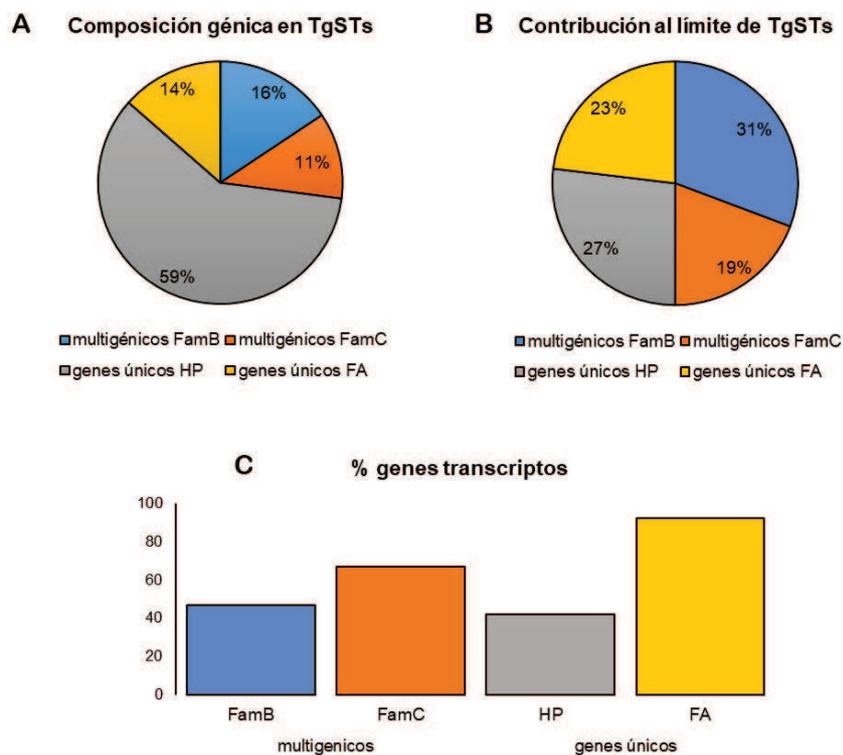
Se describe por primera vez en este análisis la presencia de genes de copia única asociados a las regiones TgSTs. En la **Tabla 4** y **Tabla 5** se describen las características de estos genes, ya sea aquellos que fueron anotados con su identidad funcional como los que son hipotéticos, respectivamente.

##### 11.1.1.1 Genes con función anotada

Se detectaron 13 posibles genes STs cuya anotación se refiere a la identidad del gen con respecto a la función asignada, cuyo análisis está detallado en la **Tabla 4**. Doce de ellos presentan datos de expresión de mRNA en las cepas ME49 y RH, siendo ADP/ATP traslocasa el único sin datos de

transcripción. Once muestran datos de proteómica. En general, los datos de transcripción acompañaron la descripción de su composición nucleosómica en la región TSS definida en el capítulo anterior.

**Fig.27. Genes subteloméricos de *T. gondii*.** **A.** Composición total de familias multigénicas y genes únicos. Los valores son en porcentaje relativo al total de los genes descritos (96). HP: hipotéticos, FA: con función asignada. **B.** Los valores son relativos al total de genes que delimitan una región ST (26). **C.** Porcentaje de genes que se encontraron asociados a datos de transcriptómica en ToxoDB (se muestra la proporción en relación al total de genes cada grupo).



Resulta muy interesante destacar los valores fenotípicos encontrados para estos genes. Los valores fenotípicos están basados en un experimento de disrupción génica masiva mediante la tecnología CRISPR/Cas9, el cual permitió identificar genes esenciales de aquellos que no lo son durante el ciclo lítico de *T. gondii in vitro* (Sidik et al. 2016). Los genes que contribuyen a la aptitud del parásito, es decir que no fueron detectados o fueron débilmente detectados luego de dos días de cultivo, estarán representados por puntuaciones negativas ya que la pérdida de la función del gen reduce la aptitud y, por lo tanto, menos parásitos sobreviven a tres ciclos líticos. Los valores positivos indican que esos genes

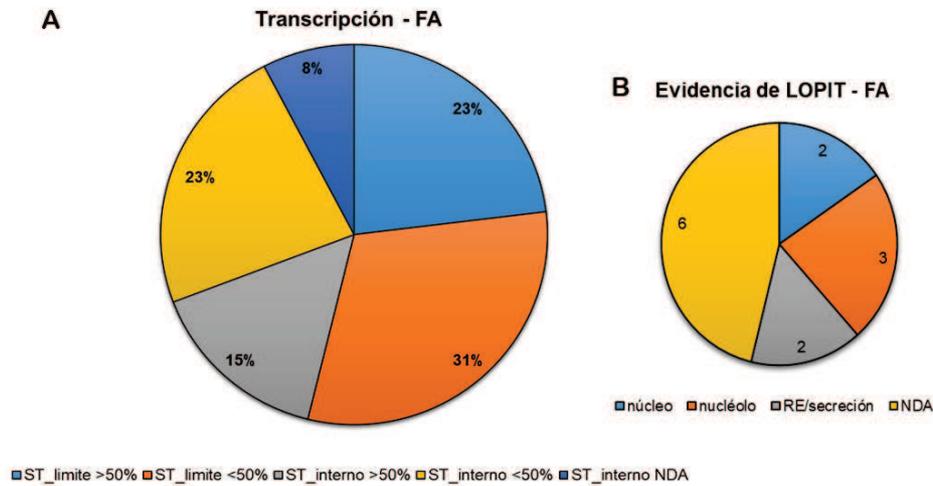
no están asociados a la supervivencia del parásito en el ciclo lítico. Las puntuaciones del fenotipo encontradas en este conjunto de datos van desde -6,89 a +2,96, indicando, los valores muy negativos, una alta probabilidad de indispensabilidad de función para dicho gen. Por lo tanto, se han observado aquí 5 genes con un valor fenotípico muy negativo en escala, correspondientes a los genes *chapA-tub* (TGME49\_295310), *otu-cp* (TGME49\_237894), *eIF3* (TGME49\_317720), *pp-i* (TGME49\_238000) y *vtc2* (TGME49\_298630) (**Tabla 4**), los cuales serían indispensables para la biología del parásito y podrían ser considerados como posibles blancos terapéuticos.

Como se ha mencionado, varios de estos genes limitarían internamente las regiones STs, mientras que otros tendrían localización interna, por lo tanto, hemos analizado si había una correspondencia entre los niveles de transcripción observados y sus localizaciones en el subtelómero. En la **Fig.28A** comparamos genes anotados (FA) bordeando las regiones STs o internos con sus porcentajes de transcripción observando que un 23% de los genes localizados en el límite tendrían niveles de transcripción elevados mientras un 31% de ellos ocuparía una región limitante tendría niveles bajos de transcripción. Esto podría indicar que algunos de los genes identificados en el borde podrían estar sujetos a TPE.

En total, solo dos de 5 genes codificarían proteínas nucleares, un factor de transcripción relacionado a la cromatina, *ap2VIIa-9* y un factor de elongación no relacionado a la cromatina sino a la traducción, *eIF3*; y al menos 3 de genes codifican proteínas detectadas en el nucléolo por la técnica experimental de marcado y localización proteica, LOPIT (Barylyuk et al. 2020), *chapA-tub*, *pp-i* y *enoil-coA* (**Tabla 4**) (**Fig.28B**). Por otro lado, dos de los 13 genes, *vtc2* y *oxidoreductina*, lo harían para proteínas del retículo endoplasmático y/o proteínas vacuolares, en el caso de *vtc2*, estaría relacionada además al Golgi y a las proteínas de membrana (**Tabla 4**) (**Fig.28B**). Estos resultados sugieren un cierto sesgo entre los genes presentes en TgSTs con respecto a la localización subcelular de sus productos, al menos los de localización nuclear.

**Fig.28. Evidencia de transcripción y localización de genes únicos con función asignada (FA).**

**A.** Evidencia de transcripción génica vs localización génica. **B.** Compartimentos asignados. RE: retículo endoplasmático, NDA: *not data assigned*. <https://toxodb.org/toxo/app>.



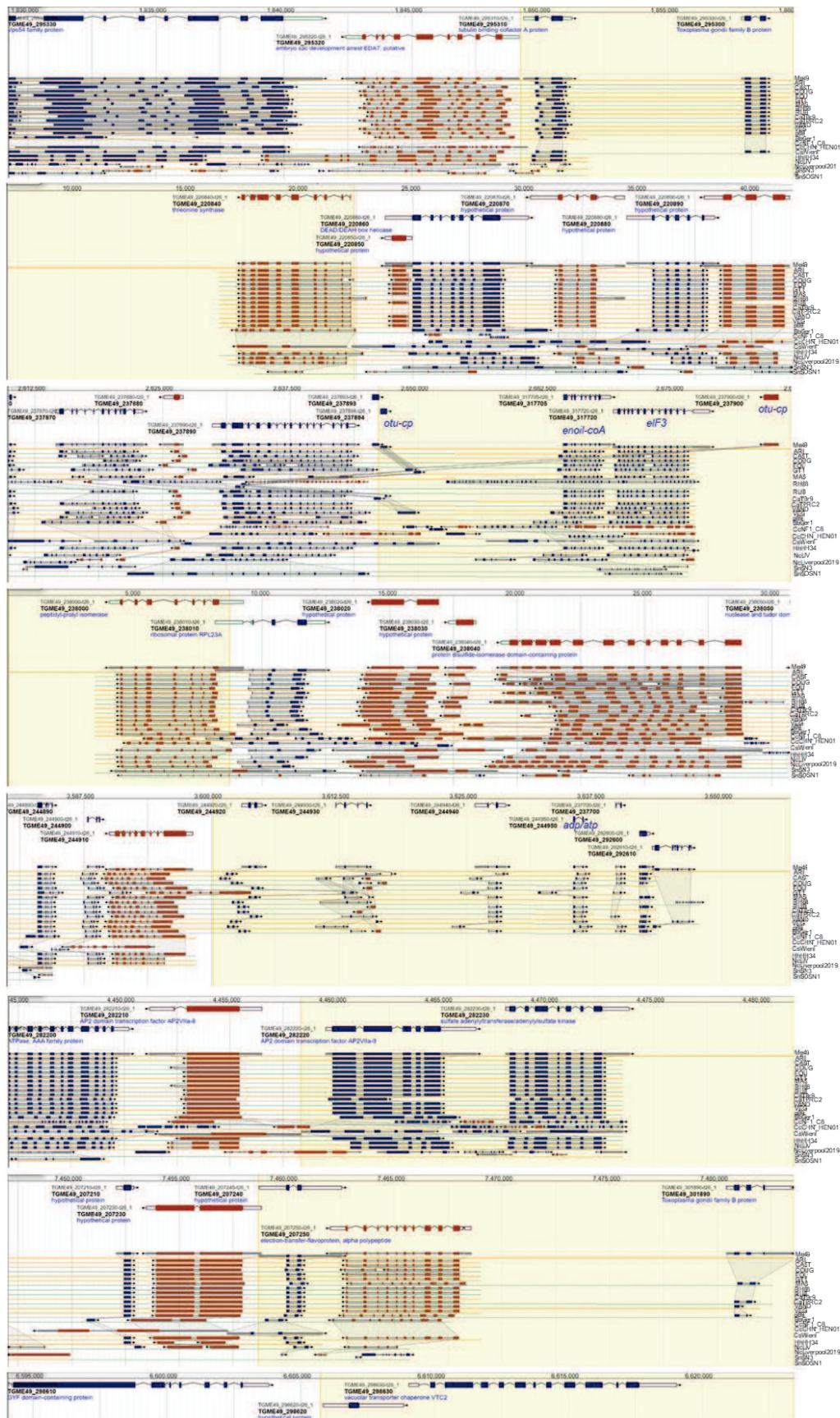
Una observación destacada con respecto a algunos de estos genes es que al analizar la sintenia (conservación de la localización génica en el cromosoma) en todas las cepas de *T. gondii* (Tg) y especies relacionadas (*H. hammondi*: Hh, *N. caninum*: Nc, *Besnoitia besnoiti*: Bb, *Cyclospora cayetanensis*: Cc, *Cystoisospora suis*: Cs, *Sarcocystis neurona*: Sn) en la base datos (ToxoDB) (Fig.29), se pudo ver que algunos de ellos están conservados en todas las cepas y especies relacionadas (TgHhNc/conservados) como *chapA-tub*, *ts*, *pp-i*, *adp/atp traslocasa*, *ap2Vlla9*, *sa-t/k* y *oxired-RE*; y/o conservados pero reorganizados (con variaciones leves en la localización genómica) como *enoil-coA*, *eIF3* y *electron-t*. Mientras que, otros genes solo se encuentran en *T. gondii* (Tg/conservados) y no en especies relacionadas (no se encontró correspondencia en *H. hammondia* o *N. caninum*) como *adp/atp traslocasa*, o, únicamente en algunas de las cepas de *T. gondii* (parcialmente conservados) y reorganizados como *otu-cp* TGME49\_237894. El gen TGME49\_237900 solo esta presente en la cepa de *T. gondii* ME49 (cepa específico). Otros estarían duplicados como *ts* (COUG) y *vtc2* (p89), apoyando

Anotación	Gen ID	ST	AA <sup>1</sup>	Motivo/Dominio	SP	TM	SP+TM	Núcleo <sup>2</sup>	Fenotipo	ME49/RH <sup>3</sup>	PR	LOPIT	Función posible
Chaperona A de unión a tubulina	_295310	la_R	189	TBCA ( <i>Tubulin Binding Cofactor A</i> )	no	no	no	si	-4,78	85.7/85.6	si	nucleolo	ensamblaje de complejos proteicos, plegamiento de proteínas, unión a proteínas citoesqueléticas
Treonina sintasa	_220840	ll_L	571	TS ( <i>Treonine Sintase</i> )	no	no	no	si	1,76	20.7/7.66	ooquiste proteoma	NDA	proceso metabólico, actividad treonina sintasa, unión de piridoxal fosfato
Cisteína proteasa de la familia OTU-I	_237894	IV_R	219	OTU-like cysteine protease	no	no	no	no	-4,37	39.9/38.8	NDA	NDA	actividad peptidasa
Enoil-CoA hidratasa / isomerasa	_317705	IV_R	371	Crotonase/Enoyl-CoA hydratase/isomerase	no	no	no	si	-0,45	93/92.3	si	nucleolo	actividad catalítica, actividad enoil-CoA hidratasa, actividad isomerasa
Subunidad 7 del factor 3 de iniciación de la traducción eucariota	_317720	IV_R	609	eIF-3_zeta	no	no	no	si	-4,87	90.8/92.7	si	núcleo- no cromatina	iniciación de la traducción, actividad del factor de iniciación de la traducción
Cisteína proteasa de la familia OTU-II	_237900	IV_R	486	OTU-like cysteine protease	no	no	no	no	NDA	39.9/38.8	si	NDA	actividad peptidasa
Peptidil-proil isomerasa	_238000	VI_L	283	Cyclophilin type peptidyl-prolyl cis-trans isomerase	no	no	no	si	-6,27	83.2/87.1	mitocondria	nucleolo	proceso de modificación de proteínas celulares, plegamiento de proteínas, actividad de isomerasa
ADP/ATP translocasa	_237700	VI_R	61	Solute carrier	no	no	no	no	-0,76	NDA	NDA	NDA	NDA
AP2Vlla-9	_282220	Vlla_R	1271	AP2 domain (2)	no	no	no	si	-0,67	32.8/19.4	si	nucleo-cromatina	regulación de la actividad del factor de transcripción de unión a ADN, plantilla de ADN, transcripción
Sulfato adenililtransferasa / adenililsulfato quinasa	_282230	Vlla_R	634	Adenylylsulphate kinase	no	no	no	no	0,94	10.2/6.1	ooquiste proteoma	NDA	asimilación de sulfato, actividad de adenililsulfato quinasa, actividad de sulfato adenililtransferasa (ATP)
Flavoproteína de transferencia de electrones	_207250	X_R	450	ETF_alpha	no	no	no	no	0,73	33.4/16.6	ooquiste proteoma	NDA	NDA
Chaperona transportadora vacuolar VTC2	_298630	XI_R	1308	SPX, VTC2/SPXD	no	2	si	si	-3,3	31.9/35.4	si	ER/Golgi/PM	Unión de polifosfato de inositol / fusión de membrana vacuolar
Oxidoreductina de retículo endoplásmico	_300380	XII_L	587	Endoplasmic Reticulum Oxidoreductin 1 (ERO1)	no	si	si	no	0,82	63.7/74.8	si	ER	proceso de oxidación-reducción, actividad de isomerasa de disulfuro de proteína, unión a dinucleótidos de flavina y adenina

**Tabla 4. Descripción de genes únicos anotados asociados a TgSTs.** ID: TGME49. <sup>1</sup>Los genes pueden tener errores de anotación. <sup>2</sup>Predicción de localización nuclear por PsortII o evidencia de LOPIT. <sup>3</sup>Los sets de datos corresponden a porcentajes de perfiles de expresión de tres líneas arquetipo según base de datos ToxoDB de las cuales solo se muestran dos: ME49 y RH (puede haber evidencia de otros datos de transcripción no analizados). SP+TM: posible proteína de secreción. NDA: *not data available*. PR: proteómica.

la idea de que estas serían regiones sujetas a recombinación genómica (**Fig.29**). Cabe destacar que la mayoría de estos genes presentan en la cepa ME49 una expansión del tamaño de los exones e intrones que componen sus estructuras génicas, como por ejemplo *chapA-tub*, *enoil-coA*, *eIF3*, *ppi*, *ap2Vlla9*, *sa-*

**Fig.29. Sintenia de genes únicos anotados asociados a TgSTs.** El recuadro amarillo denota la región ST definida en cada caso. *H. hammondi*: Hh, *N. caninum*: Nc, *Besnoitia besnoiti*: Bb, *Cyclospora cayetanensis*: Cc, *Cystoisospora suis*: Cs, *Sarcocystis neurona*: Sn.





*t/k*, *oxired-RE* y *vtc2* (**Fig.29**), lo que apoyaría un gran dinamismo evolutivo de esta región entre las diferentes cepas.

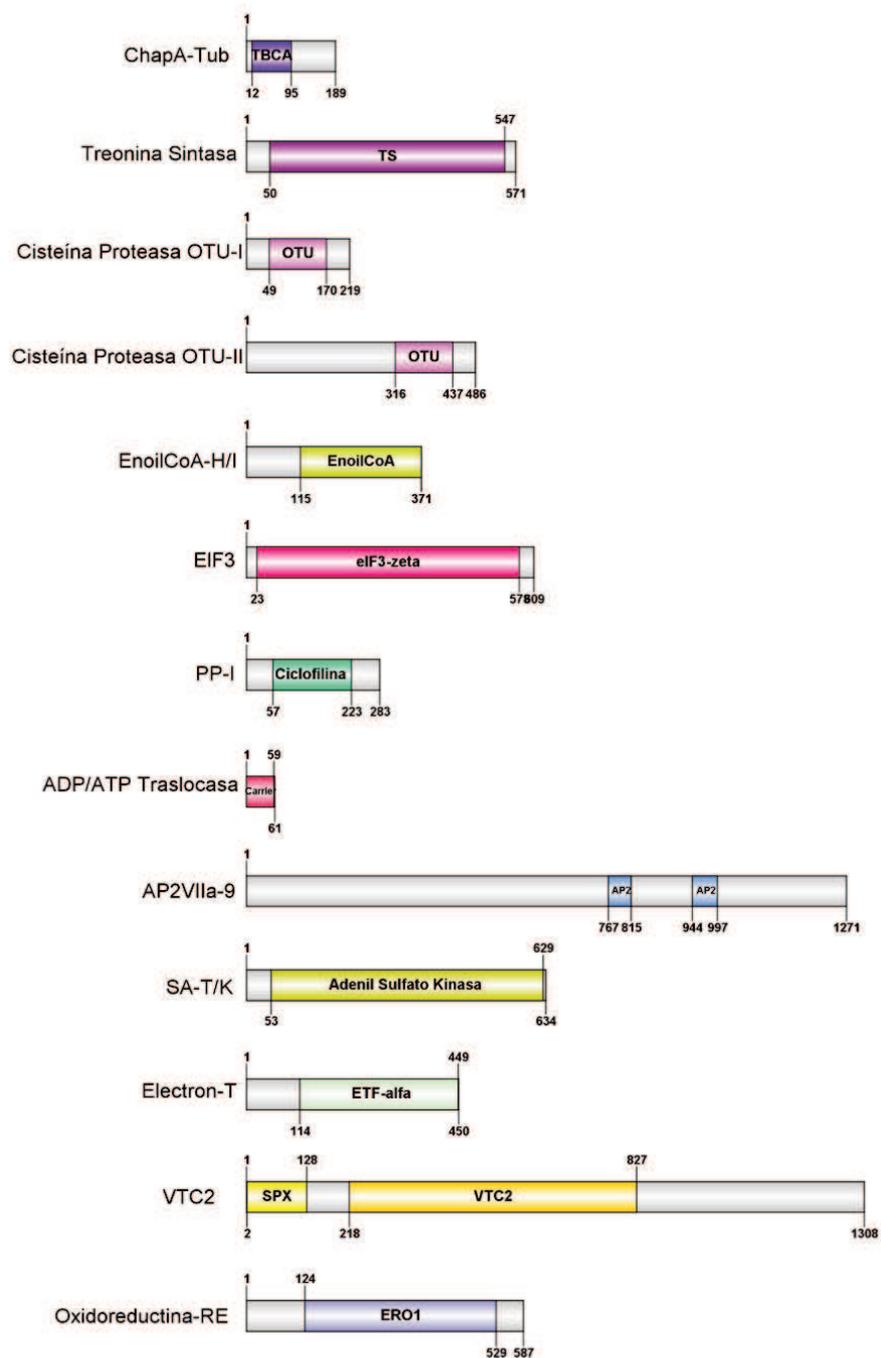
#### 11.1.1.1.1 Características, dominios y motivos de las proteínas únicas con función asignada

Los productos de los genes únicos anotados no mostraron presencia de péptido señal, aunque algunos de ellos tendrían dominios transmembrana indicando posibilidad como proteína de membrana expresada en la vía del retículo endoplasmático, RE, o el aparato de Golgi (**Tabla 4**). Los diferentes productos proteicos de los genes fueron analizados por presencia de motivos y dominios conservados empleando predictores por homología de secuencias: [https://myhits.sib.swiss/cgi-bin/motif\\_scan](https://myhits.sib.swiss/cgi-bin/motif_scan) y <https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi> respectivamente (**Tabla 4**). Para establecer la significancia de los mismos se tomó en cuenta que los dominios y motivos encontrados coincidieran en los dos predictores utilizados y que los valores de *E-value* sean similares en ambos casos. En la **Fig.30** se muestran los dominios detectados. Abajo se describen las funciones de los mismos para cada una de las proteínas identificadas.

La proteína Chaperona A de Tubulina (ChapA-Tub) (**Tabla 4**) presenta un dominio de unión cl03807: correspondiente al cofactor A de unión a  $\beta$ -Tubulina (primer cofactor identificado que participa en el proceso de plegamiento de Tubulina "*in vitro*") (Tesis Doctoral Kerman Aloria, 1990) (**Fig.30**). Se encontró que dicha proteína tendría expresión y localización asociada al nucléolo (**Tabla 4**), en

concordancia con su localización detectada en el Atlas humano de proteínas (<https://www.proteinatlas.org/>).

**Fig.30. Productos proteicos posibles de genes únicos anotados en asociación TgSTs.** Los números indican la posición de aminoácidos predicha para cada dominio según la base de datos del NCBI <https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>. Ilustrado con IBS (W. Liu et al. 2015).



La proteína Treonina sintasa presenta un dominio PRK09225 perteneciente a la super familia de TS (**Fig.30**) y participaría en procesos metabólicos de unión a fosfato piridoxal, PLP, a través de su actividad treonina sintasa, enzima dependiente de PLP que cataliza la última reacción en la síntesis de treonina a partir de aspartato (Thomazeau et al. 2001). Dicha proteína se encontró asociada a datos de proteómica del estadio de ooquiste y su localización por LOPIT es nuclear (**Tabla 4**). Cabe destacar que TS tiene una débil localización mitocondrial en el Atlas humano de proteínas (<https://www.proteinatlas.org/>).

La familia de proteasas de cisteína tipo OTU (del inglés: *Ovarian Tumor*), OTU-I y OTU-II, tendrían homología a la proteína OTU encontrada en *Drosophila*, virus y bacterias patógenas las cuales poseen actividad peptidasa en varios procesos celulares (Makarova, Aravind, & Koonin 2000) pero que sin embargo, no han podido ser ligadas aún a algún compartimento celular específico en *T. gondii* (**Tabla 4**).

En la proteína EnoilCoA (ECH-1) encontramos asociación a un dominio cl23717 perteneciente la superfamilia de Crotonasas EnoilCoA-Hidratasa/Isomerasas (**Fig.30**) las cuales poseen roles muy importantes en el metabolismo de ácidos grasos, la cual fue detectada en la fracción nucleolar (**Tabla 4**). Cabe destacar que ECH-1 tiene una localización mitocondrial en el Atlas humano de proteínas (<https://www.proteinatlas.org/>).

La subunidad 7 (D o EIF3\_zeta) del factor 3 de iniciación de la traducción eucariota, EIF3, posee un dominio eIF-3 zeta (**Fig.30**) (**Tabla 4**). El factor de iniciación eucariota 3 es un complejo de múltiples subunidades que se requiere para la unión del ARNm a las subunidades ribosómicas 40S, la estabilización de la unión del complejo ternario a las subunidades 40S y la disociación de las subunidades 40S y 60S. Estas funciones y la naturaleza compleja de EIF3 sugieren múltiples interacciones con muchos componentes de la maquinaria de traducción. En nuestro caso encontramos evidencia de localización nuclear no asociada a la cromatina (**Tabla 4**). Cabe destacar que EIF3 tiene una localización citosólica en el Atlas humano de proteínas (<https://www.proteinatlas.org/>).

La proteína PP-I, Peptidil-Protil Isomerasa, también se encontró asociada a la fracción mitocondrial y al nucléolo, presenta un dominio perteneciente a la superfamilia de las Ciclofilinas: cl00197, las cuales son enzimas chaperonas que aceleran el plegamiento proteico y tienen actividad isomerasa (**Fig.30**) (**Tabla 4**). Esta proteína tiene mayor identidad con PPI-G de humanos la cual estaría localizada en el núcleo y citosol según el Atlas humano de proteínas (<https://www.proteinatlas.org/>)

La proteína *carrier* ADP/ATP Traslocasa (TGME49\_237700) es una proteína pequeña que presenta motivos Solcar y Mito\_carrier (*carrier* de solutos asociados a la mitocondria). Esta proteína presenta un dominio cl36523: PTZ00169, perteneciente a la superfamilia de transportadores ADP/ATP o Adenilato Traslocasas (**Fig.30**). Cabe mencionar que esta superfamilia de proteínas aparecería únicamente en los Apicomplexa, y su inhibición en *P. falciparum* detendría el desarrollo del parásito en los eritrocitos (Razakantoanina, Florent, & Jaureguiberry 2008). No se encontraron datos de proteómica y/o de localización celular para esta proteína (**Tabla 4**).

AP2VIIa-9 pertenece a la superfamilia de factores de transcripción con dominio Apetala 2 de plantas, cl00033. AP2VIIa-9 presenta dos de estos dominios en su secuencia (**Fig.30**). Los datos de proteómica mostraron expresión de dicho FT en parásitos intracelulares 36 horas (h) post-infección y localización nuclear asociada a la cromatina (**Tabla 4**).

La proteína SA-T/K presenta un dominio: cl32078: PRK05537 perteneciente a la superfamilia de quinasas bifuncionales sulfato adenil transferasa quinasa (**Fig.30**) y su expresión estaría asociada al estadio de ooquiste al igual que la subunidad alfa de la flavoproteína de transferencia de electrones, Electron-T (**Tabla 4**) (**Fig.30**). Estaría localizada en vesículas según el Atlas humano de proteínas (<https://www.proteinatlas.org/>).

La proteína VTC2 (**Fig.30**) sería una proteína chaperona de transporte vacuolar que mostró señales de localización en RE, Golgi y membrana; con presencia de dos dominios (**Tabla 4**), el dominio cl34877: VTC/SPXD implicado en la acumulación de poli-fosfato vacuolar (transporte y metabolismo de

iones inorgánicos, tráfico intracelular, secreción y transporte vesicular) y el dominio cl21499, de la superfamilia SPX, donde los miembros de esta familia están involucrados en la transducción de señales.

Finalmente se muestra la estructura de la proteína Oxidorreductasa de retículo endoplásmico (*Fig.30*), la cual estaría asociada a procesos de oxidación-reducción, actividad isomerasa de disulfuro de proteínas y unión a dinucleótidos de flavina y adenina en el RE (*Tabla 4*).

#### **11.1.1.2 Genes hipotéticos**

En total se encontraron 57 genes hipotéticos asociados a las regiones TgSTs, donde 2 de ellos son compatibles con genes similares a la Familia B (*Tabla 5*). En general, la gran mayoría de las regiones TSS de estos genes se observaron enriquecidas en nucleosomas con H2A.X, es decir, en un posible estado de cromatina silente (Capítulo I). Para el 42% de los genes analizados se encontraron datos que avalan la evidencia de existencia génica como tal, debido a la presencia de datos de transcripción/proteómica (evidencia de Espectrometría de Masas) y/o LOPIT recabados de la base de datos ToxoDB, mientras que el 49% restante solo presentaría datos fenotípicos, y el 9% ningún dato (*Fig.31A*). Del 42% (24 genes), más de la mitad presenta un percentil de transcripción por debajo del 50% (*Fig.31B*). Dentro de este grupo, 1 gen codificaría un producto que tendría homología a la Familia B, TGME49\_321280 (*tgh57*); y otros 6 de ellos; *tgh4*, *tgh18*, *tgh23*, *tgh28*, *tgh29* y *tgh38* tendrían asociación a dominio TM y/o péptido señal (*Tabla 5*). Del grupo que no tiene evidencia de transcriptómica, 1 de ellos, TGME49\_321270 (*tgh56*), también codificaría un producto homólogo a la Familia B y 3 de ellos; *tgh20*, *tgh31* y *tgh43* codificarían productos cuyo único indicio es la presencia probable de un dominio putativo TM (*Tabla 5*).

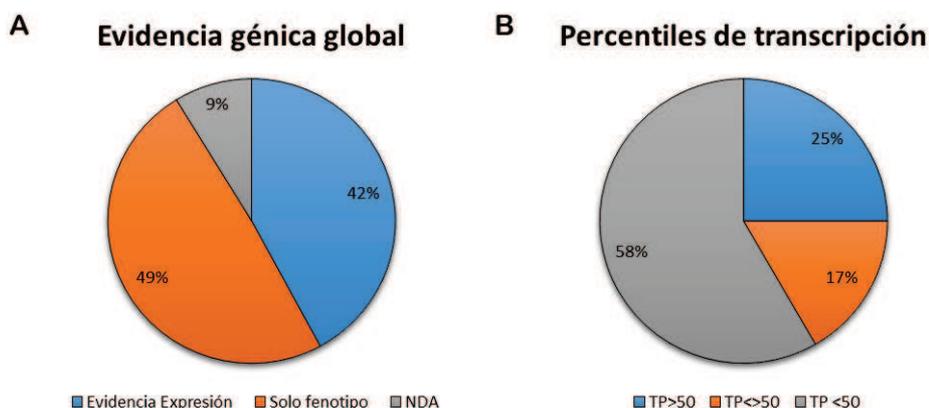
El análisis de datos fenotípicos mostró que los genes hipotéticos *tgh5* y *tgh8* presentarían valores muy negativos (-1,98 y -3,89 respectivamente) lo cual indicaría indispensabilidad de los mismos para el parásito y posibilidad de blancos terapéuticos. Ambos mostraron evidencia de expresión, donde *tgh5* estaría asociado al estadio específico de ooquiste y *tgh8* asociado a localización nuclear en la cromatina (*Tabla 5*).

Gen	ID	ST	AA <sup>1</sup>	NLS/BP	SP	TM	SP+ TM	LN	Fenotipo	ME49/RH <sup>2</sup>	PR	LOPIT	EE	S <sup>#</sup>	Otros
<b>Genes tipo-FamB</b>															
<i>tgh56</i>	_321270	lb_R	95	NDA	no	1	si	no	-0,15	ND	NDA	NDA	no	2	tipo-FamB CIM GO:0005575
<i>tgh57</i>	_321280	lb_R	79	NDA	si	1	si	no	-0,24	29.1/12.5	NDA	NDA	si	2	tipo-FamB
<b>Evidencia de transcripción</b>															
<i>tgh1</i>	_200110	XII_R	125	NDA	no	(-)	no	no	NDA	50.7/32.0	NDA	NDA	si	4	
<i>tgh4</i>	_206910	IV_L	98	NDA	si	(-)	si	no	0.53	56.8/26.4	NDA	NDA	si	3	
<i>tgh5</i>	_207240	X_R	108	NDA	no	(-)	no	no	-1,98	93.3/94.4	si, Tz/Oo	citósol	si	1	proteína de unión GO:0003674
<i>tgh8</i>	_207370	lb_L	1373	NDA	no	(-)	no	si	-3,89	75.5/82.4	si	nucleo-cromatina	si	1	dominio de interacción proteica
<i>tgh10</i>	_220070	V_L	155	NDA	no	(-)	no	no	NDA	22.9/24.0	NDA	NDA	si	2	
<i>tgh14</i>	_264230	VIIb_L	112	NDA	no	(-)	no	no	NDA	12.3/10.6	NDA	NDA	si	3	similar a _274700
<i>tgh18</i>	_274700	IX_L	220	NDA	no	1	si	no	NDA	6.0/5.2	NDA	NDA	si	3	similar a _264230
<i>tgh23</i>	_279550	VIIa_L	239	NDA	no	3	si	no	1.7	3.01/3.03	NDA	NDA	si	1	CIM GO:0005575
<i>tgh25</i>	_280810	VIIa_L	244	NDA	no	(-)	no	no	0.04	5.5/15.7	si	NDA	si	1	1 peptido en la fracción citosólica
<i>tgh26</i>	_281030	VIIa_L	153	NDA	no	(-)	no	no	0.74	42.4/38.6	NDA	NDA	si	3	
<i>tgh27</i>	_281040	VIIa_L	132	NDA	no	(-)	no	no	NDA	58.12/45.2	NDA	NDA	si	3	
<i>tgh28</i>	_292600	VI_R	209	NDA	si	(-)	si	no	NDA	84.6/85.8	NDA	NDA	si	1	
<i>tgh29</i>	_296121	V_L	904	si	si	(-)	si	no	2.28	40.7/72.1	NDA	NDA	si	3	
<i>tgh30</i>	_296231	V_L	817	NDA	no	(-)	no	no	NDA	65.0/65.8	NDA	NDA	si	3	
<i>tgh32</i>	_298830	III_R	304	NDA	no	(-)	no	no	NDA	70.0/53.1	NDA	NDA	si	3	
<i>tgh33</i>	_298840	III_R	242	NDA	no	(-)	no	no	0.25	28.1/37	NDA	NDA	si	2	
<i>tgh36</i>	_300780	XII_L	116	NDA	no	(-)	no	no	NDA	57.6/53.9	NDA	NDA	si	4	
<i>tgh37</i>	_300790	XII_L	221	NDA	no	(-)	no	no	NDA	35.0/18.6	NDA	NDA	si	4	
<i>tgh38</i>	_306680	IX_R	243	NDA	no	2	si	no	1.91	31.8/37.9	si, membrana	NDA	si	1	CIM GO:0005575 (similar aTgH31)
<i>tgh39</i>	_306688	IX_R	176	NDA	no	(-)	no	no	NDA	79.0/85.0	NDA	NDA	si	3	
<i>tgh41</i>	_306700	IX_R	207	NDA	no	(-)	no	no	1.11	40.2/38.2	NDA	NDA	si	2	
<i>tgh52</i>	_307470	V_R	174	NDA	no	(-)	no	no	NDA	21.1/20.1	NDA	NDA	si	3	
<i>tgh53</i>	_317820	IV_L	164	NDA	no	no	no	no	NDA	11.1/9.31	NDA	NDA	si	2	
<b>Fenotipo</b>															
<i>tgh31</i>	_298820	III_R	173	NDA	no	1	si	no	-0,2	NDA	NDA	NDA	no	1	CIM GO:0005575 (similar aTgH38)
<i>tgh2</i>	_200595	X_L	126	NDA	no	(-)	no	no	0.18	NDA	NDA	NDA	no	2	
<i>tgh3</i>	_200600	X_L	122	NDA	no	(-)	no	no	0.38	NDA	NDA	NDA	no	3	poca evidencia acerca de que se transcribe o se traduce
<i>tgh6</i>	_207350	lb_L	94	NDA	no	(-)	no	no	1.83	NDA	NDA	NDA	no	1	
<i>tgh7</i>	_207360	lb_L	126	NDA	no	(-)	no	no	1.83	NDA	NDA	NDA	no	2	
<i>tgh12</i>	_244940	VI_R	77	NDA	no	(-)	no	no	1.04	NDA	NDA	NDA	no	2	
<i>tgh13</i>	_244950	VI_R	96	NDA	no	(-)	no	no	0.36	NDA	NDA	NDA	no	2	diversidad funcional, adhesión
<i>tgh15</i>	_264240	VIIb_L	99	NDA	no	(-)	no	no	0.66	NDA	NDA	NDA	no	3	
<i>tgh16</i>	_264250	VIIb_L	141	NDA	no	(-)	no	no	1.4	NDA	NDA	NDA	no	3	interacciones proteína-proteína
<i>tgh17</i>	_264260	VIIb_L	134	NDA	no	(-)	no	no	0.91	NDA	NDA	NDA	no	2	
<i>tgh20</i>	_276810	XII_R	246	NDA	no	2	si	no	0	NDA	NDA	NDA	no	2	
<i>tgh21</i>	_276820	XII_R	162	NDA	no	no	no	no	0.5	NDA	NDA	NDA	no	2	
<i>tgh22</i>	_279540	VIIa_L	124	NDA	no	(-)	no	no	1.56	NDA	NDA	NDA	no	1	
<i>tgh34</i>	_298070	II_R	107	NDA	no	(-)	no	no	1.42	NDA	NDA	NDA	no	2	
<i>tgh35</i>	_298080	II_R	79	NDA	no	(-)	no	no	0.75	NDA	NDA	NDA	no	3	
<i>tgh40</i>	_306692	IX_R	79	NDA	no	(-)	no	no	0.81	NDA	NDA	NDA	no	2	
<i>tgh42</i>	_306710	IX_R	88	NDA	no	(-)	no	no	0.79	NDA	NDA	NDA	no	2	
<i>tgh43</i>	_306720	IX_R	148	NDA	no	1	si	no	0.29	NDA	NDA	NDA	no	2	
<i>tgh44</i>	_306730	IX_R	139	NDA	no	(-)	no	no	0.17	NDA	NDA	NDA	no	2	
<i>tgh45</i>	_307050	XI_L	69	NDA	no	(-)	no	no	0.67	NDA	NDA	NDA	no	3	
<i>tgh46</i>	_307060	XI_L	104	NDA	no	(-)	no	no	0.57	NDA	NDA	NDA	no	2	
<i>tgh47</i>	_307070	XI_L	136	NDA	no	(-)	no	no	0.7	NDA	NDA	NDA	no	2	
<i>tgh48</i>	_307080	XI_L	100	NDA	no	(-)	no	no	0.69	NDA	NDA	NDA	no	2	
<i>tgh49</i>	_307090	XI_L	146	NDA	no	(-)	no	no	0.57	NDA	NDA	NDA	no	2	
<i>tgh50</i>	_307270	XI_L	95	NDA	no	(-)	no	no	0.37	NDA	NDA	NDA	no	2	
<i>tgh51</i>	_307450	V_R	84	NDA	no	(-)	no	no	-0,29	NDA	NDA	NDA	no	2	
<i>tgh54</i>	_317830	IV_L	74	NDA	no	no	no	no	0,44	NDA	NDA	NDA	no	2	
<i>tgh55</i>	_320780	IV_L	103	NDA	no	no	no	no	0,05	NDA	NDA	NDA	no	2	
<b>Sin datos</b>															
<i>tgh9</i>	_220060	V_L	116	NDA	no	(-)	no	no	NDA	NDA	NDA	NDA	no	3	
<i>tgh11</i>	_244930	VI_R	103	NDA	no	(-)	no	no	NDA	NDA	NDA	NDA	no	3	
<i>tgh19</i>	_274280	IX_L	98	NDA	no	(-)	no	no	NDA	NDA	NDA	NDA	no	3	
<i>tgh24</i>	_280070	VIIa_L	226	si	no	(-)	no	si	NDA	NDA	NDA	NDA	no	3	

**Tabla 5. Descripción de genes únicos hipotéticos asociados a TgSTs.** ID: TGME49. AA: cantidad de aminoácidos, <sup>1</sup>Los genes pueden tener errores de anotación. NLS\_BP: *Bipartite nuclear localization signal profile*. NDA: *not data available*. SP: *signal peptide*. TM: *transmembrane domain*. LN: Localización nuclear estimada por PSORT II. <sup>2</sup>Los sets de datos corresponden a porcentajes de perfiles de expresión de dos líneas arquetipo según base de datos ToxoDB ME49 y RH. PR: Proteómica. EE: Evidencia de expresión (ver texto). CIM: Componente integral de membrana, GO: análisis de componente por ontología de genes (<https://toxodb.org/toxo/app>). S#: Sintenia: (1) TgHhNc\_conservados; (2) Tg\_conservados; (3) Tg\_parcialmente conservados; (4) Tg\_cepa específicos.

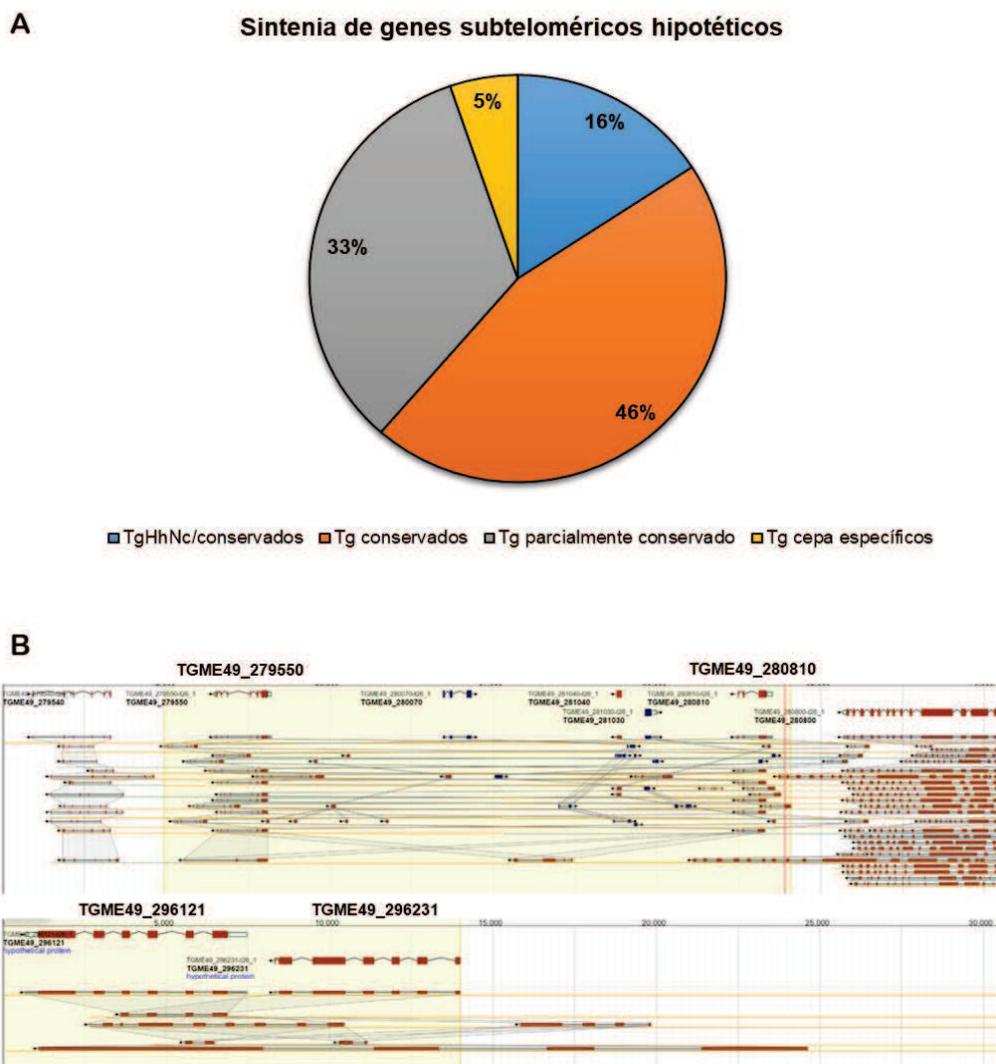
El análisis de la sintenia también se realizó para los genes hipotéticos en donde se pudo observar cuatro grupos: genes conservados en todas cepas de *T. gondii* y en las especies relacionadas *H. hammondi* y *N. caninum* (TgHhNc\_conservados: *tgh5, tgh6, tgh8, tgh22, tgh23, tgh25, tgh28, tgh31, tgh38*), genes conservados únicamente en todas las cepas de *T. gondii* (Tg\_conservados: *tgh2, tgh7, tgh10, tgh12, tgh13, tgh17, tgh20, tgh21, tgh33, tgh34, tgh40, tgh41, tgh42, tgh43, tgh44, tgh46, tgh47, tgh48, tgh49, tgh50, tgh51, tgh53, tgh54, tgh55, tgh56, tgh57*), genes conservados en algunas cepas (Tg\_parcialmente conservados: *tgh3, tgh4, tgh9, tgh11, tgh14, tgh15, tgh16, tgh18, tgh19, tgh24, tgh26, tgh27, tgh29, tgh30, tgh32, tgh35, tgh39, tgh45, tgh52*) y algunos genes específicos de la cepa ME49 (Tg\_cepa específicos: *tgh1, tgh36 y tgh37*) (Tabla 5) (Fig.32A). El 46% de estos genes se corresponden

**Fig.31. Evidencia de expresión de genes hipotéticos.** A. Evidencia génica global (57 genes analizados). Evidencia de expresión: en conjunto datos de transcriptómica, proteómica y/o LOPIT. Fenotipo: solo datos de fenotipo. NDA: sin datos. B. Percentiles de transcripción según base de datos (24 genes analizados con alguna evidencia de expresión). TP: *Transcriptional percentile*. NDA: *not data assigned*. <https://toxodb.org/toxo/app>.



con genes específicos de la especie, mientras que el 5% sería específico de la cepa ME49 analizada. Un dato interesante que surgió de este análisis fue la observación de posibles eventos de duplicación génica y reordenamiento entre los genes *tgh23* (TGME49\_279550) y *tgh25* (TGME49\_280810) como así también entre los genes específicos de especie *tgh29* (TGME49\_296121) y *tgh30* (TGME49\_296231) (**Fig.32B**). Vale aclarar que también hay genes cepa específicos no detectados en ME49 y no detallados en esta tesis.

**Fig.32. Sintenia de genes hipotéticos asociados a TgSTs. A.** Grupos de genes hipotéticos según conservación a nivel de especies relacionadas, especie específicos y cepa. **B.** Eventos de recombinación entre genes hipotéticos en regiones STs. La región resaltada en amarillo correspondería a los loci de genes con posibles eventos de recombinación génica.



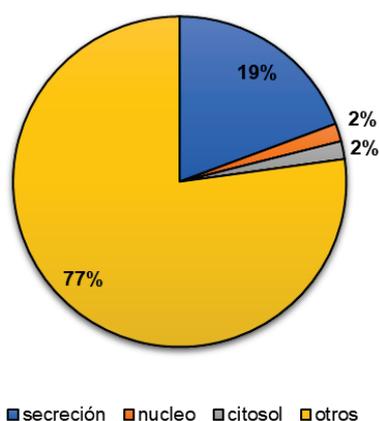
#### 11.1.1.2.1 Características, dominios y motivos de las proteínas hipotéticas, TgH

El análisis de motivos y dominios en las bases de datos no reveló estructura proteica consistente en ninguno de ellos, aunque algunos se encontraron como posibles componentes integrales de membrana o de interacción proteica mediante análisis por ontología de genes, GO (*Tabla 5*) (*Fig.33*). Como ya se mencionó, en la *Tabla 5* se describe la presencia de péptido señal, de dominios transmembrana y datos de proteómica y/o LOPIT encontrados en la base de datos ToxoDB. Los datos de LOPIT, péptido señal y dominio TM nos estarían indicando posibilidad de proteínas de secreción en algunos casos, aunque para la gran mayoría no existen datos de localización subcelular descriptos o asignados.

Las proteínas hipotéticas descritas tendrían tamaños que irían desde 69 a 1373 aminoácidos. Algunas proteínas presentarían de 1 a 3 dominios TM: como TgH18, TgH20, TgH23, TgH31, TgH38, TgH43, TgH56 y TgH57; donde en algunos casos estarían respaldadas por el análisis de GO como componentes integrales de la membrana (TgH23, TgH38, TgH31 y TgH56) (*Tabla 5*). Mientras que, el único dato de proteómica con localización específica hallado fue para la proteína TgH38, la cual se encontró asociada a la membrana.

**Fig.33. Evidencia de localización de proteínas hipotéticas.** Los compartimentos fueron asignados en base a la evidencia observada.

**Localización subcelular - HP**



### 11.1.2 Familias multigénicas

Tal como se mencionó en el Capítulo I, se identificaron dos familias multigénicas en las regiones STs, denominadas FamB y FamC. *T. gondii* presenta varias familias multigénicas en su genoma, que generalmente codifican para proteínas de membrana o secreción, algunas asociadas a virulencia como ser las familias de proteínas de las roptrias y gránulos densos, o invasión como ser las proteínas de las micronemas o de superficie SRS (Lorenzi et al. 2016). Además, los estudios genómicos detectaron otras familias multigénicas, que codificarían para posibles proteínas de membranas las cuales fueron denominadas FamA, FamB, FamC, FamD y FamE (Lorenzi et al. 2016). De ellas solo FamB y FamC fueron asociadas a los subtelómeros (Dalmaso et al. 2014; y Capítulo I).

Con el objetivo de determinar si las proteínas de las familias FamB y FamC tenían alguna relación evolutiva cercana se planteó un estudio filogenético. *T. gondii* cuenta con 33 miembros anotados pertenecientes a la familia TgA (FamA), 16 miembros de la familia TgB (FamB) y 2 de Tipo-B (reasignados en este trabajo), 14 miembros de la familia TgC (FamC), 11 miembros de la familia TgD (FamD), y 7 para la familia TgE (FamE), en la cepa de referencia ME49. Respecto a las familias FamB y FamC, en las especies relacionadas encontramos 7 miembros de la familia B y 4 de la familia C en *H. hammondi*, 4 de la familia B y 3 miembros de la familia C en *N. caninum* (nombrados en base a su gen homólogo en *T. gondii*) (ver más adelante **Tabla 7** y **Tabla 8**). En la **Fig.34** se muestra el árbol evolutivo correspondiente únicamente a las familias TgA, TgB y TgC y las proteínas ortólogas encontradas en *H. hammondi* y *N. caninum*. Debido a que las distancias entre las secuencias de las familias TgD y TgE eran muy largas y divergentes como para alinearlas (dato no mostrado), estas no se emplearon en el análisis. Por otro lado, las secuencias de la familia TgA coinciden en organización de dominios y longitud de secuencias de las familias B y C. De la familia TgA solo se incluyeron los miembros de *T. gondii*. Cabe mencionar que ninguno de los genes miembros de TgA se encuentran en asociación a las regiones STs o en cercanías como así tampoco los miembros de las familias TgD y TgE, estando todos sus genes a



Familia Multigénica ID	Número de genes			Número de Clusters <sup>+</sup>	Número de genes por clúster, localización cromosómica
	Distancia desde el final cromosómico				
	1-100 Kpb	101 -250 Kpb	>250 Kpb		
FamA	0/32	0/32	32/32	3	19 genes, chrXII 9 genes, chrVI* 3 genes, chrIX
FamB	11/15	4/15	0/15	1	3 genes, chrV
FamC	7/12	4/12	1/12 <sup>§</sup>	0	ND
FamD	0/11	0/11	11/11	1	3 genes, chrVIII
FamE	0/8	0/8	8/8	1	7 genes, chrVI*

**Tabla 6. Localización cromosómica de genes FamA a FamE.** <sup>+</sup>Clúster fueron definidos por la presencia de 3 o más genes ligados en una región cromosómica. <sup>\*</sup>Ambos clústers son cercanos. <sup>§</sup> TGME49\_300990 (*tgc12*), chrIV: 2,248,782-2,255,114 (longitud cromosómica 2,686,605). ND: no detectado.

### 11.1.2.1 Genes *famb*

Se describen 15 miembros de los genes de la familia B (*famb: tgb*) en asociación a las regiones TgSTs (**Tabla 7**). Sin embargo, por búsqueda de homología de genes se encontró un miembro más de esta familia anotado como gen hipotético, TGME49\_306870, el cual localiza por fuera de la región subtelomérica del brazo izquierdo del cromosoma XI a unos ~40,6 Kbp (ToxoDB) y por lo tanto fue excluido del análisis. La búsqueda de genes se amplió a *H. hammondi* donde se encontraron siete genes *famb* (*hfb*), seis de ellos con ortólogos en *T. gondii* y uno especie específico (*hfb16*) (**Tabla 7**). Por otra parte, para *N. caninum* se detectaron 4 genes *ncb* por búsqueda en la base de datos ToxoDB y NCBI (*Bioproject* PRJEB7872) en la cual se encontraron dos secuencias más largas (CEL68745 y CEL64819) que las existentes en ToxoDB (NCLIV\_004620 y NCLIV\_006930 respectivamente). El gen *ncb13* compartiría una alta identidad con *tgb13*, mientras que *ncb17*, *ncb18* y *ncb19* serían especie específicos. En la **Tabla 7** se detallan los miembros de esta familia junto a las características en cuanto a expresión génica encontradas en la base de datos ToxoDB. Como se pudo ver en el Capítulo I, un tercio de los genes *tgb* (5 genes, 33%) parecerían estar en un estado activo de expresión en la cepa ME49. Mientras que el análisis de los datos de transcripción en las cepas ME49 y RH indicaría que un ~46% (7 genes) lo harían (**Tabla 7**).

Por otra parte, solo *tgb10* presentó fenotipo de *fitness* negativo de -1.12, lo que podría representar algún fenotipo importante para la biología de *T. gondii* (**Tabla 7**).

Gen	ID	ST	In.	AA	SP	TM	Fenotipo	ME49/RH	PR	TSS	Sintenia	S#
<i>tgb1</i>	_200480	VIII	1	167	0.19/+	1	1.23	NDA	NDA	silente	Tg conservado	2
<i>tgb2</i>	_244920	VI	1	162	0.39	1	-0.82	NDA	NDA	activo	Tg conservado, reordenado	2
<i>tgb3</i>	_264225	VIIIb	2	208	0.72	2 <sup>c</sup>	2.03	NDA	NDA	activo	TgHh/conservado, reordenado,	1
<i>tgb4</i>	_274290	IX	1	136	<0.01	1	0.9	91.8/87.7	NDA	activo	Tg conservado, (ex) ME49, (con) ARI	2
<i>tgb5</i>	_276230	III	1	163	0.87/+	1	2.38	69.9/85.4	NDA	activo	Tg conservado, (-COUG), (ex) ME49	2
<i>tgb6</i>	_279460	IX	1	149	0.82/+	1	0.43	NDA	NDA	ambiguo	TgHh/conservado, reordenado, (ex) ME49	1
<i>tgb7</i>	_283450	V	3	265	0.35	1	1.72	67.0/25.3	NDA	ambiguo	TgHh/conservado, (-CatBr9), reordenado, (du) RUB, (ex) ME49	1
<i>tgb8</i>	_295300	la	1	158	0.87/+	1	0.42	60.7/24.7	NDA	silente	TgHh/conservado, reordenado, (ex) ME49	1
<i>tgb9</i>	_298090	II	2	172	0.91/+	1	-0.24	NDA	NDA	silente	TgHh/conservado, reordenado, (ex) CAST, GT1, RUB, VAND	1
<i>tgb10</i>	_299570	IX	1	150	0.85/+	1	-1.12	50.9/31.0	NDA	silente	Tg conservado, (ex) ME49	2
<i>tgb11</i>	_301890	X	2	165	0.012	1	NDA	91.0/82.2	NDA	ambiguo	Tg parcialmente conservado, (ex) ME49, MAS, CatPCR2, (con) VAND, VEG	3
<i>tgb12</i>	_307045	XI	2	258	0.17/+	1	1.30	63.7/37.9	NDA	activo	Tg conservado, reordenado, (ex) ME49	2
<i>tgb13</i>	_307460	V	2	212	<0.01	1	1.16	NDA	NDA	silente	Tg conservado, (-CatBr9,p89) reordenado, (ex) CAST, VEG	2
<i>tgb14</i>	_307480	V	1	177	<0.01	1	0.28	NDA	NDA	ambiguo	Tg conservado	2
<i>tgb15</i>	_320770	IV	2	182	0.89/+	1	-0.05	NDA	NDA	ambiguo	Tg conservado, (ex) ME49	2
<i>hhb3</i>	_264225	NA	1	264	0.064	2 <sup>c</sup>	NDA	NDA	NDA	NDA	contraído	1
<i>hhb6</i>	_279460	NA	1	152	0.84/+	1	NDA	NDA	NDA	NDA	conservado	1
<i>hhb7</i>	_283450	NA	1	126	<0.01	1	NDA	NDA	NDA	NDA	conservado	1
<i>hhb8</i>	_295300	NA	1	158	0.92/+	1	NDA	NDA	NDA	NDA	conservado	1
<i>hhb9</i>	_298090	NA	1	137	0.92/+	0	NDA	NDA	NDA	NDA	contraído	1
<i>hhb10</i>	_299570	NA	1	107	0.91/+	1	NDA	NDA	NDA	NDA	no comparte sintenia	4
<i>hhb16</i>	_454170	NA	1	183	0.79/+	1	NDA	NDA	NDA	NDA	no comparte sintenia, especie específico	4
<i>ncb13</i>	CEL64573*	lb**	NDA	163	0.71	1	NDA	NDA	NDA	NDA	no comparte sintenia	4
<i>ncb17</i>	_060290	XII	1	147	<0.01	1	NDA	NDA	NDA	NDA	no comparte sintenia, especie específico	4
<i>ncb18</i>	CEL66112	NDA	NDA	199	0.30	1	NDA	NDA	NDA	NDA	NDA	
<i>ncb19</i>	CEL64819*	III***	NDA	173	<0.01	1	NDA	NDA	NDA	NDA	NDA	

**Tabla 7. Descripción familia multigénica *famb*. ID: TGME49. \*Bioproject PRJEB7872 (NCBI).**

\*\*Cromosoma corresponde a la secuencia de gDNA de NCLIV\_004620, una versión truncada de 87aa.

\*\*\*Cromosoma corresponde a la secuencia de gDNA de NCLIV\_006930, una versión truncada de 113aa.

NA: *not assigned*. In.: intrones. NDA: *not data available*. AA: cantidad de aminoácidos. SP: *probability of signal peptide*; (+): predicho por PSORT II (<https://psort.hgc.jp/form2.html>). TM: *transmembrane domain*.

<sup>c</sup>Uno es un TM con un sitio de clivaje putativo. ME49/RH: transcripción, expresión en dos linajes arquetípicos. PR: proteómica. S#: Sintenia; (1) TgHhNc\_conservados; (2) Tg\_conservados; (3) Tg\_parcialmente conservados; (4) Tg\_cepa específicos. Estructura génica de tamaño expandido en algunas regiones (ex), estructura génica de tamaño contraído algunas regiones(con), gen duplicado (du).

Los genes *tgb* presentan en general un modelo de dos exones y un intrón, aunque algunos presentan exones adicionales cortos que codificarían la región C-terminal (**Tabla 7**).

En el análisis de la sintenia de este grupo génico se pudo observar que *tgb3*, *tgb6*, *tgb7* (con algunos reordenamientos), *tgb8* y *tgb9* (con algunas regiones génicas de mayor longitud en algunas cepas) poseen genes ortólogos en *H. hammondi*; mientras que el resto de los genes están conservados dentro de la especie (ToxoDB) (**Tabla 7**). La localización genómica se mantiene estrictamente conservada dentro de la especie en *tgb1* y *tgb15*. Finalmente, *tgb11* es el único gen encontrado en algunas cepas: ME49 (con regiones génicas de mayor longitud, es decir, expandidas), MAS, CatPCR2, VAND (contraído, es decir, con regiones génicas de menor longitud y/o tamaño), y VEG (ToxoDB) (**Tabla 7**).

#### 11.1.2.1.1 Características, dominios y motivos de las proteínas FamB

En la **Tabla 7** se muestra las características de los productos proteicos de la FamB (TgB, HhB y NcB). Se pudo observar, que todas las proteínas varían en longitud, desde los 107 a los 264 aminoácidos. Los miembros de esta familia poseerían en general un péptido señal en el extremo N-terminal y tendrían un dominio transmembrana codificado por el exón 2 (**Fig. 35**). En el caso de TgB3 y HhB3, hay dos dominios transmembranas, pero uno de ellos está localizado en la región N-terminal, solapando con el posible péptido señal (**Fig. 35**).

El análisis de motivos y dominios de las proteínas FamB en las bases de datos (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi> y [https://myhits.sib.swiss/cgi-bin/motif\\_scan](https://myhits.sib.swiss/cgi-bin/motif_scan)) no reveló estructura proteica consistente en ninguna de ellas.

#### 11.1.2.2 Genes *famc*

Se describen aquí 13 miembros de esta familia para *T. gondii* (*tgc*), donde 12 de ellos estarían asociados a regiones STs y uno de ellos (*tgc14*) no posee localización cromosómica asignada (**Tabla 8**). Los genes *tgc4* y *tgc10* originalmente habían sido asignados a localizaciones cromosómicas que no coincidieron con este análisis. En este análisis, el gen *tgc4* fue reasignado a la región ST en el brazo derecho del cromosoma XII, mientras que el gen *tgc10* ha sido agregado como parte de la región

**Fig. 35. Estructura proteica intrínseca de proteínas de la FamB.** La región subrayada se corresponde con la región del péptido señal, las flechas negras encierran junto a la Metionina 1 la región codificada por el primer exón, las flechas rojas encierran la región codificada por el exón 2, la región con letra roja corresponde al dominio TM.

```

>TgB1
MSPHMYMAILSSPATLAVFTFLLCNLHNCYSGSWAMAAERLETVSGIGSGSDGRHAASAVPRSAILLAEAEQQHTL
VPVHLRBSKRTSITKRSTALSRRRTKTMIAAKVVLSLVAILLGLLVASVKLQOCRRSLPESPSREAEGITGRRRLAEGG
GDDEKCVSVVPSSAL
>TgB2
MTANLSSAATVALCIFLSCNLHNYCFGSWTLAAEQLDTAPGIRGTARHAPSAPVSSAI PAEAEQQQAPVAVRLKK
LKRTSLRKNGTALSRRRTKTVTAAKVVLSLVAIVLGLLVASVKLQOCRRKLPANTSKQAAGSTGRRLAEGGSDDKPCV
SVALSSFPWYR
>TgB3
MQVTLRSAAALAFFVLSFCMDNWCIE TRAMTAAVVSSTVTSATDSIYAASPVHSSGVAVEAKHQGTRAADPIVTS
KSSPLTKKAEALRRRRITTAASVGAVLTAVLGLLVAVAKLQOCRRKLVTDNTGTTEGHTGRRLADGGDEDDCVEE
VCILGDEVSPRSQKAARTIHFVVCAILRTFYCDSRVSPALHLPVLLSRDTCGRVTL
>TgB4
MTTAILDSIPLIIRTPGSDGAIIVSSSAIPSEEKHRTEVAISLKRLRSSSHAEKQLVTLRRLKMTATMVVFMSTA
FLGLLLAAVKLRQCRQELMKNANGAIEGGTGRRLAEGGDEDECVRFVLATGHRQRAFH
>TgB5
MVNLFVSVVLTIVLLIVVSGTLSYSRVESLTMAVETMESAPATISAANATVAGSSLPFSFSPRQOTQLHDLRINSSRR
SRIHEGKSHASPRKEVPLSLVGLLVTAIVLGLLGGFFKLHQCRQOALKTNSETEGTILRRLAEGGHEENDCVGLGP
ASGDVANVVYD
>TgB6
MVNPFSAALVCFITFLYELGCGCTGSRAIVTEVLESVPAINETNVKNGETAVSSSARLVGTGNQLGRRGGGGLKKS
GRKSHAKETTAPSRRTQTIAATVGVLLASVFGLLLASAKLMQCPQGSIRGDAEGNTRRRRLAEGGYSGEKCVS
>TgB7
MNPSSAASLTSFTIFIVGQAHSVNSLTMATHVLDQTPTIHSRTNVAVATFRASATTPTEEGQKRQAGVLLKSKK
RNSPPKKTVALRRTYKKNVTLIFTLTASLVLLLVSVKLQOCRRRLRRTSSEIAGSTSRRLAEGGHEDNCTREV
RAQTVGSEEFVRRMLRFRCSSGCGRSTGRELSVCAEDYKAQVLEAIVEAGTLTGYWIRRASRTEGQPACVTCTKC
PFSNQCLRHKPSGGIHCVCLEASGACVCRDRHSGEVP
>TgB8
MILKISSANLVFVIIWLCARSCCIESLEVATSAVDSAAAHS GDNGSDAAFVMPNPMVAMTNAQHORMQLLGLLN
KSRSSPAKKKTVARRRTEKNVAAGLLLLTVIFGILLAFVKLQHCRCQKLLKSASGATEGNTTRRLAEDNCRVLPD
ARYLYL
>TgB9
MALFIGVLCALSCCSIESRAMTTGEVTSAPAVTTGAAHAGMENDSLTALRKLRRSPQSPSQKRTAVISGKRLATV
TTAAVLLTAVVGMMLMASGKLRQCLHAMMNQTTGGEIEGNKRRRLAEGNDESDKRCV NATEDRVRVPEPQRKVSVPSPR
SLGQGGAYRGCADIGERRGF
>TgB10
MIVNTCSATMLALFIEFLCALSNSSSESAMTTEEVTSAPAVTTGATLAEVKNEDSATAIKKLSRTROSP IQKKT
AAIHCKKTVTVTAGVLLTAVVGMMLMASGKLRQRLHATLNKAGGIEEGNTMRRRLAEGGDDGKCVSLVITSESER
>TgB11
MMNLPSAATAAFFTVYFGDLRDFRIRSLVGATNVVDSITAITD GKADDSGATLVASSPATQGEAEHQVTQVEIFSKS
SENIPNKKTTMALHRIKKANAITTFWLLTGVLGLMLVSVKLLKCRQDLLRNTGRGPEGNTVRRRLAEGGDYNDKCLQ
EVRSDGEAAVGGD
>TgB12
MSWLLVGFPPKCLTASLPTSQPFCSVKNRRENASLFHEYGPLHQYIRRDVSYRLLVTMMKRNRSAAATVAFFIFP
SCALHNCCVGSQAVATDAVESTQETDGEANANSATVVTTGELHPTQNDSGAAVLEKTSKETMLADSTTGLNFS
VLAILALRLAMTHLHQGEEGLTNTSDVTGGIQRRLSTRVPGNQRFASDVRI GGEGVVTGGQTRKRFPCCQADSSVL
WWIPQTFGADQSCFWLEDPSWGSYSDVEMD
>TgB13
MLWRYQPLDLEAVLSSEIHASGPFGRFVNGSRKYCIRSSTMATHVLDQIPIINSRTNGDSAAFVLSAAVRAEEK
QQRRQVDVILMKSKRSNPAKRTTTLRSTSGSNVTLIFILTAAALALLVSVLQOCRRQLRKTSGETAGSTKRR
LAEGGHEDNCRVRLAPASTKCPKACRFHKEQNSVPTRRLLTPPTLGIYSLTSQTIFFQ
>TgB14
MAKHSV DHLVHSQTRGGETEFLASNATILNEEKYQRQGDVLLKSKRSPAKRTTTLRRTNGNNAVTLIFTLTAA
LVLLLVSVKLQOCRRQGLLRTTGETAGNSTRRLAEGGHEDKCVSLATASTRCLQVACRFHEEQNSVPTRRLLTPS
IPGIYSAPTSQTSYFQWFRLPVSS
>TgB15
MTHMRYSSSAATLAFVLLLCMQSRSIWSWMTQESLDSPTVNSGFAADHAPSGTAPLATPSETDQQRTRVPL
ILKRSKRASPTFKSTALLRRTKTLTAAKVALFLTAVAAALLLAKLHQCRNQLLKDIGREAGNTGRRLAAGGGDD
EKCGSGAMARTGVEDTGFGRDITPVDRQA

```

```

>HhB3
MHVNSALLRIVSWHSLINMKVTPPSAAALAFVFLFLCDMNNCCVKTRAMAAAGVASDTV SATDGIHTASAVRSSG
PVEVKQOETHAALIMTSRSIPATKKAARIRRRKITTAVSAGAVLTAVLGLLVAVAKLQECRQKLVTDITNGTTEGH
RRRLAEGGDDEDDCVSMSSASLGRQWSACHFQPD RVSTSPALCSSTTTCRHFLPSSYFPVQLHDCAPWFSESTSSI
LRQKFSGQCDSSELLLTQCGVPWTAFPPTLHLHSSQ
>HhB6
MMNPFSAALACFTFFLCEGLSCCAGSRAMVNGVLDSPAI NGTTVNNGAAAVPSSGRPGIASQLHKRVGRAVVQI
SMRKSEVKGITASSRSTKTIAATVGVLLASILGLLIASAKLLQCP TTFQANTRGDAEGNTRRRRLAEGGNSSTKCV
<
>HhB7
MATHALDQTQTIHSWTNVAEAAFRASAAA ILTEEKQQQRQAGVLLKSKRS SPAKKT TVALRRLTNKKNAVTLIFTI
TAALLLLLVSVK LQQCRQLLRKTSGETAGSTSRRLADDGNEENCVSLTP
>HhB8
MILKLSASCASVLV FVI IWTCTARS CCLESLEVAAA AVDSATAI HSGDN GSDAAS TVPNPLTTT KAQHQRMLAGLLI
KSRGSPA EKKTVAIRRTKEKNVAAGLLLLTVVFGILLAFV KLQQCRQKLLKSASGGTEGNKTRRLAEDNCVRLEI
ARNVYL
>HhB9
MALFIGLLCALSSCSIGSRAMTTGVV TSSVPAVTTGA AHAEVKND DSIKKLSRSRQSPSQKRTAVIYGKKT VTVT
AAVLLTAVGRTIMSSGKLRQCLHAMNKAGGEIEGNKRRRLAEGGADDGKGVSLDATS VGCR
>HhB10
MMVNTCSATMVALFIGFLCALSSGSI ESRAMTTEEV TSSAPAVTTGATLAEVKNE DSATALKLSRTRQSPVQKKI
AAIHGKKTVTVTAAAVLLTAVVGMLMASGEL
>HhB16
MMFHASSAATLAF FIFL LICALSSCCIASRAMTTEVVAAS SPMTTTTAAALAE LKNKDS PAVPQRRLRRSSPTEKI
TAAKRRRMAMTVIATAVLLTAVVRMLLVSVKLRQCLVKNAE TAGAKIEGNTERRLAEGGSD EDCVSLATTSTGCI
EYVCLFDQDHSDSTLQSSRATRGH SVLSTY
>NcB13
MTAEVVDSI PVTDNGDATSAVSRTAIPAEAEQEGTRDAVKRSRRSGTKMAE PRHKKRI AALVGA VLTAAVGVLLI
AVKLHQCRQEVLKNTSDATEGTTGRRLAAGNGD GICVSVAAAGPGYRQAGSVPVVFRCTSQPLQLSASP
>NcB17
MMVNSSSATAAVVFLTFSLHPSHIWCNGSRVMTTGVLDFVPVMTSGDN GSDAASAVSSSAISAEGKQRIQVGRLP
RPRRRI PAEGRTLTLPRTQKVIPVTVGVLLTAVLGLL LASVKLHQCRQQLKNADGQTEGNTKRRRLAEGGSDSS
VSLAPASMRYSRPLSFIQRTTLPPTLHSSRRTCRHNAPS IYWAVTS
>NcB18
MKVNQSSVATLAA FVIVSSCSLGCSCNGSPALATAAVDFVPAISSQANGSNAAAMVSGTAIPGEADQOHTQVVFVVI
RSRRSNPAIGKTMASRRAEKR IAATVGLLMTAVAGLLLASVKLQQCRQALRQKTRGETEGNTRRRRLAEGGNDDEK
VSMAPARTGCR
> NcB19
MTADS INSGNGS DATSVAFGSAFQTEAQOQRTQVHDLLK GKRRSSPLERKTSALRRTKYMSAAKAGFLLATV FVI
LLASVKLRQCRHDL LKTSSETEGNTRRRRLAEGGGG GEDGKCVSLASASIGCRWIGWLTPLV KKEAGTASPTRPH
PFTARGKLVDMPPSIYWAI SF

```

ST en el brazo izquierdo del cromosoma X, ST\_X\_L. Por otro lado, se ha descrito la asociación de nuevos miembros de esta familia a TgSTs, como en el caso de *tgc13* y la aparición de un nuevo miembro, *tgc14* (TGME49\_329400) (Tabla 8). Al igual que para *famb*, la búsqueda se amplió a *H. hammondi* y *N. caninum*. En *Hammondia* se encontraron cuatro genes ortólogos a genes *tgc*: *hhc6a*, *hhc6b*, *hhc8* y *hhc11*; mientras que en *Neospora* los tres genes encontrados serían especie específicos: *ncc15*, *ncc16* y *ncc17*. Los genes *ncc15* y *ncc17* fueron detectados en NCBI *Bioproject* PRJEB7872 donde se extrajo una secuencia más larga (CEL68745) que la existente en ToxoDB (LIV\_000632200) para *ncc15* y lo mismo para CEL66113 y NCLIV\_019380 en el caso de *ncc17* (Tabla 8).

En este caso, la composición nucleosomal observada para genes *tgc* activos (Capítulo I) parecería estar en congruencia con los datos de transcripción recabados con excepción de *tgc6* para el cual no se registraron datos. Mientras que se observaron datos de transcriptómica para *tgc5*, *tgc8* y

*tgc12*, pero con valores de percentiles por debajo del 50% (**Tabla 8**), sugiriendo bajos niveles de expresión.

Por otra parte, no se observaron entre estos genes, fenotipos de *fitness* negativos extremos que demuestren ser indispensables para la biología del parásito.

Los genes *tgc* presentan un modelo general de dos exones y un intrón, aunque algunos presentan exones adicionales cortos que codificarían la región C-terminal al igual que los miembros de la familia *tgb* (**Tabla 8**).

En el análisis de la sintenia de los genes *tgc* los únicos miembros que poseen genes ortólogos en *H. hammondi* (con versiones de genes más cortos) son *tgc8* y *tgc11* (ToxoDB) (**Tabla 8**). Los genes *tgc3*, *tgc5*, *tgc10* y *tgc12* son específicos de la especie, pero no conservan localización genómica entre cepas; mientras que *tgc1*, *tgc4*, *tgc6* y *tgc13* solo están en algunas cepas y mantienen la localización genómica (ToxoDB) (**Tabla 8**). Por otra parte, *tgc2*, solo estaría presente en las cepas ME49 y CatPCR2, mientras que *tgc7* y *tgc14* solo estarían en ME49 (ToxoDB) (**Tabla 8**).

En este análisis se ha observado que los genes *tgc* que están duplicados en algunas cepas, tal es el caso de *tgc4* (CatPCR2), *tgc6* (ARI, CAST, COUG) y *tgc10* (ARI, CAST, COUG, CatPCR2) (ToxoDB) (**Tabla 8**). Cabe mencionar que dos genes de *H. hammondi*, *hhc6a* y *hhc6b* no comparten sintenia con su ortólogo en *T. gondii* lo que sugeriría mecanismos de recombinación cromosómica, reordenamiento génico y posibles eventos de duplicación especie específicos dada la pérdida de sintenia. Por otra parte, *N. caninum* también poseería miembros específicos de la especie que no comparten sintenia con ningún otro gen (*ncc16*) (ToxoDB) (**Tabla 8**).

#### 11.1.2.2.1 Características, dominios y motivos de las proteínas FamC

En la **Tabla 9** se muestra una comparación los dominios estructurales encontrados para los productos proteicos de esta familia génica C con distintas bases de datos: NCBI CDD, J-PRED y ToxoDB. Como vimos en la **Tabla 8** la única proteína TgC con datos disponibles de proteómica en la base de datos ToxoDB pertenece a la fracción de membrana y corresponde a TgC8. El análisis de dominios

Gen	ID	ST	In.	AA	SP	TM	Fenotipo	ME49/RH	PR	Sintenia	S#
<i>tgc1</i>	_321170	III	2	232	0.13	2 <sup>c</sup>	0.52	44.7/61.2	NDA	Tg parcialmente conservado, (ex) ME49, (con) p89	3
<i>tgc2</i>	_200700	X	0	139	0.05	1 <sup>c</sup>	NDA	NDA	NDA	Tg parcialmente conservado, ME49, CatPCR2	3
<i>tgc3</i>	_307260	XI	1	206	0.14+	1	0.96	75.2/64.8	NDA	Tg conservado, (-RH), reordenado, (ex) ME49	2
<i>tgc4</i>	_200130	XII	1	238	0.04+	1	NDA	NDA	NDA	Tg parcialmente conservado, (du) CatPCR2, (ex) ME49	3
<i>tgc5</i>	_298960	III	0	282	<0.01	1 <sup>c</sup>	0.06	53.9/40.5	NDA	Tg conservado, reordenado	2
<i>tgc6</i>	_200590	X	2	201	0.01+	1	2.09	NDA*	NDA	Tg parcialmente conservado, (du) ARI, CAST, COUG; (ex) ME49, (con) MAS	3
<i>tgc7</i>	_287280	Ia	3 <sup>a</sup>	225	0.06+	1	NDA	53.0/36.9	NDA	Tg cepa específico ME49	4
<i>tgc8</i>	_220080	V	2 <sup>a</sup>	306	<0.01	1	0.44	23.7/22.4	Membrana	TgHh conservado, (-CATBr9), reordenado, (ex) ME49, RH CatPCR2	1
<i>tgc10</i>	_228780	X	2 <sup>a</sup>	286	<0.01	2	1.49	15.2/23.9	NDA	Tg conservado, (-CATBr9), reordenado, (du) ARI, CAST, COUG, CatPCR2, (ex) RH	2
<i>tgc11</i>	_298060	II	16	305	<0.01	3 <sup>c</sup>	1.22	14,5/32.6	NDA	TgHh conservado, reordenado, (con) RH	1
<i>tgc12</i>	_300990	IV	3	252	<0.01	1	0.04	18.4/16.0	NDA	Tg conservado, (ex) ME49	2
<i>tgc13</i>	_292610	V	5 <sup>b</sup>	265	0.06+	1	NDA	53.0/36.9	NDA	Tg parcialmente conservado, (ex) ME49, VAND, VEG	3
<i>tgc14</i>	_329400	NA	3	225	0.06+	1	NDA	53.0/36.9	NDA	Tg cepa específico ME49	4
<i>hhc6a</i>	_321170	NA	1	216	0.02	1	NDA	NDA	NDA	no comparte sintenia	1
<i>hhc6b</i>	_200590	NA	1	193	0.06+	1	NDA	NDA	NDA	no comparte sintenia	1
<i>hhc8</i>	_220080	NA	0	180	0.004	1 <sup>c</sup>	NDA	NDA	NDA	contraído	1
<i>hhc11</i>	_298060	NA	6	330	<0.01	1	NDA	NDA	NDA	contraído	1
<i>ncc15</i>	CEL68745*	VIII**	NDA	313	0.01	2 <sup>c</sup>	NDA	NDA	NDA	NDA	
<i>ncc16</i>	_069620	NA	2	242	0.03	1	NDA	NDA	NDA	Nc conservado	4
<i>ncc17</i>	CEL66113*	VIIa***	NDA	146	0.005	1	NDA	NDA	NDA	NDA	

**Tabla 8. Descripción familia multigénica *famc*.** ID: TGME49. In.: intrones. AA: aminoácidos. (+): SP predicho por PSORT II. ME49/RH: transcripción, corresponde al porcentaje del perfil de expresión en dos linajes arquetípicos (ME49 y RH). PR: proteómica. S#: Sintenia: (1) TgHhNc\_conservados; (2) Tg\_conservados; (3) Tg\_parcialmente conservados; (4) Tg\_cepas específicos. NA: *not assigned*. NDA: *not data available*. \*Bioproject PRJEB7872 (NCBI). \*\*Cromosoma corresponde a la secuencia de gDNA de Ncaninum\_LIV\_000632200, una versión truncada de 153aa. \*\*\*Cromosoma corresponde a la secuencia de gDNA de NCLIV\_019380 una versión truncada de 122aa. <sup>a</sup>Un intrón in GT1 (*tgc8*) y VEG (*tgc7*, *tgc8*, *tgc10*). <sup>b</sup> Dos intrones en VEG. <sup>c</sup> Uno es un dominio TM en un sitio putativo de clivaje. Estructura génica de tamaño expandido en algunas regiones (ex), estructura génica de tamaño contraído algunas regiones (con), gen duplicado (du).

conservados arrojó que TgC5, TgC8 y NcC16 tendrían dominios correspondientes a la superfamilia de bombas de Na<sup>+</sup>/Ca<sup>+</sup> dependiente de K<sup>+</sup> implicadas en el transporte de cationes, iones y unión a proteínas (**Tabla 9**). Mientras que TgC7 poseería un dominio S2P correspondiente a una clase de metaloproteasas de Zinc (MEROPS *family* M50) que clivan dominios transmembrana en las proteínas sustrato. Por otra parte, TgC8 tendría además de una correspondencia con una proteína Vir de *Plasmodium* y una proteína viral Vpr de unión a proteínas implicada en el control del ciclo celular, la división celular y la división de los cromosomas (**Tabla 9**). TgC10 se encontró asociado a un dominio de quelatasa de Mg, implicada en el transporte de esta co-enzima y el metabolismo de la misma (**Tabla 8**). Otro miembro de la FamC, HhC8, se encontró asociado a un dominio estructural BDHCT. Esta familia de dominios es encontrada en el síndrome Bloom asociado a las helicasas DEAD-box en eucariotas superiores (**Tabla 9**).

En la búsqueda por predicción de estructuras secundarias con JPred y el banco de proteínas PBD se encontró que TgC1, TgC3, TgC5, TgC8, NcC15, NcC16 comparten un dominio de monometilamino metiltransferasa MTMB. Este dominio se encuentra en proteínas metiltransferasas de metilamina en bacterias metanogénicas (Burke, Lo, & Krzycki 1998). Mientras que TgC3 y TgC8, además de, TgC10, y NcC15 presentaron asociación a un dominio de la subfamilia de canales receptores transientes de cationes v miembro 2, TRPV2. Este dominio se encuentra en receptores de calcio en células humanas (<https://www.uniprot.org/uniprot/Q9Y5S1>).

En conclusión, dada la diversidad de posibles dominios e identidades, no hay evidencias firmes para establecer ninguna identidad de las proteínas FamC, aunque los datos acumulados sugieren una identidad asociada a proteínas integrales de membrana.

Se pudo observar, además, que todas las proteínas FamC varían en longitud, desde los 139 a 330 aminoácidos (**Tabla 8**). Todas las proteínas FamC presentan de 1 a 3 dominios transmembrana, y la gran mayoría posee péptido señal en el extremo N-terminal (**Tabla 8**). Nuevamente, cuando presenta más de un dominio transmembrana, uno de ellos está relacionado al extremo N-terminal, como posible

péptido señal (*Fig. 36*). A diferencia de TgB, la región de transmembrana se encuentra flanqueando la unión de los exones 1 y 2 (*Fig. 36*).

Nombre	NCBI CDD search <sup>a</sup> (interval; E-value)	Jpred 4 <sup>b</sup> PDB (E-value)	ToxoDB
TgC1	-	mtmb1 <sup>i</sup> (1e-07)	NDA
TgC2	-	-	NDA
TgC3	-	mtmb1 <sup>i</sup> (2e-07) receptor cation channel subfamily v 2 <sup>j</sup> (3e-0.7)	NDA
TgC4	-	-	NDA
TgC5	Ion dependent exchanger <sup>c</sup> (94-227; 4.88e-0.3)	mtmb1 <sup>i</sup> (2e-09)	NDA
TgC6	-	-	NDA
TgC7	Site-2 protease (S2P) <sup>d</sup> (133-176; 0.08)	-	NDA
TgC8	<i>Plasmodium</i> Vir superfamily <sup>e</sup> (65-150; 3.86e-0.3)	mtmb1 <sup>i</sup> (1e-06)	
	HIV-1 Vpr-binding protein <sup>f</sup> (45-192; 4.67e-0.3)	receptor cation channel subfamily v 2 <sup>j</sup> (5e-0.7)	LpxA enzyme <sup>k</sup> (125-171, 2.0e-05)
	Ion dependent exchanger <sup>c</sup> (53-214; 0.03)		chloride ion channels (8.8e-9)
TgC10	Mg-chelatase <sup>g</sup> (54-209, 0.05)	receptor cation channel subfamily v 2 <sup>j</sup> (9e-0.7)	NDA
TgC11	-	-	NDA
TgC12	-	-	NDA
TgC13	-	-	NDA
TgC14	-	-	NDA
HhC1	-	-	NDA
HhC6b	-	-	NDA
HhC8	BDHCT-box associated dom <sup>h</sup> (91-157, 8.75e-0.3)	-	NDA
HhC8		-	NDA
HhC11	-	-	NDA
		mtmb1 <sup>i</sup> (3e-06)	
NcC15	-	receptor cation channel subfamily v 2 <sup>j</sup> (5e-0.6)	NDA
NcC16	Ion dependent exchanger <sup>c</sup> (53-175; 2.50e-03)	-	NDA
NcC17	-	mtmb1 <sup>i</sup> (2e-07)	NDA

**Tabla 9. Dominios conservados y proteínas putativas de productos FamC según bases de datos.** <sup>a</sup>(<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>); opciones: CDD v3.17, NCBI\_curated 14908, pfam v31.0, SMART v6.0, KOG v1.0, COG v1.0, PRK v6.9, TIGR v15.0. <sup>b</sup><http://www.compbio.dundee.ac.uk/jpred/>. <sup>c</sup> *K<sup>+</sup>-dependent Na<sup>+</sup>/Ca<sup>+</sup> exchanger*. Nombre: 2A1904 super family, número de acceso: cl36772. <sup>d</sup> *Site-2 protease (S2P)*. Nombre: S2P-M50 super family, número de acceso: cl10020. <sup>e</sup> *Plasmodium Vir superfamily; provisional*. Nombre: PTZ00473 super family, número de acceso: cl26999. <sup>f</sup> *HIV-1 Vpr-binding protein*. Nombre: KOG1832, número de acceso: KOG1832. <sup>g</sup> *Cobalamin biosynthesis protein CobN, Mg-chelatase*. Nombre: CobN super family, número de acceso: cl34262. <sup>h</sup> *BDHCT-box associated domain on Bloom syndrome protein*. Nombre: BDHCT\_assoc super family, número de acceso: cl24698. <sup>i</sup> *Monomethylamine methyltransferase mtmb1*; PDB id: 1nth. <sup>j</sup> *Transient receptor potential cation channel subfamily v member 2*. Chain: a, d, b, c; PDB id: 6bo4. <sup>k</sup> *Trimeric LpxA-like enzymes*; InterPro ID: IPR011004. NDA: not data available.

Otro rasgo distintivo es que los miembros de la FamC presentan tres regiones conservadas, una extensa hacia la región N-terminal, una corta codificada por el exón 1 y otra también extensa que abarca parte del exón 1 y parte del 2, y una tercera región hipervariable hacia el extremo C-terminal (**Fig. 37**). Dentro de la región variable que esta codificada por el exón 1, antes de la región conservada 3, se puede observar la presencia de motivos repetidos variables en tamaño y secuencia en algunas TgCs, los cuales se muestran en la **Fig. 36**. Los motivos repetidos son diferentes en secuencias aminoacídica. Sin embargo, cuando se analiza a nivel de nucleótidos, se puede observar que existe una mayor similitud en las secuencias codificantes para las regiones repetitivas, lo que podría soportar la posibilidad de que las diferencias observadas a nivel proteína se deban a cambios de marco de lectura por deleciones o inserciones (**Fig.38**).

## 11.2 Las proteínas FamB y FamC y la virulencia

Debido a la posible estructura como proteínas integrales de membrana se hipotetizó una asociación como posibles familias génicas específicas participes en la interacción parásito-célula hospedadora y se decidió indagar un poco más acerca del potencial rol de estos genes como factores virulentos de *T. gondii*. Cabe agregar, que la mayoría de los factores de virulencia de *T. gondii*, *P. falciparum*, *Trypanosoma sp.*, etc, están codificados por familias multigénicas de proteínas de

**Fig. 36. Estructura proteica intrínseca de proteínas de la FamC.** Subrayado: péptido señal, flechas negras encierran junto a la Metionina 1 la región codificada por el exón 1, las flechas rojas encierran la región codificada por el exón 2, la región con letra roja corresponde al dominio TM. Las regiones resaltadas en colores denotan los motivos repetitivos. Diferentes colores delimitan la unidad de repetición o diferentes regiones de la unidad de repetición en una secuencia (TgC8). (\*) Codón stop en NCLIV\_019380. Las regiones de color verde denotan un TM extra.

```

>TgC1
MESMVEKKRSAFSTGWHLPSQLKARGLLFHHDSPILLSSVVLSLMVLEPLVSLDCSACRVCTA
LADSVSGRETDTDWSDTYRWTETGDSTVEGSGGVTEVEENPAEGNSGAIEAEDRTTVEGCGG
ATEVYGGTAEVYDGAVEVYGGAAEENYTSVGKPIESHAVEPLRERSTNPSVAAHVSRIRHCGR
RPFVVASGILAGLVLSVLLGAIVLNSGRKAHKATTPAPSRV
←→

>TgC2
MESMVEKKRSGFSTGWHLPSQLKARGLLFHHDSPILLFSVVLPLMVLAPLVSLDCFACRVCTA
LADSVSGRETDTDWSDTYRWTETGDSTVEGSGGVTEVKEGNPTEGISGAIEAEDRTTVEGSGG
ATEVYGGAAEV
←→

>TgC3
MIEEKRSGFSTGWHLHSQLKSRRGLLFHHDSPILLFSVVLCLSVLAALVSFDCSTCRVCTALAD
SVSGRETDDLNDTYSRAETGEGNPAEESSGATEVYGGAGEENCTTVGKRTESHTVGPLQERSR
YPSVAARVSRILHSGRARPVVASGILASVLSVLLGAIFLNSDRKSGKATTPAPSKTSDVHLVN
TRKSPQHSPEDMSR
←→

>TgC4
MESTIEKKRSGFSTGWYLPSQLKARGLLFHHDSPVLLFSVVLCMMVLAPLVSLDRSTCRVCTA
LADISGRETDDLSDTYSGTETEESTTAEWSGGVTEVEEGKPAEGSRGVLEAEDRTTVDGSGG
APEVLGGAAEENYTSVGKLIESHAVEPFQERSRNPSVAARVSRTRHSGRARPLVASGILASVLSV
SVLLGAIFLKSGRKSRKATTPAPSMTPVLLVNARKSLQHSLGDM
←→

>TgC5
MTRDEKRMENKVEKICSGLSIGLHLPSRRKVRRGLLSHRGSLLILLSSVALCLMVLAPFVVSFEF
STPRMCTALADSVPGLEADNDGSDAYSWTDTEDGTTVEESDDATEEDGTTVEETDGATEGEDG
TTVEETDGATEGEDGTTAEETDGATEAEDGTTSEGGGAGTETEDGTTSEGGGAGTETEGGTF
EGGGAGTAAEEGTVYEGGGAGIEAEEGTVYEGGAGAVEVYGGSAEVIYATNDKQIEQNAVEPV
QERSRKRTRVSRRVSKIQHSERDRPIVASGL
←→

>TgC6
MVNTIEKKHSGFPTAFQLPSRLKTRRGLLFCHDPRILLFSVALCLTALVPFVTFECSTRMCT
TLVDSVLGLETNDWSGTYSWTDTREGTTVGGDGGAAEADSSASELDRGATIESSSVESSQEL
STNSSVTARTSRIQRSGRAGPIVAFTILASVILSIVLVAIFLSWGRKSRKATTPAPTDSAVQL
SLRGRAHAECHA
←→

>TgC7
MMHLIQKKCPGFPPGFQLPCRLKARRGLFRHESCTMLFSVALCLTALASFVPFECSTRMRTA
WANSVLGLETDPERGGTYSGATERDGGATGVNDTSLDMVRDTNAETPLGERSKRPSVAARLSKV
QLSGQRRRIVASGILASVVLAVLLGTIFLNSGRNLRKATTPASSAALQVSTTGLQKRVCFADSE
GDMIESLFRRTSCSSVLGDMKRERFRRLLGLGEN
←→

>TgC8
MMKPTPDNVVAPRAVTNSETRMVNMEKFSAYSTGFQLPCQLNARRGLFLLRDLRILLFSVALC
LTVFASFVSFECSTYRLCPALADSVSLQTHSSTDTREEMTVEEDGDTEDDNIATEGDDKAAV
VDGGSTVDDGSTVDDGSTVDDGSTVDDGSTVDDGSTVDDGSTVDDGGFAEVDGGFAEVDGGFAE
VDDGGFAEVDGGSTGVDNGTADASDATKLDGGAAEIHSANLDEPTESAAVETMQKRMKTSNTVW
RASRLLSYGRARPGMVAGILAGVLSVLLGATFLRSGRKSRKGTPASSKSDA
←→

>TgC10
MKPTPGRVVLSPVVTRGENRLVNMIEKKLSGFPSGLQLTAQLKARRGLLFRGSSVLLLSVAL
CLVVLAPFVPFDCPTRRTCTALADAVTSLETDTGWNDSSSWTDTQDDTAVEGSADAREDGSAT
ELDGGGTEGDGSATELDGGATDVYGGATDADGGAIDVEGGVTEGMSMTLGEATESYAIESLRA
PSRSPSVAALVPRIHRSGQARPIVVSGILASVLSIILGAIFLNSDRILITETTPSPNATYVT
DDTSTLAWQLYLCTYRWRRNATRVVAIYGCHCVL
←→

>TgC11
MKHTIDRVVASRGVTNGEKIMVNVIGTNRSGFFTFPLPSQLKARRGLLLYDSRVLLFLVALYV
TVFTAFVSFECSTRICTALAHSFIGVEMDPDWSGSHSWTDIREDTTVEGDGGATGADGSAAEV
NYEALGELLIAYDALETLRNRAKAAGVLARVSRVQRAGRTCPIAASGILAGVLLSVLVGALFLK
KGQTPVQETTPPIAAPEKTEDAIDGEEAINQLATEHANPLIRKLVVAELAAIAGAVALYLMT
MPEPLLQPSQQPQLQEDGAMDVRFCHFNRRQQSSLPAGARSISPESML
←→

```

>TgC12  
MMKPAAKEIVPSRGVNDQMGMENTSEKKCFGFSVGSHPFSEPKTRPAWRFRHNSRVLLCVAL  
CLTVLSSFSVYESSTSRMYTALSAAAADVDTDSLVDTSWTESGEEGTTVDDGATEADDGAVEL  
QGDAEAEGGANELYGTDFENQPKIIRLARTWLRKDPRVAALVSKVERFGREPRMVLSGIVAGV  
ALSVILSAVFLKSGRKPGSPATQGPSSRSAESSQSALAWFRWARTTHPLLLTMLDPIDKA

>TgC13  
MMHLIQKKCPGFPPGFQQLPCRLKARRGRLFRHESCTMLFSVALCLTALASVFPFECSTRMRMTA  
WANSVLGLETDPERSGTYSGATELDGATGVNDTALDMVRDTNADTPLGERGKRPSVAARLSKI  
QLSGKRRRIVASGILASIVLAVLLGTIFLNSGRNLRKATAPASSVGIAMHDDGRLEAQGLDIRR  
SLRTIHARDGLNLMCVVPIFMDPSPPGYSGVLDVRKVVDAHNTKLLYGSEPNESLPEMKRDM  
LLRVTWLS

>TgC14  
MMHLIQKKCPGFPPGFQQLPCRLKARRGRLFRHESCTMLFSVALCLTALASVFPFECSTRMRMTA  
WANSVLGLETDPERGGTYSGATELDGGATGVNDTSLDMVRDTNAETPLGERSKRPSVAARLSKI  
QLSGQRRRIVASGILASVVLAVLLGTIFLNSGRNLRKATTPASSAALQVSTTGLQKRVCFADSE  
GDMIESLFRRTSCSSVLGDMKRERFRRLGLGEN

>HhC6a  
MSGVTSSETRMVMIEKKHLGFPTAFQLPSRLKTRRGLLFRHDPRIILFSVALCLAALVPFVFP  
DCSTRMCTALVDSVLGLETNDWSGTYSWTDTSSEGTVEGDGSGTEVDGGAAEVDGGATIESF  
AVELSQEQSKNPSVTARASRIQRSGRARTMVAYTILASVMLS AVLVAIFLSSGRKSRKATTPAP  
SKSSDALLRNAPKSLQNSPEVILQ

>HhC6b  
MIEKKHGFPTAFQLASRLRTRRGLLFSHDPRIILFSVALCLTALVPFVFPDCSTHRMCTTLVD  
SVLGLTNTDSSGTDTWTDGGATEVDGGATEVDGGATTE SVAVGLSQELRKNASVTAHASRIQR  
SGRAGPIVAETILAGVMLS VLVVAISLNSGRKSREATTPASSKSSDAFLRNARKSPONSPEVIS  
Q

>HhC8  
MKPTPDKVVASRGMSTSETRMVNMKKRSAYSAGFQLSSQLKARRGLSLFRHLRILLFSAALCLT  
MFASLVSFECSTCWMCPALADSVSGLHTSSSTDAREETTAEEEDGDTSEYSGSTDDDSWTTEVDG  
GAAEIHSTTLNQLTESGAVETLQQGINSSNVASRVSRLLNAGRANPGMVAKM

>HhC11  
MKHTLDRVVASRGTNGEKRMVNVSGKKRSGFFTGFPPLPSQLKAKFGLFLYDSRILLFLVALY  
LTVFAAFVSFECSTRRICTALADSLTGAEMDLWGSWSWTDMGEDTVEGDGSGATGADGGATG  
TVGGATGADGGAAEVNVEALGELLIAYDALETLRNRGKATGVVARVSRVQHSGRICPIVASGIL  
AGVVLVLLGAVFLMKDRKPLQATTPAPRSEGGMIGSLFRRTPCPSVVGMLRDI FCRLVPVGG  
EETINQLVNTPSLWPLSCSTSSDCRWCCALPDDTGKSAPGAVQVSSALWKPCLMVRLMSWALS  
KVCRTSLATH

>NcC15  
MKPTTSTALLIGGETGGEKVLSDQDTPRLSVAFRFRPQSKQSRRLFCYHDSRLLFLPAALSLVV  
LSIFCSDCSMPRMCTAMAAGFTALETDAEDGSTQMDATTETRFDE TETHFDG TETRFDG TETRF  
DGTETHFDGTETHFDGTETHFDGTETVGEQEGD TDAGVTEDTDYNEGIRKSSHNRRGRSKAA  
PLTVHGNSIRDSRRGRQIAILGLVTGVAAFVLLCTMLYKRGKKNSSQGARGAHGRTAETRARR  
SADTSLWAPQGCFFILMNVVFSRSQKPALVQSLQNTTAILGGGKQKQKRG IATGESA

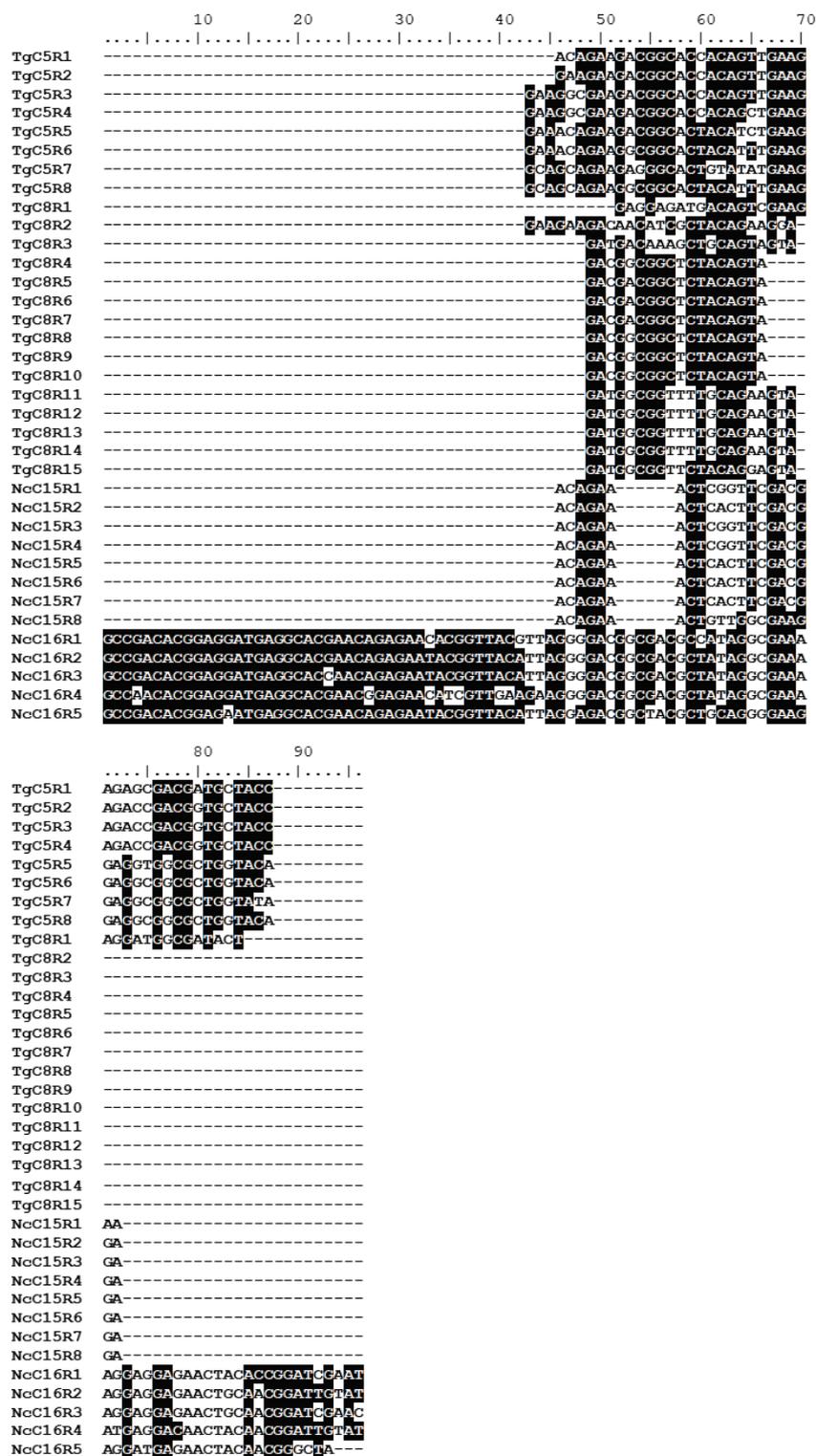
>NcC16  
MFTSVAAAVPIVETTAWDSTADTEDEARTENTVTLGDGDAIGEKEENYTGSNADTEDEARTENT  
VTLGDGDAIGEKEENCNGLYADTEDEARTENTVTLGDGDAIGEKEENCNGSNANTEDEARTENI  
VEEGDGDGAIGENEDNYNGLYADTENEARTENTVTLGDGYAAGEEDENYNGLNRSITLNTKGRS  
SLPGVARRGSTIQRSGRVQRRLSFLISVALFVLCAMLFKAGKRQIGA

>NcC17  
MCTSLAAAVSTGETTAWESTVDEDEARTENAVTLGDS DATDEEDENYNGPHADVEDDARTENM  
VTEEDGDAAGEDDKNYNGLNRSSTLKRKRQSSRSVAMRVSTIQRSGRVQEMRLGVLSVAL  
FVLCAMLFKAGKRQIGA

membranas, en algunos casos de localización subtelomérica (Lorenzi, et al. 2016; Omelianczyk et al. 2020; Ramirez 2019). Para esto se utilizó la base de datos ProtVirDB en búsqueda de asociación a factores virulentos de diferentes especies. El análisis comparativo se realizó utilizando la herramienta ViroBLAST con los parámetros por defecto. Para algunos productos proteicos se obtuvieron más de una



**Fig.38. Distintos marcos de lectura en proteínas FamC con motivos repetitivos.** Se muestran las regiones codificantes para los motivos repetitivos en *tgc5*, *tgc8* y *nnc15*. R: marco de lectura (*repeat*).



En la **Tabla 10** se listan todos los factores virulentos que se encontraron asociados a las proteínas de la familia FamB (TgB/HhB/NcB/Blike) y FamC (TgC/HhC/NcC). También se incluyeron al

menos un miembro de la FamA, FamD, FamE y la de antígenos de superficie (sin dominios transmembrana) SRS. Las proteínas de las FamA a FamE no están incluidas en las bases de datos, por lo tanto, el programa se verá forzado a identificar proteínas probables de otras familias de factores de virulencia según su similitud. Por el contrario, SAG1 pertenece a una familia de genes que, si está incorporada a la base de datos y de esta manera los *scores* mas fuertes estarán relacionados a esta familia.

Los *scores* de alineamiento no superaron el 30% de identidad para los miembros de la FamB y FamC, FamD y FamE, mientras que para un miembro de la FamA fue del 42% (*Tabla 10*). Todas ellas identificaron proteínas de la familia Trans-Sialidasas (TS) de *T. cruzi*, de establecimiento de la patogénesis de *Babesia bovis* y a través de la variación antigénica de *Trypanosoma brucei* y de *P. falciparum*, como así también moléculas de adhesión también implicadas en los mecanismos de establecimiento de la infección en *P. falciparum* (*Tabla 10*). Si bien estas proteínas fueron asociadas a algún factor patogénico, los valores de *score* e *E-value* no son suficientemente robustos para aseverar la correlación de dichas moléculas. Por otra parte, dada la diversidad de blancos, que además involucra diferentes regiones de esos antígenos según el caso (dato no mostrado) se podría inferir que su relación con dichas proteínas no es suficiente para sugerir alguna relación. Más aún, las proteínas de la familia SRS presentaron valores de *score* altos con respecto a otros miembros de dicha familia, pero también detectaron algunas de estas proteínas a *score* similar que las de la FamB y FamC (dato no mostrado). Cabe mencionar que algunas proteínas no identificaron factores virulentos entre ellas: TgB2, TgB5 y NcB13.

### **11.3. Caracterización de la familia FamC en *T. gondii*: TgCs**

Del árbol filogenético se desprendió que los miembros de la familia C serían los más conservados, agrupados de manera conjunta, lo que denota un gran porcentaje de similitud en sus secuencias (*Fig.34*). Este grado de conservación también fue visible durante los alineamientos (*Fig. 37*).

Nombre	ViroBLASTDB	EuPathDB_ID	Score	E value	Especie	Proteína	Proceso Biológico
TgB1	XP_818862	TcCLB.507875.220	25	3,2	<i>T. cruzi</i>	trans-sialidase, Group II, putative	Invasión
TgB2	NCF	(-)	(-)	(-)	(-)	(-)	(-)
TgB3	XP_001609033	BBOV_I000020	27	0.74	<i>B. bovis</i>	variant erythrocyte surface antigen-1 Bb	Establecimiento
TgB4	XP_001608796	BBOV_I000020	23	6.8	<i>B. bovis</i>	variant erythrocyte surface antigen-1 Bb	Establecimiento
TgB5	NCF	(-)	(-)	(-)	(-)	(-)	(-)
TgB6	XP_807757	TcCLB.418185.70	24	3.2	<i>T. cruzi</i>	trans-sialidase, Group II, putative	Invasión
TgB7	XP_001609754	BBOV_I000020	27	1.1	<i>B. bovis</i>	variant erythrocyte surface antigen-1 Bb	Establecimiento
TgB8	XP_812720	TcCLB.505365.60	24	4.3	<i>T. cruzi</i>	trans-sialidase, Group VIII, putative	Invasión
TgB9	XP_001351221	PF3D7_0315200	23	8.1	<i>P. falciparum</i>	circumsporozoite- and TRAP-related protein	Adhesión
TgB10	XP_820300	TcCLB.506245.60	24	4.0	<i>T. cruzi</i>	trans-sialidase, Group V, putative	Invasión
TgB11	AAS38514	TVAG_127180	28	0.31	<i>T. vaginalis</i>	CP65 -purine nucleoside phosphorylase	Cisteína proteasa
TgB12	XP_845640	Tb427.BES126.15	25	5.1	<i>T. brucei</i>	variant surface glycoprotein (VSG)	Establecimiento
TgB13	XP_821703	TcCLB.508873.60	26	1.9	<i>T. cruzi</i>	trans-sialidase, Group V	Invasión
TgB14	XP_807449	TcCLB.508581.40	27	0.90	<i>T. cruzi</i>	trans-sialidase, Group II, putative	Invasión
TgB15	XP_845135	Tb927.5.4730	24	4.7	<i>T. brucei</i>	variant surface glycoprotein (VSG, atypical)	Establecimiento
HhB3	XP_001609033	BBOV_I000010	26	2.2	<i>B. bovis</i>	variant erythrocyte surface antigen-1 Bb	Establecimiento
HhB6	XP_001610496	BBOV_I000010	27	2.3	<i>B. bovis</i>	variant erythrocyte surface antigen-1 Bb	Establecimiento
HhB7	XP_001350798	PF3D7_1240600	26	0.68	<i>P. falciparum</i>	erythrocyte membrane protein 1, PfEMP1	Establecimiento
HhB8	XP_818883	TcCLB.510729.280	25	3.0	<i>T. cruzi</i>	trans-sialidase, Group III, putative	Invasión
HhB9	XP_001611365	BBOV_I000010	27	0.54	<i>B. bovis</i>	variant erythrocyte surface antigen-1 Bb	Establecimiento
HhB10	XP_804812	TcCLB.506035.4	24	1.9	<i>T. cruzi</i>	trans-sialidase, putative (fragment)	Invasión
HhB16	XP_805548	TcCLB.509931.20	28	0.35	<i>T. cruzi</i>	trans-sialidase, Group VIII, putative	Invasión
NcB13	NCF	(-)	(-)	(-)	(-)	(-)	(-)
NcB17	XP_804432	TcCLB.509771.10	28	0.28	<i>T. cruzi</i>	trans-sialidase, Group II, putative	Invasión
NcB18	XP_804173	TcCLB.463279.20	27	0.72	<i>T. cruzi</i>	trans-sialidase, Group II, putative	Invasión
NcB19	XP_001351563	PF3D7_0426000	25	1.8	<i>P. falciparum</i>	erythrocyte membrane protein 1, PfEMP1	Establecimiento
TgH56Blike	XP_804800	TcCLB.510419.20	24	1.4	<i>T. cruzi</i>	trans-sialidase, Group VI, putative	Invasión
TgH57Blike	XP_821669	TcCLB.511255.10	27	0.14	<i>T. cruzi</i>	trans-sialidase, Group V, putative	Invasión
HhH56Blike	XP_816556	TcCLB.506757.60	22	8.5	<i>T. cruzi</i>	trans-sialidase, Group V, putative	Invasión
HhH57Blike	XP_806431	TcCLB.508853.20	26	0.47	<i>T. cruzi</i>	trans-sialidase, Group V, putative	Invasión
TgC1	XP_814156	TcCLB.506965.170	34	0.008	<i>T. cruzi</i>	trans-sialidase, Group V	Invasión
TgC2	XP_814156	TcCLB.506965.171	28	0.15	<i>T. cruzi</i>	trans-sialidase, Group VI	Invasión
TgC3	XP_821904	TcCLB.511183.230	26	1.4	<i>T. cruzi</i>	trans-sialidase, Group V	Invasión
TgC4	BAC44837	PF3D7_0102500	30	0.093	<i>T. cruzi</i>	erythrocyte binding antigen-181	Adhesión
TgC5	XP_821084	TcCLB.510279.230	32	0.056	<i>T. cruzi</i>	trans-sialidase, Group VI	Invasión
TgC6	XP_806774	TcCLB.508753.20	28	0.28	<i>T. cruzi</i>	trans-sialidase, Group V, putative	Invasión
TgC7	XP_822017	TcCLB.508221.650	24	6.2	<i>T. cruzi</i>	trans-sialidase, Group V, putative	Invasión
TgC8	XP_822026	TcCLB.508221.870	33	0.027	<i>T. cruzi</i>	trans-sialidase, Group V, putative	Invasión
TgC10	XP_001351221	PF3D7_0315200	26	2.7	<i>P. falciparum</i>	circumsporozoite- and TRAP-related protein	Adhesión
TgC11	XP_807541	TcCLB.507773.10	27	2.0	<i>T. cruzi</i>	trans-sialidase, Group IV, putative	Invasión
TgC12	XP_807798	TcCLB.509685.40	28	0.48	<i>T. cruzi</i>	trans-sialidase, Group V, putative	Invasión
TgC13	XP_822017	TcCLB.508221.650	25	4.7	<i>T. cruzi</i>	trans-sialidase, Group V, putative	Invasión
TgC14	XP_822017	TcCLB.508221.651	24	6.5	<i>T. cruzi</i>	trans-sialidase, Group V, putative	Invasión
HhC6a	XP_821405	TcCLB.504081.80	27	0.73	<i>T. cruzi</i>	trans-sialidase, Group V, putative	Invasión
HhC6b	XP_001347288	PF3D7_1000100	29	0.16	<i>P. falciparum</i>	erythrocyte membrane protein 1, PfEMP1	Establecimiento
HhC8	XP_001219219	Tb927.1.5300	29	0.16	<i>T. brucei</i>	variant surface glycoprotein (VSG), putative	Establecimiento
HhC11	XP_001350935	PF3D7_1255200	28	0.54	<i>P. falciparum</i>	erythrocyte membrane protein 1, PfEMP1	Establecimiento
NcC15	XP_829794	Tb927.11.20550	26	3.9	<i>T. brucei</i>	variant surface glycoprotein (VSG)	Establecimiento
NcC16	XP_001351561	PF3D7_0425800	28	0.5	<i>P. falciparum</i>	erythrocyte membrane protein 1, PfEMP1	Establecimiento
NcC17	XP_001351435	PF3D7_0412400	27	0.58	<i>P. falciparum</i>	erythrocyte membrane protein 1, PfEMP1	Establecimiento
TgA1	XP_001351221	PF3D7_0315200	42	7,00E-05	<i>P. falciparum</i>	circumsporozoite- and TRAP-related protein	Adhesión
TgD1	XP_821378	TcCLB.511173.80	27	2.8	<i>T. cruzi</i>	trans-sialidase, Group V	Invasión
TgE1	XP_001611366	BBOV_I0002320	28	6.2	<i>B. bovis</i>	variant erythrocyte surface antigen-1	Establecimiento
SAG1	AAO72426	TGME49_233460	523	e-150	<i>T. gondii</i>	SAG-related sequence SRS29B	Adhesión
SRS16B	AAK82887	TGME49_320180	230	1,00E-61	<i>T. gondii</i>	SAG-related sequence SRS16C	Adhesión

**Tabla 10. Análisis de factores virulentos asociados a proteínas FamB y FamC.** Búsqueda de factores virulentos a través de base de datos (<http://bioinfo.icgeb.res.in/protvirdb/blast/viroblast.php>), y comparación con base de datos específica de organismos patógenos eucariotas (<https://veupathdb.org/veupathdb/app>). NCF: *Non Coincidence Found*.

Al revisar las secuencias genómicas en la base de datos ToxoDB, se observaron diferencias de secuencias con algunas cepas. Los genes *tgc7*, *tgc8* y *tgc10* mostraban diferencias en el número de intrones entre las cepas GT1 y VEG de *T. gondii* (Tabla 8), generando principalmente diferentes

secuencias de aminoácidos en la región variable C-terminal (**Fig. 37**) y (**Fig.39**). La diferencia más importante es que la proteína TgC2 es más larga en la cepa GT1 comparada a la secuencia TgC2 de ME49 (**Fig.39**), la cual estaría codificada solo por el exón 1 de manera truncada.

La existencia de 3 genes en *N. caninum*, 4 en *H. hammondi* y 14 en *T. gondii* indicaría que en este último se desencadenó una amplificación de los genes de la familia C. Dado que las regiones centrales repetidas muestran diferencias en el nivel de la secuencia de aminoácidos que podrían surgir de cambios en el marco de codones, se utilizaron las secuencias nucleotídicas codificantes (ARNm) para establecer su relación filogenética. Este análisis mostró que la familia C en *N. caninum* está agrupada separada de *T. gondii* y *H. hammondi* indicando que *ncc16* sería el gen putativo más ancestral (**Fig. 40A**). A su vez, este gen más antiguo presenta una región central con un elevado número de motivos repetitivos. La topología del árbol filogenético obtenido muestra una diversificación en *T. gondii* en la que los genes *tgc5* y *tgc8*, ricos en motivos repetitivos, forman un grupo mientras que, otro grupo de genes con menos o ninguna presencia de repeticiones forman otro agrupamiento. En *H. hammondi*, los genes se agrupan con aquellos con el menor número de repeticiones. La explosión recíproca y el análisis filogenético nos permiten definir ortología y paralogía entre los genes de *T. gondii* y *H. hammondi* representados por los mismos nombres (**Tabla 8, Fig. 40A**). En el caso de HhC8, ToxoDB sugiere esa identidad, mientras que en el árbol filogenético aparece agrupado con TgC11 y HhC11.

En conclusión, el modelo básico de genes de la familia C está formado por al menos el exón 1 y 2, mientras que, a nivel de proteína, tiene un péptido señal, una región repetitiva y una región TM (**Fig. 40B**).

### **11.3.1 Caracterización de la expresión génica *tgc***

#### **11.3.1.1 Expresión en condiciones normales de crecimiento**

Para corroborar el análisis *in silico* realizado en las posibles regiones TSS sobre los genes *tgc*, decidimos indagar acerca de la expresión de los mismos en parásitos de la cepa de referencia tipo I: RH. En la **Fig.41A** se presenta una tabla con el detalle de cada uno de los fragmentos esperados. Como primera aproximación de expresión, realizamos RT-PCR de todos los genes *tgc* en taquizoítos

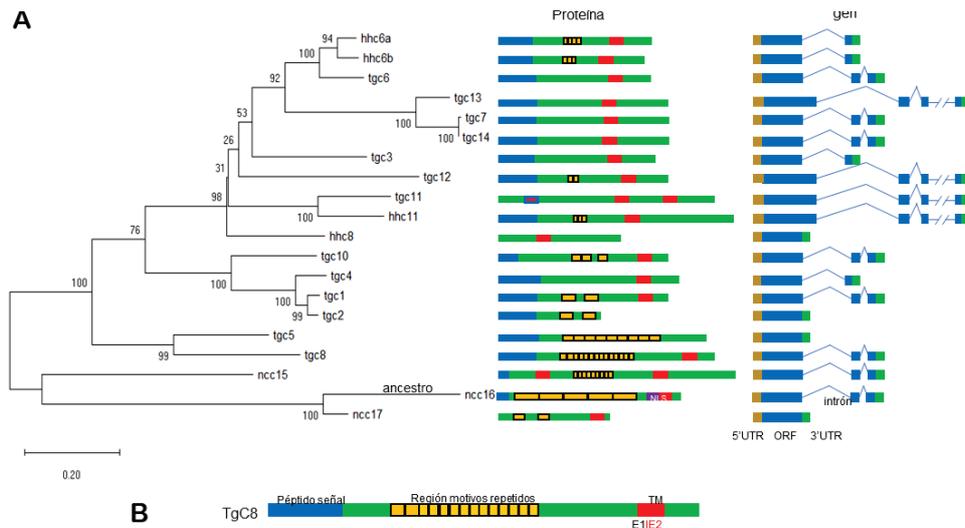


extracelulares crecidos en condiciones normales de cultivo *in vitro*. Como se puede observar en la **Fig.41B**, la mayoría de los genes se expresarían, inclusive *tgc2* y *tgc6*, para los cuales no existían datos registrados; mientras que los genes *tgc4* y *tgc12* no muestran productos de amplificación en estas condiciones normales de cultivo. El gen *tgc5* es un gen sin intrones, por lo cual el producto de su RT-PCR sería idéntico al de PCR sobre gDNA, por consiguiente, en el resultado se debería considerar algún tipo de contaminación. Sin embargo, no se observaron bandas de gDNA para otros genes de interés y controles sugiriendo que *tgc5* se expresaría. A su vez, el gen *sag1* de taquizoito y *tubulina*, constitutivos, dieron amplificación esperada para ADN copia (ADNc), mientras que para el gen *bag1*, que no se expresa en taquizoito, no se observó banda del tamaño esperado, aunque si algunas amplificaciones inespecíficas, tal vez debidas a las condiciones utilizadas en las PCRs. Es importante aclarar que para *tgc7* se obtuvo un producto de menor tamaño al esperado. Al analizar la base de datos, se observó que este gen no estaba en sintenia con las otras cepas. Por lo tanto, se realizó una búsqueda por homología de secuencias en la base de datos ToxoDB y se encontró una copia de menor tamaño del gen *tgc7* (TGRH88\_000020) en el genoma de RH en otra posición genómica, en sintenia con otras dos cepas (RUB y CatPCR2). Por lo cual, el producto génico debería secuenciarse para establecer su identidad. En la **Fig.41C** se muestra una comparación de los niveles de expresión relativos cuantificados por qPCR, relativizados al gen de expresión específica de taquizoítos, *sag1*; donde, se puede observar la expresión de los genes *tgc* evaluados en la cepa RH en condiciones normales de cultivo *in vitro*. En el caso de los genes *tgc4*, *tgc5* y *tgc12* no se observó expresión.

#### **11.3.1.2 Expresión génica en condiciones de estrés ambiental**

Debido a lo expuesto anteriormente, y a la posible implicancia que estos genes podrían tener ante la respuesta ambiental, se midió la expresión génica en parásitos que hubieran sufrido un estrés inducido. Para esto, se cuantificaron y analizaron los niveles relativos de expresión en taquizoítos extracelulares crecidos en condiciones normales en comparación con taquizoítos sometidos a distintas condiciones de estrés ambiental. Para esto se midieron los niveles de expresión de algunos genes *tgc*

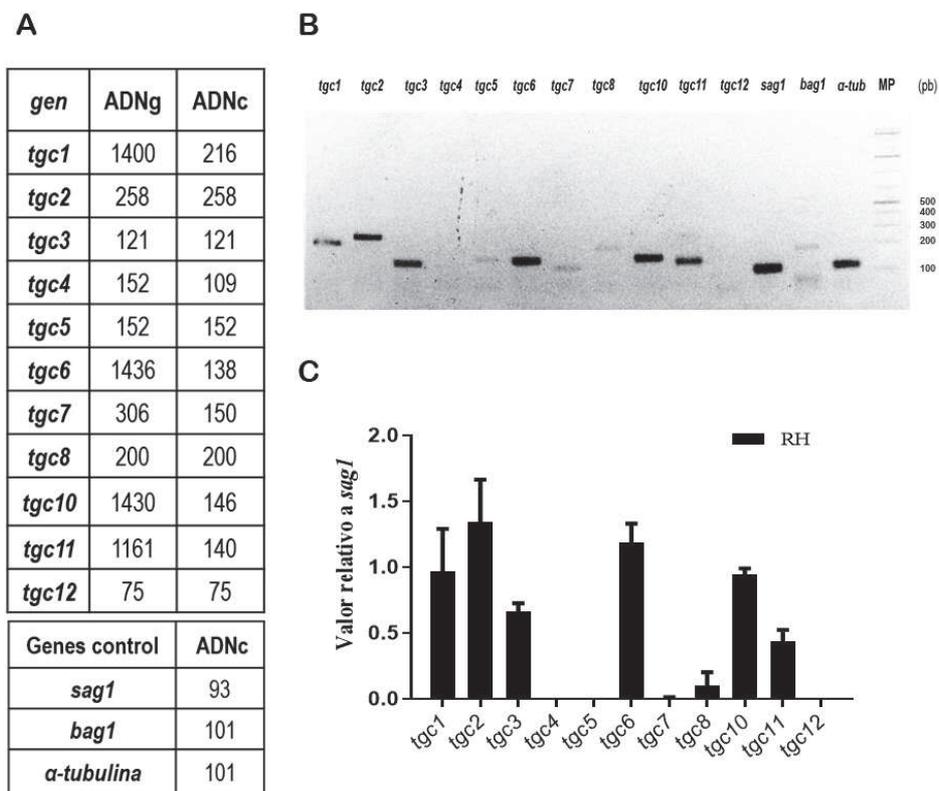
**Fig. 40. Diversidad evolutiva de los genes de la familia C. A. Izquierda,** cladograma reconstruido a partir de las secuencias nucleotídicas de los genes de la familia C. Los valores en los nodos indican el grado de similitud entre los distintos grupos tras 1000 procesos iterativos. En la regla de abajo se indica el peso del árbol, en donde cuanto más cercano a 0 sea el valor más relacionadas están las secuencias. *Medio,* modelos estructurales de las proteínas. En azul se demarca la región conservada y en rojo el dominio TM. En color amarillo los motivos repetitivos. NLS, señal de localización nuclear. *Derecha,* modelo estructural de los genes, N-terminal, péptido señal (ocre), exón 1: azul, exón 2: verde. **B.** Estructura de TgC8, E1-E2, marca la unión de los dos exones.



por qPCR en condiciones normales, tratados una hora a pH alcalino (el cual se sabe induce la diferenciación a bradizoíto), incubados una hora a un estrés térmico por exposición a una temperatura elevada a 44°C produciéndoles un *shock* térmico y una incubación en solución fisiológica (SF), una condición en la cual el parásito estaría en un ambiente de inanición por falta de nutrientes (Lin et al. 2011) (**Fig.42**). En todos los casos se activó la expresión del gen específico de bradizoíto *bag1* bajo condiciones de estrés, como había sido reportado previamente (Dalmasso et al. 2009). Se pudo observar que existen variaciones significativas en los niveles de expresión de cada gen *tgc* según el tipo de tratamiento aplicado. Ante la presencia de un estrés alcalino (**Fig.42A**) la expresión de los genes *tgc* disminuiría en todos los casos con respecto a la expresión en taquizoítos no tratados, siendo significativa en *tgc3*, *tgc5*, *tgc8*, *tgc10* y *tgc11*. En el caso del tratamiento térmico, las expresiones de *tgc5* y *tgc10* se incrementaron significativamente, mientras que la expresión de *tgc6* y *tgc8* descendería (**Fig.42B**). La

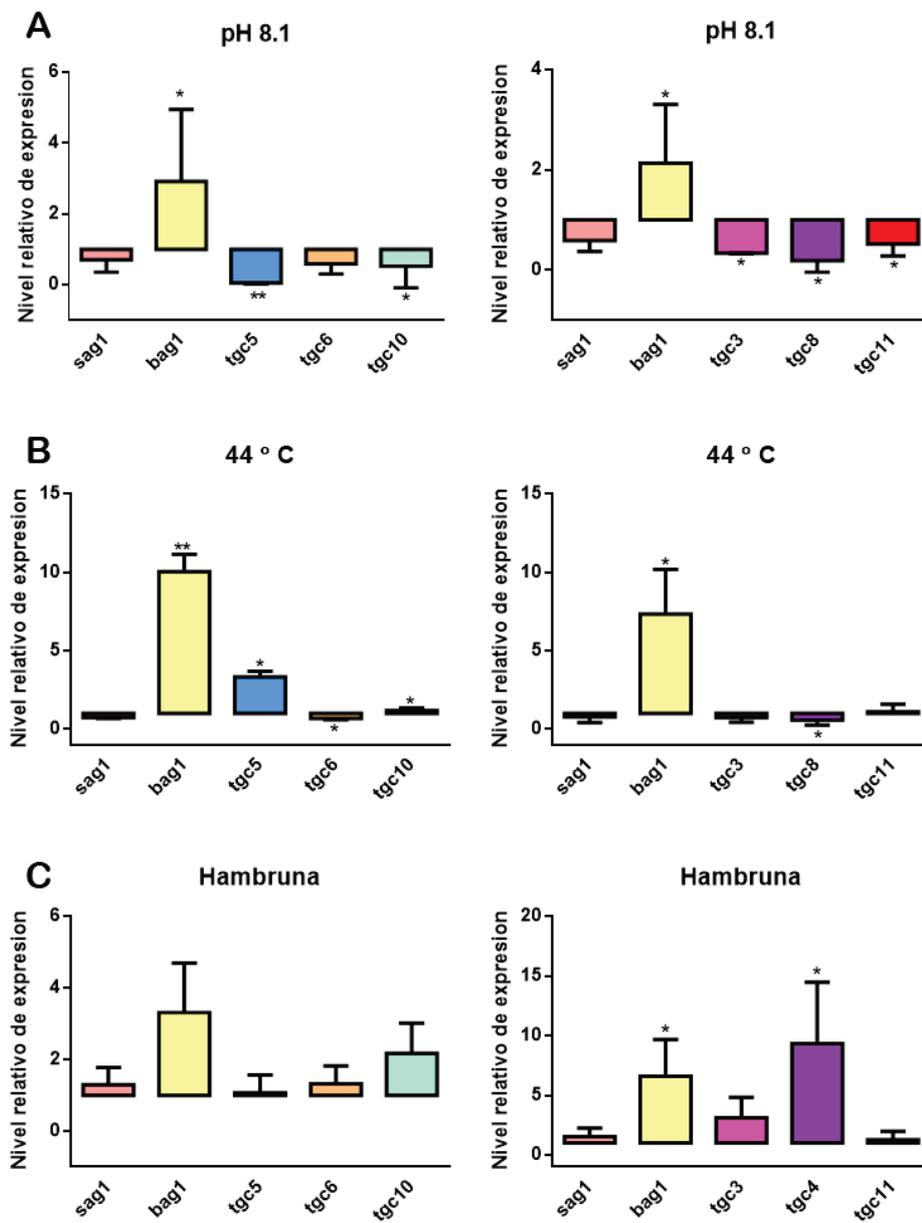
restricción de nutrientes no pareció afectar la expresión de los genes *tgc*, con excepción de *tgc4*, el cual parecía pasar de silenciado a expresarse (Fig.42C).

**Fig.41. Expresión de los genes *tgc* en condiciones normales.** A. Tabla con detalle del tamaño esperado para cada gen en pares de bases (pb). B. Expresión en la cepa RH, RT-PCR, gel de agarosa 1,5%. Figura representativa de 3 experimentos. C. Cuantificación de la expresión génica en taquizoítos extracelulares de la cepa RH relativa a la expresión del gen específico de taquizoíto, *sag1*. ANOVA de una vía, prueba de comparaciones múltiples de Sidak's  $\alpha = 0.05$ , valor de  $p < 0,0001 = ****$ . Se realizaron tres experimentos independientes con sus respectivas réplicas por gen cada uno.



Por lo tanto, podemos decir que casi todos los genes *tgc* se expresarían en condiciones normales de crecimiento en el estadio extracelular de *Toxoplasma*, pero que su expresión podría estar regulada en forma diferencial ante determinadas condiciones de estrés ambiental.

**Fig.42. Cuantificación en tiempo real de la expresión de los genes *tgc* bajo condiciones de estrés.** **A.** Expresión en taquizoítos extracelulares de la cepa RH tratados por 1 hora a pH 8.1. **B.** Expresión en taquizoítos tratados por 1 hora a 44°C. **C.** Expresión en taquizoítos tratados por 1 hora con SF (hambruna). Todas las cuantificaciones fueron normalizadas a la expresión del gen de  $\alpha$ -tubulina de expresión constitutiva y relativizadas a la expresión de los genes en taquizoítos en condiciones normales. Valor de  $p < 0,0001 = ****$ , muy significativo. Cada experimento se realizó por triplicado.



## **11.3.2 Caracterización de la expresión proteica TgC**

### **11.3.2.1 Producción de proteínas recombinantes**

Una vez confirmada la expresión de los genes *tgc*, decidimos escoger algunos de ellos para expresarlos como proteínas recombinantes, generar anticuerpos específicos para subsecuentemente determinar su localización subcelular y posible rol inmunomodulador. Los productos proteicos escogidos se seleccionaron teniendo en cuenta evidencias de expresión observada en los ensayos de RT-PCR de dichos genes.

A continuación, se describe detalladamente las características estructurales de cada una de las proteínas de interés: TgC3, TgC5, TgC6 y TgC8, como así también las particularidades que se tuvieron en cuenta para cada una de las proteínas recombinantes generadas. Cabe destacar que para este análisis utilizamos las secuencias disponibles en la base de datos ToxoDB para la cepa TGME49. Todas las proteínas mencionadas tienen evidencia de transcriptómica y/o transcripción por RT-PCR (*Tabla 8*) (*Fig.41B*).

#### **11.3.2.1.1 rTgC3**

La proteína TgC3 presentaría un sitio de clivaje entre los aminoácidos 46 y 47, y liberaría un péptido señal de ~46 aminoácidos en el extremo N-terminal. En este caso, el TM se localizaría hacia el extremo C-terminal, entre la glicina 152 y la leucina 168. Por lo tanto, se clonó de manera comercial en el vector pET14b la región comprendida entre el aminoácido 48 y el aminoácido 151, y la región C-terminal entre el aminoácido 169 y el aminoácido 206 quedando una proteína recombinante, rTgC3, codificada por un transcripto de 432 pb. El péptido recombinante final, rTgC3, con la etiqueta de 6 histidinas (6xHis) y el sitio ATG aportados por el plásmido está codificado por 495 pb y 164 aa con un peso estimado de ~17 KDa.

#### **11.3.2.1.2 rTgC5**

La proteína TgC5 presentaría un sitio de clivaje entre los aminoácidos 51 y 52, liberando un péptido señal de 51 aminoácidos. Presentaría un dominio hidrofóbico transmembrana entre la

isoleucina 43 y la valina 59 y también presentaría una secuencia señal de localización nuclear NUCDISC: PSRRKVR, en el aminoácido 26. Por lo tanto, en este caso se clonó en vector de expresión pRSET-A de manera estándar a partir del aminoácido 60 hasta el extremo C-terminal. El péptido recombinante final, rTgC5, con la etiqueta de 6xHis y la región desde el ATG aportada por el plásmido está codificado por 783 pb y 261 aa con un peso molecular de ~27 KDa.

#### *11.4.1.3 rTgC6*

La proteína TgC6 tendría un posible sitio de clivaje entre los aminoácidos 48 y 49, que parecería ser un péptido señal clivable. La predicción arrojó que la proteína presentaría un TM probable en el extremo N-terminal entre los aminoácidos 36 y 54 y un segundo entre los aminoácidos 148 y 164. Por lo tanto, la región clonada fue entre los aminoácidos 55 y 147 en el vector pET14b de manera comercial. El péptido recombinante final, rTgC6, tendría desde el ATG del plásmido, incluyendo el 6xHis un tamaño de 474 pb, posee un peso molecular de ~17 KDa.

#### *11.4.1.4 rTgC8*

De todas las proteínas de esta familia, TgC8 es la única con datos de proteómica. TgC8 presentaría un posible sitio de clivaje entre los aminoácidos 59 y 60, liberándose un péptido señal de 59 aminoácidos. Además, presentaría según la predicción dos dominios transmembrana, uno a partir de la isoleucina 55 hasta la valina 71, que incluiría el sitio de clivaje; y un segundo dominio transmembrana a partir de la metionina 267 y la alanina. Por lo tanto, se clonó de manera estándar en el vector de expresión pRSET-A la región no transmembrana comprendida entre la cisteína 75 hasta la prolina 265, quedando así un amplicon de 573 pb. La proteína recombinante final, rTgC8, tendría desde el ATG del plásmido, incluyendo el 6xHis hasta el codón terminador del plásmido un tamaño de 750 pb, estaría codificada por 249 aa y presentaría un peso molecular de ~25 KDa.

### 11.3.2.2 Producción de anticuerpos específicos

Para llevar a cabo la producción de anticuerpos policlonales específicos de cada una de las proteínas seleccionadas, expresamos de manera recombinante cada una de ellas (rTgC3, rTgC5, rTgC6 y rTgC8, respectivamente). Todas fueron clonadas en vectores de expresión en bacterias, donde rTgC5 y rTgC8 fueron obtenidas por clonado en el vector pRSET-A de manera estándar y rTgC3 y rTgC6 en el vector pET14b a partir de síntesis de manera comercial. Todas fueron expresadas mediante la inducción con IPTG y purificadas en columnas de níquel en condiciones desnaturalizantes. Todas las proteínas recombinantes se expresaron de forma soluble, y se resolvieron en geles SDS-PAGE 15% (*Fig.43A*). Los tamaños de las proteínas recombinantes no se condijeron con los tamaños esperados en el caso de rTgC5 y rTgC8. Al purificarlas observamos dos bandas, una de menor tamaño de entre 20-25 KDa y otra de mayor tamaño cercana de los 40-45 KDa en ambos casos (asteriscos negros en la *Fig.43A*). Por otra parte, realizamos una inmunodetección de las proteínas recombinantes utilizando un anticuerpo  $\alpha$ -histidina, el cual reconoce específicamente la etiqueta de poli histidina presente en cada una de ellas (*Fig.43B*). En todos los casos pudimos constatar la presencia de bandas aproximadas al tamaño esperado en el caso de rTgC3 y rTgC6, pero en el caso de rTgC5 y rTgC8, observamos solo las bandas de mayor tamaño en ambos casos. Debido a esto y para corroborar que dichas proteínas se correspondían con los productos esperados, todas las bandas se enviaron a secuenciar por espectrometría de masas.

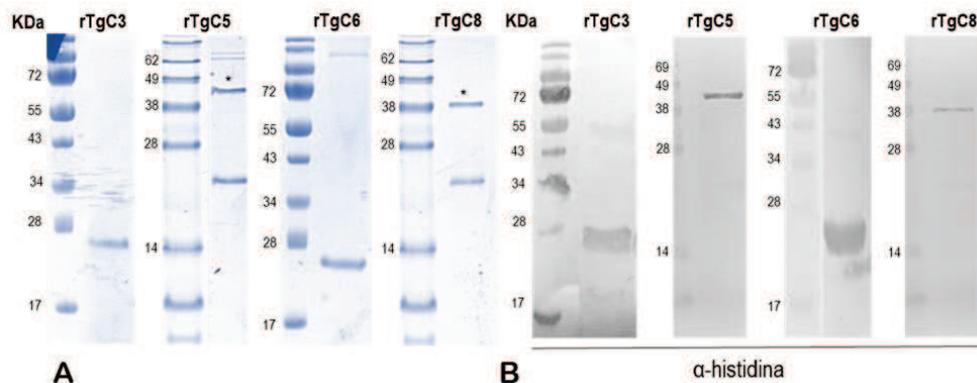
A través de los resultados de espectrometría de masa se confirmó que la banda de mayor tamaño se correspondía con las proteínas recombinantes esperadas en el caso de rTgC5 y rTgC8 mientras que la banda de menor tamaño no trajo resultados. Por otro lado, en el caso de rTgC3 y rTgC6 las bandas correspondientes a las proteínas recombinantes se condijeron con las esperadas. Las proteínas recombinantes fueron utilizadas para inmunizar ratones y de esta manera se produjeron los anticuerpos policlonales correspondientes.

### 11.3.2.3 Inmunodetección en condiciones normales de crecimiento

Analizamos la reactividad con extractos de proteína total de taquizoítos extracelulares crecidos en condiciones normales, tanto de la cepa RH como así también de otra cepa disponible en el laboratorio, la cepa tipo II, PRU (en el caso de TgC5 y TgC8). Como control probamos los anticuerpos con extractos proteicos provenientes de la célula hospedadora.

**Fig. 43. Purificación de las proteínas recombinantes (rTgCs) en condiciones desnaturalizantes.**

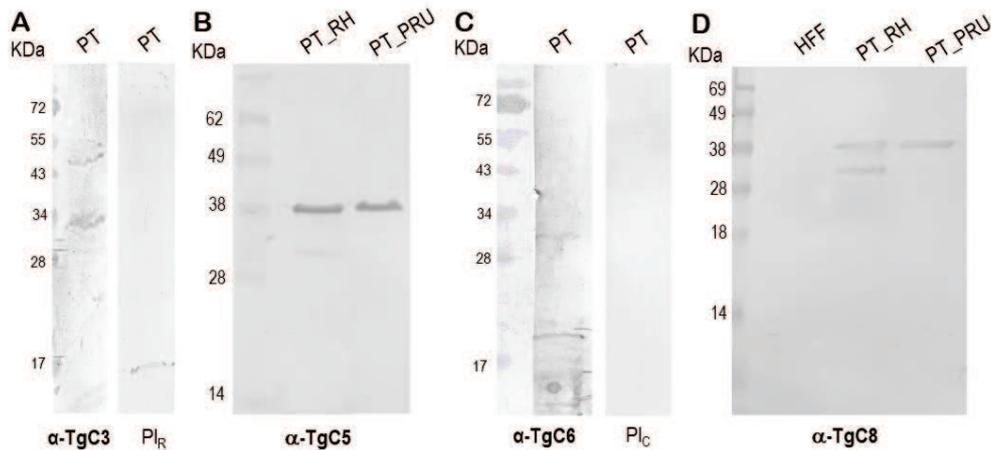
**A.** Tinción con *coomassie blue*, SDS-PAGE 15%. En cada caso se sembró 1 µg de proteína recombinante. Tamaño esperado de rTgC3 ~17 KDa, observado ~25 KDa; tamaño esperado de rTgC5 ~27 KDa, tamaño observado ~25 KDa y ~45 KDa; tamaño esperado para rTgC6 ~ 17 KDa, tamaño observado ~25 KDa; tamaño esperado para rTgC8 ~25 KDa, tamaño observado ~25 KDa y ~40 KDa.  
**B.** WB de las proteínas recombinantes, en todos los casos se sembró 1 µg de proteína recombinante.



Los extractos totales también fueron ensayados con diluciones correspondientes de los sueros pre-inmunización para cada uno de los casos, no resultando reactivos (*Fig. 44*).

Todos los sueros ensayados reconocieron productos proteicos para las respectivas proteínas TgC en taquizoítos extracelulares (*Fig. 44*). El tamaño esperado de la proteína nativa TgC3 era de 22 KDa para la cual se observaron dos bandas de distinto tamaño, una de ~30 KDa y otra de ~50 KDa. En el caso de TgC5 el tamaño esperado era de 29 KDa y se observó una proteína de ~38 KDa. El tamaño esperado para la proteína nativa TgC8 era de 31 KDa y el tamaño observado fue de ~40KDa. A diferencia del resto, para TgC6 se observaron bandas muy poco intensas. Si bien el tamaño esperado

**Fig. 44. Análisis por WB de los anticuerpos  $\alpha$ -TgC con extractos proteicos de *T. gondii*.** Se sembraron extractos de proteínas totales (PT) provenientes de  $1.10^7$  taquizoítos extracelulares de la cepa RH y de la cepa PRU, y  $1.10^7$  células provenientes de fibroblastos de prepucio humano, HFF crecidas en condiciones normales de cultivo. **A.**  $\alpha$ -TgC3, suero de ratón, dilución 1:500. **B.**  $\alpha$ -TgC5, suero de ratón, dilución 1:500. **C.**  $\alpha$ -TgC6, suero de conejo, dilución 1:1000. **D.**  $\alpha$ -TgC8, suero de ratón, dilución 1:500. Pl<sub>R</sub>: pre-inmune de ratón, Pl<sub>C</sub>: pre-inmune de conejo.



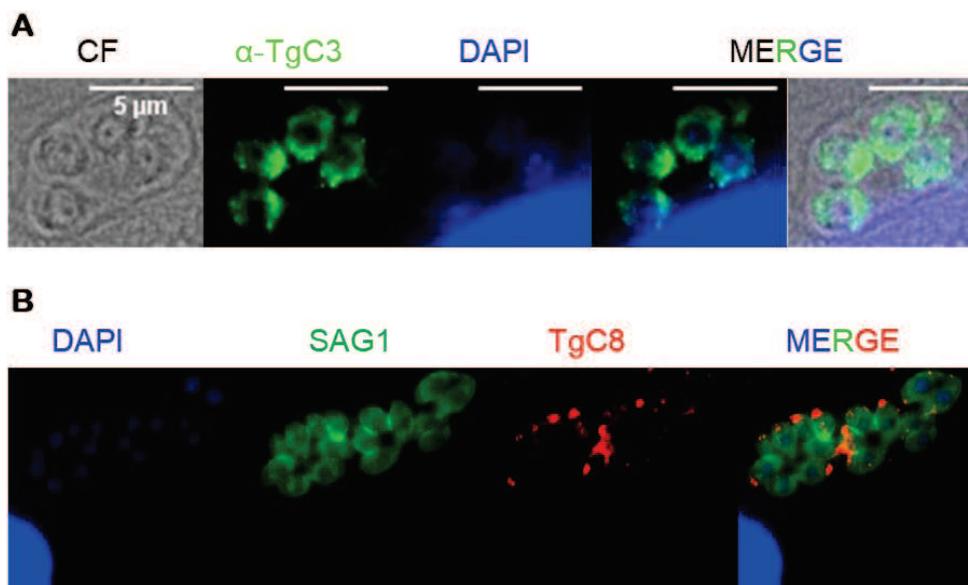
para la proteína TgC6 era de 21 KDa, la presencia de numerosas bandas no permite descartar la presencia de contaminantes, por lo tanto, la especificidad debería ser corroborada.

#### 11.3.2.4 Inmunolocalización subcelular

El siguiente paso fue indagar acerca de la localización subcelular de dichas proteínas y para esto se realizaron ensayos de IFI en taquizoítos intracelulares. En el caso de TgC5 y TgC6 no fue posible localizar a dichas proteínas en taquizoítos intracelulares, consiguiendo únicamente localizar subcelularmente a TgC3 y TgC8. TgC3 se localiza excluyendo completamente al núcleo, concentrada hacia el extremo basal de los parásitos (**Fig. 45A**) mientras que la proteína TgC8 tendría una localización compatible con el extremo apical del taquizoito intracelular (**Fig. 45A**), teniendo en cuenta que el taquizoito se replica y desarrolla con el ápice mirando hacia la periferia en la VP. Cabe destacar que no todos los parásitos están marcados con el anticuerpo anti-TgC8, sugiriendo que el anticuerpo puede ser deficiente para su uso en la técnica de inmunofluorescencia o que no todos los taquizoítos expresan TgC8 al mismo

tiempo. Todos estos hallazgos deben confirmarse con marcadores adecuados (anticuerpos anti-proteínas de las micronemas, Mic; para marcar la región apical y/o transfección con el gen *tgc8* etiquetado, etc).

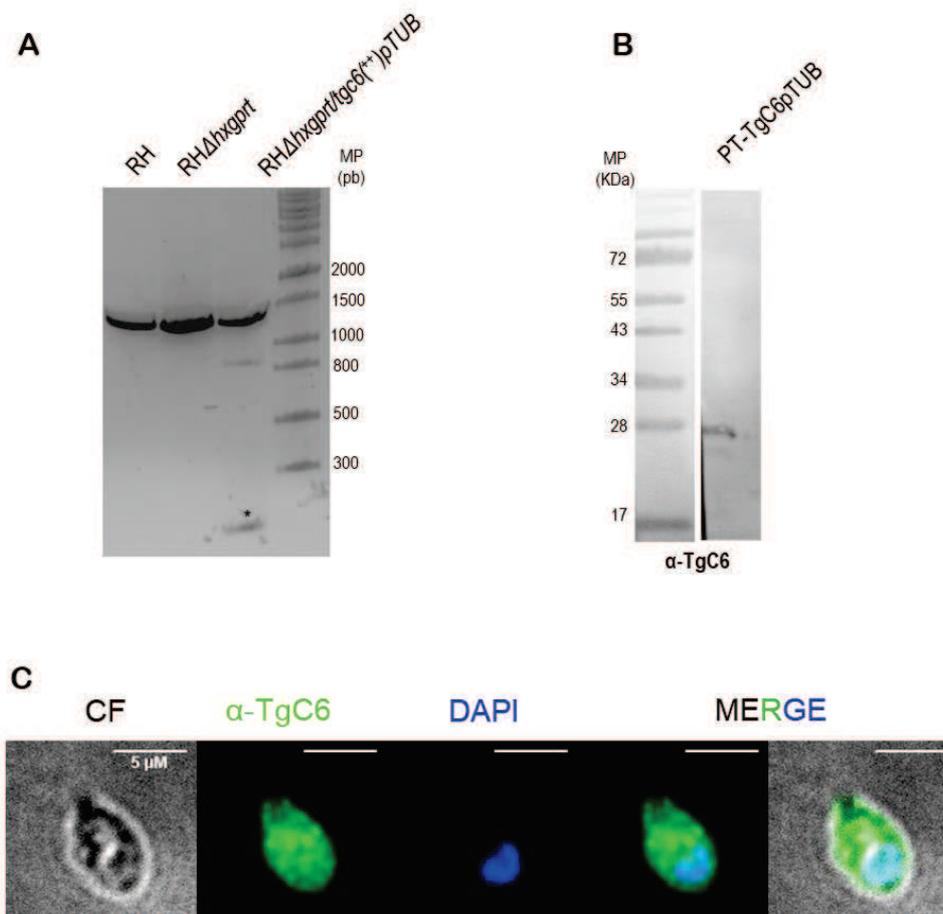
**Fig. 45. Localización subcelular por inmunofluorescencia de proteínas TgC.** A. Localización de TgC3 en taquizoítos intracelulares de la cepa RH, dilución 1:100. B. Localización intracelular de TgC8, dilución 1:500. SAG1: antígeno de superficie 1.



#### 11.3.2.5 Sobreexpresión de TgC6

Con el fin de complementar el análisis de expresión de los genes *tgc* decidimos generar líneas de parásitos que sobreexpresen algunos de estos genes. Comenzamos por generar una línea celular que sobreexpresen el gen *tgc6*. Para esto transfectamos parásitos de la cepa  $RH\Delta hvgprt$  que poseen un cassette de selección para xantina y ácido micofenólico lo cual nos permitió crecer parásitos y seleccionar solo aquellos que incorporaron el plásmido con la copia extra del gen junto a la selección. En primer lugar, evaluamos por PCR la correcta inserción del transgen en el genoma de la cepa parental, en la **Fig. 46A** podemos ver una banda correspondiente a una región del transcripto de la copia extra del gen *tgc6*. Luego evaluamos por WB la presencia de posibles productos proteicos (**Fig. 46B**). Se pudo

**Fig. 46. Sobreexpresión de la proteína TgC6.** **A.** PCR de ADN genómico poblacional de la cepa salvaje (RH), la cepa parental (RH $\Delta$ hxgprt) y la sobreexpresante (RH $\Delta$ hxgprt-tgc6(++)pTUB). Los fragmentos amplificados se corresponden con una región del locus del gen *tgc6*, banda de 1433 pb, y una de menor tamaño de 138 pb (asterisco) correspondiente a una región del transcripto extra del gen. MP: marcador de peso. **B.** WB poblacional de taquizoítos RH $\Delta$ hxgprt-TgC6(++)pTUB. Se sembraron extractos de proteínas totales (PT) provenientes de  $1.10^7$  taquizoítos extracelulares.  $\alpha$ -TgC6 de conejo dilución 1:1000. **C.** Localización de TgC6 en taquizoítos extracelulares de la cepa RH $\Delta$ hxgprt-TgC6(++)pTUB.  $\alpha$ -TgC6 1:100.



observar la presencia de una banda de ~30-31 KDa que se correspondería con el tamaño de la proteína esperada. A nivel celular en taquizoítos extracelulares observamos que la proteína presentaría una marca en todo el parásito (*Fig. 46C*). Es importante resaltar, además, que esta línea sobreexpresante presenta un crecimiento lento y no se pudo obtener clones por lo que se considera necesario realizar más ensayos de caracterización de la cepa y confirmar su localización por otros métodos microscópicos.

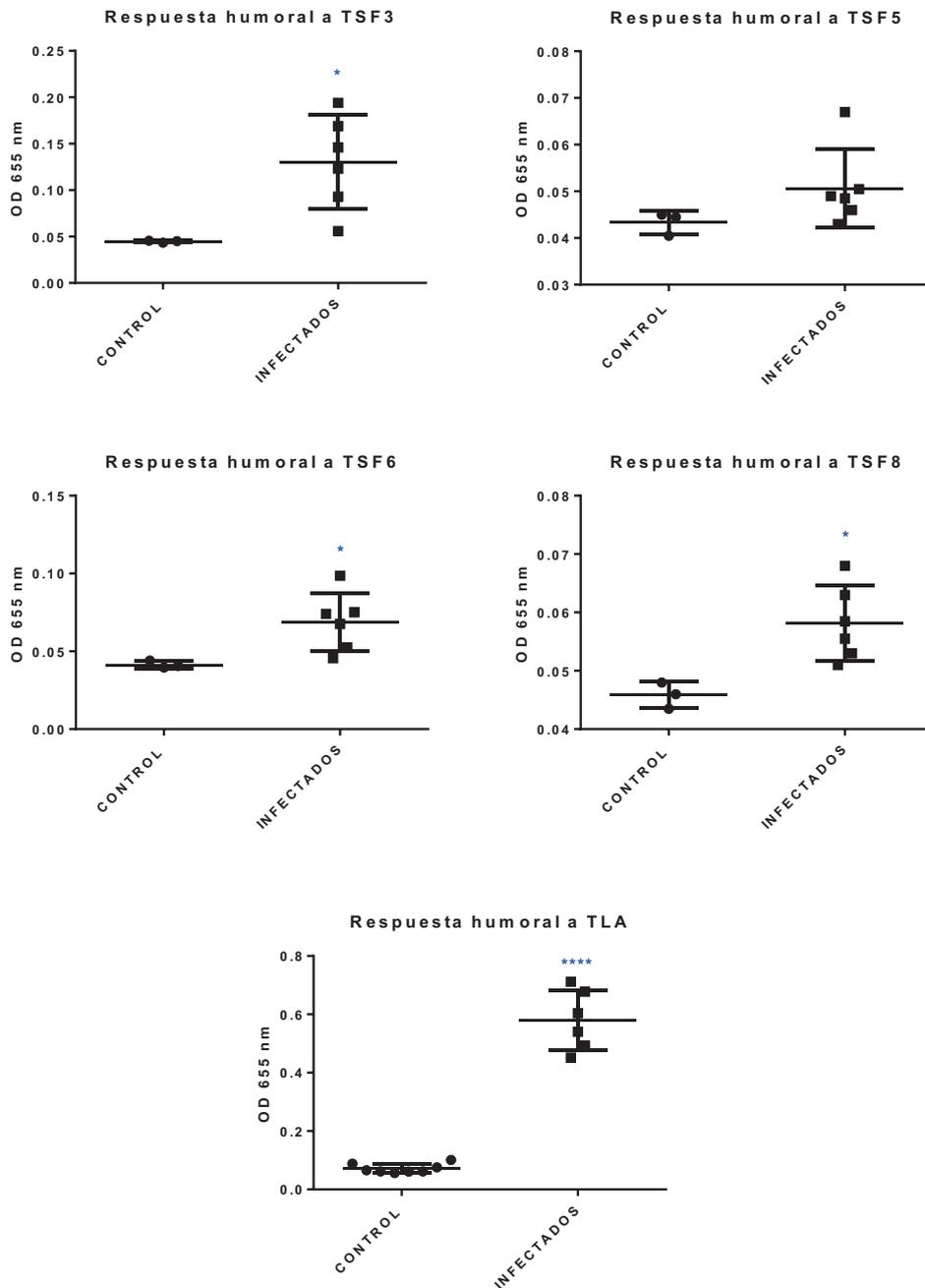
### 11.3.2.6 Ensayos de inmunogenicidad

Con el objetivo de saber si las proteínas TgC recombinantes podrían ser reconocidas como antígenos durante la infección, se evaluó entonces, la capacidad de las mismas en ser detectadas durante la respuesta humoral en ratones infectados con *T. gondii*. Para esto, infectamos ratones intraperitonealmente con taquizoítos de *T. gondii* de la cepa ME49 y extrajimos sangre a los 30 días para analizar la respuesta humoral midiendo inmunoglobulinas totales (IgG) por ELISA contra TgC3 (TSF3), TgC5 (TSF5), TgC6 (TSF6) y TgC8 (TSF8). En la **Fig. 47** podemos observar la respuesta obtenida para cada una de las proteínas como así también para el control positivo, TLA. En todos los casos observamos una reactividad superior al control no infectado, sugiriendo una posible respuesta humoral contra las proteínas TgC aquí evaluadas (**Fig. 47**). En el caso de TgC3, TgC6 y TgC8 la respuesta fue significativa. Estos resultados parecerían indicar una estimulación de la respuesta humoral durante a infección por parte de las proteínas TgC aquí estudiadas, aunque muy leve.

La proteína recombinante rTgC8, es una variante con numerosas unidades de motivos repetidos en su secuencia (**Fig. 36 y Fig. 40**), mientras que TgC6 recombinante (rTgC6), tiene el mínimo de exón 1 y 2 de la estructura proteica, sin motivos repetidos en su secuencia (**Fig. 36 y Fig. 40**). Se evaluó entonces, la reactividad de dichas proteínas ante sueros de ratones infectados de forma experimental con 20 quistes de la cepa ME49 de *T. gondii* por vía oral. Los sueros se obtuvieron en diferentes días después de infección y fueron analizados por IgG-ELISA sensibilizados con rTgC8 o rTgC6. En el caso de TgC8, un pico de reactividad se pudo observar el día 15 post-infección, aunque con valores bajos de OD, que disminuyeron rápidamente en el día 21 y posteriores (**Fig. 48A**). Por el contrario, la reactividad contra rTgC6 fue baja, pero aumentó a lo largo del tiempo de la infección (**Fig. 48A**). En conjunto, rTgC8 parecería ser más sero-reactiva durante la infección aguda, mientras que rTgC6 sería reactiva en caso de infección crónica, en el modelo ratón.

Del mismo modo, se analizaron sueros de mujeres embarazadas agrupadas en negativas y con infecciones toxoplasmáticas crónicas o agudas (**Fig. 48B**). Para ambas proteínas recombinantes,

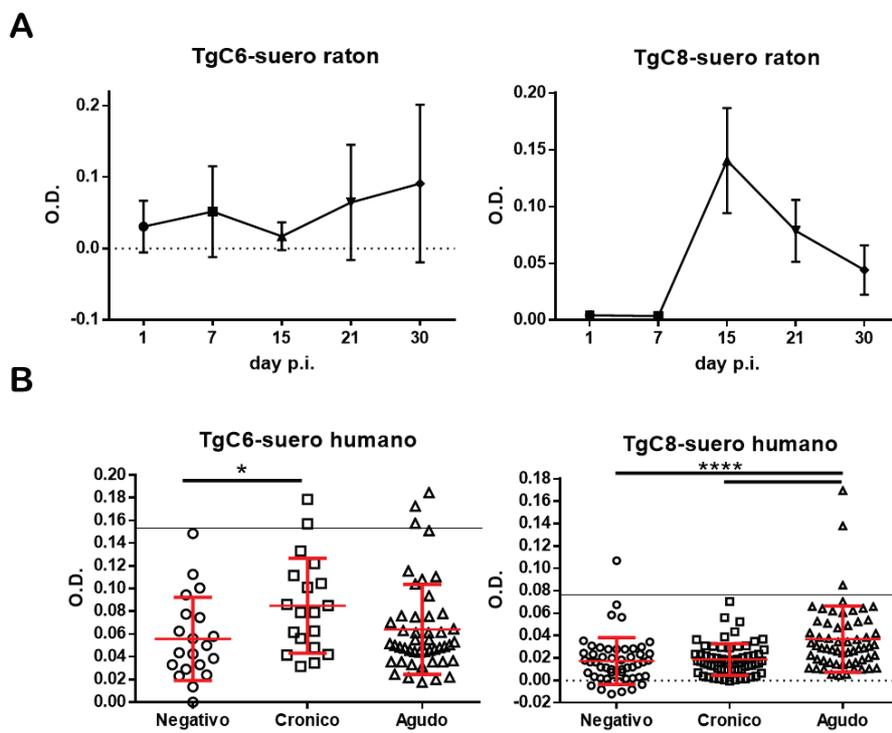
**Fig. 47. Medición de IgG total en la respuesta humoral a proteínas TgC.** Control: ratones no infectados. OD: Absorbancia a 655 nm. *Unpaired t-test*. Valor de  $p < 0,0001 = ****$ , muy significativo.



rTgC6 y rTgC8, solo unos pocos sueros parecerían ser reactivos, superando el valor de corte (**Fig. 48B**). No obstante, los promedios de OD fueron significativamente mayores en el grupo de sueros agudos para rTgC8 y sueros crónicos para rTgC6 (**Fig. 48B**), en la misma línea de lo observado en el modelo de infección en ratones. Estos resultados, sugieren un valor antigénico pobre para estas proteínas durante

la infección en humanos y/o la presencia de diferentes cepas y/o estímulos ambientales en cada individuo.

**Fig. 48. ELISA  $\alpha$ -IgG con TgC6 y TgC8.** La placa se sensibilizó con 5  $\mu$ g/ml por pocillo. Los sueros se utilizaron en una dilución de 1:100. El anticuerpo secundario fue  $\alpha$ -IgG humana de conejo conjugado con peroxidasa, utilizado en una dilución 1:20000 (*Santa Cruz Biotechnology*). El valor de corte se calculó como el promedio de las de los sueros negativos más 1 desviación estándar.



## 12. Discusión del capítulo II

Las regiones subteloméricas son pobres en genes, pero en muchos casos los genes subteloméricos responden a aspectos relevantes de la biología del organismo: variación antigénica, patogenicidad, respuesta a estrés, etc. Por dicha razón, consideramos pertinente conocer en profundidad el perfil génico de las regiones STs de *T. gondii* descritas en el capítulo anterior. Pudimos observar que, de los genes únicos, solo 13 genes tienen una anotación funcional, mientras que el resto (57) fueron anotados como hipotéticos. Dentro de los genes únicos funcionales hay 5 de 13 que codifican para productos cuya localización por LOPIT es nuclear/nucleolar. Es un número modesto, y debería corroborarse su localización. De hecho, algunos parecen tener otras localizaciones en sus contrapartes humanas. Aún así, parecieran focalizarse en tres subcompartimientos: núcleo, membrana/vesículas, mitocondria. Estos datos podrían sugerir un cierto sesgo a la presencia de genes asociados a una vía biológica común debido a posibles arreglos evolutivos. Ese sesgo no sería raro ya que las regiones STs suelen agrupar genes de alguna ruta metabólica o aspecto biológico en particular. En avena se observó la presencia de 12 genes relacionados a una vía biosintética de la defensa contra patógenos, arreglados en forma de *clúster* en una región subtelomérica (Li et al. 2021). Otro aspecto importante fue la observación de que varios genes presentan diferencias en el número entre las diferentes cepas de *T. gondii*, como así también modificaciones en el tamaño de la estructura génica. Esta observación también indicaría que estos genes podrían estar sujetos a cambios asociados a la plasticidad genómica, probablemente asociada a su ubicación en regiones STs. Existen evidencias que sugieren que las regiones subteloméricas son altamente dinámicas con una mayor tasa de duplicación génica que el resto del genoma (Brown, Murray, & Verstrepen 2010).

Los genes hipotéticos surgen de programas que definen posibles exones e intrones. La presencia de ARNm permite una confirmación de su identidad génica. Dentro de los genes hipotéticos,

solamente el 42% de los genes (24 genes) presenta alguna evidencia de su naturaleza génica: proteómica (incluyendo LOPIT) y/o transcripción. De los genes hipotéticos solo dos codificarían productos con homología a la Familia B, TGME49\_321280 (*tgh57*) y TGME49\_321270 (*tgh56*); mientras que otros 9 de ellos (*tgh4*, *tgh18*, *tgh20*, *tgh23*, *tgh28*, *tgh29*, *tgh31*, *tgh38* y *tgh43*) solo tendrían asociación a un dominio TM y/o péptido señal, aunque deberían realizarse más experimentos para su confirmación. Por otra parte, el 49% tendrían únicamente datos de *score* de *fitness* relacionado al fenotipo, pero sin evidencia de expresión alguna; y el restante 9% no tendría ninguna evidencia adicional, además de estar silenciados. Esta posibilidad no los descarta como genes funcionales ya que las regiones subteloméricas estarían bajo la influencia de silenciamiento por el efecto de posicionamiento TPE (Ottaviani, Gilson, & Magdinier 2008; Hernández-Rivas et al. 2010; Scherf, Figueiredo, & Freitas-junior 2001; Harari et al. 2013; revisado por Kwapisz & Morillon 2020). Del análisis de la sintenia de este grupo génico también se desprendió la observación de posibles eventos de recombinación y reordenamiento entre algunos genes resaltando nuevamente la versatilidad de las regiones STs y su dinámica. En resumen, cerca de la mitad de ellos ameritan considerarlos genes reales realizando su interés para comprender sus perfiles de expresión, funciones, localización subcelular, etc; con el fin de ahondar sobre la función general de los STs de *T. gondii*. Un aspecto interesante en este sentido surgió en un trabajo publicado recientemente y discutido al final de esta sección.

Considerando que los genes *tgb* y *tgc* han sido descriptos como importantes para entender como funcionan las interacciones huésped-patógeno (Lorenzi et al. 2016), además de ser buenos candidatos para entender las diferencias entre especies en cuanto a colonización del nicho (Reid et al. 2012), resulta atractivo estudiar este tipo de familias para poder entender los mecanismos claves de supervivencia y adaptación en este tipo de parásitos. Estos genes claramente fueron asociados a familias multigénicas localizadas en regiones subteloméricas. Básicamente, la existencia de dos familias multigénicas de proteínas que expresarían proteínas integrales de membranas en los STs de *T. gondii*, sin una asociación con alguna función molecular, recuerda a las familias multigénicas de variación

antigénica de *Plasmodium sp.*, *T. brucei* y de algunos hongos (Omelianczyk et al, 2020; Ramirez 2019; Cormack et al. 1999, De Las Penas et al. 2003; Hoyer 2001). Sin embargo, no se nos debería escapar que otras familias multigénicas de *T. gondii*, donde también sus genes parecieran codificar para proteínas integrales de membrana (FamA, FamD y FamE), no están localizadas en regiones ST del parásito, así como otros genes de proteínas de superficie SRS (Lorenzi et al. 2016). Por lo tanto, la presencia subtelomérica de FamB y FamC es aún intrigante. Solo poder discernir sus funciones podrá arrojar alguna luz a ese tema.

Algunos genes, tanto de las familias FamB como FamC, mostraron datos de transcriptómica sugiriendo que estos se transcriben, aunque en menores niveles que otros genes (46% y 60% respectivamente en comparación al 96% de los genes con FA). Aquí solo hemos analizado por RT-PCR genes de la familia FamC y pudimos corroborar que la mayoría de ellos se transcribe. Resultó atrayente el hecho de que en la mayoría de ellos se encontró correlación entre en los datos de transcriptómica, nucleosomas activos y RT-PCR. Sin embargo, para algunos de ellos (*tgc5*, *tgc6*, *tgc8* y *tgc12*) no coinciden los datos nucleosómicos, con los de transcripción masivos y la RT-PCR de nuestro laboratorio. Eso puede deberse a que justamente los datos masivos podrían analizar una enorme cantidad de datos y están sujetos a menores niveles de reproducibilidad inter ensayos. De hecho, a veces no coinciden los datos de transcriptoma de RNA-seq con los de *microarray* (van der Kloet et al. 2020).

Los datos observados respaldarían en parte, la asociación a marcadores de sitios activos de la transcripción presentados en la sección 11.1.2.2 (**Tabla 8**) estableciendo una correlación entre la presencia de la marca de histona H3K4me3 y el doble nucleosoma de variantes (H2A.Z/H2B.Z) en la región TSS con la expresión génica activa. Por otra parte, los genes silenciados se asociarían a la presencia de H2A.X (Dalmasso et al. 2009; Bogado et al. 2014; Nardelli et al. 2022) y tal vez la combinación de H2A.X con algunas de las otras histonas podría estar asociada con promotores ambiguos como ya se mencionó. Al analizar los nucleosomas presentes en el promotor de los genes *tgc2* y *tgc5*, la distribución observada sería compatible con genes silenciados (**Tabla 8, Fig.41**). Los

resultados de la RT-PCR serían más sensibles sobretodo en caso de que en la población existan parásitos que expresen el producto mientras otros lo estén silenciando a nivel de célula individual, como también podrían ser cepa específicos.

Los genes *tgc4* y *tgc12* en cambio, podrían asociarse con genes pausados ya que no solo muestran enriquecimientos en H2A.X, sino también al menos en uno de los marcadores de promotores activos por lo que podrían considerarse regiones ambiguas (**Tabla 8, Fig.41**). Por último, cabe mencionar el caso de *tgc7*, el cual no se condijo con lo propuesto para su región TSS y datos de transcriptómica. Cuando se revisó la sintenia se pudo observar que *tgc7* tendría una localización distinta en el genoma de RH que en ME49 y además con un tamaño génico menor, según el análisis comparativo (dato no mostrado). El hecho de que los tamaños de los fragmentos amplificados no se condijeron con los esperados podría corresponderse a las diferencias entre cepas. Mientras que la no detección de su expresión podría deberse a la amplificación inespecífica, ya que las mayores diferencias en la estructura génica de *tgc7* entre las cepas ME49 y RH se dan en las regiones 5' y 3' del gen (esta última donde se diseñaron los oligonucleótidos utilizados con el genoma de ME49).

En este sentido, consideramos que más importante que determinar su expresión génica, cuyos transcritos pueden estar sujetos a regulación traduccional, es determinar su expresión a nivel de proteína. Más aún, lo más relevante es conocer la presencia de los productos proteicos de esos genes de la FamB y FamC a nivel celular, para establecer si hay diferencias en sus perfiles de expresión. El análisis de WB es compatible con la expresión de las proteínas TgC8 (Tsf8), TgC5 (Tsf5) y TgC3 (Tsf3), mientras que el resultado para TgC6 (Tsf6) pareciera mostrar una banda muy débil, indistinguible de posibles contaminaciones menores que pueden ser observadas con sueros policlonales generados a partir de proteínas recombinantes purificadas de bacterias. La marcación por IFI en parásitos intracelulares muestra patrones diferentes en la localización de las proteínas TgC3 y TgC8. Mientras que la primera pareciera ser citosólica con localización predominante hacia región posterior del taquizoito, TgC8 mostraría una marca más apical, y no en todos los taquizoítos. Podría ser que se trate de diferentes

localizaciones de las proteínas o que algún anticuerpo este dando reacciones inespecíficas. Para ello deberíamos generar taquizoítos con etiquetados endógenos que nos permitan resolver la localización de ambas proteínas con anticuerpos anti-etiqueta. No obstante, cabe destacar que las proteínas de las variantes multigénicas Rifin, de *P. falciparum*, presentan diferentes localizaciones subcelulares, siendo algunos de ellos localizados en la región apical y otros citosólicos en el merozoíto (Petter et al. 2007).

En el análisis realizado en búsqueda de posibles factores virulentos asociados a las proteínas TgC, encontramos que muchos de ellos parecieran estar relacionados a proteínas integrales de membranas codificadas por familias multigénicas, tal como ocurre en *T. cruzi*, *T. brucei*, *B. bovis* y *P. falciparum*. Entre ellas las Transialidasas y los antígenos variables de superficie. Dado que se observó una amplia variación entre las regiones que presentan cierta identidad, el bajo *score*, lo limitado de la región de identidad detectada (entre 43 y 60 aminoácidos semi-conservados) probablemente este resultado este mostrando más la gran diversidad de genes variables de esas especies, las cuales terminan encontrando cierto nivel de identidad a nivel proteína por las cientos a miles de versiones de proteínas integrales que conforman a dichas familias, más que una relación filogenética. La presencia de genes FamB y FamC en las tres especies de Apicomplexa, y no en otros Apicomplexas, sugiere que esos genes no tienen origen génico, y se crearon a partir del ancestro común a *N. Caninum*, *H. hammondi*/*T. gondii*, lo cual descartaría un origen común con genes de variación antigénica de otros parásitos del phylum Apicomplexa, pero no descartar un rol similar. Un aspecto interesante es que con respecto a *N. caninum*, esos genes se expandieron en *H. hammondi* y sobre todo en *T. gondii*. Esto debería sugerir una mayor dinámica genómica en *T. gondii* y/o una relevancia biológica de estos genes, aún no determinada. Una vez más, un estudio reciente pareciera arrojar alguna luz sobre este aspecto, el cual es discutido al final de esta sección.

La respuesta humoral contra TgC8 y TgC6 fue baja en absorbancia y, en el caso de humanos, en el número de sueros individuos con toxoplasmosis reactivos. En ese sentido, la inmunoreactividad de los sueros humanos con Malaria contra proteínas VSA de *Plasmodium* también es baja

(entre 10-30%) (Oliveira et al. 2006; Albrecht et al. 2014)). En *Plasmodium vivax*, los genes *vir* no tendrían exclusión alélica, no habiendo expresión clonal de las proteínas Vir en la superficie de los reticulocitos infectados. Más aun, los individuos infectados por primera vez con *P. vivax* mostrarían niveles de IgG anti-Vir, revelando una respuesta humoral adquirida naturalmente (Fernandez-Becerra et al. 2005; Oliveira et al. 2006). Estos datos son consistentes con un papel importante de los genes *vir* en la evasión de la respuesta inmune, pero son inconsistentes con un rol en el sentido estricto de la variación antigénica. Estos datos están de acuerdo, además, con la opinión de que los genes *vir* podrían estar implicados en otro mecanismo de evasión de la respuesta inmune aún no descubierta (del Portillo et al. 2004; Fernandez-Becerra et al. 2009). En este sentido, pudimos ver que los genes *tgc* subteloméricos en *Toxoplasma* no estarían sujetos a una expresión monoalélica y quizás sus bajos niveles de absorbancia y/o inmuno-reactividad estén relacionados con su baja expresión. Más estudios se deberían realizar para dilucidar ese aspecto, incluyendo estudios tipo “*single cell*”. Por ejemplo, sería interesante analizar cómo impacta en la respuesta humoral la sobreexpresión de dichos genes con promotores constitutivos en el modelo ratón.

Un aspecto interesante, es que algunas proteínas FamC poseen motivos repetidos en su región variable amino terminal. El gen ancestral propuesto, *ncc16*, y la proteína TgC8 son ejemplos de los más ricos en esas regiones repetitivas. De manera interesante, esas regiones repetitivas son diferentes en secuencia aminoacídica y parecieran responder a modificaciones del marco de lectura en su evolución. En nuestro caso, podríamos considerar a las proteínas NcC15, NcC16, HhC6a, HhC6b, TgC5 y TgC8 como proteínas del tipo *tándem repeat* (TR) ya que cuando analizamos las secuencias correspondientes a las regiones inter-dominios transmembrana, vimos que estas presentaban patrones repetitivos en tándem y, además, que estos patrones repetitivos presentaban diferencias con respecto a la cantidad dentro de las mismas entre distintas cepas. Dichas observaciones nos darían la pauta no solo de que estas proteínas podrían estar asociadas a la respuesta inmune, sino que también la cantidad de motivos repetitivos en las regiones hipervariables de las proteínas FamC analizadas podrían ser

determinantes en su rol biológico actual, aunque aún desconocido. En un trabajo realizado por Moraes Barros y col., (Moraes Barros et al. 2012), demostraron que entre las familias génicas asociadas a la evasión de la respuesta inmune en *T. cruzi*, asociadas a las regiones subteloméricas, se encuentran las TS, involucradas en la invasión de células hospedadoras, cuya región C-terminal que contiene repeticiones de aminoácidos inmuno-dominantes definen el antígeno SAPA, mientras que el dominio N-terminal contiene la supuesta región para la actividad enzimática. El antígeno SAPA induce una fuerte respuesta humoral detectada poco después de la infección, tanto en humanos como en ratones. Esta respuesta está dirigida al dominio inmuno-dominante, pero es irrelevante en términos de neutralización de la actividad TS. Al comienzo de la infección en el ratón parecieran existir anticuerpos capaces de neutralizar la actividad TS, lo que llevó a la sugerencia de que la función *in vivo* de estas repeticiones es alterar la respuesta humoral al dominio de repetición hasta que se establezca la infección (Leguizamón et al. 1994). Por otra parte, en *T. brucei* el mecanismo de respuesta humoral está mediado por una única glicoproteína variante de superficie: VSG, codificada por más de 1000 genes *vsg* de expresión monoalélica, los cuales cambian entre sí, generando el mecanismo de variación antigénica (Field & Boothroyd 1995). La expresión de los genes *vsg* se produce a partir de un sitio de expresión telómero-proximal (ES). El gen *vsg* presente en un ES puede ser sobre escrito por un proceso de recombinación (Pays et al. 1983), resultando en el reemplazo de la secuencia en el ES por la secuencia de un gen *vsg* donante (Cross 1990; Pays, Vanhamme, & Berberof 1994). Alternativamente, se puede activar un nuevo ES o intercambiar secuencia entre dos ES para permitir la expresión de un nuevo gen *vsg*. Aproximadamente 1-2 Kbp corriente arriba del gen *vsg* en el ES se encuentra un número variable de repeticiones de 76 pb, que han sido mapeadas en el límite 5' de la recombinación (Campbell, Thornton, & Boothroyd 1984), aguas arriba de genes *vsg* dadores (Liu et al. 1983), y cercano o dentro del extremo 3' del marco de lectura abierto (Bernards et al. 1981; Liu et al. 1983); sugiriendo que estas secuencias repetitivas podrían actuar como un límite de movilización de secuencias para la recombinación de genes de la familia 117 VSG (Field & Boothroyd 1995).

Por lo tanto, estas observaciones concuerdan con lo mencionado anteriormente para los genes *vir* de *P. vivax*, y con nuestras observaciones de expresión de los genes *tgc*, como así también para con el valor inmunogénico observado para TgC6 y TgC8. Dado que todo lo que sea motivos repetitivos en proteínas genera un patrón de reconocimiento. Las regiones repetitivas en los genomas constituyen una fracción abundante de los genomas eucariotas ya que aportan variación genética cuantitativa en las estructuras y funciones de las proteínas, ofreciendo beneficios adaptativos para los organismos en distintos nichos ecológicos (Mendes et al. 2013). Hay un trabajo donde se encontró una buena correlación entre la presencia de regiones repetitivas y genomas de parásitos eucariotas versus los eucariotas de vida libre. Este trabajo encontró que las proteínas conteniendo dichas regiones estarían implicadas en la supervivencia de los parásitos dentro de la célula hospedadora (Gaillard et al. 1991). De hecho, varias proteínas con regiones repetitivas están involucradas en la interacción huésped-patógeno, como ser la entrada a la célula hospedadora y la adaptación para la supervivencia intracelular (Gaillard et al. 1991; Kędzierski et al. 2004). Mendes y col., (2013) demostraron que los parásitos intracelulares poseen más proteínas de superficie y/o secretadas con motivos repetitivos que aquellos que son extracelulares o de vida libre y resaltan además la presencia de secuencias señal en su estructura (Mendes et al. 2013). Por lo tanto, los estudios descriptos en esta tesis son el comienzo para dilucidar el posible rol biológico de estos genes de la FamB y FamC en *T. gondii*.

Se ha especulado que la preferencia de los genes TS, al igual que todos los genes subteloméricos de especies parasíticas, por las regiones teloméricas-subteloméricas está íntimamente conectada con la generación de variantes de proteínas para sortear el sistema inmune vía recombinación no homologa (Moraes Barros et al. 2012). Un hallazgo reciente, demostró que las regiones subteloméricas de *T. cruzi* son un punto importante de reparación por la vía de recombinación homóloga, y aumentarían enormemente la variabilidad de las familias multigénicas asociadas a estas regiones (Chiurillo et al. 2016). En este sentido resulta interesante resaltar el caso de genes como *tgc4*, *tgc6* y *tgc10* en los cuales existen eventos de duplicación génica específica dentro de cepas de *T. gondii*,

inclusive con cambios en la sintenia de los genes (ToxoDB), lo que sugiere a estos loci también, como puntos posibles de recombinación génica.

Finalmente, Primo y col., (2021) publicaron un análisis de evolución y datos de resecuenciación de ARNm de taquizoítos crecidos *in vitro* por 1500 generaciones donde se publicó una base de datos de expresión diferencial, DEA, en parásitos extra e intracelulares (Primo et al. 2021). Solo un gen ST de 12 genes fue detectado en el análisis DEA de parásitos intracelulares. Por otra parte, se ha visto que tanto miembros de las familias TgB como TgC como así también algunos genes hipotéticos internos de ST, y uno con función asignada, también interno de ST, incrementaron sus expresiones en la versión extracelular (**Tabla 11**). La mayoría de los genes hipotéticos que incrementaron su expresión en las condiciones del ensayo no presentaban evidencias de transcripción en la base de datos, con excepción del gen TGGT1\_306688 correspondiente a TGME49\_306688 (79.0/85.0; **Tabla 5**). Interesante, las expresiones de los genes FamA, FamD y FamE solo mostraron reducir sus niveles de expresión mientras que no ocurre lo mismo con FamB y FamC, a excepción de un gen famB que delimita el ST (TGGT1\_276230). En total son 15 genes STs internos de 435 genes que incrementaron sus expresiones bajo presión de selección durante su ciclo lítico *in vitro*. Los genes que incrementan su expresión en estas condiciones podrían estar indicando una adaptabilidad del parásito a una vida extracelular *in vitro*, que no esta presente a nivel fisiológico. Es así que se considera que dichas proteínas podrían tener un rol durante el mecanismo de invasión de la célula hospedadora (análisis de datos de Primo et al. 2021). Respecto de las familias multigénicas, pareciera interesante el hecho de que podrían tener funciones de adaptabilidad diferentes entre el grupo subtelomérico con respecto a las familias presentes en la región de eucromatina FamA, FamE y FamD.

DEA intra	Gen	Localización	DEA extra up	Gen	Localización	DEA extra down	Gen	Localización
TGGT1_264260	HP	ST ( <i>tgh17</i> -interno)	TGGT1_317720	eIF3	ST-FA (interno)	TGGT1_220840	TS	ST-FA (borde)
TGGT1_213620			TGGT1_301890	FamB	ST ( <i>tgb11</i> -interno)	TGGT1_237894	ST-OTU	ST-FA (borde)
TGGT1_213635			TGGT1_361230	FamB	ST (GT1-NA) 97% identidad <i>tgb11</i> , interno	TGGT1_276230	FamB	ST ( <i>tgb5</i> -borde)
TGGT1_216290B			TGGT1_411240	FamB	ST (GT1-NA) 97% identidad TGGT1_361230	TGGT1_243190	FamA	Otra
TGGT1_254840			TGGT1_200700	FamC	ST ( <i>tgc2</i> -interno)	TGGT1_278390	FamA	Otra
TGGT1_262590			TGGT1_307260	FamC	ST ( <i>tgc3</i> -interno)	TGGT1_278400	FamA	Otra
TGGT1_262600			TGGT1_411160	FamC	ST (GT1-NA) 59% identidad <i>tgc13</i> , interno	TGGT1_278410	FamA	Otra
TGGT1_264240			TGGT1_200595	HP	ST (interno)	TGGT1_410900	FamA	Otra
TGGT1_280400			TGGT1_220060	HP	ST (interno)	TGGT1_411550	FamA	Otra
TGGT1_280500			TGGT1_264240	HP	ST (interno)	TGGT1_315560	FamD	Otra
TGGT1_409050			TGGT1_264260	HP	ST (interno)	TGGT1_316670	FamD	Otra
TGGT1_411800			TGGT1_276810	HP	ST (interno)	TGGT1_408330	FamD	Otra
			TGGT1_306688	HP	ST (interno)	TGGT1_240325	FamE	Otra
			TGGT1_306692	HP	ST (interno)	TGGT1_240340A	FamE	Otra
			TGGT1_307270	HP	ST (interno)	TGGT1_240370	FamE	Otra

**Tabla 11. Análisis de expresión diferencial en genes subteloméricos.** Análisis de 988 genes expresados diferencialmente intracelulares (intra) y extracelulares (extra); *up*: genes que aumentan su expresión; *down*: genes que disminuyen su expresión. ST: subtelómero, HP: genes hipotéticos, FA: genes anotados funcionales, borde: ubicación del gen en el borde del ST, interno: gen interno subtelomérico, GT1: cepa de *Toxoplasma* utilizada en el análisis de Primo et al. 2021.

### 13. Conclusiones del capítulo II

En este capítulo caracterizamos 96 genes STs de *T. gondii*, de los cuales 57 son hipotéticos y 26 comprenden dos familias multigénicas. No se pudo establecer una función biológica para los genes hipotéticos ni para los productos de las familias multigénicas. Pero la gran mayoría mostró datos de expresión haciéndolos relevantes para futuros estudios. De hecho, algunos mostraron estar asociados a adaptabilidad a condiciones extracelulares prolongadas en experimentos *in vitro*. Respecto de las familias multigénicas FamB y FamC, se comprobó que estas presentan una expansión en *T. gondii*, sugiriendo un posible rol relevante para esta especie. Se pudo encontrar algunas características que las emparentan con otras familias multigénicas asociadas a invasión, patogenicidad y/o escape inmunológico descritas en otros parásitos del Phylum (*Babesia* y *Plasmodium*): localización en ST, ser proteínas integrales de membrana, presentar regiones repetitivas (FamC), y mostrar perfiles de reactividad humoral, aunque de baja respuesta. La expresión individual de cada gen de la FamC a diferentes situaciones de estrés, incluyendo la observada en el experimento de Evolución y resecuenciación también pareciera indicar un rol de estas a diferentes situaciones, incluyendo invasión de la célula hospedadora y virulencia. No obstante, se requiere más trabajo para dilucidar este aspecto.

## CAPÍTULO III | PROTEÍNAS INTERACTORAS DE SUBTELÓMEROS

#### 14. Proteínas asociadas a telómeros y subtelómeros

En los organismos eucariotas, el ADN genómico se envuelve alrededor de las histonas para conformar la cromatina. Como se mencionó, la eucromatina se encontraría generalmente asociada a histonas en un estado hiperacetilado y a marcas de metilaciones en histonas como H3K4me1/2/3, mientras que la heterocromatina estaría asociada a histonas en un estado de hipoacetilación y metilaciones en H3K9me3 (heterocromatina constitutiva), o a H3K27me2/3 en la heterocromatina facultativa (Beisel & Paro 2011; Wang, Jia, & Jia 2016).

En los territorios heterocromáticos subteloméricos de *T. gondii*, TgSTs, las regiones TSS de los genes activos al momento fueron asociadas a la marca H3K4me3 y el doble nucleosoma de histonas variantes H2A.Z/H2B.Z, mientras que las regiones silentes, lo harían únicamente a H2A.X y H3.3 y probablemente a H4K31me1 y H3K4me1 como se mencionó en el Capítulo I. Como se mencionó, los telómeros de *Toxoplasma* están compuestos por la secuencia repetitiva TTTAGGG (Dalmasso et al. 2014) y estarían provistos de nucleosomas, al menos enriquecidos en las histonas H2A.X y H3.3, los cuales se propagan en gran parte en los subtelómeros adyacentes (Capítulo I). A su vez, las regiones subteloméricas presentan picos de la marca H3K4me3 en la mayoría de los extremos delimitando con los telómeros.

Los telómeros de *Plasmodium* están organizados en dos estructuras: una región interna con al menos tres nucleosomas y una región terminal, desprovista de nucleosomas denominada telosoma (Figueiredo et al. 2000; Scherf, Figueiredo, & Freitas-junior 2001). El telosoma también fue muy estudiado en levaduras debido a su implicancia en la homeostasis de la heterocromatina subtelomérica (Kupiec 2014). En *Plasmodium* se describieron *in silico* varios ortólogos de proteínas telosómicas de *S. cerevisiae*, las cuales incluyen la histona deacetilasa SIR2: PfSIR2, la telomerasa: PfTERT, una proteína similar a la Miosina 1: PfMlp1; la proteína RAP1: PfRAP1; y las proteínas Rif1: PfRIF1, Taz1: PfTaz1 (Scherf, Figueiredo, & Freitas-junior 2001). Sin embargo, solo se comprobó experimentalmente la asociación de algunas de las proteínas mencionadas: la proteína PfTERT (Figueiredo et al. 2005), de la

subunidad 1 del complejo de reconocimiento de origen de replicación, ORC1: PforC1 (Mancio-Silva et al. 2008), la cual complementa la actividad de SIR3 de levaduras (Varunan et al. 2013) y de PfsIR2A en el silenciamiento génico (Tonkin et al. 2009). PforC1 no solo cumpliría un rol en la replicación del ADN uniéndose a este y determinando los sitios de pre-inicio de la replicación (Gupta, Mehra, & Dhar 2008; Gupta et al. 2009), sino que también cooperaría con PfsIR2 en la coordinación del desplazamiento de la heterocromatina sobre el subtelómero y la regulación de los genes *var* (Deshmukh et al. 2012). Además, PforC1 participaría en la formación de la estructura del T-loop del telómero, lo cual evitaría la fusión entre los extremos de los cromosomas (Mancio-Silva et al. 2008) y uniría directamente el ADN telomérico en su extremo N-terminal en la periferia nuclear (Deshmukh et al. 2012).

En mamíferos, el complejo telosómico se denomina en inglés *Shelterins* y posee seis proteínas: TRF1 y TRF2, POT1, TIN2, el ortólogo de levaduras RAP1 y TPP1 (de Lange 2005). Este conjunto de proteínas específicas previene que los extremos de ADN cromosómico sean reconocidos por el mecanismo de reparación del daño al ADN (Denchi 2009); y participan en muchas funciones teloméricas como la formación del T-loop, la replicación, la regulación de la longitud del ADN telomérico, la formación de la estructura heterocromática y el anclaje de los telómeros a la periferia nuclear (Palm & de Lange 2008).

En células de mamífero, la formación de la heterocromatina se inicia en los centros de nucleación, regiones pericentroméricas asociadas a elementos repetitivos satelitales (Bulut-Karslioglu et al. 2012) mediante proteínas de unión a ADN específicas o también ARN no codificantes largos (Volpe et al. 2002; Bierhoff et al. 2014), los cuales reclutan histonas deacetilasas, HDACs, y HMTs de la familia Suv39 y Suv4-20. Esta nucleación da como resultado una hipocetilación de las histonas e hipermetilación de H3K9 y H4K20 en estos sitios (Wang, Jia, & Jia 2016), en donde Suv39h cataliza la tri-metilación de H3K9 (H3K9me3) y esta a su vez facilita la unión de proteínas HP1 que interactúan con Suv4-20h y establecen la marca sobre H4K20me3 (Schotta et al. 2004; García-Cao et al. 2004; Benetti et al. 2007). Estas acciones combinadas conducen a la propagación de la heterocromatina a lo largo de un gran

dominio de heterocromatina (Hall et al. 2002; Talbert and Henikoff 2006; Zhang et al. 2008; Al-Sady, Madhani, & Narlikar 2013; Wang et al. 2014).

Durante la replicación del ADN, las histonas parentales que contienen modificaciones, se incorporan al azar en ambas cadenas hijas detrás de la horquilla de replicación (Probst, Dunleavy, & Almouzni 2009). La marca H3K9me recluta nuevamente HMTs de H3K9 y restaura los dominios de heterocromatina en ambas cadenas hijas, lo que conlleva al mantenimiento estable de este estado epigenético a través de las generaciones (Hall et al. 2002; Reddy et al. 2011; Rangunathan, Jih, & Moazed 2015; Audergon et al. 2015).

En la levadura *S. cerevisiae* el mecanismo operante de ensamblaje de la heterocromatina sería ejercido principalmente a través de la deacetilación de las histonas asociadas al ADN subtelomérico utilizando el complejo SIR2, SIR3 y SIR4 (Rusche, Kirchmaier, & Rine 2003; Kueng, Oppikofer, & Gasser 2013).

Las proteínas SIR2 se estudiaron en *Plasmodium*, PfSIR2A y PfSIR2B, y se sabe que su actividad está íntimamente ligada a la represión de genes en las regiones PFTAS (Tonkin et al. 2009). Tonkin y col., demostraron, además, que PfSIR2A cumpliría un rol fundamental en la protección del extremo telomérico (Tonkin et al. 2009). Las regiones TAS de *Plasmodium* están compuestas de seis TAREs diferentes, dispuestos consecutivamente hacia el centrómero. La región codificante está localizada directamente después del TARE6 y posee genes de la familia *var*, *rif*, *stevor* y *pfmc-2tm* (Freitas-Junior et al. 2000; Kaviratne et al. 2002; Petter et al. 2007) responsables de la variación antigénica y la citoadherencia (Figuereido et al, 2000; Scherf, Figueiredo, & Freitas-junior 2001; Hernández-Rivas et al. 2010).

Por otra parte, también se encontró que las regiones PFTAS están asociadas a la proteína de heterocromatina HP1, PfHP1 (Pérez-Toledo et al. 2009). PfHP1 y la marca H3K9me3 las cuales fueron encontradas asociadas a los genes *var* en las regiones PFTAS en la periferia nuclear (Lopez-Rubio, Mancio-Silva, & Scherf 2009). En este trabajo también encontraron que la falta de SIR2 causaría cambios

en la localización de H3K9me3 y generaría disrupción en la transcripción monoalélica de los genes *var*, sugiriendo la existencia de centros perinucleares represivos de la transcripción asociados al control de la expresión de genes involucrados en la variación fenotípica de *Plasmodium* (Lopez-Rubio, Mancio-Silva, & Scherf 2009). También se ha demostrado la asociación de PfORC1 en unión a TARE3 (Mancio-Silva et al. 2008).

Por otra parte, se demostró la asociación de proteínas de unión al ADN como, PfAlba3 a la región TARE6 de PFTAS, demostrando su implicancia en la unión a ADN como así también a moléculas de ARN simple cadena (Chêne et al. 2012). La proteína PfAlba3, también se encontró asociada a PfSIR2A en la periferia nuclear, donde PfSIR2A parecería interactuar con PfAlba3 deacetilando su extremo N-terminal y favoreciendo su unión al ADN (Goyal et al. 2012). Esta interacción no sería exclusiva de los Apicomplexa ya que ha sido descrita previamente en Arqueas, en donde SIR2 regularía el silenciamiento a través de la deacetilación de la principal proteína de unión al ADN, Alba (Bell 2002). Por otra parte, Gupta y col., (2018) demostraron que la proteína CAF-1, una chaperona que carga a las histonas H3 y H4 en el ensamblaje del nucleosoma después de la respuesta al daño del ADN, DDR, co-localiza con PfHP1 en la periferia nuclear y también con PfAlba3 entre otras (Gupta et al. 2018).

En Arqueas, las proteínas Alba son una de las principales componentes de la cromatina y median la represión transcripcional a través de la interacción con ácidos nucleicos. Esta interacción requiere la acetilación de lisinas para la dimerización de las proteínas (Bell 2002). Por lo tanto, este constituiría un mecanismo ancestral conservado de regulación génica.

Gissot y col., (2013) estudiaron la presencia de estas proteínas en *Toxoplasma*, TgAlba1 y TgAlba2 (ortóloga a PfAlba3) y encontraron que estas tendrían una localización mayormente citoplasmática en taquizoítos intracelulares (Gissot et al. 2013), mientras que su posicionamiento sería perinuclear en taquizoítos extracelulares, co-localizando con gránulos de ARN formados en respuesta a estrés (Lirussi & Matrajt 2011), coincidiendo, además, con la localización de TgSIR2A (Nardelli, comunicación personal) y de PfSIR2A (Mancio-Silva et al. 2013).

En conclusión, las regiones teloméricas y subteloméricas poseen una cromatina única con el reclutamiento de numerosas proteínas, alguna de ellas con capacidad de modulación de la cromatina. Claramente para poder mantener un estado de heterocromatina y silenciamiento en el telómero y subtelómero es importante la participación de deacetilasas y metiltransferasas.

### 15. Control epigenético de la cromatina subtelomérica

El éxito de la infección por *Toxoplasma* radica en la capacidad que tiene el taquizoíto para infectar rápidamente cualquier célula nucleada durante la etapa asexual de vida (Tenter, Heckeroth, & Weiss 2000) y comenzar a dividirse en sucesivas rondas de replicación dentro de la VP, en el citoplasma de la misma (Striepen et al. 2007). Pero para llevar a cabo una replicación exitosa, los taquizoítos deben vincular el censado ambiental con un complejo mecanismo de regulación de la expresión génica (Cleary et al. 2002; Radke et al. 2005; Behnke et al. 2010). Actualmente se sabe que las PTMs sobre las histonas generan un código epigenético asociado a la respuesta ambiental permitiendo cambios conformacionales rápidos en la estructura de la cromatina. Estas PTMs son esenciales para controlar la expresión génica, la replicación y reparación del ADN (Jezek et al. 2019).

La acetilación de histonas es una de las PTMs más importantes que influyen en la expresión génica (Basu et al. 2009; Choudhary et al. 2009; Feng & Jacobsen 2011). Por lo tanto, las enzimas acetiltransferasas y deacetilasas de histonas han sido ampliamente estudiadas como posibles blancos farmacológicos en diferentes especies parasitarias de humanos (Vergnes et al. 2005; Vanagas, Jeffers, & Bogado 2012; Ferreira et al. 2014; Campo 2017; Hailu et al. 2017).

Por otra parte, hay varios estudios que han demostrado que las regiones subteloméricas y sus proteínas vinculadas regularían la expresión génica de los genes asociados (Figuereido et al. 2002; Hernandez-Rivas et al. 2010). Uno de los principales mecanismos que regulan la expresión génica en estas regiones sería la deacetilación de los residuos de lisina (K) en el extremo N-terminal de las histonas. En los telómeros y subtelómeros de hongos y protozoos como *Trypanosoma* y *Plasmodium* este mecanismo sería orquestado por la actividad de la histona deacetilasa de tipo III, SIR2; afectando desde

la reparación del ADN, la replicación, la inactivación de la transcripción de genes asociados y el silenciamiento general de la cromatina de dicha región cromosómica (Freitas-Junior et al. 2005; Xu, Parmigiani, & Marks 2007; Alsford et al. 2007; Mancio-Silva et al. 2008; Tonkin et al. 2009; Hoggard et al. 2018; Price et al. 2019).

Las proteínas Sirtuinas (SIR2 y sus proteínas relacionadas dependientes de NAD<sup>+</sup>) deacetilan histonas actuando sobre la lisina 9 de la histona (H)3 (Shankaranarayana et al. 2003), la lisina 4 de H3 (H3K4) (Guillemette et al. 2011), la lisina 18 (H3K18) (Mishra et al. 2022), la lisina 56 de H3 (H3K56) (Maas et al. 2006; Vempati 2012; Mishra et al. 2022) y la lisina 16 de H4 (H4K16) (Vaquero et al. 2004; Vaquero, Sternglanz, & Reinberg 2007; Mishra et al. 2022). Curiosamente, las Sirtuinas también se dirigen a muchas proteínas que no son histonas. Entre los blancos no relacionados con la cromatina no histonas más representativos se encuentran los factores de transcripción específicos implicados en la supervivencia al estrés, la inhibición de la apoptosis, la proliferación celular, el metabolismo y muchos otros (Bosch-Presegué & Vaquero 2015).

En *T. gondii* se han descrito dos deacetilasas SIR2 del tipo III: TGME49\_227020 y TGME49\_267360, ortólogas de SIR2A y SIR2B en *P. falciparum* respectivamente (Vanagas, Jeffers, & Bogado 2012). En *P. falciparum*, las proteínas SIR2 tienen un papel importante en la regulación de los genes *var* y otros genes asociados a PFTAS ligados al mecanismo de virulencia de variación antigénica, tal como ya se mencionó (Tonkin et al. 2009).

La actividad mediada por SIR2 ha sido ampliamente estudiada en especies de levaduras como parte del mecanismo de silenciamiento TPE junto a proteínas como RAP1 y KU. La base molecular del TPE estaría mediada por estas proteínas, en donde RAP1 y KU se unirían el ADN telomérico, y reclutarían a SIR2, la cual se dispersaría desde el telómero hacia el interior de la heterocromatina reprimiendo genes STs a través de su interacción con las colas de las histonas. Según este modelo, SIR2 removería los grupos acetilo de las K16 de H4 (H4K16ac), lo cual permitiría unirse a SIR3 y SIR4 a las colas de histonas no acetiladas (Grunstein et al. 2013). La deacetilación de H4K16 por parte de

SIR2 estimularía entonces la propagación del complejo RAP1/KU/SIR, mientras que la proteína Sas2, antagonizaría este proceso acetilando la K16 de H4. Este mecanismo opuesto entre SIR2 y Sas2 generaría un gradiente de acetilación sobre H4K16 que a su vez marcaría el límite entre la cromatina activa e inactiva en los dominios subteloméricos (Suka, Luo, and Grunstein 2002).

Estructuralmente, un modelo más actual en levaduras también propone que el complejo RAP1/KU/SIR2, entre otras proteínas, formarían una superestructura. En dicho análisis se demostró que las proteínas SIR (SIR2, SIR3 y SIR4) se concentrarían en las regiones silenciadoras, pero no en las regiones promotoras de los genes subteloméricos reprimidos (*hml* y *hmr*) formando estructuras tipo bucle, creando de esta manera, un límite dentro del cual se dispersaría el efecto de represión de la cromatina (Thurtle & Rine 2013).

Por otra parte, hay trabajos donde se ha informado que la acetilación-deacetilación en H4K16 es un punto de control importante durante la fase S. Esta marca ha demostrado tener un pico en la fase S que desaparece después de la replicación del ADN, promoviendo la deposición de mono-metilación en K20 de H4 (Serrano et al. 2012). La acetilación sobre H4K16 también se ha asociado a la acetilación en variantes de histonas. Estudios en levaduras han demostrado que la acetilación sobre H4K16 es un pre-requisito para la incorporación dependiente de NAD<sup>+</sup> de Htz1 (Zhang, Roberts, & Cairns 2005; Shia, Li, & Workman 2006), la contraparte de H2A.Z de mamíferos, involucrada como se mencionó, en los sitios activos de la transcripción. Por otra parte, estudios en células HeLa demostraron que H2A.Z regularía epigenéticamente la activación de los orígenes de replicación tempranos manteniendo el *timing* de replicación a través del eje H4K20me<sub>2</sub>-Suv420h1-ORC1 (Long et al. 2019).

En *Plasmodium* también se ha observado que SIR2 actuaría como límite para la deposición de H2A.Z en la formación de la cromatina (Petter et al. 2011). Como así también se encontró que PfSIR2A deacetilaría H3 y removería H2A.Z de los sitios activos de transcripción de los genes *var*. Mientras que en parásitos deficientes de PfSIR2A ( $\Delta$ SIRA) H2A.Z sería relocalizada junto con H3ac en los sitios activos de la transcripción, no así, en parásitos deficientes de PfSIR2B ( $\Delta$ SIR2B) (Petter et al. 2011).

Notablemente, en *T. gondii* se encontró que la proteína acetiltransferasa MYST-B (ortóloga a Sas2 y TIP60 de mamíferos) acetilaría a H4 en K16 facilitando el mecanismo de respuesta a daño de ADN mediado por la ATM quinasa (Vonlaufen et al. 2010).

Por lo tanto, considerando el rol que tendrían las deacetilasas tipo III en la estructuración de la heterocromatina subtelomérica en otros modelos de estudio, generando tal vez una respuesta ambiental rápida a través de la regulación de la expresión de genes específicos y por ello activando o no la replicación celular es, que consideramos a las proteínas SIR2 de *T. gondii* como muy buenos candidatos para su estudio y asociación las regiones subteloméricas. A su vez, estas enzimas son blancos de drogas que se están experimentando actualmente, lo cual permite realizar experimentos de tratamiento con drogas que afecten su función.

## 16. Drogas anti- deacetil transferasas

En *Toxoplasma* se han ensayado varios inhibidores de histonas deacetilasas de clase I y II y se ha encontrado que tienen actividad anti-proliferativa contra *T. gondii* (Hailu et al. 2017; Strobl et al. 2007), mientras que el inhibidor de deacetilasas de clase III, nicotinamida, no ha mostrado ningún efecto sobre la capacidad proliferativa de *T. gondii* (Strobl et al. 2007).

El Sirtinol, SIRT, es un inhibidor de Sirtuinas, altamente específico contra las Sirtuinas de humanos, hSIR-1 y hSIR-2 con un IC<sub>50</sub> de 131 μM y 38 μM respectivamente (Grozinger et al. 2001). El SIRT presentó actividad antitumoral y quimio-sensibilidad mejorada a los agentes que dañan el ADN colapsado en la horquilla de replicación (Ota et al. 2006; Kojima et al. 2008). Se encontró, además, que inhibe el crecimiento axigénico de amastigotes de *L. infantum* (etapa replicativa del parásito intracelular), independientemente de la concentración utilizada (Vergnes et al. 2005). Por otra parte, el SIRT se probó en las etapas de vida de *T. cruzi* donde se vió que causó efectos específicos del estadio al aumentar la infectividad de los tripomastigotes y bloquear la metacicloogénesis, pero sin afectar el crecimiento del epimastigote, ni la etapa replicativa del amastigote (Campo 2017).

Resveratrol, un polifenol derivado de las uvas, a diferencia de SIRT, resultó ser un componente activador de la actividad de SIR2 (Kaeberlein et al. 2005). El Resveratrol es una fitoalexina polifenólica producida en plantas que puede alterar las actividades de las HDAC I y II, de las metiltransferasas de ADN (DNMT) y activar la función de SIR2, Sirtuina perteneciente a HDAC III (ver Angel et al. 2020). RSV actúa como antioxidante en dosis bajas, pero es pro-oxidante y genera daño en el ADN a dosis elevadas (Shaito et al, 2020). El tratamiento de líneas celulares humanas con RSV indujo la expresión de Rad9, un actor clave en la respuesta DDR, como así también indujo la fosforilación de H2A.X en el motivo SQE generando  $\gamma$ H2A.X, un marcador de rotura de doble hebra, DSB (Chen et al. 2018).

La administración de RSV durante el estadio de infección aguda causada por *T. gondii* en ratones conferiría protección al huésped (Bereswill et al. 2010), regularía la inflamación de los tejidos y reduciría la replicación del parásito (Bottari et al. 2015).

Chen y col., (2018) observaron que la incubación de taquizoítos extracelulares con RSV durante 24 hs afectaba su viabilidad. También observaron que el RSV reducía el crecimiento intracelular de taquizoítos y promovía la autofagia en macrófagos infectados (Chen et al. 2018). Por su parte, Adeyemi y col., (2019) identificaron a RSV como un posible candidato a fármaco contra la toxoplasmosis (Adeyemi et al. 2019). Sin embargo, ninguno de estos estudios ha investigado el (los) mecanismo(s) por el cual RSV afectaría a los taquizoítos de *T. gondii*.

Consecuentemente, resulta más que relevante poder elucidar posibles interactores de las regiones STs y establecer cuáles son las PTMs que rigen el mecanismo epigenético subyacente a la cromatina de los subtelómeros de *Toxoplasma*, para finalmente dar con putativos blancos terapéuticos.

## 17. Hipótesis del capítulo III

Debido a lo expuesto, nuestra hipótesis es que también existe una relación de las proteínas SIR2 con las regiones subteloméricas de *T. gondii* y que estas estarían modulando la cromatina de esa región hacia un estado silente, alterando varias funciones biológicas de al menos el taquizoíto. A su vez, que esta enzima podría ser un excelente blanco terapéutico. En dicho proceso, el blanco principal sería la marca de histona H4K16ac, lo cual impactaría en el estado de la cromatina, afectando las marcas de histonas y la activación de genes en los subtelómeros, la regulación de la replicación del ADN, como así también los procesos de daño del ADN.

De esta manera nos propusimos dentro de los objetivos comenzar una caracterización de TgSIR2A y TgSIR2B. El uso de drogas que tienen como blanco SIR2 podría ser de valor para estudiar efectos sobre las histonas, al mismo tiempo se analiza el valor terapéutico de las mismas. Los estudios de funcionalidad génica, permitirían confirmar que las drogas estudiadas tendrían como blanco a TgSIR2A y/o TgSIR2B, como también determinar la localización genómica de las mismas y determinar si alguno o ambas enzimas regulan el estado de la cromatina subtelomérica y la biología de *T. gondii*.

## 18. Objetivo del capítulo III

Indagar acerca de las proteínas interactoras de las regiones teloméricas-subteloméricas y su posible mecanismo de acción

### 18.1 Objetivos específicos del capítulo III

- ✓ Buscar *in silico* proteínas ortólogas descritas en otras especies
- ✓ Evaluar efectos de drogas con acción sobre sirtuinas, Resveratrol y Sirtinol, en la replicación de *T. gondii*
- ✓ Evaluar los efectos de RSV y SIRT sobre los niveles de acetilación de H3 y H4K16
- ✓ Evaluar el efecto de RSV y SIRT sobre el ciclo celular de *T. gondii*.
- ✓ Evaluar el efecto de RSV sobre el daño al ADN midiendo los niveles de  $\gamma$ H2A.X.
- ✓ Analizar la localización genómica y subcelular de TgSIR2A y TgSIR2B junto a su rol biológico

## 19. Resultados del capítulo III

### 19.1 Búsqueda *in silico* de proteínas ortólogas descritas en otras especies

Para tener una aproximación más abarcativa de las proteínas que participan en conformación de las regiones teloméricas/STs de *T. gondii* se realizó una búsqueda *in silico* de proteínas que integrarían la estructura de la heterocromatina telomérica-subtelomérica en la especie parasítica Apicomplexa relacionada, *Plasmodium* sp; en las levaduras, tanto en *S. cerevisiae* como en *S. pombe*; como así también las que componen el complejo Shelterinas en humanos. En la **Tabla 12**, se muestran todas las putativas proteínas ortólogas encontradas tanto para *Toxoplasma* como para *Plasmodium* en relación a sus contrapartes en levaduras y humanos donde han sido ampliamente estudiadas (Ottaviani, Gilson, & Magdinier 2008; Vanagas, Contreras, & Angel 2019).

En el caso de *Toxoplasma* pudimos encontrar una proteína con un porcentaje de 31% de similitud a PfORC1/SIR3, la ATPAsa-AAA (TGME49\_283900), además de las descritas previamente para SIR2 (Vanagas et al, 2012) y TgChromo1 (Gissot et al. 2012).

Otra proteína encontrada se corresponde a una proteína hipotética (TGME49\_248160) de 1102 aa correspondiente a la subunidad KU70 del heterodímero KU70/KU80, denominada aquí TgKU70. Este heterodímero se encontró asociado a múltiples funciones como ser, la unión al ADN doble y simple cadena, la recombinación y reparación del ADN y la asociación a los orígenes de replicación (Tuteja & Tuteja 2000). En *T. gondii* la subunidad KU80, TgKU80 participaría en la vía de reparación del daño al ADN por recombinación no homóloga NHEJ, ya que los parásitos *knockouts* de esta proteína, favorecerían el proceso reparación por recombinación homóloga, HRR, pero presentarían una mayor sensibilidad a las roturas de doble hebra en el ADN inducidas por el tratamiento con Fleomicina o irradiación gamma (Fox et al. 2009). El análisis de la secuencia proteica de TgKU70 en búsqueda de dominios conservados en la base de datos NCBI mostró que además del dominio KU característico de unión al ADN (Posición: 713-955, *E-value*: 1.09e-36), TgKU70 posee un dominio ATP ADN helicasa II entre los aminoácidos 342 y 1090 (*E-value*: 6.47e-36). Mientras que el mismo análisis en la proteína

TgKU80 de 1107 aa, reveló la presencia de 3 dominios conservados, uno N-terminal correspondiente al dominio VWA de 162 aa (Posición: 84-246, *E-value*: 5.13e-05) encontrado principalmente en moléculas adhesión como la proteína MIC2 de *T. gondii*; el dominio KU característico (Posición: 445-624, *E-value*: 6.71e-22) y un dominio C-terminal de asociación a DNA-PKs durante la reparación por la vía NHEJ (Posición: 929-1070, *E-value*: 5.88e-21).

Por otra parte, no se encontraron proteínas de unión al telómero RAP1, Rif1, Rif2 y Taz1 de levaduras; ni tampoco a las del complejo Shelterinas de mamíferos: TRF1, TRF2, TIN2; ni del complejo de proteínas de unión de simple cadena del telómero: Stn1, Ten1, Cdc13, TPP1, POT1, Pif1 (Vanagas, Contreras, & Angel 2019) (*Tabla 12*). Cabe destacar que solo se tuvieron en cuenta para el análisis, las

Levaduras	Mamíferos	<i>T. gondii</i> (ME49)	<i>P. falciparum</i> (PF3D7)
Sir2 (P06700)	Sir2 (Q8IXJ6)	Sir2A (_227020)	Sir2A (1328800)
		Sir2B (_267360)	Sir2B (1451400)
Sir3 (P06701)	Sir3 (Q9NTG7)	ATPase, AAA (_283900)	Orc1/Sir3 (_1203000)
Sir4 (P11978)	Sir4 (Q9Y6E7)	NF	NF
RAP1 (P11938)	RAP1 (Q9NYB0)	NF	NF
RIF1 (P29539)	RIF1 (Q5UIP0)	NF	NF
RIF2 (Q06208)	NF	NF	NF
Ku70 (P32807)	XRCC6 (B1AHC9)	Ku70 (_248160)	NF
Ku80 (Q04437)	XRCC5 (P13010)	Ku80 (_312510)	NF
Taz1 (P79005)	TRF1 (P54274)	NF	NF
NF	TRF2 (Q15554)	NF	NF
NF	TIN2 (Q9BSI4)	NF	NF
NF	HP1 (Q13185)	Chromo1 (_268280)	HP1 (_1220900)
Stn1 (P38960)	NF	NF	NF
Ten1 (Q07921)	NF	NF	NF
Cdc13 (P32797)	NF	NF	NF
NF	TPP1 (Q96AP0)	NF	NF
NF	POT1 (Q9NUX5)	NF	NF
Pif1 (P07271)	Pif1 (Q9H611)	NF	NF

**Tabla 12. Proteínas de Apicomplexas ortólogas a proteínas de levaduras y humanos asociadas a telómeros STs.** NF: no encontrada (del inglés: *Not Found*). PfOrc1 complementa la actividad de Sir3 de levaduras (Petter et al. 2007). La proteína ortóloga en *T. gondii* fue hallada en la base de datos ToxoDB, utilizando la secuencia de PfOrc1.

proteínas que presentaban un porcentaje de identidad en sus secuencias de al menos un 25-30% (Vanagas, Contreras, & Angel 2019). En este sentido, la secuencia aminoacídica correspondiente a la proteína putativa TgORC1 (TGME49\_283900) de 1264 aa (ToxoDB) se analizó en búsqueda de homología de dominios en la base de datos del NCBI, encontrando un dominio correspondiente a la superfamilia ORC1 cl36513: PTZ00112 de 334 aa, entre los aminoácidos 736 y 1070 (*E value*: 6.45e-82).

En base a estos hallazgos se proponen a las proteínas TgSIR2s, TgORC1, TgChromo1 y el heterodímero TgKU70/KU80 como posibles actores en la estructuración y dinámica de la heterocromatina telomérica-subtelomérica de *T. gondii* según lo observado en otras especies.

## **19.2 Efecto del activador e inhibidor de Sirtuinas en la replicación de *T. gondii***

### **19.2.1 Efecto del RSV sobre la replicación**

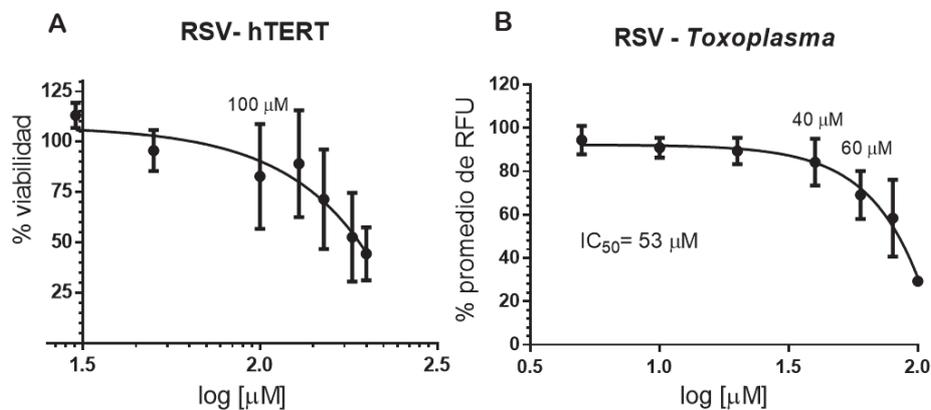
El RSV es una droga multifunción (Harikumar & Aggarwal 2008), entre las cuales se ha detectado que es un potente activador de Sirtuinas (Kaeberlein et al. 2005). Para probar el efecto de RSV en nuestro modelo de estudio primero testeamos la citotoxicidad sobre las células hospedadoras. Para esto se incubaron células hTERT no infectadas durante 24 hs con diferentes dosis de la droga (**Fig.49A**). Casi el 80% de las células hTERT no infectadas permanecieron viables a una concentración de RSV de 100  $\mu$ M (**Fig.49A**). Luego, para determinar el impacto del RSV en el ciclo lítico de *T. gondii*, se cultivaron intracelularmente taquizoítos transgénicos RH\_RFP durante 24 hs en presencia de RSV. El RSV mostró un efecto dosis-dependiente sobre la replicación intracelular de *T. gondii* (**Fig.49AB**), produciendo un valor de IC<sub>50</sub> de  $53 \pm 4$   $\mu$ M. Dicha concentración es la que se determinó como necesaria para reducir *in vitro* el crecimiento poblacional en un 50% (Contreras et al. 2021).

### **19.2.1 Efecto de SIRT en la replicación**

Con el fin de definir la dosis máxima de este fármaco en la que las células fibroblásticas hTERT permanecían viables en nuestros experimentos con *T. gondii* y SIRT, se realizó la prueba de toxicidad. La incubación de 24 hs mostró que la concentración máxima de trabajo era de 100  $\mu$ M para

experimentos con taquizoítos intracelulares (**Fig.50A**). También se analizó el efecto del inhibidor de Sirtuinas, SIRT, sobre la replicación de los taquizoítos. Cuando los taquizoítos RH\_RFP intracelulares se trataron durante 24 hs, se observó un efecto dosis-dependiente sobre la replicación de los parásitos con un  $IC_{50}$  de  $27 \pm 2,5 \mu M$  (**Fig.50B**). Este resultado indicaría que SIRT tendría un efecto sobre la replicación de los taquizoítos.

**Fig.49. Efectos del RSV sobre la replicación de *T. gondii*.** **A.** Viabilidad de las células hospedadoras. La citotoxicidad de RSV se estimó mediante la reducción de MTT en presencia de la droga que se presenta en relación con la viabilidad de los controles no tratados (definida como viabilidad del 100%). Los resultados presentados son representativos de tres experimentos independientes. **B.** Curva de crecimiento dependiente de la dosis (0-100  $\mu M$ ) relativa a la fluorescencia de taquizoítos RH\_RFP, RFU, después de 24 hs de tratamiento. Tres experimentos independientes fueron realizados por triplicado cada vez y los datos se presentan como media  $\pm$  desviación estándar (DE).



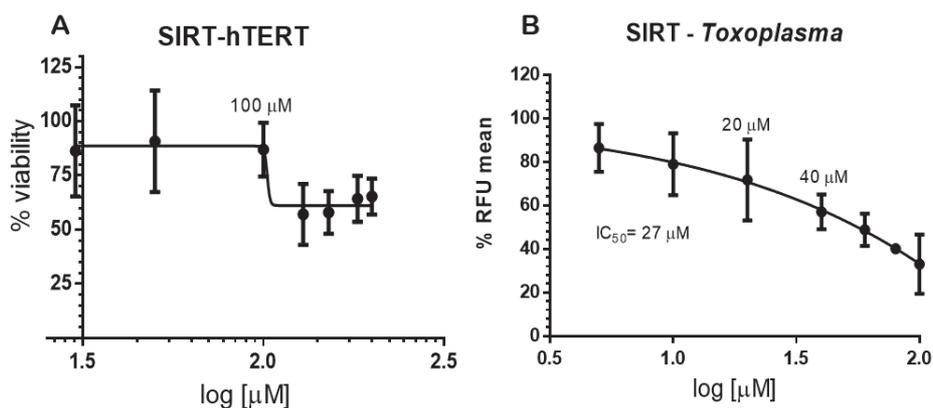
### 19.3 Efectos de activador e inhibidor de Sirtuinas sobre PTMs de *T. gondii*

#### 19.3.1 Efecto de RSV sobre las acetilaciones en H3 y H4

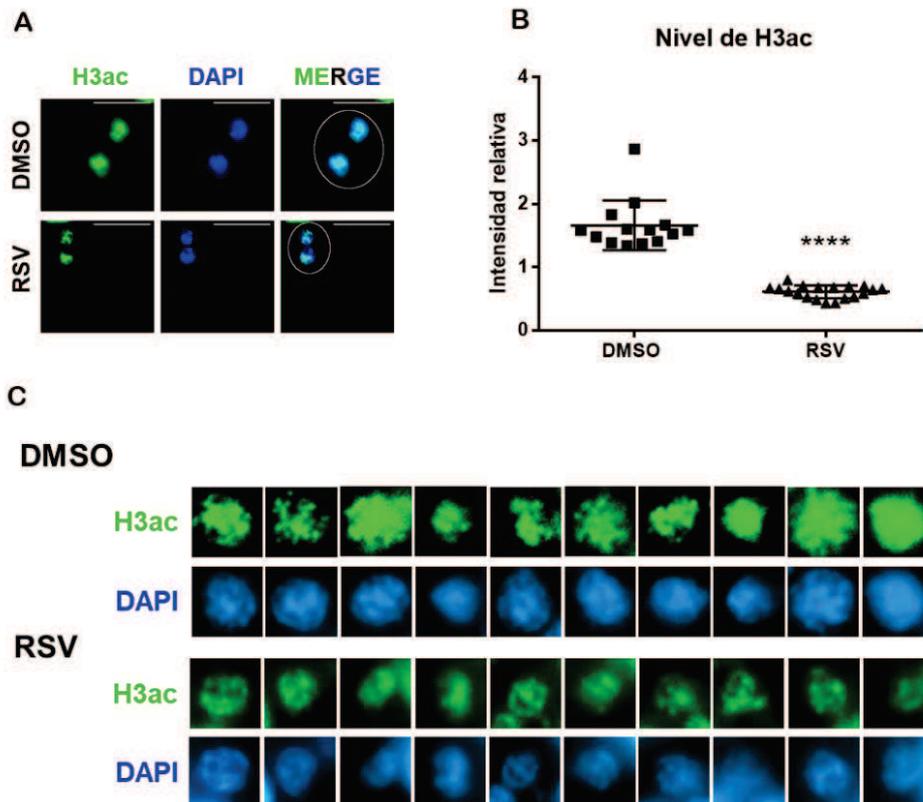
Dado que las Sirtuinas son uno de los múltiples objetivos de RSV, analizamos los niveles de acetilación de H3 y H4K16 en los núcleos de taquizoítos intracelulares tras el tratamiento con la droga mediante IFI. Los taquizoítos tratados con RSV presentaron una disminución significativa en el marcaje

de H3ac (**Fig.51**) ( $p \leq 0.05$ ) y de H4K16ac (**Fig.52**) ( $p \leq 0.0001$ ) en comparación con los taquizoítos tratados con DMSO. Los taquizoítos sin tratar y a los que se les aplicó DMSO (control, vehículo de la droga) no difirieron en las intensidades de sus fluorescencias (**Fig.51** y **Fig.52**). Por otra parte, lo observado por IFI se corroboró por intensidad de bandas proteicas mediante WB. En la **Fig.53** se puede ver la comparación del tratamiento con el control DMSO. En este caso se pudo ver que el tratamiento con RSV redujo el nivel de las intensidades H3ac y H4K16ac a dosis mínimas de 70  $\mu\text{M}$  ( $p \leq 0,0001$ ) y 50  $\mu\text{M}$  ( $p \leq 0,05$ ), respectivamente. Estos resultados concuerdan con un estudio anterior en el que se descubrió que el tratamiento con RSV disminuía el nivel de acetilación sobre H3 y H4 en *T. cruzi* (Campo 2017).

**Fig.50. Efectos del SIRT sobre la replicación de *T. gondii*.** **A.** Viabilidad de las células post 24 hs en ausencia o presencia de SIRT. Los resultados presentados son representativos de tres experimentos independientes. **B.** Curva de crecimiento dependiente de la dosis (0-100  $\mu\text{M}$ ) RFU de taquizoítos RH-RFP después de 24 hs de tratamiento. Tres experimentos independientes fueron realizados por triplicado cada vez, y los datos se presentan como media  $\pm$  desviación estándar (DE).



**Fig.51. Efecto de RSV sobre H3ac en taquizoítos intracelulares.** **A.** IFI con anticuerpo anti-H3ac (1: 200). Se eligieron VPs con dos parásitos (círculos blancos) para comparar el tratamiento y el vehículo, DMSO. Las líneas blancas de la parte superior derecha corresponden a la barra de escala: 10  $\mu$ m. Los núcleos se tiñeron con DAPI. Los taquizoítos intracelulares se trataron con RSV (100  $\mu$ M) o DMSO al 0,5% durante 24 hs. **B.** Las intensidades de fluorescencia de H3ac y de DAPI se cuantificaron a partir de 10 núcleos (**C**) y se representaron gráficamente como intensidades relativas. El panel y el gráfico son representativos de tres experimentos independientes con resultados similares. El análisis estadístico se realizó con ANOVA unidireccional y prueba de comparación múltiple de Tukey (\*  $p \leq 0.05$  y \*\*\*\*  $p \leq 0.0001$ ).

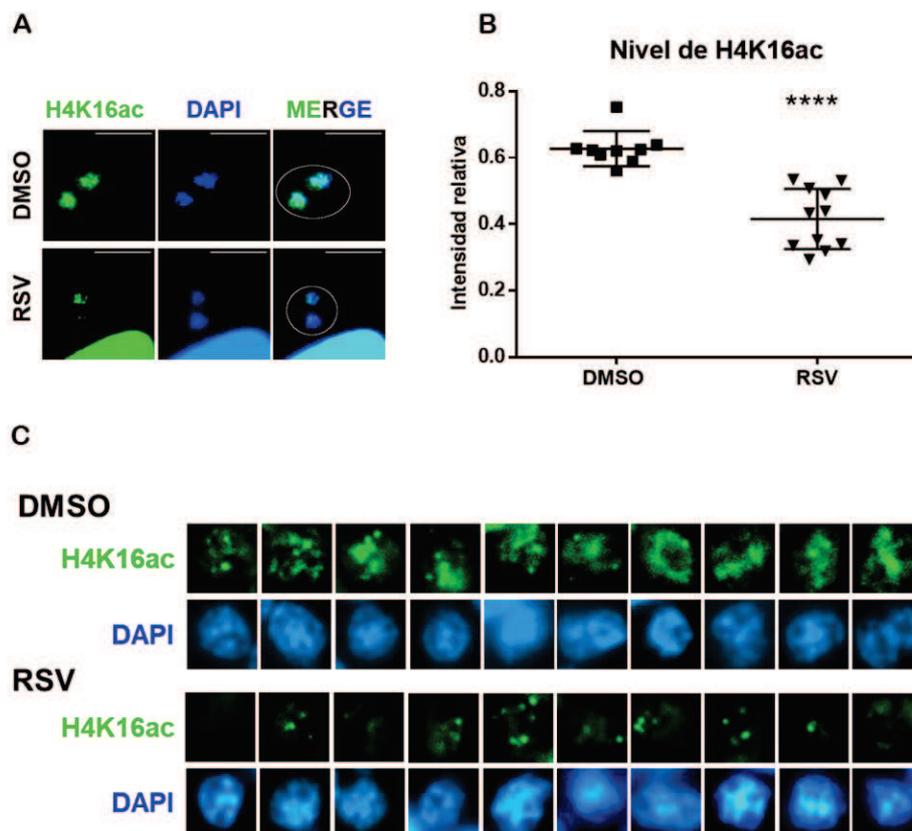


### 19.3.2 Efecto de SIRT sobre las acetilaciones en H3 y H4

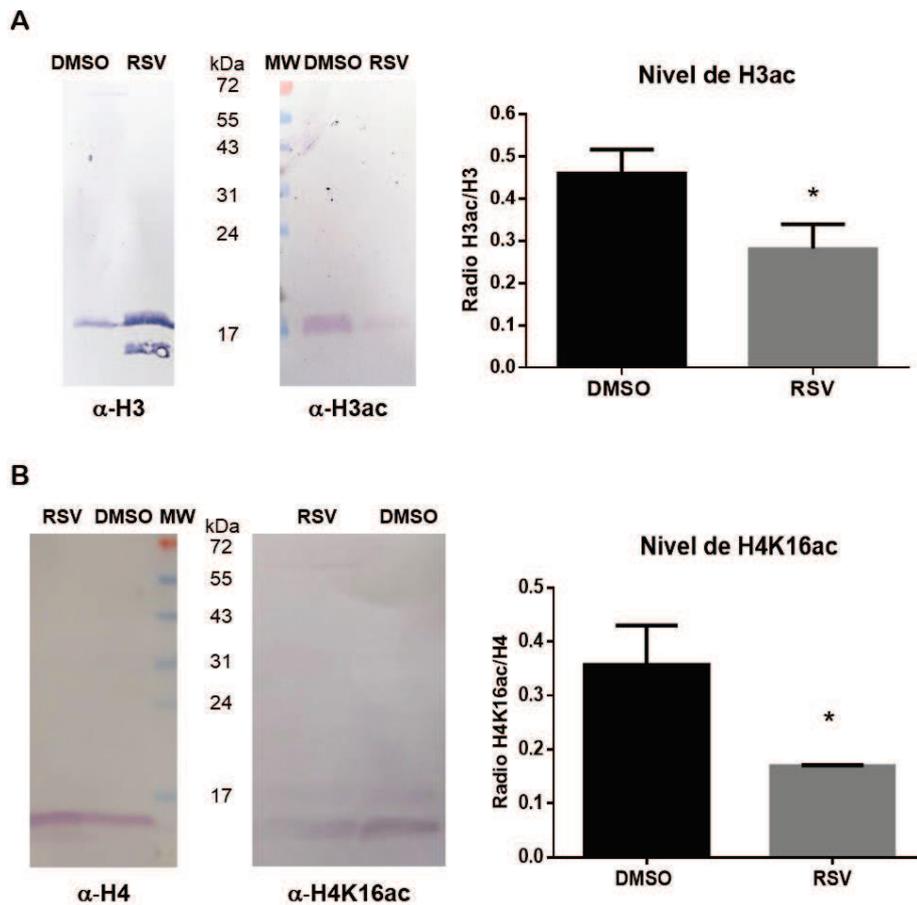
El SIRT es un fármaco descrito como inhibidor de la actividad de las histonas deacetilasas de tipo III como SIR2. En este sentido, se espera que el tratamiento con SIRT provoque un aumento de la acetilación en las histonas H3 y H4 al afectar a SIR2. Los taquizoítos tratados con SIRT no mostraron cambios de los niveles en el marcaje de H3ac y H4K16ac por IFI en comparación con los taquizoítos tratados con DMSO (*Fig.54* y *Fig.55*). Los taquizoítos sin tratar y a los que se les aplico DMSO no difirieron

en las intensidades de sus fluorescencias (*Fig.54* y *Fig.55*). Sin embargo, por WB observamos un pequeño aumento significativo en el nivel de acetilación de H4K16 en comparación con el control a la dosis máxima de 100  $\mu$ M ( $p \leq 0,05$ ), pero no de H3ac (*Fig.56*). Estos resultados resultan ser preliminares ya que solo se pudieron realizar 2 ensayos.

**Fig.52. Efecto de RSV sobre H4K16ac en taquizoítos intracelulares.** **A.** IFI con anticuerpo anti-H4K16ac (1: 200). Se eligieron VPs con dos parásitos (círculos blancos) para comparar el tratamiento y el vehículo, DMSO. Las líneas blancas de la parte superior derecha corresponden a la barra de escala: 10  $\mu$ m. Los núcleos se tiñeron con DAPI. Los taquizoítos intracelulares se trataron con RSV (100  $\mu$ M) o DMSO al 0,5% durante 24 hs. **B.** Las intensidades de fluorescencia de H4K16ac y de DAPI se cuantificaron a partir de 10 núcleos (**C**) y se representaron gráficamente como intensidades relativas. El panel y el gráfico son representativos de tres experimentos independientes con resultados similares. El análisis estadístico se realizó con ANOVA unidireccional y prueba de comparación múltiple de Tukey (\*  $p \leq 0.05$  y \*\*\*\*  $p \leq 0.0001$ ).



**Fig.53. WB de lisados de proteínas totales de *T. gondii*. tratados con RSV. A.** Anti-H3 (1: 250) o anti-H3ac (1: 500). Los lisados se obtuvieron a partir de taquizoítos intracelulares purificados tratados con RSV (70  $\mu$ M) o DMSO 0,5% durante 24 hs. Las intensidades de la banda de H3ac se cuantificaron en relación con la intensidad de la banda de histona H3 y se representaron gráficamente. **B.** Anti-H4 (1: 250) o anti-H4K16ac (1: 500). Se obtuvieron lisados de taquizoítos intracelulares purificados tratados con RSV (50  $\mu$ M) o DMSO al 0,5% durante 24 hs. Las intensidades de la banda de H4K16ac se cuantificaron en relación con la intensidad de la banda de la histona H4. El análisis estadístico se realizó con ANOVA de una vía y prueba de comparación múltiple de Tukey (\*  $p \leq 0.05$ ).

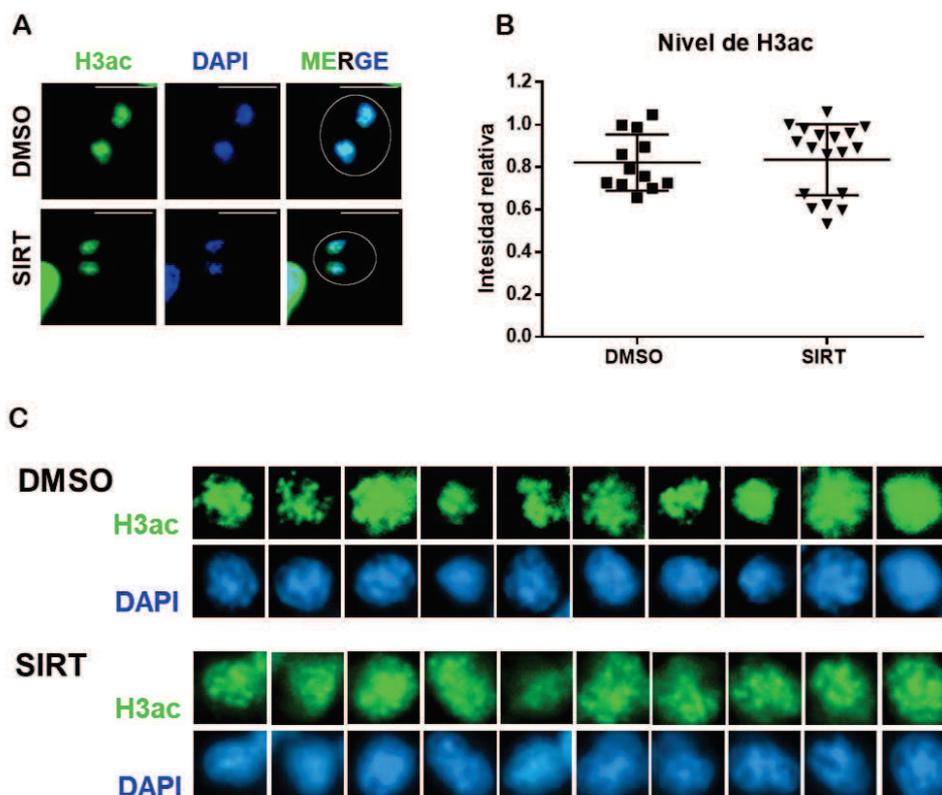


#### 19.4 Efectos del activador e inhibidor de Sirtuinas sobre el ciclo celular

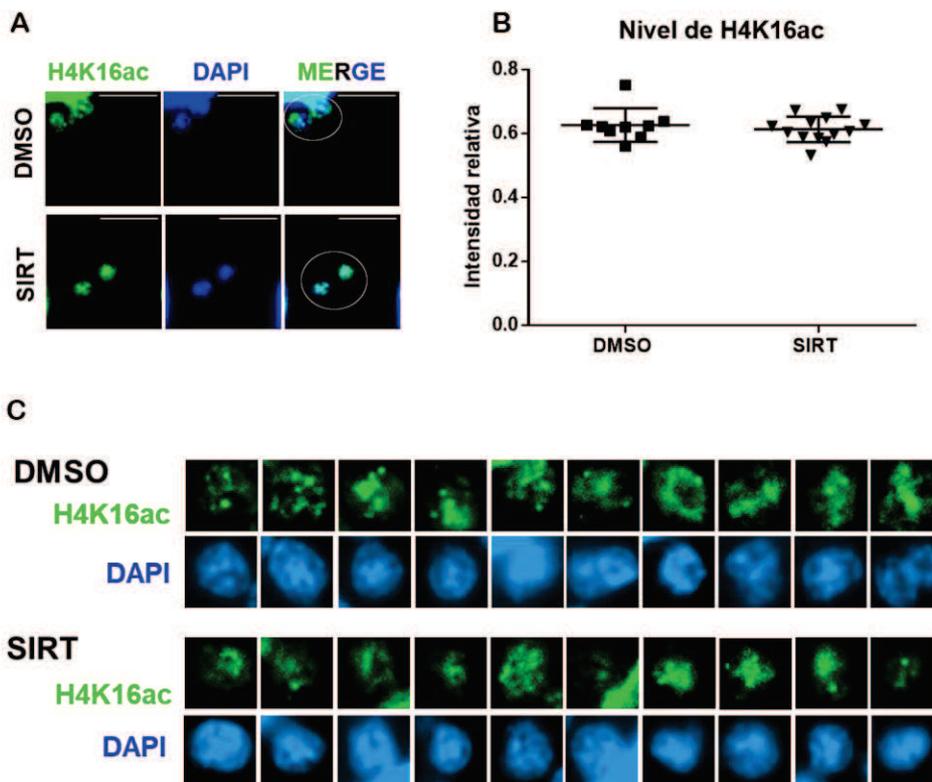
Para explorar un poco más las bases del efecto antitoxoplásmico del activador RSV y también de la contraparte SIRT como inhibidor de SIR2, se realizó un análisis por citometría de flujo, FACS, mediante el cual se midió el contenido de ADN durante el ciclo celular de los taquizoítos intracelulares en presencia de las drogas. Para esto, se cultivaron taquizoítos intracelulares durante 24 hs en presencia de 100  $\mu$ M de RSV y SIRT para luego evaluar el contenido de ADN durante el ciclo

celular en comparación al control (DMSO) e Hidroxiurea, HU, la cual se sabe detiene el ciclo celular específicamente en la fase G1 o S temprana causando una disminución y/o desaparición del pico M/G2 (de Melo, Mayerhoffer, & de Souza 2000). El control DMSO presentó un perfil de contenido de ADN característico con dos picos bien definidos (Radke et al. 2001) (*Fig. 57A*). El primer pico corresponde a un contenido 1N de ADN propio de taquizoítos en fase G1 y el segundo de menor tamaño a un contenido

**Fig.54. Efecto de SIRT sobre H3ac en taquizoítos intracelulares.** **A.** IFI con anticuerpo anti-H3ac (1: 200). Se eligieron VPs con dos parásitos (círculos blancos) para comparar el tratamiento y el vehículo, DMSO. Las líneas blancas de la parte superior derecha corresponden a la barra de escala: 10  $\mu$ m. Los núcleos se tiñeron con DAPI. Los taquizoítos intracelulares se trataron con SIRT (100  $\mu$ M) o DMSO al 0,5% durante 24 hs. **B.** Las intensidades de fluorescencia de H3ac y de DAPI se cuantificaron a partir de 10 núcleos (**C**) y se representaron gráficamente como intensidades relativas. El panel y el gráfico son representativos de tres experimentos independientes con resultados similares. El análisis estadístico se realizó con ANOVA unidireccional y prueba de comparación múltiple de Tukey sin detectar diferencias significativas.



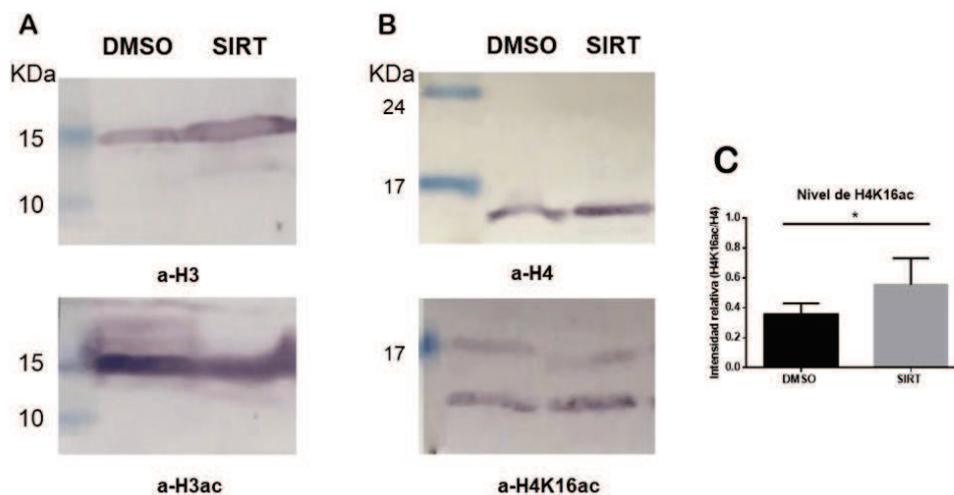
**Fig.55. Efecto de SIRT sobre H4K16ac en taquizoítos intracelulares.** **A.** IFI con anticuerpo anti-H4K16ac (1: 200). Se eligieron VPs con dos parásitos (círculos blancos) para comparar el tratamiento y el vehículo, DMSO. Las líneas blancas de la parte superior derecha corresponden a la barra de escala: 10  $\mu$ m. Los núcleos se tiñeron con DAPI. Los taquizoítos intracelulares Se trataron con SIRT (100  $\mu$ M) o DMSO al 0,5% durante 24 hs. **B.** Las intensidades de fluorescencia de H4K16ac y de DAPI se cuantificaron a partir de 10 núcleos (**C**) y se representaron gráficamente como intensidades relativas. El panel y el gráfico son representativos de tres experimentos independientes con resultados similares. El análisis estadístico se realizó con ANOVA unidireccional y prueba de comparación múltiple de Tukey sin detectar diferencias significativas.



~1.8 de ADN correspondiente a parásitos que están en fase S o mitosis/G2. Mientras que, el tratamiento con RSV mostró una disminución en el contenido de ADN en estadio M/G2 lo que nos daría un indicio de arresto del ciclo celular en la fase S, similar a la inducida por la HU (*Fig. 57A*). En contraste, SIRT, mostró un perfil de ADN similar al de DMSO (*Fig. 57A*) sugiriendo que su actividad anti-*T. gondii* no estaría asociada con una alteración en el ciclo celular. En la *Fig. 57B* se muestra una comparación de la

distribución del contenido de ADN para los cuatro tratamientos y se puede observar que las diferencias entre dicha distribución serían significativas entre el control y RSV. Por lo tanto, podríamos decir que el tratamiento con RSV a dosis elevadas (100  $\mu\text{M}$ ) afectaría la maquinaria de replicación del ADN causando un arresto celular y, en consecuencia, afectando la división celular como hemos demostrado.

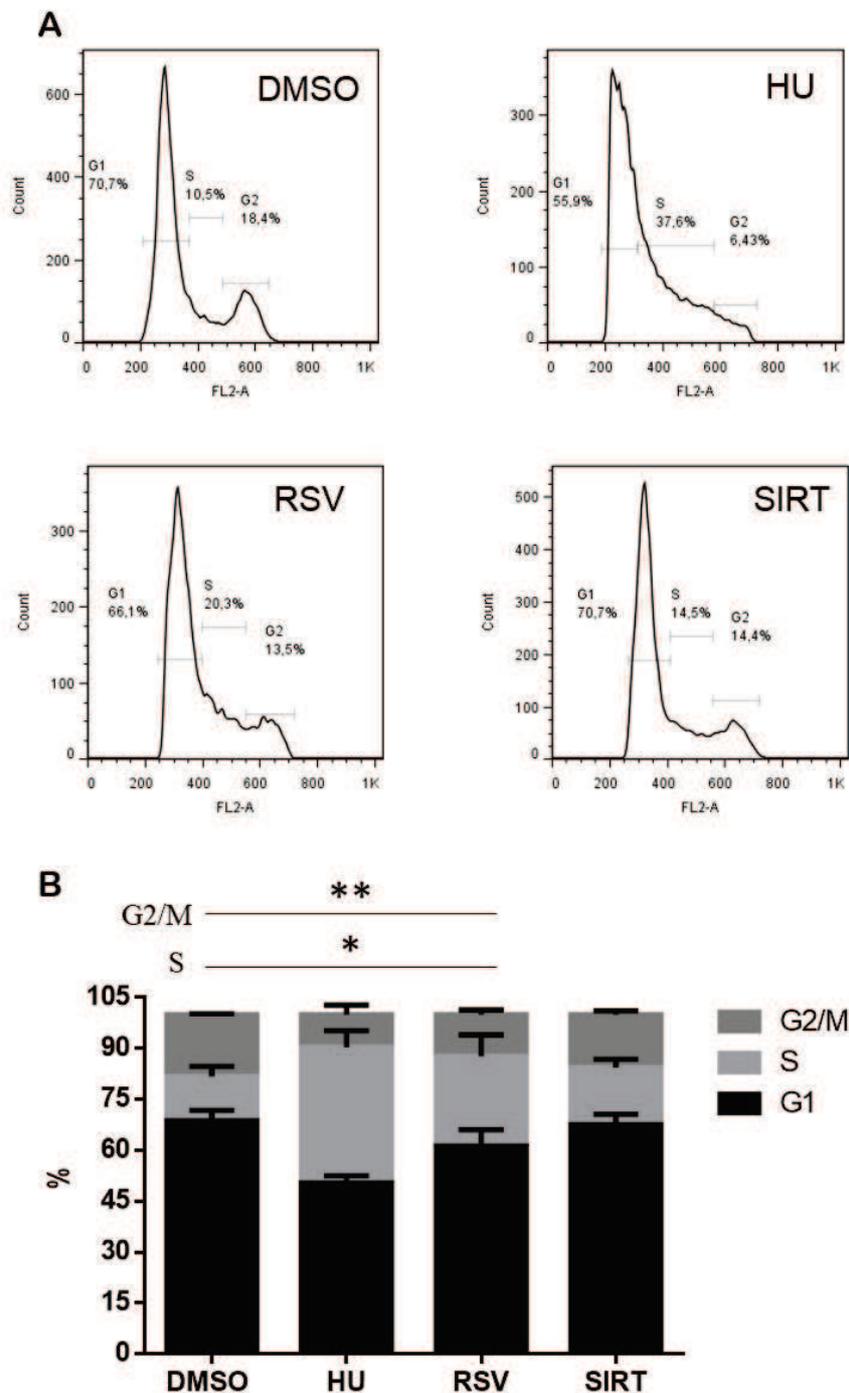
**Fig.56. WB de lisados de proteínas totales de *T. gondii* tratados con SIRT. A.** Anti-H3 (1: 250) o anti-H3ac (1: 500). Los lisados se obtuvieron a partir de taquizoítos intracelulares purificados tratados con SIRT (100  $\mu\text{M}$ ) o DMSO 0,5% durante 24 hs. **B.** Anti-H4 (1: 250) o anti-H4K16ac (1: 500). Se obtuvieron lisados de taquizoítos intracelulares purificados tratados con RSV (50  $\mu\text{M}$ ) o DMSO al 0,5% durante 24 hs. **C.** Las intensidades de la banda de H4K16ac se cuantificaron en relación con la intensidad de la banda de la histona H4. El análisis estadístico se realizó con ANOVA de una vía y prueba de comparación múltiple de Tukey (\*  $p \leq 0.05$ ).



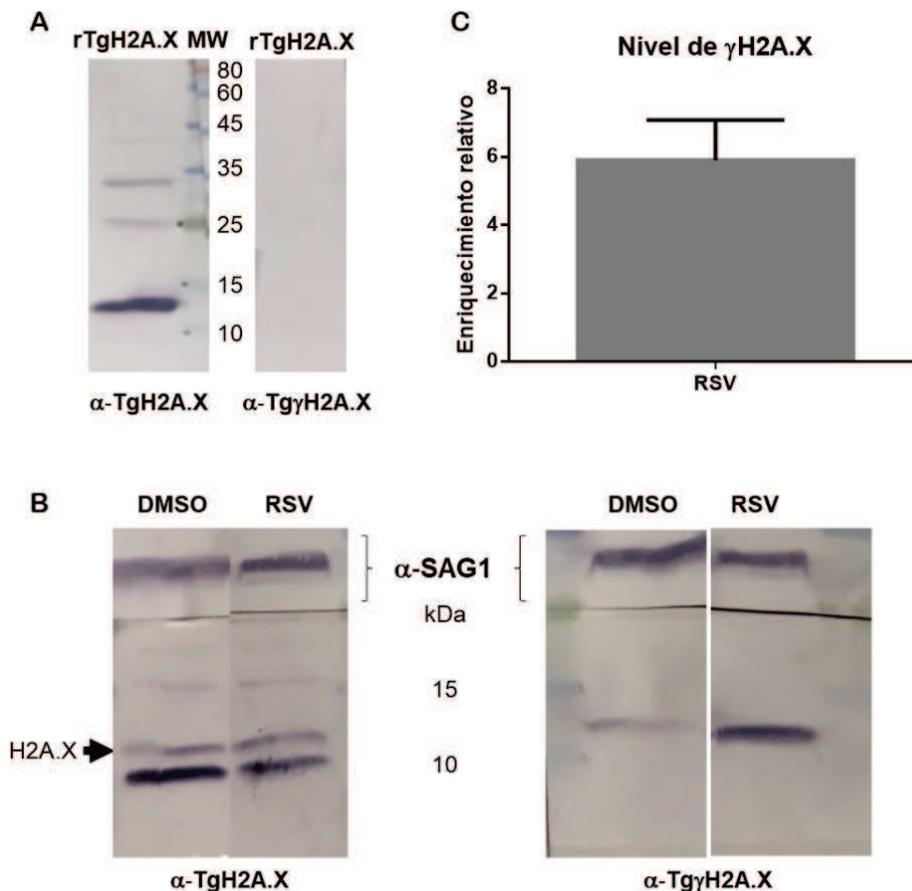
### 19.5 Efecto de RSV sobre el daño al ADN

Para determinar si el efecto de RSV estaba asociado con posibles daños en el ADN, es decir, a roturas de doble cadena, DSB, se determinó el nivel de  $\gamma\text{H2A.X}$  (un marcador del nivel de DSB) mediante WB. Se utilizó un anticuerpo específico hecho en conejo anti-*T. gondii* contra un péptido fosforilado en el dominio C-terminal SQE de H2A.X de *T. gondii* el cual sería indicativo de la presencia de daño.

**Fig. 57. Distribución del contenido de ADN en el ciclo celular de taquizoítos intracelulares tratados con RSV y SIRT. A.** Perfil de citometría de flujo de taquizoítos asincrónicos de la cepa RH que se replicaron durante 24 hs sin droga y luego 24 hs con 100  $\mu$ M en presencia de las mismas. Los resultados se representan en histogramas mediante 20000 eventos contados en función de la fluorescencia FL2-A. **B.** Distribución en el ciclo celular del contenido de ADN por tratamiento, (M-C: mitosis-cariocinesis). Valor de  $p$  (\*  $p \leq 0.05$  y \*\* $p \leq 0.01$ ), las diferencias son significativas en comparación con la condición control DMSO. ANOVA de una vía, prueba de comparaciones múltiples de Tukey,  $\alpha = 0.05$ . Solo se muestran las diferencias significativas entre DMSO y RSV.



**Fig.58. Efecto de la exposición intracelular a RSV sobre la fosforilación de H2A.X ( $\gamma$ H2AX) en *T. gondii*.** **A.** WB de H2A.X recombinante (rH2A.X; 200 ng/carril). Anti-rTgH2A.X-conejo ( $\alpha$ -TgH2A.X, 1:5000) y anti-*T. gondii* péptido fosforilado ( $\alpha$ -Tg $\gamma$ H2A.X, 1:100). La secuencia del péptido fosforilado fue NH<sub>2</sub>-C+ GKHGV-S<sub>(P03H2)</sub>-QEF-COOH. **B.** WB de proteínas totales de *T. gondii* con anti-SAG1 (1: 500), anti-TgH2A.X o anti-Tg $\gamma$ H2A.X. Los lisados se obtuvieron a partir de taquizoítos intracelulares purificados, previamente tratados con RSV (50  $\mu$ M) o DMSO al 0,5% durante 24 hs. Se cuantificaron y normalizaron las intensidades de las bandas H2A.X (flecha) y  $\gamma$ H2A.X a las intensidades de la banda SAG1. A continuación, se calculó la intensidad relativa de las bandas ( $\gamma$ H2A.X / H2A.X) para cada tratamiento. La señal normalizada a SAG1 se calculó a partir de valores relativos en comparación con el control DMSO. Los resultados son la media de tres réplicas  $\pm$  DE. Este panel corresponde al experimento 3 (ver Contreras et al., 2021, Fig. Supp2). **C.** Cuantificación del aumento de veces en el nivel de  $\gamma$ H2A.X después del tratamiento con RSV.



Para confirmar la especificidad del anticuerpo generado se realizó un ensayo de WB contra proteínas recombinantes H2A.X de *T. gondii* no fosforiladas. El anticuerpo anti-Tg $\gamma$ H2A.X de conejo no reconoció a H2A.X de *T. gondii* (Fig.58A) pero si reaccionó con una banda esperada de 14,5 KDa en el

lisado de proteínas totales del parásito después del tratamiento con RSV (**Fig.58B**), la señal de  $\gamma$ H2A.X detectada se incrementó en comparación con el control DMSO, alcanzado bajo tratamiento un incremento de aproximadamente 5,90 veces (**Fig.58C**) demostrando que el efecto de RSV estaría asociado a DSB.

### **19.6 Localización genómica y subcelular de TgSIR2A y TgSIR2B**

Mediante el análisis en la base de datos (ToxoDB) se pudo constatar que ambos genes se encuentran conservados en todas las cepas de *T. gondii*. En ambos casos, la sintenia de *tgSir2a* y *tgSir2b* se mantiene, aunque se observaron duplicaciones génicas para ambos genes. El gen *tgSir2a* estaría duplicado en la cepa CAST, mientras que *tgSir2b* lo estaría en las cepas ARI, COUG, FOU, GT1, MAS y CatBr9; presentando rearrreglos en todos los casos.

#### **19.6.1 Análisis del rol biológico**

El RSV es una droga con múltiples blancos, entre ellos las deacetilasas de la familia de las Sirtuinas, SIR2. El efecto observado de esta droga sobre *T. gondii*, incluyendo el grado de acetilación de las histonas H3 y la marca sobre H4K16, respalda la posibilidad de que el RSV esté afectando a alguna o las dos SIR2 de *T. gondii*. Con el objetivo de estudiar la asociación directa de las Sirtuinas de *T. gondii* con los subtelómeros y poder establecer un nexo con los efectos del RSV y SIRT decidimos producir anticuerpos anti-SIR2A y –SIR2B. De esta manera, si el anticuerpo funcionara correctamente, podríamos incluso hacer experimentos en diferentes cepas de *T. gondii* si fuera necesario.

#### **19.6.2 Análisis in silico de Sirtuinas**

Las proteínas Sirtuinas pertenecen un grupo conservado de proteínas remodeladoras de la cromatina presentes desde bacterias hasta humanos. El dominio catalítico de SIR2 compartido entre todas las Sirtuinas, consta de dos dominios distintos, uno que se une a NAD<sup>+</sup> y otro al sustrato de acetil-lisina, respectivamente. Además del dominio catalítico, las Sirtuinas eucariotas contienen extensiones N-terminal y C-terminal variables que regulan tanto sus localizaciones subcelulares como su actividad catalítica (North & Verdin 2004). El dominio característico es de aproximadamente 260 aminoácidos y

representa a una familia que se divide en cinco clases (I-IV y U) (Frye 2000). El genoma de *S. cerevisiae* codifica cinco Sirtuinas, SIR2p y cuatro proteínas adicionales denominadas Hst1p-Hst4p (Brachmann et al. 1995). El genoma humano codifica siete Sirtuinas, con representantes de las clases I-IV (Frye 2000).

Mediante un análisis con en la base de datos NCBI y la herramienta de búsqueda de dominios conservados (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) se pudo constatar la presencia del dominio Sirtuina y su localización tanto en las proteínas SIR de humanos como en las proteínas de *T. gondii* y *P. falciparum*. En la **Tabla 13** se muestran las características de las Sirtuinas de humanos y levaduras como así también las de los organismos Apicomplexa.

La proteína TgSIR2A tendría un tamaño de 360 aa a diferencia de su contraparte PfSIR2A la cual posee un tamaño de 271 aa (**Tabla 13**). Dicha proteína presentaría una región hidrofóbica compatible con péptido señal, con posible sitio de clivaje en el aminoácido 17 predicho por el programa SignalP 2.0 (probabilidad 0.982). Mientras que TgSIR2B sería una proteína de 1715 aa en comparación con PfSIR2B la cual tiene un tamaño de 1304 aa (**Tabla 13**). A diferencia de TgSIR2A, TgSIR2B no presentaría péptido señal o región hidrofóbica terminal en su secuencia. El dominio catalítico sirtuina de TgSIR2A sería conservado y se encontraría entre los aminoácidos 97 y 360 (263 aa) (**Tabla 13**), mientras que, el dominio catalítico sirtuina de TgSIR2B estaría parcialmente conservado entre los aminoácidos 189 y 503, similar al dominio sirtuina de PfSIR2B entre los aminoácidos 48 y 415; ambos partidos por una inserción de 74 y 151 aa respectivamente, las cuales dividen en dos regiones al dominio sirtuina haciéndolo mucho más grandes que en sus contrapartes (**Tabla 13**).

En este análisis se pudo ver que en general las Sirtuinas comparadas muestran un tamaño de dominio catalítico de entre 212 y 367 aa (**Tabla 13**), presentando TgSIR2B y PfSIR2B los dominios de mayor tamaño debido a la inserción más arriba mencionada. La diferencia en tamaño del dominio también se ve reflejada en el tamaño de las proteínas, siendo TgSIR2B y PfSIR2B particularmente grandes en comparación a TgSIR2A y PfSIR2A y sus contrapartes en humanos. En el análisis realizado por Tonkin y col., (2009), TgSIR2A y PfSIR2A quedaron agrupadas junto a la proteína mitocondrial hSIR-5 de

Sirtuina	Localización	Dominio catalítico	Sitio activo	Longitud dominio	Longitud proteína
hSIR-1	citoplasmática	254-489	363	235	747
hSIR-2	nuclear	77-331	187	254	389
hSIR-3	mitocondrial	138-373	248	235	399
hSIR-4	mitocondrial	47-308	161	261	314
hSIR-5	mitocondrial citoplasmática nuclear	51-301	158	250	310
hSIR-6	nuclear	45-257	133	212	355
hSIR-7	nuclear nucléolo	100-314	187	214	400
TgSir2A	citoplasmática perinuclear <sup>1</sup>	97-360	228	263	360
TgSir2B	¿?	189-257... 333-503 <sup>2</sup>	385	314	1715
PfSir2A	nuclear perinuclear	1-271	133	271	271
PfSir2B	¿?	48-78... 229-415 <sup>2</sup>	294	367	1304

**Tabla 13. Comparación de proteínas sirtuinas presentes en humanos y Apicomplexas.** Las secuencias correspondientes a las proteínas de humanos (hSIR-1 a 7) fueron extraídas de la base de datos *Uniprobe*, en la versión de isoforma 1. hSIR-1: Q96EB6, hSIR-2: Q8IXJ6, hSIR-3: Q9NTG7, hSIR-4: Q9Y6E7, hSIR-5: Q9NXA8, hSIR-6: Q8N6T7, hSIR-7: Q9NRC8; ToxoDB, TgSir2A: TGME49\_227020, TgSir2B: TGME49\_267360; PlasmoDB, PfSir2A: PF13\_0152, PfSir2B: PF14\_0489. El sitio activo se corresponde con un aminoácido aceptor de protones y se corresponde con el aminoácido en el cual se encuentra conservado. Todos los tamaños están expresados en cantidad de aminoácidos. <sup>1</sup> TgSir2A se localizaría principalmente en el citoplasma y cerca de la periferia nuclear según Nardelli y col., comunicación personal del laboratorio de la Dra Kami Kim. <sup>2</sup> TgSir2B y PfSir2B poseen un dominio sirtuina cortado por una región de 74 y 151 aa respectivamente.

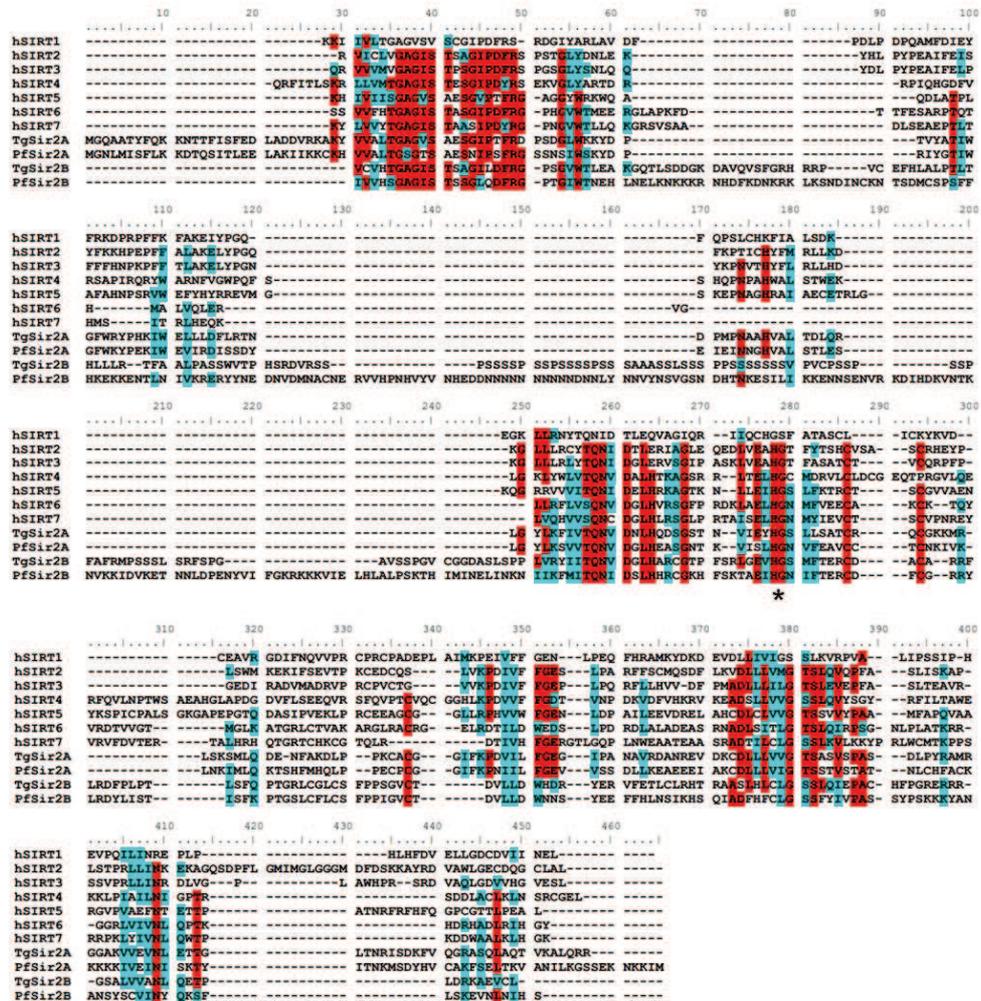
humanos como Sirtuinas de clase III (con un soporte moderado) y TgSIR2B y PfSIR2B en otro clado de Sirtuinas de tipo IV junto a las contrapartes humanas hSIR-6 y hSIR-7 (con un muy buen soporte). La Sirtuina de clase III es la forma más ampliamente distribuida en procariotas, por lo que puede ser la versión más antigua del gen sirtuina (Frye 2000). Las Sirtuinas de tipo IV agrupa a Sirtuinas con funciones diversas en la señalización celular, el mantenimiento de la cromatina telomérica (hSIR-6) (Michishita et al. 2008; Mostoslavsky et al. 2006) y la activación de la ARN polimerasa I (hSIR-7) (Ford et al. 2006).

Un alineamiento estructural de las Sirtuinas de *Toxoplasma* y *Plasmodium* con los dominios Sirtuinas de humanos (hSIR-1 a hSIR-7) mostró que todas presentan una región homóloga conservada correspondiente al dominio catalítico Sirtuina, con varios aminoácidos también conservados (**Fig.59**). Existen varios motivos cortos de aminoácidos conservados presentes dentro del dominio central sirtuina compatibles con los descritos por Frye (2000). Estos incluyen GAGISXXXGIPXXR, PXXXH, TQNID, HG, CXXC el cual puede ser un posible dominio de dedo de zinc, FGE, GTS y N. En las Sirtuinas de clase III, el motivo GAGISXXXGIPXXR suele ser GAGISAESGIPTFR, mientras que en las Sirtuinas de la clase IV tienen GIPD dentro de este motivo (Frye 2000). En secuencias de clase IV hay un motivo GVWTL presente cuatro residuos C-terminal al motivo GAGISTXXGIPDFR. En secuencias de clase III, el motivo PXXXH es PNXXH mientras que en las secuencias de clase IV una (T) o (S) suele seguir a la (P). El motivo HG es de interés porque la mutación puntual sobre la histidina en HG a tirosina (YG) causa pérdida de ribosilación de ADP mediada por sirtuina y deacetilación (Frye 1999; Tanny et al. 1999; Guarente 2000) y convierte a SIR2 de levaduras en un gen negativo dominante (Tanny 1999). El motivo HG es estrictamente conservado en todas las Sirtuinas conocidas por lo tanto se la considera indispensable para la actividad de las mismas (Frye 2000).

Como se muestra en la **Fig.59** las Sirtuinas de tipo III TgSIR2A y PfSIR2B conservan en parte el motivo GAGISAEGIPTFR característico, en donde TgSIR2A presenta un motivo GAGVSAESGIPTFR, en el cual el aminoácido isoleucina (I) es sustituido por una valina (V). Mientras que la Sirtuina tipo IV,

TgSIR2B presenta una secuencia dentro del motivo: GLD en vez de GIPD, donde ocurre una sustitución de prolina (P) por leucina (L) y también posee el motivo GVWTL hacia la región C-terminal.

**Fig.59. Alineamiento estructural de dominios sirtuinas.** Las secuencias aminoacídicas de sirtuinas de humanos se obtuvieron a partir de la base de datos *Uniprobe*, (hSIR-1 a 7), de las cuales se utilizaron las correspondientes a la isoforma 1 de cada una. hSIR-1: Q96EB6, hSIR-2: Q8IXJ6, hSIR-3: Q9NTG7, hSIR-4: Q9Y6E7, hSIR-5: Q9NXA8, hSIR-6: Q8N6T7, hSIR-7: Q9NRC8; ToxoDB, TgSir2A: TGME49\_227020, TgSir2B: TGME49\_267360; PlasmoDB, PfSir2A: PF13\_0152, PfSir2B: PF14\_0489. El alineamiento se realizó utilizando el programa *MEGA-X* con el método *CLUSTAL-W* y se editó en el programa *Bioedit*. En color rojo se muestran los residuos conservados. En color celeste se muestran las regiones con sustituciones similares o sinónimas. El asterisco denota los residuos HG conservados, donde la H sería esencial para la actividad catalítica.



El motivo PNXXH de Sirtuinas tipo III también se encuentra conservado en TgSIR2A, pero no en PfSIR2A, donde (P) es sustituido por (I). Mientras que no lo haría en TgSIR2B y PfSIR2B. Por otra parte, los motivos HG, CXXC y N se encontraron totalmente conservados en las proteínas de Apicomplexa. El motivo TQNID se encontró totalmente conservado en PfSIR2B y sustituido en TgSIR2A, TgSIR2B y PfSIR2A (TQNVD). Del motivo GTS solo se conserva la glicina (G) en todas ellas y en TgSIR2A y PfSIR2A todo el motivo. Mientras que el motivo FGE, solo se conserva en las proteínas TgSIR2A y PfSIR2A (*Fig.59*).

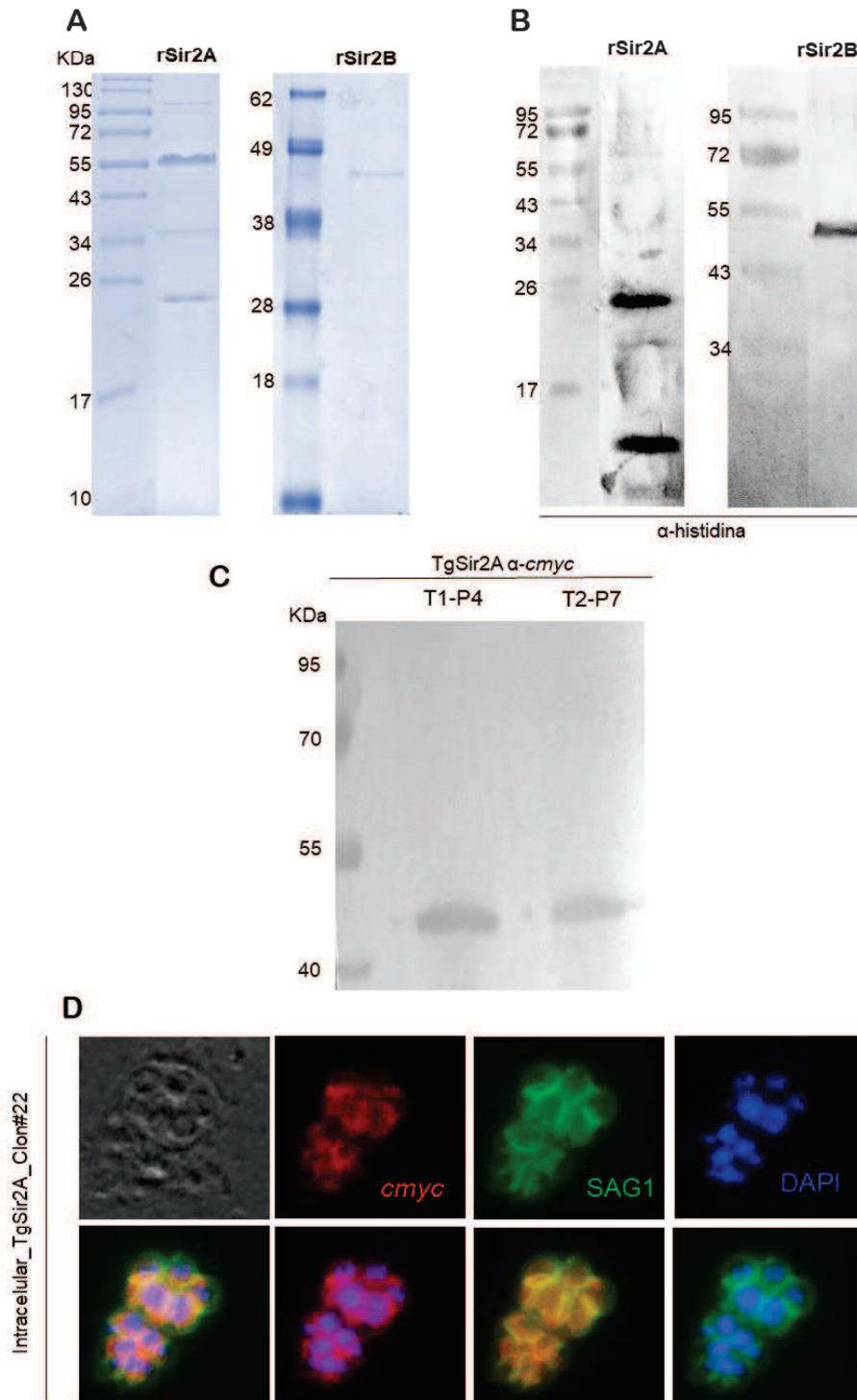
### **19.6.3 Producción de proteínas recombinantes, generación de anticuerpos específicos e inmunodetección**

Para poder llevar a cabo los experimentos de asociación de TgSIR2s a las regiones TgSTs decidimos generar proteínas recombinantes para luego producir anticuerpos policlonales específicos y poder caracterizar la expresión y localización de dichas proteínas.

#### **19.6.3.1 TgSIR2A recombinante (rSIR2A)**

Para esto se expresó de manera recombinante una primera versión truncada de TgSIR2A sin la región hidrofóbica N-terminal de 18 aa (rTgSIR2A\_truncada), quedando una secuencia de 324 aa que junto a la etiqueta de histidinas y el aporte *linker* del plásmido se esperaba una versión de 381 aa y un peso estimado de ~42 KDa. El constructo codificante para rTgSIR2A\_truncada fue clonado en el vector de expresión en bacterias pRSET-A de manera estándar, confirmado por secuenciación, e inducido a expresión con IPTG. La proteína recombinante rSIR2A fue purificada por afinidad en columnas de níquel en condiciones desnaturizantes como se describe en el Capítulo IV y se resolvió en geles SDS-PAGE 12%, observándose varias bandas, ninguna del tamaño esperado por tinción con *coomasie blue* (*Fig.60A*). Luego, se realizó una inmunodetección utilizando el anticuerpo  $\alpha$ -histidina, el cual

**Fig.60. Purificación de las proteínas Sir2 recombinantes en condiciones desnaturalizantes. A.** Tinción con *coomasie blue*, SDS-PAGE 12%. En cada caso se sembró 1µg de proteína recombinante. Tamaño esperado de rSir2A (rTgSir2A\_truncada) ~42 KDa; tamaño esperado de rSir2B (rTgSir2B\_DC) ~42 KDa, tamaño observado ~48 KDa. **B.** WB de las proteínas recombinantes, en todos los casos se sembró 1µg de cada una de ellas. **C.** WB anti-*cmyc* (1:3000) con proteínas totales de parásitos transfectados con la copia extra del gen *tgSir2a-cmyc*. **D.** IFI de taquizoitos intracelulares con anti-*cmyc* conejo 1:300 y anti-SAG1 ratón 1:200 del clon estable TgSir2A-*cmyc*\_clon#22.



reconoce la etiqueta de polihistidina (6xHis) presente en la construcción rTgSIR2A\_truncada (**Fig.60B**). En este caso, no pudimos constatar la presencia de la banda esperada de ~42 KDa. Por otra parte, se construyó una segunda versión de TgSIR2A contenida entre los sitios de restricción Nsil y PacI, N\_rTgSIR2A\_P, clonada de manera comercial en pRSET-A para posterior subclonado en vector de expresión específico en *T. gondii*, MCS pTUB8mycGFPPftailTy-HX, el cual nos permitiría agregar una versión extra del gen con una etiqueta *c-myc* en la copia de proteína, TgSIR2A\_*c-myc*, y de esta manera localizarla subcelularmente. En este caso el fragmento recombinante N\_rTgSIR2A\_P junto al aporte del plásmido de expresión en bacterias pRSET-A codificarían una versión de 401 aa con un PI de 7,2 y un peso estimado de ~45 KDa la cual aún no fue ensayada. En el caso de N\_rTgSIR2A\_P se prosiguió con el subclonado en el vector de expresión en *Toxoplasma* pudiendo obtener clones de parásitos positivos para TgSIR2A\_*c-myc* en pasajes estables en dos transfecciones distintas: transfección 1, pasaje 4 (T1-P4) y transfección 2, pasaje 7 (T2-P7) conteniendo la doble copia del gen como se observa en la **Fig.60C**. En el extracto total de proteínas pudimos reconocer la proteína del tamaño esperado de ~42 KDa correspondiente a la copia extra del gen. Uno de los clones estables obtenido por selección (clon#22-T2) se testeó mediante IFI con el anticuerpo anti- etiqueta *c-myc* en el cual se pudo reconocer a la proteína en el citoplasma de parásitos intracelulares (**Fig.60D**). Al parecer la proteína excluiría el núcleo en el estadio de replicación intracelular lo cual se condice con lo observado por Nardelli y col., en su comunicación personal como se menciona en la **Tabla 13**.

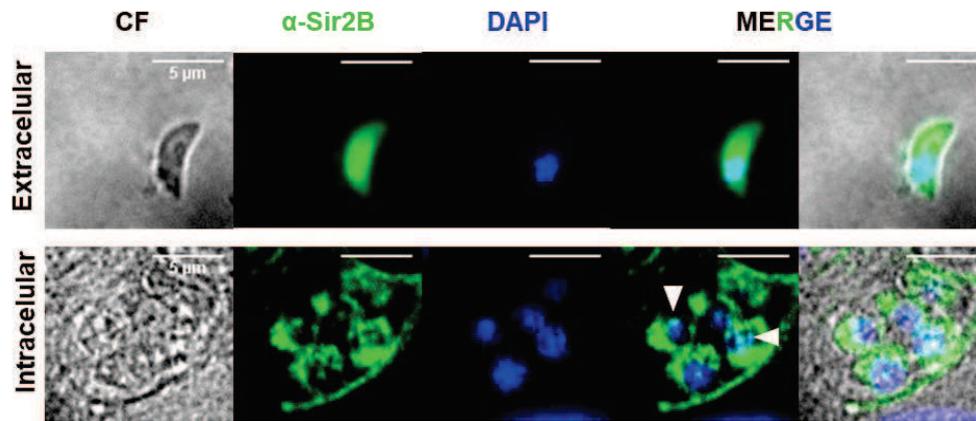
#### **19.6.3.2 TgSIR2B recombinante (rTgSIR2B)**

En el caso de la proteína TgSIR2B se produjo una versión recombinante conteniendo el dominio catalítico (DC) de TgSIR2B. En este caso, realizamos una versión recombinante de TgSIR2B (rTgSIR2B\_DC o rSIR2B) desde el aa (F) 176 al aa (P) 522 generando una versión del dominio catalítico de 347 aa que junto al aporte plasmídico suman un total de 390 aa con un peso estimado de ~42 KDa (rSIR2B). La versión recombinante se clonó en pET14b de manera comercial. Fue expresada mediante la inducción con IPTG y purificada en columna de níquel en condiciones desnaturizantes como se

describe en el Capítulo IV. Se logró expresar una proteína de forma soluble, y se resolvió en geles SDS-PAGE 12% con tinción con *coomassie blue* (**Fig.60A**). Durante la purificación de rSIR2B se reconoció una banda de ~45 KDa (**Fig.60A**), mientras que cuando se realizó una inmunodetección con el anticuerpo anti-etiqueta de histidina se detectó una banda de ~48 KDa (**Fig.60B**). La proteína recombinante correspondiente a rSIR2B fue utilizada para inmunizar ratones según el protocolo descrito en el capítulo IV y de esta manera se produjeron los anticuerpos policlonales correspondientes.

Como siguiente aproximación y para testear de manera preliminar el anticuerpo generado se realizó una inmunodetección de TgSIR2B mediante WB con extractos de proteínas total de taquizoítos extracelulares, sin lograr reconocer con éxito a la misma. Esta ausencia de proteína podía deberse a factores técnicos dado que la proteína posee un tamaño de 187 KDa y es necesario realizar corridas electroforéticas en condiciones especiales a estandarizar: geles largos, porcentajes de acrilamida/bis acrilamida por debajo de 6%, etc; que quedaron pendientes para futuros trabajos. Por lo tanto, se realizó un ensayo de IFI con el anticuerpo generado en taquizoítos extracelulares e intracelulares y de esta manera se pudo localizar subcelularmente la expresión tentativa de esta proteína. En la **Fig.61** se puede observar que TgSIR2B presentaría señal de localización en todo el parásito, incluyendo el núcleo, en taquizoítos extracelulares; mientras que en taquizoítos intracelulares su localización sería principalmente citoplasmática con localización puntual dentro del núcleo (flechas blancas), sugiriendo que su presencia nuclear podría estar relacionada solo a regiones genómicas discretas.

**Fig.61. Localización subcelular por inmunofluorescencia de TgSir2B.** Localización en taquizoítos de la cepa RH extracelulares (arriba) e intracelulares (abajo), una VP con 4 parásitos. Las flechas blancas indican focos puntuales en el núcleo.



## 20. Discusión del capítulo III

En el trabajo de Bunnik y col., (2019) demostraron que en la arquitectura global del genoma de *T. gondii* los telómeros no estarían agrupados, lo que apoyaría la idea de extremos teloméricos en contexto de cromatina separados dentro del núcleo. En este sentido, habría telómeros en contexto de heterocromatina silente localizados tal vez en la periferia nuclear y otros orientados posiblemente hacia territorios nucleares más propicios para la transcripción. De hecho, algunos telómeros enriquecidos en H3K4me3 (como se mencionó en el Capítulo I), son activos y se transcriben a lncRNA TERRA (no se muestran los datos). Mientras que, en relación a los distintos territorios nucleares, los cromosomas de *T. gondii* mostraron una tendencia a la territorialización similar a la del genoma humano con contactos de entre 50 y 500 Kpb correspondientes a TADs o dominios topológicos; los cuales se saben están enriquecidos en interacciones intra-dominios (Bunnik et al. 2019). Las proteínas interactoras son factores claves para entender cuáles son los mecanismos que acompañan a esta arquitectura de los telómeros/STs ya que son las que escanean, establecen y efectúan con su actividad el funcionamiento de estas regiones tan particulares del ADN. Como un aporte para comenzar a dilucidar el intrincado mecanismo molecular que subyace asociado a las regiones subteloméricas de *T. gondii* realizamos una búsqueda *in silico* de posibles proteínas asociadas a dichas regiones. De manera notable, la mayoría de las proteínas detectadas en los telómeros/subtelómeros de levaduras y mamíferos no están presentes ni en *T. gondii* ni en *Plasmodium spp.* Si bien, estos parásitos podrían tener un repertorio menor de proteínas en estas regiones, es necesario realizar experimentos más precisos para develar el proteoma de los STs y telómeros. Aunque hemos identificado algunas de estas aún no se ha podido demostrar la asociación directa a estas regiones como, por ejemplo, de SIR2. Por lo tanto, de manera provisoria y según lo estudiado en otras especies se proponen a las proteínas TgSIR2A, TgSIR2B, TgORC1, TgKU70/TgKU80 y TgChromo1 como posibles componentes de los TgSTS, responsables de su estructuración y funcionalidad.

Los resultados obtenidos en este trabajo muestran que el RSV, activador de SIR2, tendría un efecto importante sobre el crecimiento y la replicación de *T. gondii* a un IC<sub>50</sub> de 53,4 µM. Estos resultados coinciden con los reportados por Chen y col., (2018), quienes encontraron que a 50 µM el RSV bloquea el crecimiento intracelular de *T. gondii*. Mientras que, Adeyemi y col., (2019); identificaron a RSV como uno de 62 compuestos con efectos anti-*T. gondii*, que produjeron un valor de IC<sub>50</sub> de 1,03 µg/ml (4,4 µM) después de 72 hs de incubación en comparación con Pirimetamina. Los autores también observaron que las células hospedadoras eran 100% viables en una concentración de RSV de 2 µg/ml (8,5 µM) tras dicho período de tiempo. Las diferencias en los valores de IC<sub>50</sub> entre los estudios pueden deberse a diferencias en la duración de los tiempos de exposición a la droga.

Cuando se compara el efecto observado en *T. gondii* con otras especies de parásitos vemos que dicho efecto en la replicación del taquizoíto fue a un valor de IC<sub>50</sub> inferior al observado sobre la replicación del estadio de epimastigote (forma replicativa) de *T. cruzi* (IC<sub>50</sub> 250 µM) pero mayor al observado en el modelo de *Leishmania amazonensis* (IC<sub>50</sub> 27 µM) (Vergnes et al. 2005). El estudio realizado por la Dra. Campo (Campo 2017) también ha demostrado que el tratamiento extracelular de los promastigotes de *T. cruzi* con RSV reduce el porcentaje de parásitos vivos con un IC<sub>50</sub> de 50.3 µM. Liu y col., (2013) también observaron que el RSV tendría actividad anti-plasmodial con un IC<sub>50</sub> de 70 µM. Tomados en conjunto, estos datos indican que el RSV sería efectivo contra una amplia variedad de parásitos protozoarios, incluyendo a *T. gondii*, en valores de IC<sub>50</sub> similares (Liu et al. 2013).

Dado que el RSV es un fármaco que posee varios blancos, es difícil establecer qué objetivos específicos en *T. gondii* se ven afectados por su acción. Uno de los blancos podría ser una histona acetiltransferasa, pero también se sabe que muestra actividad activadora de la Sirtuina SIR2 (Gertz et al. 2013). Por lo tanto, es posible que las Sirtuinas se encuentren entre los blancos de RSV lo que conduciría a la deacetilación excesiva de las histonas en las regiones STs, entre ellas la marca H4K16ac; como así también a una desregulación en la vía de reparación íntimamente ligada a la regulación epigenética como se vio en levaduras (Liu et al. 2020)

Aquí, observamos que RSV podría estar asociado a cambios en la marca H3ac y H4K16ac en la cromatina del taquizoíto. Un objetivo de la lisina deacetilasa de levaduras SIR2 es H4K16, cuyo estado de acetilación es importante para regular la propagación de la heterocromatina (Wang, Jia, & Jia 2016; Chen, Yu, & Chen 2016). La señal débil de H4K16ac observada en presencia de RSV fue evidente. Este resultado concuerda con otro estudio reciente en el que el tratamiento con RSV redujo los niveles de H4K16ac (Hoggard et al. 2018). La presencia de RSV también afectó la tasa de crecimiento de *T. gondii* y su arresto en la fase S de su ciclo celular. Este resultado acerca del ciclo celular concuerda con lo observado en otros modelos en los que RSV presentó actividad anti proliferativa e indujo un arresto en la fase G1/S o S en células cancerosas (Sgambato et al. 2001; Singh et al. 2017). También se ha encontrado que RSV detiene el ciclo celular en promastigotes de *Leishmania* en la fase G1 (Ferreira et al. 2014). En levaduras, SIR2 controla negativamente la activación de los orígenes de la replicación del ADN dentro de la heterocromatina y la eucromatina deacetilando H4K16 (Cao et al. 2017), pudiendo ser esta la causa del efecto en la replicación del parásito. En este sentido el uso de RSV estaría bloqueando no solo la replicación sino también podría estar afectando la cascada de reconocimiento y señalización de DDR, causando en consecuencia, detención de la progresión en fase S y arresto celular con aumento significativo de  $\gamma$ H2AX como se observó. Similar a lo que se observó con el RSV, se ha informado que KU-55933, un inhibidor de ATM, alteró la replicación de los taquizoítos, detuvo el ciclo celular en la fase S y bloqueó la fosforilación de la histona H2A.X, una marca de histona de la respuesta al daño del ADN (Munera López et al. 2019).

Por el contrario, el tratamiento con SIRT no parecería interferir en el ciclo celular. Del mismo modo, se ha encontrado anteriormente que el SIRT no muestra ningún efecto sobre el ciclo celular en células cancerosas (Peck et al. 2010). Como así también se ha encontrado que en los amastigotes de *Leishmania*, SIRT no causaría cambios en el perfil de contenido de ADN medido por FACS, sino más bien una reducción de los picos G0/G1, S y M a una concentración de 60  $\mu$ M, indicando una fragmentación marcada del ADN (Vergnes et al. 2005). Basándonos en estos datos, podríamos

considerar que, SIRT causaría un efecto similar en los taquizoítos de *T. gondii*, explicando el resultado observado en el análisis de replicación en el que se observó un efecto brusco solo cuando se expusieron a 100  $\mu$ M de la droga.

El SIRT es un fármaco que se identificó como candidato contra el cáncer entre más de 1.000 compuestos contra hSIR-1 con la mitad de la concentración inhibitoria máxima, IC<sub>50</sub>, de 40  $\mu$ M (Grozinger et al. 2001). La proteína PfSIR2A no demostró ser sensible al SIRT *in vitro*, pero fue inhibida por Nicotinamida (Merrick & Duraisingh 2007), un fármaco que no tuvo efecto sobre el crecimiento de *T. gondii* (Strobl et al. 2007). El efecto del SIRT sobre los taquizoítos intracelulares también fue un efecto importante sobre el crecimiento con IC<sub>50</sub> 27  $\mu$ M, y a dosis elevadas (100 $\mu$ M) el efecto anti-*T. gondii* fue cercano al 70% después de 24 hs de tratamiento. Por otra parte, en un modelo de cáncer *in vitro* se observó que el SIRT presentó un efecto anti-proliferativo y apoptótico a través del factor de transcripción FOXO3a y la quinasa AKT (Fong et al. 2014). La inhibición de la vía de señalización a través de las quinasas PI3K/AKT de la célula huésped redujo la proliferación de *T. gondii* (Zhou et al. 2013). Dado que no observamos cambios en los niveles de acetilación de H4K16 ni H3 por el tratamiento con SIRT, ni sobre el ciclo celular de *T. gondii* no podemos descartar que el efecto observado se deba a un efecto sobre la célula hospedadora infectada.

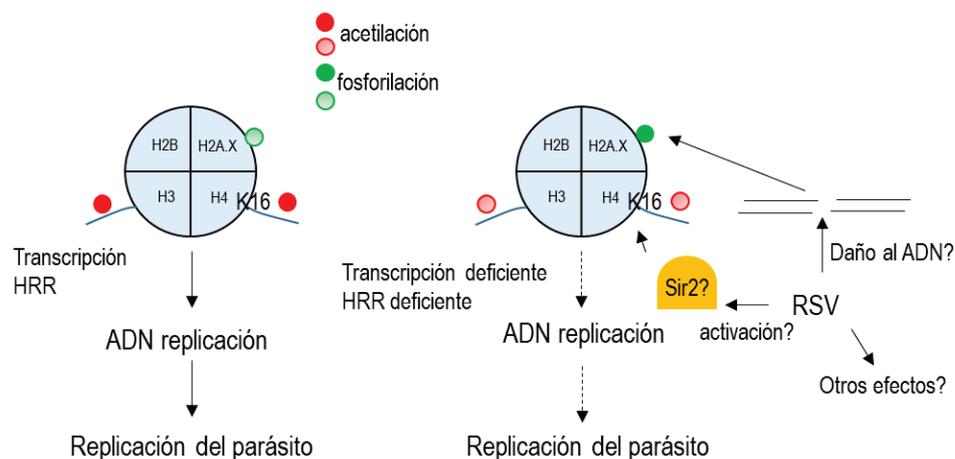
En el presente estudio, se encontró que el nivel de  $\gamma$ H2A.X, marcador del daño presente en el ADN, aumentó en presencia de RSV. En otro modelo de estudio, el uso de 50  $\mu$ M de RSV indujo daño en el ADN, detención en la fase S y niveles aumentados de  $\gamma$ H2A.X en un panel de líneas de carcinoma de células escamosas de cabeza y cuello (Tyagi et al. 2011). Además, la presencia de H4K16ac en el ADN dañado por la rotura de la doble hebra es capaz de reclutar proteínas de reparación del ADN asociadas a la vía de reparación HRR (Gao et al. 2019), entre ellas la quinasa ATM que inicia la respuesta al daño del ADN (Chen et al. 2007). De manera similar a lo que se observó con el RSV, KU-55933, inhibidor de ATM, como ya se mencionó; altera el crecimiento de taquizoítos y detiene el ciclo celular en la fase S (Munera López et al. 2019). La fosforilación del motivo SQE en H2A.X ( $\gamma$ H2A.X) es una marca

sensible de DSB y estrés de replicación del ADN (Rogakou et al. 1998; Gagou, Zuazua-Villar, & Meuth 2010). Aquí observamos que los niveles de  $\gamma$ H2A.X aumentaron en presencia de RSV a 50  $\mu$ M y se indujo daño en el ADN y detención de la fase S al doble de concentración. Además, el RSV se asoció con la inhibición de la HRR a través de la inactivación de la Tirosil-tRNAsintetasa, una enzima asociada a la síntesis de proteínas, pero también a la regulación de genes en la HRR cuando se transloca al núcleo (Cao et al. 2017; Gao et al. 2019).

Tomando en cuenta todos los hallazgos mencionados, el RSV podría inducir una deficiencia en la acetilación de H3 y H4K16, lo que conllevaría a niveles de transcripción alterados, el inicio de la replicación del ADN y, potencialmente, la activación de la vía HRR. Estos procesos podrían afectar la precisión de la replicación del ADN y la progresión en el ciclo celular manteniendo  $\gamma$ H2A.X en niveles muy elevados. A su vez, el RSV podría inducir daño en el ADN, posiblemente por oxidación, generando colapso en las horquillas de replicación y DSB (Branzei & Foiani 2008) presentando una dificultad de reparación mediante HRR debido a la ausencia de H4K16ac (Fig. 62).

Una cuestión muy importante a tener en cuenta es que el RSV a bajas dosis tiene efectos beneficiosos en una amplia variedad de enfermedades humanas; sin embargo, en dosis altas y bajo aplicación a largo plazo, puede tener efectos adversos (Shaito et al. 2020), lo cual resulta un aspecto que

**Fig. 62. Modelo putativo de acción de RSV que conduce a una disminución en la replicación de *T. gondii*.** Las líneas punteadas indican actividades o procesos reducidos. Para acetilaciones y fosforilaciones, intensidad del color representa el nivel de la marca PTM.



debe tenerse en cuenta al considerar esta droga como una terapia anti-*T. gondii*. Aquí, observamos que *in vitro*, el valor de IC<sub>50</sub> de RSV fue menor que el umbral de toxicidad tolerado por las células hTERT. Quizás la dosis y la duración del tratamiento podría manejarse utilizando tratamiento combinados con otros medicamentos (por ejemplo, inhibidores de HRR) para lograr un efecto sinérgico.

En el trabajo de Gupta y col., (2018) se demostró que CAF-1 también co-localiza con H3K56ac, una señal de finalización re-ensamblaje de la cromatina después de la DDR (Masumoto et al., 2005). Resulta interesante el hecho de que H3K56ac es blanco de SIR2 (Maas et al. 2006; Vempati 2012; Mishra et al. 2022) apuntando a la posibilidad de que la cromatina subtelomérica sea un punto de nucleación importante para la replicación y reparación del ADN a través de la acción de SIR2.

*T. gondii* presenta dos Sirtuinas, TgSIR2A y TgSIR2B para las cuales no se han reportado aún funciones específicas (por ejemplo, sus roles en la deacetilación de H4K16), ni sus localizaciones genómicas cuando están presentes en el núcleo, etc. Esta falta de conocimiento perjudica nuestra capacidad para inferir si los efectos observados de RSV están asociados con estas Sirtuinas o no. Nuestros resultados y los de otros muestran que RSV exhibe actividad anti-*T. gondii in vitro*. Se considera pertinente seguir evaluando la acción de dicha droga en estudios con ratones para evaluar más a fondo el potencial del RSV como un fármaco terapéutico contra la toxoplasmosis.

En este trabajo realizamos por primera vez una descripción de la posible localización de la proteína TgSIR2B. Si bien existe una proteína homóloga descrita en *Plasmodium*, PfSIR2B, para la cual se vio asociación a promotores *UpsB* de genes subteloméricos y por ende con localización nuclear, su localización por inmuno-localización subcelular aún no ha sido descrita. Mientras que la localización de PfSIR2A si ha sido descrita y esta parecería tener una ubicación perinuclear, y ocasionalmente se translocaría al núcleo, además de encontrarse asociada a los *clústers* teloméricos periféricos y elementos repetitivos presentes en las regiones PFTAS, como así también enriquecida en el nucléolo (Freitas et al. 2005; Mancio-Silva et al. 2008; 2013). El grupo colaborador de la Dra. Kami Kim marcó endógenamente a TgSIR2A-HA y encontró que esta se localizaría principalmente en el citoplasma y cerca de la periferia

nuclear. Por otra parte, mediante inmunoprecipitación de la cromatina seguida de hibridación de microarrays (ChIP-ChIP), lograron localizar a TgSIR2A distribuida por todo el genoma, pero, observaron un mayor enriquecimiento de la misma en la región centromérica (región heterocromática silenciada), esto se confirmó mediante co-localización con CenH3 (Dra. Kami Kim, comunicación personal). Además, observaron que la línea *knockout*  $\Delta$ SIR2A crecería más rápido al igual que lo observado para  $\Delta$ PfSIR2A en el estadio intraeritrocítico (Mancio-Silva et al. 2013), y que durante la diferenciación a bradizoíto, los quistes se formarían más rápidamente en dichos parásitos, reflejando un fenotipo relacionado con la replicación del parásito (Nardelli, comunicación personal). PfSIR2A deacetilaría H3 y removería H2A.Z de los sitios activos de transcripción de los genes *var*. Mientras que en parásitos deficientes de PfSIR2A ( $\Delta$ SIR2A), H2A.Z quedaría relocalizada en los TSS activos y no así, en parásitos deficientes de PfSIR2B ( $\Delta$ SIR2B) (Petter et al. 2011). Los genes *var* regulados por PfSIR2A ocupan el límite de heterocromatina/eucromatina tanto en los grupos cromosómicos internos como subteloméricos; por el contrario, PfSIR2B regula la mayoría de los genes *var* de tipo *UpsB* proximales a los telómeros separados del resto del cromosoma por los genes *var UpsA* subteloméricos regulados por PfSIR2A (Duraisingh et al. 2005; Tonkin et al. 2009). Ahora bien, en el análisis de localización de TgSIR2B vimos que ocuparía regiones en el núcleo ligando a esta deacetilasa con un probable blanco que regule la acetilación de H4H16 y H3. Habría que hacer más experimentos para confirmar esta posibilidad.

Sabemos que la estructura de la cromatina, particularmente la heterocromatina, es una estructura dinámica y polifacética que podría estar desempeñando un rol muy importante en el proceso de replicación de los parásitos Apicomplexas, articulando dicho proceso con la regulación de la transcripción génica, proporcionando un ambiente subnuclear compartimentalizado y adecuado asegurando, además, la correcta expresión de genes específicos durante todo el transcurso del ciclo celular. Si bien este trabajo no logró establecer una relación directa de las proteínas TgSIR2A y TgSIR2B con las regiones TgSTs creemos que futuros estudios genómicos podrían dar luz sobre este tema e ir ampliando nuestro conocimiento sobre la dinámica de la cromatina en las regiones TgSTs.

## 21. Conclusiones del capítulo III

Las regiones TgSTs podrían estar asociadas a proteínas involucradas en el mecanismo de posición telomérica dirigido principalmente por las Sirtuinas de clase III, SIR2. Las proteínas TgSIR2A y TgSIR2B localizarían en el citoplasma de taquizoítos intracelulares, aunque TgSIR2B también ocuparía puntos específicos dentro del núcleo. El RSV, activador de SIR2, posee actividad antireplicativa contra *T. gondii*, afectando de manera significativa los niveles de PTMs específicas sobre H3K9/14 y H4K16 (posibles blancos de SIR2). El RSV perturba el ciclo celular y causa daño sobre el ADN de los taquizoítos. Por lo tanto, el RSV podría considerarse como posible tratamiento contra la toxoplasmosis durante la fase aguda de la infección.

## 22. Conclusiones finales

Se detectó la longitud y estructura de los STs de *T. gondii* en cada extremo cromosómico en función de varios factores: densidad génica, estado de la cromatina y presencia de ADN satélite.

Se definieron 26 STs que oscilan entre 7,8 y 232,4 Kpb, con un promedio de 74,4 Kpb, lo que nos permite generar un modelo representativo subtelomérico de *T. gondii* (Fig. 63).

El análisis detallado del enriquecimiento de H2A.X, H3.3, H3K9me3 y H4K20me3 en toda la región subtelomérica respalda los resultados obtenidos anteriormente por Nardelli y col., (2022).

Se detectó la presencia de H3K4me3 en los extremos cromosómicos solo o con H2A.X y H3.3. En casi todos los casos se detectó solo en los extremos cromosómicos en los que se detectaron repeticiones teloméricas, lo que sugiere una estrecha relación de esta marca con los telómeros y la transcripción de lncRNA TERRA.

Se observó la presencia de ADN tándem repetitivo, compatible con ADN satélite, en la mayoría de los TgSTs. En general, el elemento repetitivo más consistente como ADN satelital corresponde a la familia sat350.

La marca H3.3 acompaña a la marca H2A.X en los TgSTs como así también las PTMs H4K31me1 y H3K4me1, particularmente en la mayoría de TSS de genes STs inactivos.

Las marcas H4K31me1 y H3K4me1 ocupan putativas regiones compatibles con lncRNAs subTERRA.

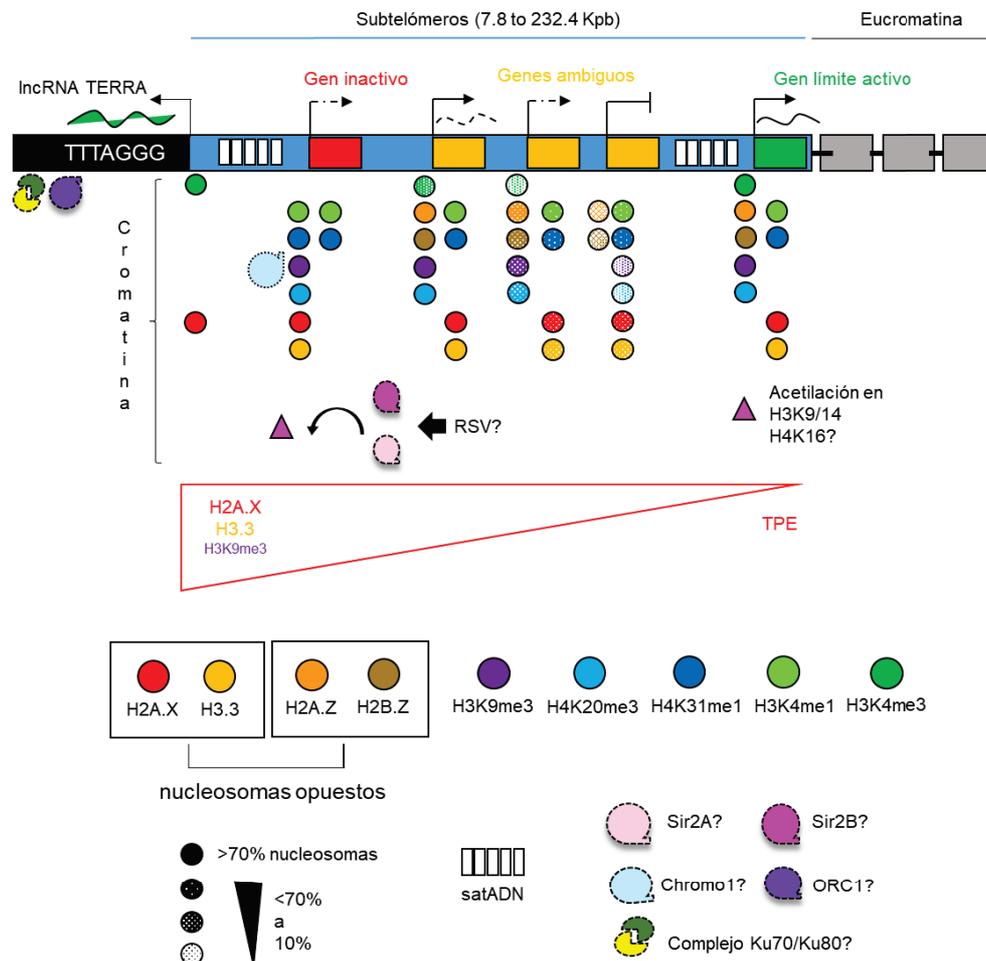
Los TgSTs abarcan uno o más genes en sus territorios. Un total de 96 genes pudieron vincularse a ellos, un gran número asociado con cromatina inactiva y evidencia de transcripción baja o nula, como se esperaba para los genes bajo TPE. Setenta de estos genes corresponden a genes únicos, muy pocos con anotaciones funcionales relacionados con rutas metabólicas. El resto de los genes detectados se corresponden con dos familias multigénicas integrales de membrana FamB y FamC.

Los TgSTs también poseen genes en estado activo y ambiguo. Los promotores de los genes activos están enriquecidos en H3K4me3/H2A.Z/H2B.Z. Sin embargo, a veces los genes silenciados o

pobremente expresados también fueron detectados asociados al dímero H2A.Z/H2B.Z sin H3K4me3 en un aparente estado de cromatina ambigua.

Tanto los genes *famb* como *famc* mostraron evidencia de transcripción, aunque a niveles de percentil más bajos que otros genes. La expresión de TgC8 se confirmó por WB e IFI. TgC8 mostraría una marca apical pero no en todos los taquizoítos.

**Fig. 63. Modelo representativo de las características de los subtelómeros de *T. gondii*.** Se muestra la cromatina asociada con el efecto de posición telomérica, TPE, los promotores de genes activos, ambiguos y silenciados y el cuerpo de los mismos. El triángulo rojo indica la descomposición del TPE hacia la eucromatina. Los círculos con diferentes grados de relleno de color se asocian con el porcentaje relativo en que se observaron en esa posición. Se muestran las posibles proteínas interactoras listadas en la **Tabla 12**. Las proteínas hipotéticas son representadas con líneas puntuadas. Las preguntas son abiertas debido a que estas proteínas todavía no se han confirmado. El, RSV, activador de Sir2 podría estar actuando sobre alguna o las dos sirtuinas de *T. gondii*.



La respuesta humoral frente a TgC8 y TgC6 fue baja en absorbancia en el modelo murino y humano, así como en el número de sueros individuales con toxoplasmosis reactiva.

Los STs de *T. gondii* podrían estar asociados a un mecanismo de posición telómerica dirigido principalmente por las deacetilasas TgSIR2s entre otras. Ambas Sirtuinas, TgSIR2A y TgSIR2B presentarían localización citoplasmática, aunque TgSIR2B ocuparía puntos específicos dentro del núcleo.

El RSV, activador de SIR2, posee actividad antireplicativa contra *T. gondii*, afectando de manera significativa los niveles de PTMs blancos de SIR2, perturbando el ciclo celular y causando daño sobre el ADN de los taquizoítos. Por lo tanto, el RSV podría considerarse como posible tratamiento contra la toxoplasmosis durante la fase aguda de la infección.

### 23. Perspectivas futuras

Dadas las conclusiones abordadas por este trabajo, se propone seguir indagando en los siguientes aspectos:

Ampliar el estudio en cuanto a la conformación estructural y funcional de la cromatina subtelomérica, involucrando aspectos de ocupación en territorios nucleares específicos, mecanismos epigenéticos asociados a ARNs no codificantes como así también la presencia de sitios de origen de replicación y/o reparación del ADN y su asociación al ciclo celular.

Evaluar el rol de los ADN satelitales en cuanto a plasticidad y evolución del genoma de *T. gondii*.

Profundizar en el estudio de la estructura y función de genes STs con valor fenotípico adaptativo ya sea como blancos terapéuticos durante el ciclo lítico o como moléculas inmunogénicas.

Establecer colaboración con la Dra Kami Kim para evaluación de líneas mutantes de proteínas TgSIR2s y posicionamiento de las PTMs blancos en relación al ciclo lítico.

Desarrollar el tratamiento de la fase aguda de la infección de *T. gondii* con RVS en modelos murinos e *in vitro* en asociación sinérgica con otras drogas.

## CAPÍTULO IV | METODOLOGÍA

## 24. Análisis de la cromatina

Los datos de ChiP-Seq fueron obtenidos de la base de datos NCBI ([https://bit.ly/Nardelli\\_GEO](https://bit.ly/Nardelli_GEO), [https://bit.ly/Sindikubwabo\\_GEO](https://bit.ly/Sindikubwabo_GEO)). Las lecturas del posicionamiento genómico fueron alineadas a la versión 53 del genoma de referencia de la cepa ME49 en colaboración con la Dra. Josefina Ocampo. Para el análisis de los datos procesados se utilizó el programa IGV versión 2.4 (<https://software.broadinstitute.org/software/igv/>) y se establecieron los rangos de datos de manera arbitraria con valores de 0 a 5 (Robinson et al. 2011). Se visualizó el enriquecimiento de CenH3 y se contrastó con las localizaciones descritas para los centrómeros según ToxoDB. Se visualizó el enriquecimiento de H2A.X, H3.3, H3K9me3, H4K20me3, H4K31me1, H3K4me1, H3K4me3, H2A.Z y H2B.Z junto a la presencia de genes en ambos extremos cromosómicos sobre las posiciones genómicas descritas para TgTAS por Dalmaso y col., (2014). Se constató la presencia de genes *Tsf* previamente descritos (Dalmaso et al. 2014) y se procedió a la redefinición de las regiones subteloméricas. Con las posiciones subteloméricas definidas se descargaron de la base de datos ToxoDB las regiones correspondientes desde la posición 1 hasta el límite indicado en el brazo izquierdo, y desde la posición interna del lado derecho hasta el final cromosómico para luego analizarse mediante *SnapGene* versión 2.3.2 en búsqueda de las repeticiones teloméricas TTTAGGG.

## 25. Definición de LncRNAs

El análisis de la posición de los lncRNAs en los STs de *T. gondii* se estableció por los parámetros establecidos por Sindikudwabo y col., (2017): presencia de marcas de cromatina activa: H4K31ac, H3K14ac y H3K4me3 en la región TSS y H3K4me1/H4K31me1 a lo largo de la longitud transcrita en una región no génica (CGen). En este análisis dichas marcas se correspondieron también con regiones TSS de genes con cromatina activa (H3K4me3/H2A.Z/H2B.Z). Se tomaron como lncRNA de TgSTs putativos todas aquellas regiones por fuera de los CGen enriquecidas en H3K4me3/H4K31me1 en las cuales se encontro de manera cercana a alguno de los extremos, una región con enriquecimiento de marcadores de cromatina activa. La longitud aproximada de cada lncRNA fue establecida desde el

inicio en la región TSS con cromatina activa mas cercana hasta el final determinado por el enriquecimiento de las marcas H3K4me3/H4K31me1.

## 26. Análisis bioinformático

Las búsquedas en base de datos, el análisis de comparación de secuencias y el análisis de los dominios putativos se realizaron utilizando los programas BLAST-N, BLAST-X del NCBI ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) y UniPROBE (<https://bit.ly/UniPROBE>). La búsqueda de proteínas ortólogas en organismos Apicomplexa se realizó utilizando las bases de datos disponibles en VEUPATH (<https://bit.ly/VEUPATHDB>): PlasmoDB (<https://bit.ly/PLASMODB>) y ToxoDB (<https://bit.ly/TOXODB>). La búsqueda y localización de motivos se realizó con el programa *Motifcan* (<https://bit.ly/MOTIFSCAN>), y de dominios funcionales se realizó utilizando las bases de datos de dominios conservados del NCBI: CDD (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>), J-Pred (<http://www.compbio.dundee.ac.uk/jpred/>) y ToxoDB. La búsqueda de péptido señal, sitios de clivaje y los dominios transmembrana fueron predichos por PSORT II (<https://bit.ly/PSORTII>), la probabilidad del péptido señal se confirmó por SignalP 5.0 (<https://bit.ly/SIGNALP>). Las regiones de ADN repetitivas se determinaron mediante análisis de diagrama de puntos, Dot plot, en la página web de YASS (<https://bioinfo.lifl.fr/yass/yass.php>). El análisis se extendió a los primeros 100 Kbp, excepto en los casos donde el borde de los subtelómeros propuestos se extendiera más allá de los 100 Kbp. La identidad de los satélites de ADN se realizó mediante una combinación del uso del programa YASS y BLAST-N (<https://blast.ncbi.nlm.nih.gov/>). Los análisis de densidad génica se realizaron en ToxoDB, mediante un escaneo manual del *Genome Browser*. También de forma manual, se determinó la sintenia de los genes detectados en la cepa ME49 en las 14 cepas restantes, incluida ME49 en la sección de ToxoDB, ortología y sintenia. Todos los análisis se realizaron con los parámetros por defecto.

El alineamiento múltiple de secuencias de aminoácidos se realizó utilizando los programas MEGA-X (*Molecular Evolutionary Genetics Analysis Software Version X*) (Kumar et al. 2018) y el editor de alineamiento de secuencias *BioEdit* (<https://bioedit.software.informer.com/7.2/>) mediante el

alineamiento múltiple Clustal W. El análisis de secuencias de ADN y el diseño de los oligonucleótidos se realizaron con el programa *Snapgene* versión 2.3.2 (<https://www.snapgene.com/>).

La búsqueda de asociación a factores virulentos se realizó utilizando la base de datos ProtVirDB (<https://bit.ly/PROTVIRDB>) utilizando la herramienta ViroBLASTDB dentro de la misma con los parámetros por defecto; junto a la base de datos VEUPATH para la constatación de los números de identificación y anotación de las proteínas asociadas con *score* más alto, *E-value* más pequeño y porcentaje de identidad más alto.

## 27. Análisis evolutivo

Se construyó un árbol filogenético con MEGA-X. El soporte interno se midió usando 500-1000 réplicas de la opción de arranque de búsqueda heurística. La historia evolutiva se infirió utilizando el método de Máxima Verosimilitud. El árbol se condensó (valor de corte: 50%). Las secuencias de aminoácidos se alinearon mediante alineación múltiple Clustal W (penalización por apertura de *gap* 10,00, penalización por extensión de *gap* 0,10, penalización por apertura de *gap* de alineación múltiple 10,00, penalización por extensión de *gap* 0,20, uso de matriz negativa desactivada, retraso de corte divergente 30 %).

## 28. Cultivo de parásitos y reactivos

Los experimentos fueron realizados utilizando dos cepas salvajes de *T. gondii*: RH y PRU y respectivas cepas parentales derivadas: RH $\Delta$ *hxgp**prt*, RH\_RFP. Para la obtención de taquizoítos, todas las cepas fueron cultivadas *in vitro* en condiciones estándares: se infectaron células HFF crecidas en monocapa y se incubaron con medio DMEM (Gibco) suplementado con 10% o 1% de suero fetal bovino (SFB), penicilina (100 UI/ml; Gibco) y estreptomycin (100  $\mu$ g/ml; Gibco) a 37°C y 5% de atmósfera de CO<sub>2</sub>.

Los parásitos RH\_RFP expresan de manera estable y constitutiva la proteína fluorescente roja en el citoplasma. Los parásitos RH $\Delta$ *hxgp**prt* son deficientes en la enzima Hipoxantina-Xantina-

Guanina-Fosforibosil-Transferasa. Esta enzima permite seleccionar positiva y/o negativamente los parásitos (ver sección “Sobreexpresión de proteínas en *Toxoplasma*” y Donald et al. 1996).

Los taquizoítos extracelulares fueron recolectados frescos del medio de cultivo. Las células conteniendo los taquizoítos intracelulares fueron tripsinadas y luego pasadas a través de agujas de distinto calibre para poder romperlas y obtener los parásitos. Una vez obtenidos fueron purificados de los restos celulares por filtros de 3 mm de poro (*Nucleopore*), contados, centrifugados durante 10 minutos a 2000 rpm y lavados con solución de fosfato salino, PBS 1%.

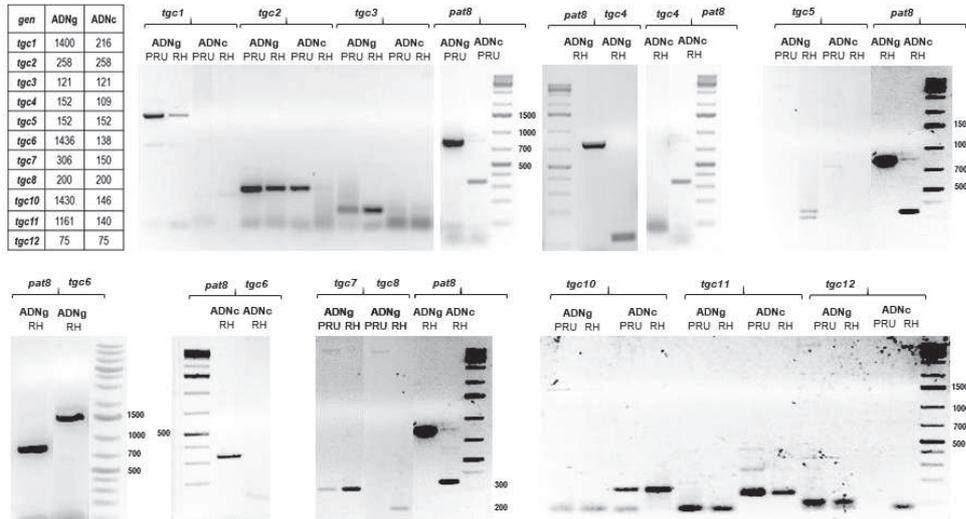
La solución madre de Resveratrol se almacenó a 10 mg/ml y la solución de Sirtinol a 10 M, ambas diluidas en DMSO.

## 29. Análisis de la expresión de genes por RT-PCR y qPCR

Para esto se diseñaron oligonucleótidos específicos para amplificar productos codificantes pequeños de entre 100 y 200 pb de cada uno de los genes. Estos oligonucleótidos utilizados para realizar las RT-PCR, fueron originalmente diseñados para ser utilizados en análisis por qPCR. Dichos oligonucleótidos fueron diseñados teniendo en cuenta que los mismos tuvieran una  $T^{\circ}$  de *melting* de  $\sim 60^{\circ}\text{C}$ , que pegaran cerca de la región 3' del gen y que cada par de oligonucleótidos amplificaran fragmentos entre dos exones y un intrón en el caso de los genes que poseen intrones. En algunos casos no fue posible diseñar sobre la unión exón-exón a pesar de tener intrones debido a la similitud de las secuencias, es por esto que en algunos genes como *tgc3*, *tgc8* y *tgc12* se amplificaron secuencias de igual tamaño entre ADNg y ADNc. Mientras que en el caso de *tgc2* y *tgc5*, los cuales no poseen secuencias intrónicas, también amplificamos ADNg y ADNc del mismo tamaño. Por dicho motivo realizamos controles para asegurar la no existencia de contaminación genómica en el preparado de ADNc (*Fig.64*). En la *Fig.64* se muestra el detalle de amplificación y especificidad de cada uno de los oligonucleótidos testeados con ADNg y ADNc.

**Fig.64. Prueba de especificidad de oligonucleótidos diseñados para amplificar los genes *tgc*.**

Se muestra la tabla con detalle del tamaño esperado para cada gen en pb y por otra parte las amplificaciones esperadas en el caso de ADNg y/o ADNc. En algunos casos incluimos muestras provenientes de parásitos de la cepa RH y PRU, gel de agarosa 1,5%. Control interno de amplificación *pat8*: *palmitoil transferasa 8*, ADNg ~800 pb, ADNc ~350 pb.



Todos los oligonucleótidos fueron testados *in silico* para determinar su especificidad ya que estos genes tienen regiones muy conservadas entre sí. Se cultivaron parásitos extracelulares ( $1.10^7$  por muestra) para obtener el ARNm total mediante la utilización del protocolo establecido para el reactivo Trizol® (Invitrogen), luego se realizó la retro-transcripción de los mismos utilizando la retrotranscriptasa reversa M-MLV® (Invitrogen) según protocolo del fabricante para obtener el ADNc. Las muestras fueron tratadas con DNasa® (Invitrogen) para evitar posibles contaminaciones con ADNg. A continuación, se realizó en primer lugar las PCR a punto final, RT-PCR, probando los oligonucleótidos con muestras de ADNg y ADNc, en algunos casos de las cepas PRU y RH, incluyendo además un control de amplificación de ADNg *pat8*; dicho gen amplifica una región de ~800 pb en presencia de ADNg y una región de ~350 pb en el caso de estar libre del mismo (**Fig.64**).

Luego se probaron en PCR a tiempo real, qPCR, para las cuantificaciones. Las reacciones de amplificación se llevaron a cabo en un volumen final de 10 µl conteniendo SYBR Green PCR Master Mix (Applied Biosystems, CA), 0,5 µM de cada oligonucleótido específico y 1 µl de una dilución 1:10 del

ADNc. Los oligonucleótidos utilizados para amplificar son los mismos utilizados para la RT-PCR (*Tabla 14*). Como control interno se utilizaron oligonucleótidos que amplifican una región de 101 pb de la región promotora del gen constitutivo *b-tubulina*. Como controles de transcripción estadio específica se analizó el nivel de expresión del gen específico de taquizoito *sag1* y del gen específico de bradizoito *bag1*. Las secuencias de cada uno de los oligonucleótidos se encuentran en la *Tabla 14*, el amplicon de *sag1* se corresponde con un fragmento de 93 pb de la región promotora del mismo, mientras que el amplicon de *bag1* se corresponde con un fragmento 101 pb de región promotora del gen. de Todas las reacciones se realizaron en un ciclador StepOnePlus™ Real-Time PCR System (Applied Biosystems™).

Para cuantificar la expresión génica se utilizó el método comparativo  $2^{-\Delta\Delta Ct}$  (Schmittgen & Livak 2008). Donde se tomó como parámetro de medición el Ct (*threshold cycle*), el cual se define como el ciclo de la PCR donde la señal de fluorescencia emitida por el colorante reportero supera el umbral establecido. Este método comparativo asume que la eficiencia de la PCR es cercana a 1 y que la eficiencia de la PCR de los genes blancos es similar a la eficiencia de la PCR del gen control interno. Para esto calculamos las eficiencias de las PCRs para todos los genes y encontramos que todas eran cercanas a 1. Además, se realizó una curva de disociación para cada par de oligonuclétidos, confirmando que éstos amplificaban un único producto. Los niveles de ARNm de cada gen (medidos como Ct) fueron medidos en 3 muestras independientes para cada condición, y la reacción de qPCR fue llevada a cabo por duplicado. Los valores de Ct fueron promediados y luego fueron normalizados a los niveles de *b-tubulina*. Los datos de Ct se colectaron con el programa del equipo, se analizaron en Microsoft Excel y luego se compararon con el programa InfoStat (<http://www.infostat.com.ar>). El nivel de expresión de cada uno de los genes en las condiciones de estrés se comparó a su respectiva muestra control en condiciones normales de crecimiento.

Nombre	Sentido 5'-3'	Nombre	Anti-sentido 5'-3'	Eficiencia
qPCRtsf1	GGTACCGCGGAAGTATACGA	qPCRtsf1	TCCTTGCAGGCCTTGACTT	83.1
qPCRtsf2	GTTGTGCTCCCTCTGATGGT	qPCRtsf2	GCACTACGGTTGAAGGAAGC	92.7
qPCRtsf3	TACAGAAGTATACGGCGGTGC	qPCR_tsf3	CAAACGGACTGAGTCGCATA	95.8
qPCRtsf4	GCCTTGTGCTTTCTGCTC	qPCRtsf4	CCAGTTGTGCTTCTCGTGAA	84.6
qPCRtsf5	TCTGCCTGATGGTCTTGC	qPCRtsf5	GGACAGACACAGAAGACGGC	76.6
qPCRtsf6	CGAATCATCTCAAGAACTAAGCAC	qPCRtsf6	CTATCGTCTCGTCGCCAT	91.9
qPCRtsf7	GCGTGATACAAATGCCGAAAC	qPCRtsf7	CGATAATGAACGATGAACGATGC	79.8
qPCRtsf8	GCTGTTCCCTTCGTGACTTG	qPCRtsf8	CGAAGAGGATGGCGATACTGAAG	70
qPCRtsf10	GCGAGCACCGAGCAGAAG	qPCRtsf10	GCCATCTTTCTCAACTCGGACA	93.6
qPCRtsf11	GGTGTTCTCCTTTCTGTTCTTG	qPCRtsf11	GGAGAGGAAGCAATCAACCA	89.9
qPCRtsf12	CGGTGCGTATCGTGTGATC	qPCRtsf12	ACTATGTCGTATCGTCAGAATCC	85.7
qSAG1_F	AACTTCCCACACGAGGCATT	q-SAG1_R	GATCGGCCCTAATTCAGC	82.1
qBAG1_F	TCCTCCCCTGGATCTTCTCC	q-BAG1_R	TTGCACAAAACGGCAAAGG	81.3
qTUB_F	TCCTTTTCTTCATGCACTCGC	q-TUB_R	GTCGTTCGAATTCTCTGCC	81.6
rTSF5BamHI	GGATCCGTTTCTTCGAGTTTCTACACCCC	rTSF5HindIII	AAGCTTCACAGTCCAGAGGCCACTATTG	(-)
rTSF8BamHI	GGATCCTGTTCCACATATCGGC	rTSF8HindIII	AAGCTTTTAAGGACGAGCAGC	(-)
rSir2AXhol	CTCGAGGAGCCTCATGAACAGTTTCTCG	rSir2AHindIII	TAAGAAGCTTCTATCTCCTTGCAGCGC	(-)

**Tabla 14. Oligonucleótidos utilizados para la amplificación de fragmentos de ADN.**

## 30. Obtención de proteínas recombinantes

### 30.1 Clonado de proteínas

#### 30.1.1. Proteínas TgCs

##### 30.1.1.1 TgC3

Para la construcción de la proteína rTgC3 decidimos deletar todo el extremo N-terminal hasta la alanina en posición 46 con parte de la región del primer dominio TM. Luego, unimos las dos regiones no hidrofóbicas desde el aminoácido 46 hasta la serina en posición 151 y desde la asparagina en el sitio 169 al aminoácido final 206 en un transcripto único para producir el péptido recombinante. Este transcripto de 432 pb se clonó entre los sitios NdeI y XhoI dentro del plásmido comercial pET14b (*Genescrip Co.,Ltd.*) de 4671 pb. Dicho transcripto codificaría 144 aminoácidos, con un punto isoeléctrico, PI de 6,11 y un peso estimado de ~16 KDa, calculados con la herramienta ExPASy [pi//Mw](https://web.expasy.org/compute_pi/) ([https://web.expasy.org/compute\\_pi/](https://web.expasy.org/compute_pi/)). El péptido recombinante final con la etiqueta de 6xHis más el ATG del plásmido de expresión pET14b estaría codificado por un transcripto de 495 pb y 164 aa con un peso estimado de ~17,4 KDa y posee un PI de 7,2 (*Fig.65A*).

#### **30.1.1.2 TgC5**

Al igual que para la construcción de rTgC3, en rTgC5 obviamos la región N-terminal con péptido señal y dominio TM. La secuencia codificante de la proteína recombinante quedo comprendida desde de la serina 60 hasta el aminoácido final, quedando un amplicón correspondiente a la región no hidrofóbica de 675 pb. Esta región se amplificó mediante oligonucleótidos específicos conteniendo los sitios BamHI e HindIII para posterior inserción en vector de clonado pGEM<sup>®</sup>T ([https://bit.ly/pGEMT\\_easy](https://bit.ly/pGEMT_easy)) y luego al vector de expresión o de salida pRSET-A de 2897 pb (Invitrogen, Cat. no. V351-20) (*Fig.65B*). El amplicón fue resuelto y extraído de geles de agarosa 1% para luego ser purificado. Se introdujo al vector de clonado, se secuenció y una vez confirmada la secuencia fue clonada en el vector de expresión. La región clonada codificaría una proteína de 225 aminoácidos de aproximadamente 22,73 KDa y con un PI de 3,74. Mientras que la proteína recombinante final, rTgC5, con la etiqueta de histidina y el aporte plasmídico estaría codificada por un transcrito de 783 pb y 261 aa con un peso estimado de 26,8 KDa y un PI de 4,02.

#### **30.1.1.3 TgC6**

Para rTgC6, al igual que en los casos anteriores, deletamos toda la región N-terminal conteniendo el péptido señal y el primer dominio TM hasta la leucina 49, unimos las dos regiones no hidrofóbicas desde el sitio 49 hasta la prolina 147 y desde la alanina 165 al aminoácido 201 en único transcrito para producir el péptido recombinante. El fragmento se clonó entre los sitios NdeI y XhoI en el vector pET14b de manera comercial (*Fig.65C*). Este transcrito de 411 pb, codificante de 137 aminoácidos tendría un PI de 5,3 y un peso estimado de 15,1 KDa. El péptido recombinante final, rTgC6, tendría desde el ATG del plásmido, incluyendo el 6xHis un tamaño de 474 pb, estaría codificado por 157 aa con un peso estimado de 16,6 KDa y un PI de 6,75.

#### **30.1.1.4 TgC8**

En el caso de rTgC8 también decidimos clonar la región no transmembrana comprendida entre estas dos regiones hidrofóbicas transmembrana a partir de la cisteína 75 hasta la prolina 265,

quedando así un amplicon de 573 pb. Esta región se amplificó mediante oligonucleótidos específicos conteniendo los sitios BamHI e HindIII para posterior inserción en vector de clonado pGEM®T y luego al vector de expresión pRSET-A (**Fig.65D**). El amplicón fue resuelto y extraído de geles de agarosa 1% para luego ser purificado. Se introdujo al vector de clonado, se secuenció y una vez confirmada la secuencia fue clonada en el vector de expresión. Al clonarlo en pRSET-A se deleccionaron 35 pb del plásmido previos al sitio de corte HindIII (construcción total 3435 pb). Este amplicon codificaría una proteína de 191 aminoácidos y de aproximadamente 19,23 KDa y un PI de 3,84. Mientras que la proteína recombinante final, rTgC8, tendría desde el ATG del plásmido, incluyendo el 6xHis hasta el codón terminador del plásmido un tamaño de 750 pb, estaría codificada por 249 aa y presentaría un peso estimado de 25,6 KDa con un PI de 4,06.

Las secuencias recombinantes de TgC5 y TgC8 fueron amplificadas mediante oligonucleótidos específicos (**Tabla 14**), para luego introducirlas mediante el uso de enzimas específicas en el vector de clonado pGEMT-easy y luego al vector de expresión pRSET-A. Mientras que TgC3 y TgC6 fueron directamente sintetizadas en el vector comercial de expresión pET14b (*Genescrip Co.,Ltd.*). Las enzimas utilizadas se detallan en cada caso a continuación: rTgC3, NheI y XhoI; rTgC5, BamHI y HindIII; rTgC6, NheI y XhoI; rTgC8, BamHI y HindIII (**Fig.65A-D**). Los fragmentos esperados para *tgc5* y *tgc8* se obtuvieron por PCR a partir de ADN genómico de la cepa PRU con las siguientes condiciones de reacción: 95°C 2', 95°C 30'', 55°C 30'', 72°C 20'', 72°C 5' durante 30 ciclos.

### **30.1.2 Sirtuinas**

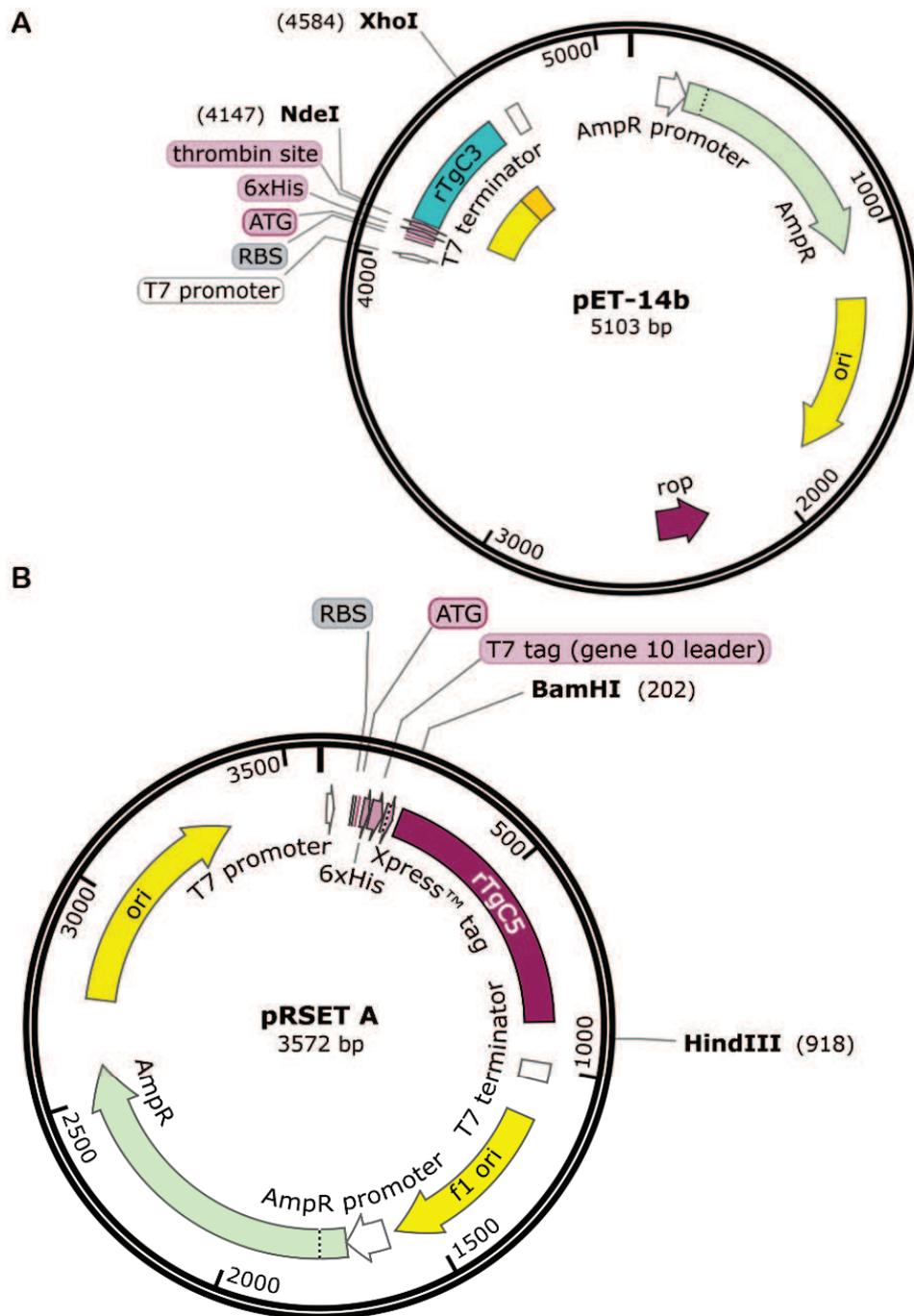
#### **30.1.2.1 TgSIR2A**

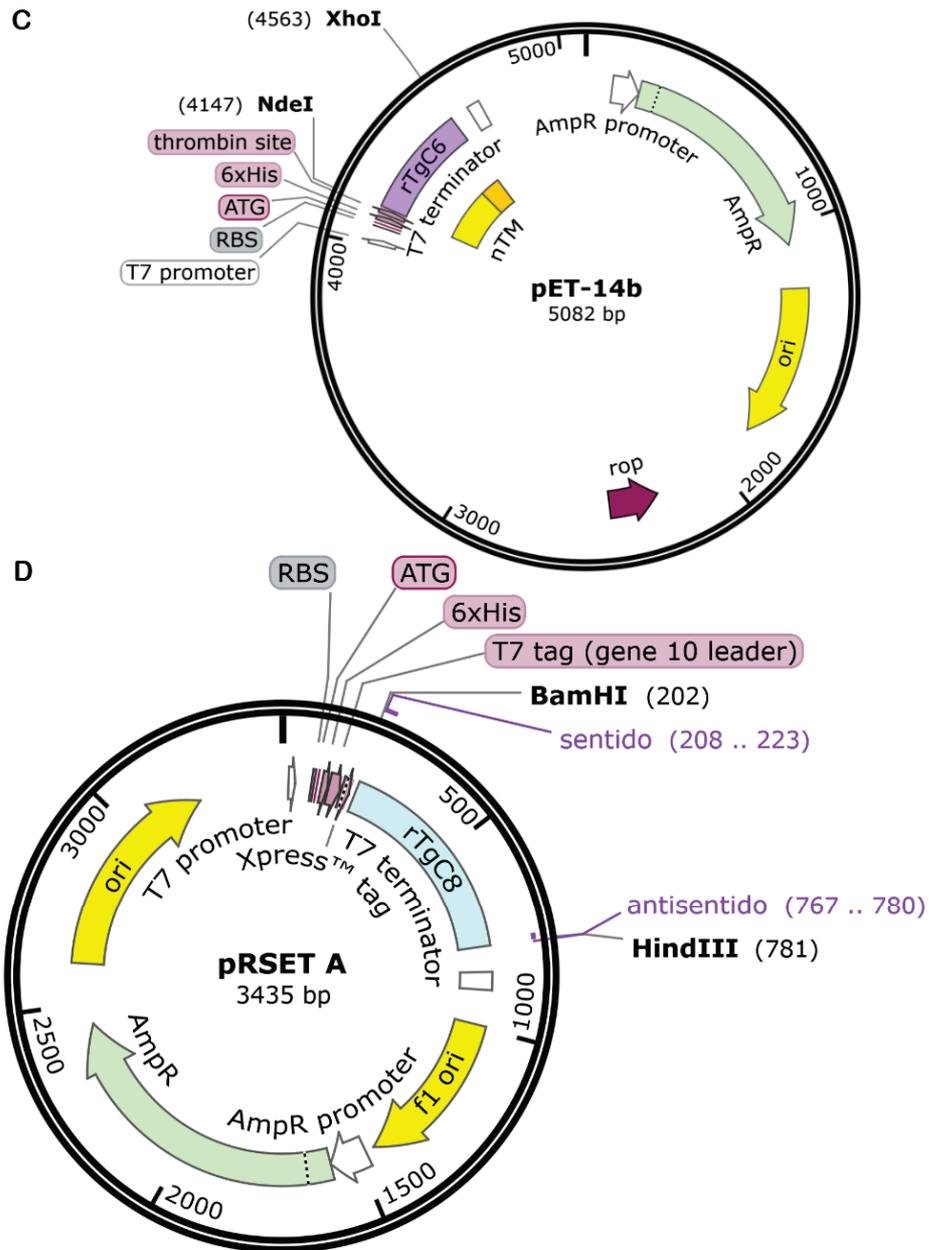
El gen *tgsir2a* (TGME49\_227020) tiene una secuencia genómica de 2633 pb incluidos los UTR 5' y 3'. Codificaría a TgSIR2A mediante un transcrito de 1083 pb (CDC, **Fig.66A**) generando una proteína de 360 aa con un peso estimado de 40 KDa y un PI de 8.5. Su dominio catalítico sirtuina (DC\_SIR2A) sería conservado y se encontraría entre los aminoácidos 115 y 360 (245 aa) (**Fig.66A**). Dicha proteína presenta un péptido señal, con un posible sitio de clivaje en el aminoácido 17 predicho por el

programa SignalP 5.0 (probabilidad 0.982). Debido a la presencia de esta posible región hidrofóbica en el extremo N-terminal, decidimos deleccionar la proteína TgSIR2A y generar una versión truncada (rTgSIR2A\_truncada) sin los primeros 18 aa codificada por los primeros 54 pb, quedando un amplicón de 1029 pb (**Fig.66B**). Para amplificar el constructo se diseñaron oligonucleótidos con los sitios de restricción específicos XhoI (oligonucleótido sentido 5'-3') e HindIII (oligonucleótido antisentido 5'- 3') (**Tabla 14**). Al cebador conteniendo el sitio de restricción HindIII se le agregaron unos nucleótidos al azar (TAAG) por fuera del sitio de restricción para permitir el correcto clonado dentro del vector de clonado pGEM®T. El fragmento esperado se amplificó por PCR mediante el uso de oligonucleótidos específicos a partir de ADN copia de la cepa PRU con las siguientes condiciones de reacción: 95°C 2', 95°C 30'', 55°C 30'', 72°C 50'', 72°C 7' durante 30 ciclos. El amplicón fue resuelto y extraído de geles de agarosa 1% para luego ser purificado. Se introdujo al vector de clonado, se secuenció y una vez confirmada la secuencia se clonó en el vector de expresión. (**Fig.66B**). La construcción final posee un tamaño de 1146 pb desde el ATG incluyendo la etiqueta 6xHis y codificaría para 381 aa de rSIR2A. El peso estimado del péptido recombinante de SIR2A (rTgSIR2A\_truncada) (**Fig.66B**) junto al ATG del plásmido y la etiqueta de histidinas tendría un peso teórico estimado de 42,5 KDa y un PI de 7.2.

Por otra parte, a la CDC de 1083 pb de TgSIR2A se le agregaron los sitios de restricción NsiI en el extremo N-terminal y PacI en el extremo C-terminal, N\_rTgSIR2A\_P, por fuera de la secuencia codificante para por un lado introducirlo entre los sitios XhoI e HindIII en el vector de expresión pRSET-A de manera comercial (*Genewiz Genomics Headquarters*) (**Fig.66A**) y por el otro permitir posterior subclonado en vector de expresión específico de *T. gondii* pTUB8mycGFPPftailTy-HX (ver sección 31). La secuencia recombinante N\_rTgSIR2A\_P junto al aporte del plásmido de expresión en bacterias pRSET-A tiene un tamaño de 1206 pb y codificarían una versión de 401 aa con un PI de 7,2 y un peso estimado de ~45 KDa cuya versión recombinante aún no fue ensayada. Mientras que la versión extra de la proteína etiquetada con *c-myc* tendría un tamaño aproximado de 42 KDa.

**Fig.65. Mapas plasmídicos de proteínas recombinantes TgCs producidas en este estudio. A.** pET14b-rTgC3. RBS (del inglés: *Ribosomal Binding Site*). ATG: inicio de la transcripción. *Thrombin site*, sitio de clivaje de la etiqueta de histidina. AmpR: gen codificante para resistencia a Ampicilina, selección del plásmido. Ori: pUC ori, permite la alta replicación de copias y crecimiento en *E. coli*. **B.** pRSETA-rTgC5. *T7 promoter* y *T7 terminator*, oligonucleótidos comerciales utilizados para la secuenciación del fragmento esperado. f1 ori: del inglés: *f1 origin of replication* para mutagénesis. **C.** pET14b-rTgC6. nTM: *no transmembrane domain*. **D.** pRSETA-rTgC8. Se muestran los sitios de pegado de los oligonucleótidos utilizados en la amplificación del fragmento esperado.





### 30.1.2.2 TgSIR2B

La proteína TgSIR2B (TGME49\_267360) estaría codificada por un transcripto de 6033 pb, posee 5' y 3' UTRs y su marco abierto de lectura estaría codificado por 5148 pb generando una proteína de 1715 aa con un peso estimado de ~187 KDa y un PI de 6.8. El dominio catalítico conservado sirtuina estaría entre los aminoácidos 176 y 522 (346aa). A diferencia de TgSIR2A, TgSIR2B no presenta péptido señal en su secuencia, además de presentar diferencias en cuanto al tamaño de la región C-terminal, como la observada en la *Fig.59*. En este caso, también realizamos una versión recombinante de TgSIR2B

pero dentro del plásmido comercial pET14b (*Genescrip Co.,Ltd.*) entre los sitios NheI y NdeI. Para realizar la versión recombinante decidimos clonar solamente el dominio funcional, con lo cual enviamos a sintetizar un fragmento correspondiente a esta región de 1041 pb. Por lo tanto, la proteína recombinante rSIR2B (DC\_rTgSIR2B) junto al ATG del plásmido, la etiqueta de histidinas y una región C-terminal del plásmido tendría un tamaño de 1147 pb (**Fig.66C**) y codificaría una proteína recombinante esperada de 390 aa con un peso estimado de 41,5 KDa y un PI de 9,1. La secuencia proteica se confirmó por espectrometría de masas.

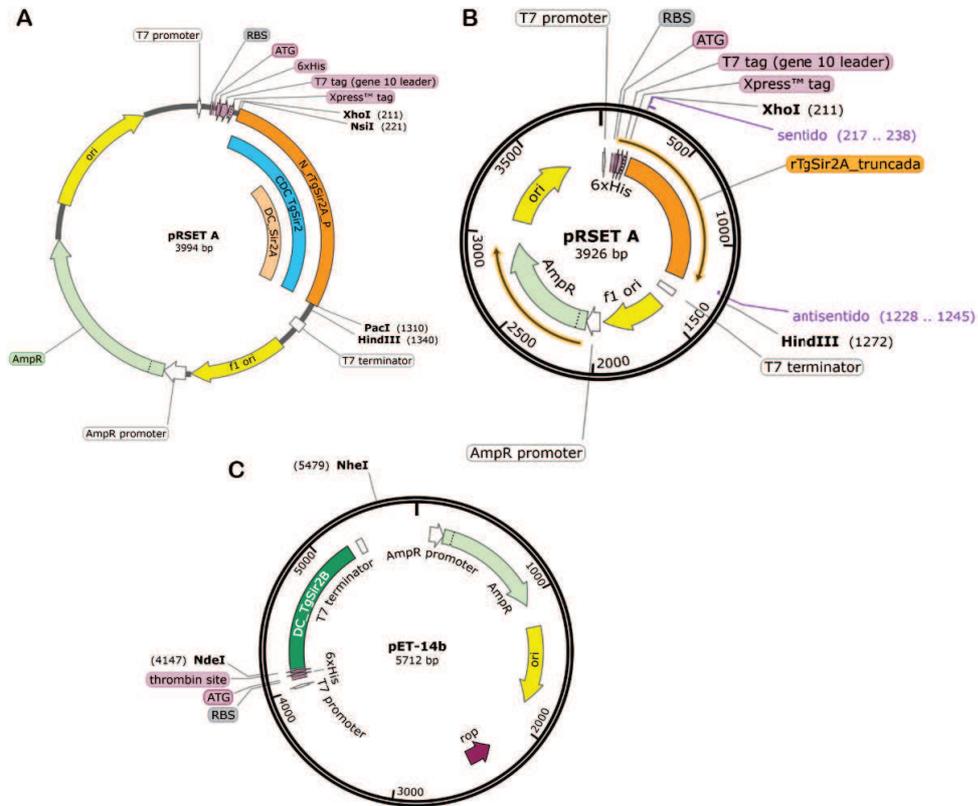
### **30.2. Expresión en bacterias**

Una vez obtenidas las construcciones se transformaron bacterias *Escherichia coli* competentes químicas BL21 (DE3) pLysS, las cuales fueron sometidas a selección por antibióticos, ampicilina y cloranfenicol y BL21 START en el caso de rSIR2B, la cual se seleccionó solo con ampicilina según recomendaciones de los fabricantes de los vectores expresión en cada caso. Una vez obtenidos los clones positivos se testearon mediante inducción con IPTG 1 mM a baja escala de cultivo para determinar la inducción proteica correspondiente a distintos tiempos. En todos los casos, se determinó que la expresión óptima era a 37°C y a las 4 hs post inducción con 1mM de IPTG.

### **30.3. Purificación de proteínas recombinantes**

Las proteínas se purificaron por afinidad en columna de níquel según instrucciones del fabricante en condiciones desnaturizantes mediante el uso de la etiqueta de histidina que favorece el intercambio y pegado a dicha columna. Se observó la presencia de todas las proteínas recombinantes en geles SDS-PAGE 12% (rSIR2A y rSIR2B) y 15% (rTgCs). Por otra parte, la purificación de rSIR2B fue más complicada, ya que intentamos purificarla en condiciones nativas sin éxito, luego en condiciones desnaturizantes, pero esta se iba a la fracción no soluble. Finalmente logramos purificarla en condiciones desnaturizantes, pero induciendo los cultivos a 18°C durante toda la noche con 0,5mM de IPTG.

**Fig.66. Mapas plasmídicos de proteínas recombinantes TgSir2s producidas en este estudio. A.** pRSETA-N\_rTgSir2A\_P. **B.** pRSETA-rTgSir2A\_truncada. Se muestran los sitios de pegado de los oligonucleótidos utilizados en la amplificación del fragmento esperado. **C.** pET14b-DC\_rTgSir2B.



### 31. Generación de anticuerpos policlonales

Con las proteínas purificadas se inmunizaron 2 ratones Balb-c machos por proteína, de aproximadamente 8 semanas cada uno. El esquema de inmunización se diagramo en 4 dosis de 15 µg de proteína recombinante, administrada por vía subcutánea cada 15 días en presencia de adyuvante completo de Freund las dos primeras dosis y adyuvante incompleto de Freund en las dos subsiguientes. Se extrajo el suero pre-inmune y se procedió a la dosificación del antígeno. En el caso de TgC6 se produjo el anticuerpo en conejo, inmunizando el animal con 3 dosis de proteína recombinante, una primera con 0,5 µg y adyuvante completo, y otras dos dosis de 0,25 µg de proteína recombinante, dosificadas en intervalos de 30 días. Transcurrido el periodo de inmunización se extrajo el suero total de los animales inmunizados, y se procedió a medir mediante ELISA indirecto la titulación de los anticuerpos

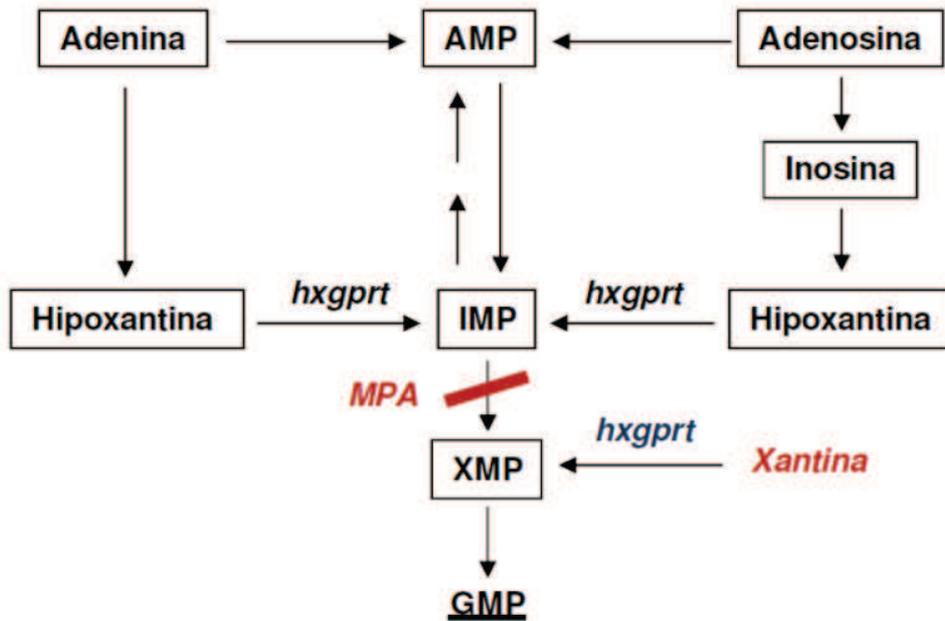
producidos. El título de los anticuerpos contra las proteínas rTgC superó los valores de 1:100000 y el título del anticuerpo contra rSIR2B fue de 1:32000. Se sacrificaron los ratones y se obtuvo el suero total de los mismos.

### 32. Sobreexpresión de proteínas en *Toxoplasma*

Para generar líneas mutantes que llevan una copia extra del gen en cuestión; en el laboratorio contamos con el vector de sobreexpresión pTUB8mycGFPPftailTy-HX específico de *Toxoplasma* el cual lleva una copia del gen *hxgprrt*, fusionado al promotor de *tubulina* (pTUB) el cual dirige su expresión y la línea parental de parásitos RH $\Delta$ *hxgprrt* deficiente para el mismo. Este sistema nos permite seleccionar parásitos de esta línea que hayan incorporado eficientemente el plásmido de sobreexpresión (Donald et al. 1996). En la **Fig.67** se muestra un diagrama de la vía de síntesis y su función en el proceso de selección. El ácido micofenólico es un potente inhibidor de la enzima IMP deshidrogenasa y por lo tanto bloquea la síntesis *de novo* de guanina a partir de AMP. Mediante la adición de xantina en el medio de cultivo, solo los parásitos que hayan incorporado el plásmido, conteniendo el gen de selección *hxgprrt*, serán capaces de sintetizar guanina a partir de xantina (**Fig.67**). Una vez obtenidas las poblaciones estables, se los debería clonar por dilución múltiple y aislar varios clones. Estos deben ser analizados por PCR, IFI y WB utilizando  $\alpha$ -*cmyc* o el anticuerpo disponible para cada caso.

Para llevar a cabo la sobreexpresión de *Tsf6* (RH $\Delta$ *hxgprrt-tgc6*<sup>(++)</sup>pTUB, se clonó de manera comercial la región codificante del gen (*Genescrip Co., Ltd*) entre los sitios Nsil-Pacl en el vector pTUB8mycGFPPftailTy-HX. Mientras que para *tgSir2a\_c-myc* se subclonó desde el vector pRSET-A con las enzimas correspondientes. Una vez obtenidas las construcciones se transformaron bacterias de clonado, se obtuvieron clones positivos y se enviaron a secuenciar. En el caso de *Tsf6* solo se transfectó y seleccionó a nivel poblacional la cepa parental pero no se llegó a la obtención clonal (**Fig. 46**, Capítulo II), mientras que para *tgSir2a\_c-myc* se obtuvo un clon estable (TgSIRA\_c-myc#22) (**Fig.60**, Capítulo III) seleccionado positivamente con xantina y ácido micofenólico.

**Fig.67. Vía de síntesis de guanosina monofosfato (GMP) y selección HXGPRT.** Se utilizan parásitos  $\Delta$  para *hxgprt*, los cuales son capaces de sintetizar GMP a través de la interconversión entre adenosina monofosfato (AMP) e inosina monofosfato (IMP). Al agregarse ácido micofenólico (MPA) se bloquea este paso, por lo que sólo los parásitos que hayan incorporado el plásmido que contiene *hxgprt* serán capaces de sintetizar GMP a través de la xantina agregada al medio. Tomado de (Tesis Doctoral Dalmasso MC, 2009)



### 33. Transfección de parásitos

Las construcciones realizadas en el plásmido pTUB8mycGFPPftailTy-HX fueron utilizadas para transfectar por electroporación parásitos  $RH\Delta hxgprt$ . Las electroporaciones se llevaron a cabo en un electroporador BTX EMC 630 utilizando  $2 \cdot 10^7$  taquizoítos con entre 30-50  $\mu\text{g}$  de ADN plasmídico a 1500 voltios,  $50\Omega$  y 25  $\mu\text{F}$ , todo resuspendido en CYTOMIX (120 mM KCl; 0,15mM  $\text{CaCl}_2$ ; 10 mM  $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$  pH: 7,6; 25 mM HEPES pH: 7,6; 2 mM EDTA pH: 7,6; 5 mM  $\text{MgCl}_2$ ; 3 mM ATP; 3 mM Glutación), en cubetas de 2 mm (BTX). Los parásitos electroporados fueron sometidos a selección después del primer pasaje, durante 3 o 4 ciclos utilizando 25  $\mu\text{g}/\text{ml}$  de ácido micofenólico y 50  $\mu\text{g}/\text{ml}$  de xantina como drogas de selección.

### 34. Ensayos de especificidad por WB

Para determinar la especificidad de los anticuerpos se ensayó con muestras de proteínas recombinantes purificadas y con extractos de proteínas totales. Se sembró por 1 µg de cada proteína recombinante en cada caso, y por otra parte un extracto total de proteínas totales provenientes de 1 a  $2 \cdot 10^7$  taquizoítos extracelulares para luego resolverlas en geles SDS-PAGE. Las proteínas fueron luego transferidas desde el gel de poliacrilamida a una membrana de PDVF durante una hora a voltaje de 80V durante los primeros 10 minutos y a 100V durante 50 minutos restantes para proteínas recombinantes, y 1 hora 30 minutos en el caso de proteínas totales de gran tamaño. Las membranas fueron incubadas con leche descremada 5% p/v en TBS-Twen (TBS-T 1%) por una hora, luego con los anticuerpos primarios generados en distintas diluciones, por una hora. Posteriormente se lavaron 3 veces con TBS-T 1% y un lavado con TBS 1% para finalmente incubar por una hora con el anticuerpo secundario anti-ratón o anti-conejo en una dilución de 1:30000. Se repitió el esquema de lavado y se procedió a revelar mediante método de detección colorimétrico con fosfatasa alcalina y NBT-BCIP según instrucciones del fabricante.

### 35. Ensayos de IFI

Con el fin de determinar la posible localización de las proteínas de interés se procedió a realizar ensayos de Inmuno Fluorescencia Indirecta, IFI. Para esto se crecieron las células hospedadoras HFF sobre vidrios en placas de 24 pocillos a confluencia, luego fueron infectadas con taquizoítos de la cepa RH y se dejó replicar a los taquizoítos intracelularmente por 24 hs más. Transcurrido el tiempo los vidrios fueron lavados y fijados con metanol durante 10 minutos en frío. En el caso de taquizoítos en el estadio extracelular, estos fueron colectados, contados, centrifugados y lavados con PBS 1%; luego fueron resuspendidos directamente en un volumen de 100 µl de paraformaldehído 4% y colocados sobre los vidrios en placas de 24 pocillos para permitir durante 10 minutos su fijación. En ambos casos los vidrios fueron lavados dos veces con PBS 1% transcurrida la fijación. Las células fijadas se permeabilizaron con 0,2% Tritón X-100 en PBS 1% por 10 minutos (con excepción de las proteínas TSF)

y se bloquearon con BSA 3% por 30 minutos. Luego se las incubó con la dilución conveniente de los anticuerpos primarios específicos diluidos en BSA 0,5% por 1 hora. Después de varios lavados con PBS 1%, se incubaron con el correspondiente anticuerpo secundario. Luego se repitió el esquema de lavado y los vidrios se montaron sobre cubreobjetos con medio de montaje (Mowbiol) conteniendo 10  $\mu\text{g}/\mu\text{l}$  de DAPI (*Southern Biotechnology Associates*) para marcar los núcleos. Algunos preparados fueron analizados por microscopía de epifluorescencia con un microscopio Nikon Eclipse E600 equipado con un objetivo de 100X y una cámara Nikon (Dc1qc) monocromática acoplada a un brazo Z y los colores emergentes de cada filtro fueron registrados separadamente y luego mezclados usando el programa ImageJ. Otros preparados fueron analizados con un microscopio ZEISS Modelo AxioObserver7, equipado con un objetivo 63X y una cámara de video digital ZEISS modelo 503 monocromática. Las imágenes fueron tomadas con el programa ZEN Pro 2.3 y analizadas con el programa Fiji (ImageJ) y Adobe Photoshop.

Los anticuerpos comerciales de inmunofluorescencia se diluyeron 1:500, en el caso de  $\alpha$ -H3ac (06-599B), o 1: 200 en el caso de  $\alpha$ -H4K16ac (ab109463). Los anticuerpos hechos en el laboratorio se utilizaron como se describe en cada caso, ver pie de figuras. Los anticuerpos secundarios Alexa flúor hechos en cabra se diluyeron 1: 4000,  $\alpha$ -mouse 488 (A11001) y  $\alpha$ -rabbit 595 (A11037) (*Invitrogen Life Technologies*).

### 36. Titulación de los anticuerpos

La titulación de los anticuerpos producidos en este trabajo se midió mediante ensayos de ELISA. Para ello, se sensibilizaron placas de 96 pocillos (*Immuno Plate Maxisorp, Nunc*) con 100  $\mu\text{l}$ /pocillo de una solución con 5  $\mu\text{g}/\text{mL}$  de proteínas recombinantes diluidas en *buffer*  $\text{NaHCO}_3$  50 mM pH 9,5 y se incubaron durante toda la noche a temperatura ambiente. Al día siguiente, las placas se bloquearon con solución de bloqueo (PBS-Twen + leche descremada 5 %) durante 1 hora a 37°C. Luego, se agregaron los anticuerpos primarios producidos en un volumen de 100  $\mu\text{l}$ /pocillo diluidos serialmente desde 1:100 hasta ~1:100000 en solución de bloqueo, como anticuerpo secundario se empleó el

anticuerpo monoclonal anti-IgG total de ratón y o conejo (1:10000, SIGMA) conjugado a HRP. Todos los anticuerpos se incubaron 1 hora a 37°C y se lavaron 3 veces con PBS-T y 2 veces con PBS. Los inmunocomplejos formados se revelaron empleando 100 µl/pocillo de TMB (Invitrogen). La reacción colorimétrica se midió a 655 nm en un lector para placas de ELISA (Synergy H1, BioTek). El *cut-off* se calculó como el promedio de las absorbancias de los sueros pre-inmunes más dos desvíos estándar.

### 37. Ensayo de inmunogenicidad

La capacidad de ser reconocidos en la respuesta humoral de las proteínas TgCs se midió mediante ensayos de ELISA acoplado a IgG, IgG-ELISA. Para ello, se sensibilizaron placas de 96 pocillos (*Immuno Plate Maxisorp, Nunc*) con 100 µl/pocillo de una solución 50 µg/mL de TLA o 5 µg/mL de proteínas recombinantes diluidas en *buffer* NaHCO<sub>3</sub> 50 mM pH 9,5 y se incubaron durante toda la noche a 4°C. Al día siguiente, las placas se bloquearon con solución de bloqueo (TBS-T + leche descremada 5 %) durante 1 hora a 37°C. Luego, se agregaron como anticuerpo primario 100 µl/pocillo conteniendo sueros de ratones infectados o no infectado diluidos en solución de bloqueo 1 en 100, como anticuerpo secundario se empleó el anticuerpo monoclonal anti-IgG total de ratón (1:10000, Sigma) conjugado a HRP. Todos los anticuerpos se incubaron 1 hora a 37°C y se lavaron 3 veces con TBS-T y 2 veces con TBS. Los inmunocomplejos formados se revelaron empleando 100 µl/pocillo TMB (Invitrogen). La reacción colorimétrica se midió a 655 nm en un lector para placas de ELISA (Synergy H1, BioTek). El *cut-off* se calculó como el promedio de las absorbancias de los sueros pre-inmunes más dos desvíos estándar.

### 38. Ensayos de citotoxicidad y crecimiento con drogas

#### 38.1 Ensayo de toxicidad celular

La viabilidad celular se determinó en base a la reducción del bromuro de 3-(4,5 dimetil-2-tiazoil)-2,5-difeniltetrazol (MTT, Sigma) en monocapas de células hTERT incubadas durante 24 hs en ausencia o presencia de RSV o SIRT. La absorbancia se midió a 540 nm desde el fondo de la placa utilizando un lector de placas BioTek Synergy (BioTek, Argentina). Los gráficos muestran la viabilidad de

las células en relación con la viabilidad de las células control tratadas con 0,5% v/v de DMSO (concentración vehículo de las drogas) en medios de cultivo (viabilidad del 100%). Las determinaciones se realizaron por triplicado.

### **38.2 Ensayos de replicación intracelular**

Para la exposición intracelular con las drogas, se infectaron células hTERT crecidas en placas de 96 pocillos con  $1 \times 10^4$  parásitos RH\_RFP/pocillo. Una mitad de la placa se usó para el tratamiento control (células hospedadoras + drogas a diferentes concentraciones) y la otra para la replicación de los parásitos. Después de 3 hs de infección, el medio se reemplazó con medio de cultivo fresco 1% que contenía una de varias concentraciones las drogas o el vehículo. A continuación, se permitió que los parásitos se replicaran durante 24 hs en las células hospedadoras. Los valores de unidades fluorescentes relativas, RFU, de RFP se recopilaron utilizando un lector de microplacas multimodo híbrido Synergy H1 (Biotek, Argentina). Los datos se recolectaron a una longitud de emisión de 545 nm hasta una longitud máxima de 590 nm (Alomar et al. 2013). Cada concentración de droga se midió por octuplicado y se realizaron 3 ensayos independientes. La fluorescencia basal (células con DMEM más drogas) se estimó en  $234 \pm 2,3$  RFU. También calculamos el valor  $IC_{50}$  con GraphPad Prism 6 (San Diego, EE. UU.) siguiendo los métodos descritos anteriormente (Campo 2017). Los datos se normalizaron con 0 como el valor más pequeño y se transformaron a una escala semilogarítmica ( $x = \log(x)$ ). Luego, los datos se analizaron como un parámetro de regresión no lineal: dosis-respuesta inhibición-log (inhibidor) frente a la pendiente variable de respuesta normalizada durante 24 hs.

### **39. Medición del contenido de ADN por citometría de flujo**

Para medir el contenido de ADN en parásitos tratados con DMSO o  $100 \mu\text{M}$  de HU, Resveratrol o Sirtinol, se realizaron 3 ensayos independientes por triplicado cada uno. Cada experimento se realizó en placas de 6 pocillos conteniendo cada uno  $1 \cdot 10^6$  células HFF en confluencia. Estas fueron luego infectadas por 24 hs con  $1 \cdot 10^4$  taquizoítos y tratadas por 24 hs más con cada una de las drogas. Una vez transcurrido el tiempo, las células fueron lisadas, y pasadas por agujas de distinto diámetro para

extraer los taquizoítos intracelulares. Estos fueron lavados con PBS 1% y fijados con etanol 100% durante toda la noche para permitir la fijación del ADN. Luego de la fijación los parásitos fueron tratados con 100  $\mu\text{g}/\mu\text{l}$  de RNAsa y luego teñidos con Ioduro de Propidio para permitir la tinción del ADN. Las poblaciones de parásitos fueron separadas por su contenido de ADN en cada una de las muestras utilizando los marcadores de área en un citómetro de flujo (Becker Dickinson), para el análisis de las muestras se utilizó el programa FlowJo y para el análisis estadístico el programa GradPad (versión 6.1).

## **Bibliografía**

- Adeyemi, Oluyomi Stephen, Olubunmi Atolani, Oluwakemi Josephine Awakan, Tomilola Debby Olaolu, Charles Obiora Nwonuma, Omokolade Alejlowo, David Adeiza Otohinoyi, Damilare Rotimi, Akinyomade Owolabi, and Gaber El-Saber Batiha. 2019. "In Vitro Screening to Identify Anti-Toxoplasma Compounds and In Silico Modeling for Bioactivities and Toxicity." *The Yale Journal of Biology and Medicine* 92 (3): 369–83. <http://www.ncbi.nlm.nih.gov/pubmed/31543702>.
- Albrecht, Letusa, Davide Angeletti, Kirsten Moll, Karin Blomqvist, Davide Valentini, Fabio Luiz D’Alexandri, Markus Maurer, and Mats Wahlgren. 2014. "B-Cell Epitopes in NTS-DBL1 $\alpha$  of PfEMP1 Recognized by Human Antibodies in Rosetting *Plasmodium Falciparum*." *PloS One* 9 (12). <https://doi.org/10.1371/JOURNAL.PONE.0113248>.
- Alomar, Maria L, Federico A O Rasse-Suriani, Agustina Ganuza, Verónica M Cóceres, Franco M Cabrerizo, and Sergio O Angel. 2013. "In Vitro Evaluation of  $\beta$ -Carboline Alkaloids as Potential Anti-Toxoplasma Agents." *BMC Research Notes* 6 (1): 193. <https://doi.org/10.1186/1756-0500-6-193>.
- Al-Sady, Bassem, Hiten D. Madhani, and Geeta J. Narlikar. 2013. "Division of Labor between the Chromodomains of HP1 and Suv39 Methylase Enables Coordination of Heterochromatin Spread." *Molecular Cell* 51 (1): 80–91. <https://doi.org/10.1016/j.molcel.2013.06.013>.
- Alsford, Sam, Taemi Kawahara, Cyril Isamah, and David Horn. 2007. "A Sirtuin in the African Trypanosome Is Involved in Both DNA Repair and Telomeric Gene Silencing but Is Not Required for Antigenic Variation." *Molecular Microbiology* 63 (January): 724–36. <https://doi.org/10.1111/j.1365-2958.2006.05553.x>.
- Altschul, Stephen F., John C. Wootton, Elena Zaslavsky, and Yi-Kuo Yu. 2010. "The Construction and Use of Log-Odds Substitution Scores for Multiple Sequence Alignment." Edited by Adam Siepel. *PLoS Computational Biology* 6 (7): e1000852. <https://doi.org/10.1371/journal.pcbi.1000852>.
- Alvarez, Carmelo A., and Elena S. Suvorova. 2017. "Checkpoints of Apicomplexan Cell Division Identified in *Toxoplasma Gondii*." Edited by Isabelle Coppens. *PLOS Pathogens* 13 (7): e1006483. <https://doi.org/10.1371/journal.ppat.1006483>.
- Angel, Sergio O., Laura Vanagas, Diego M. Ruiz, Constanza Cristaldi, Ana M. Saldarriaga Cartagena, and William J. Sullivan. 2020. "Emerging Therapeutic Targets Against *Toxoplasma Gondii*: Update on DNA Repair Response Inhibitors and Genotoxic Drugs." *Frontiers in Cellular and Infection Microbiology* 10 (June). <https://doi.org/10.3389/FCIMB.2020.00289>.

- Audergon, Pauline N C B, Sandra Catania, Alexander Kagansky, Pin Tong, Manu Shukla, Alison L Pidoux, and Robin C Allshire. 2015. "Epigenetics. Restricted Epigenetic Inheritance of H3K9 Methylation." *Science (New York, N.Y.)* 348 (6230): 132–35. <https://doi.org/10.1126/science.1260638>.
- Bannister, Andrew J, and Tony Kouzarides. 2011. "Regulation of Chromatin by Histone Modifications." *Cell Research* 21 (3): 381–95. <https://doi.org/10.1038/cr.2011.22>.
- Bártfai, Richárd, Wieteke A. M. Hoeijmakers, Adriana M. Salcedo-Amaya, Arne H. Smits, Eva Janssen-Megens, Anita Kaan, Moritz Treeck, Tim-Wolf Gilberger, Kees-Jan François, and Hendrik G. Stunnenberg. 2010. "H2A.Z Demarcates Intergenic Regions of the *Plasmodium Falciparum* Epigenome That Are Dynamically Marked by H3K9ac and H3K4me3." Edited by Kami Kim. *PLoS Pathogens* 6 (12): e1001223. <https://doi.org/10.1371/journal.ppat.1001223>.
- Barylyuk, Konstantin, Ludek Koreny, Huiling Ke, Simon Butterworth, Oliver M. Crook, Imen Lassadi, Vipul Gupta, et al. 2020. "A Comprehensive Subcellular Atlas of the *Toxoplasma* Proteome via HyperLOPIT Provides Spatial Context for Protein Functions." *Cell Host & Microbe* 28 (5): 752-766.e9. <https://doi.org/10.1016/J.CHOM.2020.09.011>.
- Basu, Amrita, Kristie L Rose, Junmei Zhang, Ronald C Beavis, Beatrix Ueberheide, Benjamin A Garcia, Brian Chait, et al. 2009. "Proteome-Wide Prediction of Acetylation Substrates." *Proceedings of the National Academy of Sciences of the United States of America* 106 (33): 13785–90. <https://doi.org/10.1073/pnas.0906801106>.
- Beck, Josh R, Imilce A Rodriguez-Fernandez, Jessica Cruz de Leon, My-Hang Huynh, Vern B Carruthers, Naomi S Morrisette, and Peter J Bradley. 2010. "A Novel Family of *Toxoplasma* IMC Proteins Displays a Hierarchical Organization and Functions in Coordinating Parasite Division." Edited by Dominique Soldati-Favre. *PLoS Pathogens* 6 (9): e1001094. <https://doi.org/10.1371/journal.ppat.1001094>.
- Behnke, Michael S., John C. Wootton, Margaret M. Lehmann, Josh B. Radke, Olivier Lucas, Julie Nawas, L. David Sibley, and Michael W. White. 2010. "Coordinated Progression through Two Subtranscriptomes Underlies the Tachyzoite Cycle of *Toxoplasma Gondii*." *PLoS ONE* 5 (8). <https://doi.org/10.1371/journal.pone.0012354>.
- Beisel, Christian, and Renato Paro. 2011. "Silencing Chromatin: Comparing Modes and Mechanisms." *Nature Reviews Genetics* 12 (2): 123–35. <https://doi.org/10.1038/nrg2932>.
- Bell, S. D. 2002. "The Interaction of Alba, a Conserved Archaeal Chromatin Protein, with Sir2 and Its Regulation by Acetylation." *Science* 296 (5565): 148–51. <https://doi.org/10.1126/science.1070506>.

- Benetti, Roberta, Susana Gonzalo, Isabel Jaco, Gunnar Schotta, Peter Klatt, Thomas Jenuwein, and María A. Blasco. 2007. "Suv4-20h Deficiency Results in Telomere Elongation and Derepression of Telomere Recombination." *The Journal of Cell Biology* 178 (6): 925–36. <https://doi.org/10.1083/jcb.200703081>.
- Bereswill, Stefan, Melba Muñ Oz, André Fischer, Rita Plickert, Lea-Maxie Haag, Bettina Otto, Anja A Kü HI, Christoph Loddenkemper, Ulf B Gö Bel, and Markus M Heimesaat. n.d. "Anti-Inflammatory Effects of Resveratrol, Curcumin and Simvastatin in Acute Small Intestinal Inflammation." Accessed April 24, 2018. <https://doi.org/10.1371/journal.pone.0015099>.
- Berná, Luisa, Pablo Marquez, Andrés Cabrera, Gonzalo Greif, María E. Francia, and Carlos Robello. 2021. "Reevaluation of the *Toxoplasma Gondii* and *Neospora Caninum* Genomes Reveals Misassembly, Karyotype Differences, and Chromosomal Rearrangements." *Genome Research* 31 (5): 823–33. <https://doi.org/10.1101/GR.262832.120>.
- Bernards, André, H. T. Lex, A. van der Ploeg, Carlos C. Frasch, Piet Borst, John C. Boothroyd, Sharon Coleman, and George A.M. Cross. 1981. "Activation of Trypanosome Surface Glycoprotein Genes Involves a Duplication-Transposition Leading to an Altered 3' End." *Cell* 27 (3 Pt 2): 497–505. [https://doi.org/10.1016/0092-8674\(81\)90391-3](https://doi.org/10.1016/0092-8674(81)90391-3).
- Bierhoff, Holger, Ingrid Grummt, Gunnar Schotta, Marcel Andre Dammert, Silvia Dambacher, and David Brocks. 2014. "Quiescence-Induced LncRNAs Trigger H4K20 Trimethylation and Transcriptional Silencing." *Molecular Cell* 54 (4): 675–82. <https://doi.org/10.1016/j.molcel.2014.03.032>.
- Blader, Ira J, Bradley I Coleman, Chun-Ti Chen, and Marc-Jan Gubbels. 2015. "Lytic Cycle of *Toxoplasma Gondii*: 15 Years Later." *Annual Review of Microbiology* 69 (August): 463–85. <https://doi.org/10.1146/annurev-micro-091014-104100>.
- Bogado, Silvina S., María C. Dalmasso, Agustina Ganuza, Kami Kim, William J. Sullivan, Sergio O. Angel, and Laura Vanagas. 2014. "Canonical Histone H2Ba and H2A.X Dimerize in an Opposite Genomic Localization to H2A.Z/H2B.Z Dimers in *Toxoplasma Gondii*." *Molecular and Biochemical Parasitology* 197 (1–2): 36–42. <https://doi.org/10.1016/j.molbiopara.2014.09.009>.
- Bosch-Presegué, Laia, and Alejandro Vaquero. 2015. "Sirtuin-Dependent Epigenetic Regulation in the Maintenance of Genome Integrity." *The FEBS Journal* 282 (9): 1745–67. <https://doi.org/10.1111/febs.13053>.
- Bottari, Nathieli B., Matheus D. Baldissera, Alexandre A. Tonin, Virginia C. Rech, Vivian S K Nishihira, Gustavo R. Thomé, Giovana Camillo, et al. 2015. "Effects of Sulfamethoxazole-Trimethoprim Associated to Resveratrol on Its Free Form and Complexed with 2-Hydroxypropyl-??-Cyclodextrin on Cytokines Levels

of Mice Infected by *Toxoplasma Gondii*." *Microbial Pathogenesis* 87: 40–44.  
<https://doi.org/10.1016/j.micpath.2015.07.013>.

Brachmann, Carrie Baker, Joyce M. Sherman, Scott E. Devine, Elizabeth E. Cameron, Lorraine Pillus, and Jef D. Boeke. 1995. "The SIR2 Gene Family, Conserved from Bacteria to Humans, Functions in Silencing, Cell Cycle Progression, and Chromosome Stability." *Genes & Development* 9 (23): 2888–2902.  
<https://doi.org/10.1101/GAD.9.23.2888>.

Branzei, Dana, and Marco Foiani. 2008. "Regulation of DNA Repair throughout the Cell Cycle." *Nature Reviews. Molecular Cell Biology* 9 (4): 297–308. <https://doi.org/10.1038/NRM2351>.

Braun, Laurence, Dominique Cannella, Philippe Ortet, Mohamed Barakat, Céline F. Sautel, Sylvie Kieffer, Jérôme Garin, Olivier Bastien, Olivier Voinnet, and Mohamed Ali Hakimi. 2010. "A Complex Small Rna Repertoire Is Generated by a Plant/Fungal-like Machinery and Effected by a Metazoan-like Argonaute in the Single-Cell Human Parasite *Toxoplasma Gondii*." *PLoS Pathogens* 6 (5): 1–20.  
<https://doi.org/10.1371/journal.ppat.1000920>.

Britten, R. J., and D. E. Kohne. 1968. "Repeated Sequences in DNA. Hundreds of Thousands of Copies of DNA Sequences Have Been Incorporated into the Genomes of Higher Organisms." *Science (New York, N. Y.)* 161 (3841): 529–40. <https://doi.org/10.1126/SCIENCE.161.3841.529>.

Broadbent, Kate M, Daniel Park, Ashley R Wolf, Daria van Tyne, Jennifer S Sims, Ulf Ribacke, Sarah Volkman, et al. 2011. "A Global Transcriptional Analysis of *Plasmodium Falciparum* Malaria Reveals a Novel Family of Telomere-Associated LncRNAs." <http://genomebiology.com/2011/12/6/R56>.

Brooks, Carrie F, Maria E Francia, Mathieu Gissot, Matthew M Croken, Kami Kim, and Boris Striepen. 2011. "*Toxoplasma Gondii* Sequesters Centromeres to a Specific Nuclear Region throughout the Cell Cycle." *Proceedings of the National Academy of Sciences* 108 (9): 3767–72.  
<https://doi.org/10.1073/pnas.1006741108>.

Brown, Chris A., Andrew W. Murray, and Kevin J. Verstrepen. 2010. "Rapid Expansion and Functional Divergence of Subtelomeric Gene Families in Yeasts." *Current Biology* 20 (10): 895–903.  
<https://doi.org/10.1016/J.CUB.2010.04.027>.

Brynska, A, E Tomaszewicz-Libudzic, and T Wolanczyk. 2001. "Obsessive-Compulsive Disorder and Acquired Toxoplasmosis in Two Children." *European Child & Adolescent Psychiatry* 10 (3): 200–204.  
<http://www.ncbi.nlm.nih.gov/pubmed/11596821>.

- Bulut-Karslioglu, Aydan, Valentina Perrera, Manuela Scaranaro, Inti Alberto de La Rosa-Velazquez, Suzanne van de Nobelen, Nicholas Shukeir, Johannes Popow, et al. 2012. "A Transcription Factor-Based Mechanism for Mouse Heterochromatin Formation." *Nature Structural & Molecular Biology* 19 (10): 1023–32. <https://doi.org/10.1038/NSMB.2382>.
- Bunnik, Evelien M., Aarthi Venkat, Jianlin Shao, Kathryn E. McGovern, Gayani Batugedara, Danielle Worth, Jacques Prudhomme, et al. 2019. "Comparative 3D Genome Organization in Apicomplexan Parasites." *Proceedings of the National Academy of Sciences* 116 (8): 3183–92. <https://doi.org/10.1073/pnas.1810815116>.
- Burke, Stephen A., Sam L. Lo, and Joseph A. Krzycki. 1998. "Clustered Genes Encoding the Methyltransferases of Methanogenesis from Monomethylamine." *Journal of Bacteriology* 180 (13): 3432–40. <https://doi.org/10.1128/JB.180.13.3432-3440>.
- Campbell, David A., Deborah A. Thornton, and John C. Boothroyd. 1984. "Apparent Discontinuous Transcription of *Trypanosoma Brucei* Variant Surface Antigen Genes." *Nature* 311 (5984): 350–55. <https://doi.org/10.1038/311350A0>.
- Campo, Vanina A. 2017. "Comparative Effects of Histone Deacetylases Inhibitors and Resveratrol on *Trypanosoma Cruzi* Replication, Differentiation, Infectivity and Gene Expression." <https://doi.org/10.1016/j.ijpddr.2016.12.003>.
- Cao, Xuanye, Chaoqun Li, Siyu Xiao, Yunlan Tang, Jing Huang, Shuan Zhao, Xueyu Li, Jixi Li, Ruilin Zhang, and Wei Yu. 2017. "Acetylation Promotes TyrRS Nuclear Translocation to Prevent Oxidative Damage." *Proceedings of the National Academy of Sciences of the United States of America* 114 (4): 687–92. <https://doi.org/10.1073/pnas.1608488114>.
- Carrier, Yves, Carine Truyens, Philippe Deloron, and François Peyron. 2012. "Congenital Parasitic Infections: A Review." *Acta Tropica* 121 (2): 55–70. <https://doi.org/10.1016/j.actatropica.2011.10.018>.
- Caslini, Corrado, James A. Connelly, Amparo Serna, Dominique Broccoli, and Jay L. Hess. 2009. "MLL Associates with Telomeres and Regulates Telomeric Repeat-Containing RNA Transcription." *Molecular and Cellular Biology* 29 (16): 4519–26. <https://doi.org/10.1128/MCB.00195-09>.
- Chen, Chang Shi, Yu Chieh Wang, Hsiao Ching Yang, Po Hsien Huang, Samuel K. Kulp, Chih Cheng Yang, Yen Shen Lu, Shigemi Matsuyama, Ching Yu Chen, and Ching Shih Chen. 2007. "Histone Deacetylase Inhibitors Sensitize Prostate Cancer Cells to Agents That Produce DNA Double-Strand Breaks by Targeting Ku70 Acetylation." *Cancer Research*. American Association for Cancer Research Inc. <https://doi.org/10.1158/0008-5472.CAN-06-3996>.

- Chen, C.-T., and M.-J. Gubbels. 2013. "The *Toxoplasma Gondii* Centrosome Is the Platform for Internal Daughter Budding as Revealed by a Nek1 Kinase Mutant." *Journal of Cell Science* 126 (15): 3344–55. <https://doi.org/10.1242/jcs.123364>.
- Chêne, Arnaud, Shruthi S. Vembar, Loïc Rivière, José Juan Lopez-Rubio, Aurelie Claes, T. Nicolai Siegel, Hiroshi Sakamoto, Christine Scheidig-Benatar, Rosaura Hernandez-Rivas, and Artur Scherf. 2012. "PfAlbas Constitute a New Eukaryotic DNA/RNA-Binding Protein Family in Malaria Parasites." *Nucleic Acids Research* 40 (7): 3066–77. <https://doi.org/10.1093/nar/gkr1215>.
- Chen, Guo Dong, Wei Dong Yu, and Xiao Ping Chen. 2016. "SirT1 Activator Represses the Transcription of TNF- $\alpha$  in THP-1 Cells of a Sepsis Model via Deacetylation of H4K16." *Molecular Medicine Reports* 14 (6): 5544–50. <https://doi.org/10.3892/MMR.2016.5942>.
- Cheng, Zhukuan, C. Robin Buell, Rod A. Wing, Minghong Gu, and Jiming Jiang. 2001. "Toward a Cytological Characterization of the Rice Genome." *Genome Research* 11 (12): 2133–41. <https://doi.org/10.1101/GR.194601>.
- Chen, Jian-Ping, Da-Li Chen, Han-Xiao Qin, Yi-Kai Yang, Zhi-Wan Zheng, Kai Dong, Qi-Wei Chen, Jiao Li, and Jin-Lei He. 2018. "The Direct and Indirect Inhibition Effects of Resveratrol against *Toxoplasma Gondii* Tachyzoites in Vitro". *Antimicrobial Agents and Chemotherapy*, no. December. <https://doi.org/10.1128/aac.01233-18>.
- Chiurillo, Miguel A., Roberto R. Moraes Barros, Renata T. Souza, Marjorie M. Marini, Cristiane R. Antonio, Danielle R. Cortez, María Curto, et al. 2016. "Subtelomeric I-SceI-Mediated Double-Strand Breaks Are Repaired by Homologous Recombination in *Trypanosoma Cruzi*." *Frontiers in Microbiology* 7 (DEC): 1–8. <https://doi.org/10.3389/fmicb.2016.02041>.
- Choudhary, Mani Kant, Debarati Basu, Asis Datta, Niranjana Chakraborty, and Subhra Chakraborty. 2009. "Dehydration-Responsive Nuclear Proteome of Rice (*Oryza Sativa L.*) Illustrates Protein Network, Novel Regulators of Cellular Adaptation, and Evolutionary Perspective." *Molecular & Cellular Proteomics: MCP* 8 (7): 1579–98. <https://doi.org/10.1074/mcp.M800601-MCP200>.
- Cleary, Michael D, Upinder Singh, Ira J Blader, Jeremy L Brewer, and John C Boothroyd. 2002. "*Toxoplasma Gondii* Asexual Development: Identification of Developmentally Regulated Genes and Distinct Patterns of Gene Expression." *Eukaryotic Cell* 1 (3): 329–40. <http://www.ncbi.nlm.nih.gov/pubmed/12455982>.
- Clemente, Marina, Natalia de Miguel, Veronica v. Lia, Mariana Matrajt, and Sergio O. Angel. 2004. "Structure Analysis of Two *Toxoplasma Gondii* and *Neospora Caninum* Satellite DNA Families and Evolution of Their

- Common Monomeric Sequence.” *Journal of Molecular Evolution* 58 (5): 557–67. <https://doi.org/10.1007/S00239-003-2578-3>.
- Contreras, Susana M., Agustina Ganuza, María M. Corvi, and Sergio O. Angel. 2021. “Resveratrol Induces H3 and H4K16 Deacetylation and H2A.X Phosphorylation in *Toxoplasma Gondii*.” *BMC Research Notes* 14 (1). <https://doi.org/10.1186/S13104-020-05416-4>.
- Cormack, Brendan P., Nafisa Ghori, and Stanley Falkow. 1999. “An Adhesin of the Yeast Pathogen *Candida Glabrata* Mediating Adherence to Human Epithelial Cells.” *Science (New York, N.Y.)* 285 (5427): 578–82. <https://doi.org/10.1126/SCIENCE.285.5427.578>.
- Cross, G. A.M. 1990. “Cellular and Genetic Aspects of Antigenic Variation in Trypanosomes.” *Annual Review of Immunology* 8: 83–110. <https://doi.org/10.1146/ANNUREV.IY.08.040190.000503>.
- Dalmasso, María C, Santiago J Carmona, Sergio O Angel, and Fernán Agüero. 2014. “Characterization of *Toxoplasma Gondii* Subtelomeric-like Regions: Identification of a Long-Range Compositional Bias That Is Also Associated with Gene-Poor Regions.” *BMC Genomics* 15 (1): 21. <https://doi.org/10.1186/1471-2164-15-21>.
- Dalmasso, Maria Carolina, William Joseph Sullivan, and Sergio Oscar Angel. 2011. “Canonical and Variant Histones of Protozoan Parasites.” *Frontiers in Bioscience (Landmark Edition)* 16 (June): 2086–2105. <http://www.ncbi.nlm.nih.gov/pubmed/21622164>.
- Dalmasso, Maria C., David O. Onyango, Arunasalam Naguleswaran, William J. Sullivan, and Sergio O. Angel. 2009. “*Toxoplasma* H2A Variants Reveal Novel Insights into Nucleosome Composition and Functions for This Histone Family.” *Journal of Molecular Biology* 392 (1): 33–47. <https://doi.org/10.1016/j.jmb.2009.07.017>.
- Dalmasso, MC. Tesis Doctoral, 2009. Caracterización y análisis funcional de las histonas H2A y H2B de *Toxoplasma gondii*. UNSaM.
- Dalmasso, María C., Pablo C. Echeverria, María P. Zappia, Ulf Hellman, Jean François Dubremetz, and Sergio O. Angel. 2006. “*Toxoplasma Gondii* Has Two Lineages of Histones 2b (H2B) with Different Expression Profiles.” *Molecular and Biochemical Parasitology* 148: 103–7. <https://doi.org/10.1016/j.molbiopara.2006.03.005>.
- De Las Peñas, Alejandro, Shih Jung Pan, Irene Castaño, Jonathan Alder, Robert Cregg, and Brendan P. Cormack. 2003. “Virulence-Related Surface Glycoproteins in the Yeast Pathogen *Candida Glabrata* Are

Encoded in Subtelomeric Clusters and Subject to RAP1- and SIR-Dependent Transcriptional Silencing.” *Genes & Development* 17 (18): 2245–58. <https://doi.org/10.1101/GAD.1121003>.

De Leon, Jessica Cruz, Nicole Scheumann, Wandy Beatty, Josh R. Beck, Johnson Q. Tran, Candace Yau, Peter J. Bradley, Keith Gull, Bill Wickstead, and Naomi S. Morrisette. 2013. “A SAS-6-Like Protein Suggests That the *Toxoplasma* Conoid Complex Evolved from Flagellar Components.” *Eukaryotic Cell* 12 (7): 1009–19. <https://doi.org/10.1128/EC.00096-13>.

Deitsch, Kirk W., and Ron Dzikowski. 2017. “Variant Gene Expression and Antigenic Variation by Malaria Parasites.” *Annual Review of Microbiology* 71 (September): 625–41. <https://doi.org/10.1146/ANNUREV-MICRO-090816-093841>.

Denchi, Eros Lazzerini. 2009. “Give Me a Break: How Telomeres Suppress the DNA Damage Response.” *DNA Repair* 8: 1118–26. <https://doi.org/10.1016/j.dnarep.2009.04.013>.

Deshmukh, Abhijit S., Sandeep Srivastava, Susann Herrmann, Ashish Gupta, Pallabi Mitra, Tim Wolf Gilberger, and Suman Kumar Dhar. 2012. “The Role of N-Terminus of *Plasmodium Falciparum* ORC1 in Telomeric Localization and Var Gene Silencing.” *Nucleic Acids Research* 40 (12): 5313–31. <https://doi.org/10.1093/nar/gks202>.

Dimitri, Patrizio, Ruggiero Caizzi, Ennio Giordano, Maria Carmela Accardo, Giovanna Lattanzi, and Giuseppe Biamonti. 2009. “Constitutive Heterochromatin: A Surprising Variety of Expressed Sequences.” *Chromosoma* 118 (4): 419–35. <https://doi.org/10.1007/s00412-009-0211-y>.

Donald, R G, D Carter, B Ullman, and D S Roos. 1996. “Insertional Tagging, Cloning, and Expression of the *Toxoplasma Gondii* Hypoxanthine-Xanthine-Guanine Phosphoribosyltransferase Gene. Use as a Selectable Marker for Stable Transformation.” *The Journal of Biological Chemistry* 271 (24): 14010–19. <http://www.ncbi.nlm.nih.gov/pubmed/8662859>.

Dubey, J. P., D. S. Lindsay, and C. A. Speer. 1998. “Structures of *Toxoplasma Gondii* Tachyzoites, Bradyzoites, and Sporozoites and Biology and Development of Tissue Cysts.” *Clinical Microbiology Reviews* 11 (2): 267. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC106833/>.

Echeverria, P C, P A Rojas, V Martin, E A Guarnera, V Pszenny, and S O Angel. 2000. “Characterisation of a Novel Interspersed *Toxoplasma Gondii* DNA Repeat with Potential Uses for PCR Diagnosis and PCR-RFLP Analysis.” *FEMS Microbiology Letters* 184 (1): 23–27. <http://www.ncbi.nlm.nih.gov/pubmed/10689160>.

- Eichler, E. E. 2001. "Segmental Duplications: What's Missing, Misassigned,  $\zeta$  and Misassembled and Should We Care?" *Genome Research* 11 (5): 653–56. <https://doi.org/10.1101/GR.188901>.
- Farrell, Megan, and Marc-Jan Gubbels. 2014. "The *Toxoplasma Gondii* Kinetochore Is Required for Centrosome Association with the Centrocone (Spindle Pole)." *Cellular Microbiology* 16 (1): 78–94. <https://doi.org/10.1111/cmi.12185>.
- Feng, Suhua, and Steven E Jacobsen. 2011. "Epigenetic Modifications in Plants: An Evolutionary Perspective." *Current Opinion in Plant Biology* 14 (2): 179–86. <https://doi.org/10.1016/j.pbi.2010.12.002>.
- Fernandez-Becerra, Carmen, Oliver Pein, Tatiane Rodrigues de Oliveira, Marcio Massao Yamamoto, Antonio Carlos Cassola, Claudia Rocha, Irene S Soares, Carlos A de Bragança Pereira, and Hernando A del Portillo. 2005. "Variant Proteins of *Plasmodium Vivax* Are Not Clonally Expressed in Natural Infections." *Molecular Microbiology* 58 (3): 648–58. <https://doi.org/10.1111/j.1365-2958.2005.04850.x>.
- Fernandez-Becerra, Carmen, Marcio M. Yamamoto, Ricardo Z.N. Vêncio, Marcus Lacerda, Anna Rosanas-Urgell, and Hernando A. del Portillo. 2009. "*Plasmodium Vivax* and the Importance of the Subtelomeric Multigene Vir Superfamily." *Trends in Parasitology* 25 (1): 44–51. <https://doi.org/10.1016/j.pt.2008.09.012>.
- Ferreira, Christian, Deivid Costa Soares, Michelle Tanny Cunha do Nascimento, Lucia Helena Pinto-da-Silva, Carolina Galvão Sarzedas, Luzineide Wanderley Tinoco, and Elvira Maria Saraiva. 2014. "Resveratrol Is Active against *Leishmania Amazonensis*: In Vitro Effect of Its Association with Amphotericin B." *Antimicrobial Agents and Chemotherapy* 58 (10): 6197–6208. <https://doi.org/10.1128/AAC.00093-14>.
- Field, Mark C., and John C. Boothroyd. 1995. "*Trypanosoma Brucei*: Molecular Cloning of Homologues of Small GTP-Binding Proteins Involved in Vesicle Trafficking." *Experimental Parasitology* 81 (3): 313–20. <https://doi.org/10.1006/EXPR.1995.1122>.
- Figueiredo, L M, L A Pirrit, A Scherf, and L A Pirritt. 2000. "Genomic Organisation and Chromatin Structure of *Plasmodium Falciparum* Chromosome Ends." *Molecular and Biochemical Parasitology* 106 (1): 169–74. <http://www.ncbi.nlm.nih.gov/pubmed/10743621>.
- Figueiredo, Luisa M., Eduardo P.C. Rocha, Liliana Mancio-Silva, Christine Prevost, Danièle Hernandez-Verdun, and Artur Scherf. 2005. "The Unusually Large *Plasmodium* Telomerase Reverse-Transcriptase Localizes in a Discrete Compartment Associated with the Nucleolus." *Nucleic Acids Research* 33 (3): 1111–22. <https://doi.org/10.1093/NAR/GKI260>.
- Fischle, Wolfgang, Yanming Wang, and C David Allis. 2003. "Histone and Chromatin Cross-Talk." *Current Opinion in Cell Biology* 15 (2): 172–83. [https://doi.org/10.1016/S0955-0674\(03\)00013-9](https://doi.org/10.1016/S0955-0674(03)00013-9).

- Fong, Yao, Yin-Chieh Lin, Chang-Yi Wu, Hui-Min David Wang, Li-Li Lin, Han Lin Chou, Yen-Ni Teng, Shyng-Shiou Yuan, and Chien-Chih Chiu. 2014. "The Antiproliferative and Apoptotic Effects of Sirtinol, a Sirtuin Inhibitor on Human Lung Cancer Cells by Modulating Akt/ $\beta$ -Catenin-Foxo3a Axis." *TheScientificWorldJournal* 2014: 937051. <https://doi.org/10.1155/2014/937051>.
- Ford, Ethan, Renate Voit, Gregory Liszt, Cornelia Magin, Ingrid Grummt, and Leonard Guarente. 2006. "Mammalian Sir2 Homolog SIRT7 Is an Activator of RNA Polymerase I Transcription." *Genes & Development* 20 (9): 1075–80. <https://doi.org/10.1101/GAD.1399706>.
- Fox, Barbara A., Jessica G. Ristuccia, Jason P. Gigley, and David J. Bzik. 2009. "Efficient Gene Replacements in *Toxoplasma Gondii* Strains Deficient for Nonhomologous End Joining." *Eukaryotic Cell* 8 (4): 520–29. <https://doi.org/10.1128/EC.00357-08>.
- Francia, Maria E., Sheila Bhavsar, Li Min Ting, Matthew M. Croken, Kami Kim, Jean Francois Dubremetz, and Boris Striepen. 2020. "A Homolog of Structural Maintenance of Chromosome 1 Is a Persistent Centromeric Protein Which Associates With Nuclear Pore Components in *Toxoplasma Gondii*." *Frontiers in Cellular and Infection Microbiology* 10 (July): 295. <https://doi.org/10.3389/FCIMB.2020.00295/FULL>.
- Francia, Maria E, Jean-Francois Dubremetz, and Naomi S Morrissette. 2016. "Basal Body Structure and Composition in the Apicomplexans *Toxoplasma* and *Plasmodium*" 5: 3. <https://doi.org/10.1186/s13630-016-0025-5>.
- Francia, Maria E, and Boris Striepen. 2014. "Cell Division in Apicomplexan Parasites." *Nature Reviews. Microbiology* 12 (January): 125–36. <https://doi.org/10.1038/nrmicro3184>.
- Francia, Maria E., Carly N. Jordan, Jay D. Patel, Lilach Sheiner, Jessica L. Demerly, Justin D. Fellows, Jessica Cruz de Leon, Naomi S. Morrissette, Jean François Dubremetz, and Boris Striepen. 2012. "Cell Division in Apicomplexan Parasites Is Organized by a Homolog of the Striated Rootlet Fiber of Algal Flagella." *PLoS Biology* 10 (12). <https://doi.org/10.1371/journal.pbio.1001444>.
- Fransz, Paul F., Susan Armstrong, J. Hans de Jong, Laurence D. Parnell, Cees van Drunen, Caroline Dean, Pim Zabel, Ton Bisseling, and Gareth H. Jones. 2000. "Integrated Cytogenetic Map of Chromosome Arm 4S of *A. Thaliana*: Structural Organization of Heterochromatic Knob and Centromere Region." *Cell* 100 (3): 367–76. [https://doi.org/10.1016/S0092-8674\(00\)80672-8](https://doi.org/10.1016/S0092-8674(00)80672-8).
- Freitas-Junior, Lucio H., Rosaura Hernandez-Rivas, Stuart A. Ralph, Dvorak Montiel-Condado, Omar K. Ruvalcaba-Salazar, Ana Paola Rojas-Meza, Liliana Mâncio-Silva, et al. 2005. "Telomeric Heterochromatin Propagation and Histone Acetylation Control Mutually Exclusive Expression of Antigenic Variation Genes in Malaria Parasites." *Cell* 121 (1): 25–36. <https://doi.org/10.1016/j.cell.2005.01.037>.

- Freitas-Junior, L H, E Bottius, L a Pirrit, K W Deitsch, C Scheidig, F Guinet, U Nehrbass, T E Wellems, and a Scherf. 2000. "Frequent Ectopic Recombination of Virulence Factor Genes in Telomeric Chromosome Clusters of *P. Falciparum*." *Nature* 407 (6807): 1018–22. <https://doi.org/10.1038/35039531>.
- Frye R A. 2000. "Phylogenetic Classification of Prokaryotic and Eukaryotic Sir2-like Proteins." *Biochemical and Biophysical Research Communications* 273 (2): 793–98. <https://doi.org/10.1006/BBRC.2000.3000>.
- Frye, Roy A. 1999. "Characterization of Five Human cDNAs with Homology to the Yeast SIR2 Gene: Sir2-like Proteins (Sirtuins) Metabolize NAD and May Have Protein ADP-Ribosyltransferase Activity." *Biochemical and Biophysical Research Communications* 260 (1): 273–79. <https://doi.org/10.1006/BBRC.1999.0897>.
- Gagou, Mary E., Pedro Zuazua-Villar, and Mark Meuth. 2010. "Enhanced H2AX Phosphorylation, DNA Replication Fork Arrest, and Cell Death in the Absence of Chk1." *Molecular Biology of the Cell* 21 (5): 739–52. <https://doi.org/10.1091/MBC.E09-07-0618>.
- Gaillard, J. L., P. Berche, C. Frehel, E. Gouln, and P. Cossart. 1991. "Entry of *L. Monocytogenes* into Cells Is Mediated by Internalin, a Repeat Protein Reminiscent of Surface Antigens from Gram-Positive Cocci." *Cell* 65 (7): 1127–41. [https://doi.org/10.1016/0092-8674\(91\)90009-N](https://doi.org/10.1016/0092-8674(91)90009-N).
- Gaji, Rajshekhar Y., Michael S. Behnke, Margaret M. Lehmann, Michael W. White, and Vern B. Carruthers. 2011. "Cell Cycle-Dependent, Intercellular Transmission of *Toxoplasma Gondii* Is Accompanied by Marked Changes in Parasite Gene Expression." *Molecular Microbiology* 79 (1): 192–204. <https://doi.org/10.1111/j.1365-2958.2010.07441.x>.
- Galinski, M. R., S. A. Lapp, M. S. Peterson, F. Ay, C. J. Joyner, K. G. le Roch, L. L. Fonseca, and E. O. Voit. 2018. "*Plasmodium Knowlesi*: A Superb in Vivo Nonhuman Primate Model of Antigenic Variation in Malaria." *Parasitology* 145 (1): 85–100. <https://doi.org/10.1017/S0031182017001135>.
- Gao, Piaoyang, Na Li, Kaihua Ji, Yan Wang, Chang Xu, Yang Liu, Qin Wang, et al. 2019. "Resveratrol Targets TyrRS Acetylation to Protect against Radiation-Induced Damage." *FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology* 33 (7): 8083–93. <https://doi.org/10.1096/FJ.201802474RR>.
- García-Cao, Marta, Roderick O'Sullivan, Antoine H F M Peters, Thomas Jenuwein, and María A Blasco. 2004. "Epigenetic Regulation of Telomere Length in Mammalian Cells by the Suv39h1 and Suv39h2 Histone Methyltransferases." *Nature Genetics* 36 (1): 94–99. <https://doi.org/10.1038/ng1278>.
- Garrido-Ramos, Manuel A. 2017. "Satellite DNA: An Evolving Topic." *Genes* 8 (9). <https://doi.org/10.3390/GENES8090230>.

- Gertz, Melanie, Frank Fischer, Giang Thi Tuyet Nguyen, Mahadevan Lakshminarasimhan, Mike Schutkowski, Michael Weyand, and Clemens Steegborn. 2013. "Ex-527 Inhibits Sirtuins by Exploiting Their Unique NAD<sup>+</sup>-Dependent Deacetylation Mechanism." *Proceedings of the National Academy of Sciences of the United States of America* 110 (30). <https://doi.org/10.1073/PNAS.1303628110>.
- Gissot, Mathieu, Robert Walker, Stephane Delhay, Tchilabalo Dilezitoko Alayi, Ludovic Huot, David Hot, Isabelle Callebaut, Christine Schaeffer-Reiss, Alain van Dorsselaer, and Stanislas Tomavo. 2013. "Toxoplasma Gondii Alba Proteins Are Involved in Translational Control of Gene Expression." *Journal of Molecular Biology* 425 (8): 1287–1301. <https://doi.org/10.1016/j.jmb.2013.01.039>.
- Gissot, Mathieu, Robert Walker, Stephane Delhay, Ludovic Huot, David Hot, and Stanislas Tomavo. 2012. "Toxoplasma Gondii Chromodomain Protein 1 Binds to Heterochromatin and Colocalises with Centromeres and Telomeres at the Nuclear Periphery." *PloS One* 7 (3): e32671. <https://doi.org/10.1371/journal.pone.0032671>.
- Gottschling, Daniel E., Oscar M. Aparicio, Barbara L. Billington, and Virginia A. Zakian. 1990. "Position Effect at S. Cerevisiae Telomeres: Reversible Repression of Pol II Transcription." *Cell* 63 (4): 751–62. [https://doi.org/10.1016/0092-8674\(90\)90141-Z](https://doi.org/10.1016/0092-8674(90)90141-Z).
- Goyal, Manish, Athar Alam, Mohd Shameel Iqbal, Sumanta Dey, Samik Bindu, Chinmay Pal, Anindyajit Banerjee, Saikat Chakrabarti, and Uday Bandyopadhyay. 2012. "Identification and Molecular Characterization of an Alba-Family Protein from Human Malaria Parasite Plasmodium Falciparum." *Nucleic Acids Research* 40 (3): 1174–90. <https://doi.org/10.1093/nar/gkr821>.
- Grewal, Shiv I. S., and Songtao Jia. 2007. "Heterochromatin Revisited." *Nature Reviews Genetics* 8 (1): 35–46. <https://doi.org/10.1038/nrg2008>.
- Grozinger, C M, E D Chao, H E Blackwell, D Moazed, and S L Schreiber. 2001. "Identification of a Class of Small Molecule Inhibitors of the Sirtuin Family of NAD-Dependent Deacetylases by Phenotypic Screening." *The Journal of Biological Chemistry* 276 (42): 38837–43. <https://doi.org/10.1074/jbc.M106779200>.
- Grunstein, Michael, Susan M Gasser, David Allis, Marie-Laure Caparros, Thomas Jenuwein, and Danny Reinberg. 2013. "Epigenetics in Saccharomyces cerevisiae." *Cold Spring Harb Perspect Biol* 5: 17491–92. <https://doi.org/10.1101/cshperspect.a017491>.
- Guarente, Leonard. 2000. "Sir2 Links Chromatin Silencing, Metabolism, and Aging." *Genes & Development* 14 (9): 1021–26. <https://doi.org/10.1101/gad.14.9.1021>.

- Gubbels, Marc-Jan, Michael White, and Tomasz Szatanek. 2008. "The Cell Cycle and *Toxoplasma Gondii* Cell Division: ¿Tightly Knit or Loosely Stitched?" *International Journal for Parasitology* 38 (12): 1343–58. <https://doi.org/10.1016/J.IJPARA.2008.06.004>.
- Gubbels, Marc-Jan, Shipra Vaishnava, Nico Boot, Jean-François Dubremetz, and Boris Striepen. 2006. "A MORN-Repeat Protein Is a Dynamic Component of the *Toxoplasma Gondii* Cell Division Apparatus." *Journal of Cell Science* 119 (Pt 11): 2236–45. <https://doi.org/10.1242/jcs.02949>.
- Guillemette, Benoit, Paul Drogaris, Hsiu Hsu Sophia Lin, Harry Armstrong, Kyoko Hiragami-Hamada, Axel Imhof, Éric Bonneil, Pierre Thibault, Alain Verreault, and Richard J. Festenstein. 2011. "H3 Lysine 4 Is Acetylated at Active Gene Promoters and Is Regulated by H3 Lysine 4 Methylation." *PLoS Genetics* 7 (3). <https://doi.org/10.1371/JOURNAL.PGEN.1001354>.
- Guo, Bing, Cora A. Styles, Qinghua Feng, and Gerald R. Fink. 2000. "A *Saccharomyces* Gene Family Involved in Invasive Growth, Cell-Cell Adhesion, and Mating." *Proceedings of the National Academy of Sciences of the United States of America* 97 (22): 12158–63. <https://doi.org/10.1073/PNAS.220420397>.
- Gupta, Ashish, Parul Mehra, Abhijeet Deshmukh, Ashraf Dar, Pallabi Mitra, Nilanjan Roy, and Suman Kumar Dhar. 2009. "Functional Dissection of the Catalytic Carboxyl-Terminal Domain of Origin Recognition Complex Subunit 1 (PfORC1) of the Human Malaria Parasite *Plasmodium Falciparum*." *Eukaryotic Cell* 8 (9): 1341–51. <https://doi.org/10.1128/EC.00170-09>.
- Gupta, Ashish, Parul Mehra, and Suman Kumar Dhar. 2008. "Plasmodium Falciparum Origin Recognition Complex Subunit 5: Functional Characterization and Role in DNA Replication Foci Formation." *Molecular Microbiology* 69 (3): 646–65. <https://doi.org/10.1111/J.1365-2958.2008.06316.X>.
- Gupta, Mohit Kumar, Meetu Agarawal, Khadija Banu, K. Sony Reddy, Deepak Gaur, and Suman Kumar Dhar. 2018. "Role of Chromatin Assembly Factor 1 in DNA Replication of *Plasmodium Falciparum*." *Biochemical and Biophysical Research Communications* 495 (1): 1285–91. <https://doi.org/10.1016/j.bbrc.2017.11.131>.
- Hailu, Gebremedhin S., Dina Robaa, Mariantonietta Forgione, Wolfgang Sippl, Dante Rotili, and Antonello Mai. 2017. "Lysine Deacetylase Inhibitors in Parasites: Past, Present, and Future Perspectives." *Journal of Medicinal Chemistry* 60 (12): 4780–4804. <https://doi.org/10.1021/acs.jmedchem.6b01595>.
- Halme, Adrian, Stacie Bumgarner, Cora Styles, and Gerald R. Fink. 2004. "Genetic and Epigenetic Regulation of the FLO Gene Family Generates Cell-Surface Variation in Yeast." *Cell* 116 (3): 405–15. [https://doi.org/10.1016/S0092-8674\(04\)00118-7](https://doi.org/10.1016/S0092-8674(04)00118-7).

- Harari, Yaniv, Gal-Hagit Romano, Lior Ungar, and Martin Kupiec. 2013. "Nature vs Nurture Interplay between the Genetic Control of Telomere Length and Environmental Factors." *Cell Cycle* 12: 3465–70. <https://doi.org/10.4161/cc.26625>.
- Harikumar, Kuzhuvilil B., and Bharat B. Aggarwal. 2008. "Resveratrol: A Multitargeted Agent for Age-Associated Chronic Diseases." *Cell Cycle* 7 (8): 1020–35. <https://doi.org/10.4161/cc.7.8.5740>.
- Hartmann, Jan, Ke Hu, Cynthia Y He, Laurence Pelletier, David S Roos, and Graham Warren. 2006. "Golgi and Centrosome Cycles in *Toxoplasma Gondii*." *Molecular and Biochemical Parasitology* 145 (1): 125–27. <https://doi.org/10.1016/j.molbiopara.2005.09.015>.
- Hasin, Yehudit, Tsviya Olender, Miriam Khen, Claudia Gonzaga-Jauregui, Philip M. Kim, Alexander Eckehart Urban, Michael Snyder, Mark B. Gerstein, Doron Lancet, and Jan O. Korbel. 2008. "High-Resolution Copy-Number Variation Map Reflects Human Olfactory Receptor Diversity and Evolution." *PLOS Genetics* 4 (11): e1000249. <https://doi.org/10.1371/JOURNAL.PGEN.1000249>.
- Heintzman, Nathaniel D., Gary C. Hon, R. David Hawkins, Pouya Kheradpour, Alexander Stark, Lindsey F. Harp, Zhen Ye, et al. 2009. "Histone Modifications at Human Enhancers Reflect Global Cell-Type-Specific Gene Expression." *Nature* 459:7243–7249. <https://doi.org/10.1038/nature07829>.
- Henikoff, Steven, and Yamini Dalal. 2005. "Centromeric Chromatin: ¿What Makes It Unique?" *Current Opinion in Genetics & Development* 15 (2): 177–84. <https://doi.org/10.1016/J.GDE.2005.01.004>.
- Hernandez-Rivas, Rosaura, Karla Pérez-Toledo, Abril Marcela Herrera Solorio, Dulce María Delgadillo, and Miguel Vargas. 2010. "Telomeric Heterochromatin in *Plasmodium Falciparum*." *Journal of Biomedicine and Biotechnology* 2010. <https://doi.org/10.1155/2010/290501>.
- Hoggard, Timothy A, FuJung Chang, Kelsey Rae Perry, Sandya Subramanian, Jessica Kenworthy, Julie Chueng, Erika Shor, et al. 2018. "Yeast Heterochromatin Regulators Sir2 and Sir3 Act Directly at Euchromatic DNA Replication Origins." Edited by Gregory P. Copenhaver. *PLoS Genetics* 14 (5): e1007418. <https://doi.org/10.1371/journal.pgen.1007418>.
- Hoyer, Lois L. 2001. "The ALS Gene Family of *Candida Albicans*." *Trends in Microbiology* 9 (4): 176–80. [https://doi.org/10.1016/S0966-842X\(01\)01984-9](https://doi.org/10.1016/S0966-842X(01)01984-9).
- Hu, Ke. 2008. "Organizational Changes of the Daughter Basal Complex during the Parasite Replication of *Toxoplasma Gondii*." *PLoS Pathogens* 4 (1): e10. <https://doi.org/10.1371/journal.ppat.0040010>.
- Hu, Ke, David S Roos, and John M Murray. 2002. "A Novel Polymer of Tubulin Forms the Conoid of *Toxoplasma Gondii*." *The Journal of Cell Biology* 156 (6). <https://doi.org/10.1083/jcb.200112086>.

- Janouskovec, Jan, Ales Horák, Miroslav Oborník, Julius Lukes, and Patrick J Keeling. 2010. "A Common Red Algal Origin of the Apicomplexan, Dinoflagellate, and Heterokont Plastids." *Proceedings of the National Academy of Sciences of the United States of America* 107 (24): 10949–54. <https://doi.org/10.1073/pnas.1003335107>.
- Janssen, Aniek, Serafin U. Colmenares, and Gary H. Karpen. 2018. "Heterochromatin: Guardian of the Genome." *Annual Review of Cell and Developmental Biology* 34 (1): 265–88. <https://doi.org/10.1146/annurev-cellbio-100617-062653>.
- Jezek, Meagan, Erin Green, Meagan Jezek, and Erin M. Green. 2019. "Histone Modifications and the Maintenance of Telomere Integrity." *Cells* 8 (2): 199. <https://doi.org/10.3390/cells8020199>.
- Jin, Chunyuan, Chongzhi Zang, Gang Wei, Kairong Cui, Weiqun Peng, Keji Zhao, and Gary Felsenfeld. 2009. "H3.3/H2A.Z Double Variant-Containing Nucleosomes Mark 'nucleosome-Free Regions' of Active Promoters and Other Regulatory Regions." *Nature Genetics* 2009 41:8 41 (8): 941–45. <https://doi.org/10.1038/ng.409>.
- Montero JJ, López de Silanes I, Graña O, and Blasco MA. 2016. "Telomeric RNAs Are Essential to Maintain Telomeres." *Nature Communications* 7 (August). <https://doi.org/10.1038/NCOMMS12534>.
- Jofuku, K D, B G den Boer, M van Montagu, and J K Okamoto. 1994. "Control of *Arabidopsis* Flower and Seed Development by the Homeotic Gene APETALA2." *The Plant Cell* 6 (9): 1211–25. <http://www.ncbi.nlm.nih.gov/pubmed/7919989>.
- John, Bernard, Max King, Dieter Schweizer, and Maria Mendelak. 1985. "Equilocality of Heterochromatin Distribution and Heterochromatin Heterogeneity in Acridid Grasshoppers." *Chromosoma* 91 (3–4): 185–200. <https://doi.org/10.1007/BF00328216>.
- Kaeberlein, Matt, Thomas McDonagh, Birgit Heltweg, Jeffrey Hixon, Eric A. Westman, Seth D. Caldwell, Andrew Napper, et al. 2005. "Substrate-Specific Activation of Sirtuins by Resveratrol." *Journal of Biological Chemistry* 280 (17): 17038–45. <https://doi.org/10.1074/jbc.M500655200>.
- Katris, Nicholas J., Giel G. van Dooren, Paul J. McMillan, Eric Hanssen, Leann Tilley, and Ross F. Waller. 2014. "The Apical Complex Provides a Regulated Gateway for Secretion of Invasion Factors in *Toxoplasma*." *PLoS Pathogens* 10 (4). <https://doi.org/10.1371/JOURNAL.PPAT.1004074>.
- Kaviratne, M, S M Khan, W Jarra, and P R Preiser. 2002. "Small Variant STEVOR Antigen Is Uniquely Located within Maurer's Clefts in *Plasmodium Falciparum*-Infected Red Blood Cells." *Eukaryotic Cell* 1 (6): 926–35. <http://www.ncbi.nlm.nih.gov/pubmed/12477793>.

- Kędzierski, Łukasz, Jacqui Montgomery, Joan Curtis, and Emanuela Handman. 2004. "Leucine-Rich Repeats in Host-Pathogen Interactions." *Archivum Immunologiae et Therapiae Experimentalis* 52 (2): 104–12. <https://pubmed.ncbi.nlm.nih.gov/15179324/>.
- "Kerman Aloria – Tesis Doctoral. 1990. <https://dialnet.unirioja.es/servlet/autor?codigo=2337201&msckid=9f853fa6ad5811ecba0075f0b1ebb5d3>.
- Kim, Kami, and Louis M Weiss. 2008. "Toxoplasma: The next 100years." *Microbes and Infection* 10 (9): 978–84. <https://doi.org/10.1016/j.micinf.2008.07.015>.
- Kloet, Frans M. van der, Jeroen Buurmans, Martijs J. Jonker, Age K. Smilde, and Johan A. Westerhuis. 2020. "Increased Comparability between RNA-Seq and Microarray Data by Utilization of Gene Sets." *PLOS Computational Biology* 16 (9): e1008295. <https://doi.org/10.1371/JOURNAL.PCBI.1008295>.
- Kojima, Keitaro, Riyako Ohhashi, Yasunori Fujita, Nanako Hamada, Yukihiro Akao, Yoshinori Nozawa, Takashi Deguchi, and Masafumi Ito. 2008. "A Role for SIRT1 in Cell Growth and Chemoresistance in Prostate Cancer PC3 and DU145 Cells." *Biochemical and Biophysical Research Communications* 373 (3): 423–28. <https://doi.org/10.1016/j.bbrc.2008.06.045>.
- Kravetz, Jeffrey D., and Daniel G. Federman. 2005. "Toxoplasmosis in Pregnancy." *American Journal of Medicine* 118 (3): 212–16. <https://doi.org/10.1016/J.AMJMED.2004.08.023>.
- Kueng, Stephanie, Mariano Oppikofer, and Susan M. Gasser. 2013. "SIR Proteins and the Assembly of Silent Chromatin in Budding Yeast." *Annual Review of Genetics* 47 (1): 275–306. <https://doi.org/10.1146/annurev-genet-021313-173730>.
- Kumar, Sudhir, Glen Stecher, Michael Li, Christina Knyaz, and Koichiro Tamura. 2018. "MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms." *Molecular Biology and Evolution* 35 (6): 1547–49. <https://doi.org/10.1093/MOLBEV/MSY096>.
- Kupiec, Martin. 2014. "Biology of Telomeres: Lessons from Budding Yeast." *FEMS Microbiology Reviews* 38: 144–71. <https://doi.org/10.1111/1574-6976.12054>.
- Kwapisz, Marta, and Antonin Morillon. 2020. "Subtelomeric Transcription and Its Regulation." *Journal of Molecular Biology* 432 (15): 4199–4219. <https://doi.org/10.1016/J.JMB.2020.01.026>.
- Kwapisz, Marta, Myriam Ruault, Erwin van Dijk, Stephanie Gourvennec, Marc Describes, Angela Taddei, and Antonin Morillon. 2015. "Expression of Subtelomeric LncRNAs Links Telomeres Dynamics to RNA Decay in *S. Cerevisiae*." *Non-Coding RNA* 1 (2): 94–126. <https://doi.org/10.3390/NCRNA1020094>.

- Laliberte, Julie, and Vernon B Carruthers. n.d. "Host Cell Manipulation by the Human Pathogen *Toxoplasma Gondii*." Accessed June 30, 2018. <https://doi.org/10.1007/s00018-008-7556-x>.
- Lange, Titia de. 2005. "Shelterin: The Protein Complex That Shapes and Safeguards Human Telomeres." *Genes & Development* 19 (18): 2100–2110. <https://doi.org/10.1101/gad.1346005>.
- Leguizamón, M. S., O. Campetella, G. Russomando, M. Almiron, I. Guillen, S. M. González Cappa, and A. C.C. Frasch. 1994. "Antibodies Inhibiting *Trypanosoma Cruzi* Trans-Sialidase Activity in Sera from Human Infections." *The Journal of Infectious Diseases* 170 (6): 1570–74. <https://doi.org/10.1093/INFDIS/170.6.1570>.
- Lekutis, Christine, David J.P. Ferguson, Michael E. Grigg, Manel Camps, and John C. Boothroyd. 2001. "Surface Antigens of *Toxoplasma Gondii*: Variations on a Theme." *International Journal for Parasitology* 31 (12): 1285–92. [https://doi.org/10.1016/S0020-7519\(01\)00261-2](https://doi.org/10.1016/S0020-7519(01)00261-2).
- Lescault, Pamela J., Ann B. Thompson, Veerupaxagouda Patil, Dario Lirussi, Amanda Burton, Juan Margarit, Jeffrey Bond, and Mariana Matrajt. 2010. "Genomic Data Reveal *Toxoplasma Gondii* Differentiation Mutants Are Also Impaired with Respect to Switching into a Novel Extracellular Tachyzoite State." Edited by Alexander Idnurm. *PLoS ONE* 5 (12): e14463. <https://doi.org/10.1371/journal.pone.0014463>.
- Levine, N D, J O Corliss, F E Cox, G Deroux, J Grain, B M Honigberg, G F Leedale, et al. 1980. "A Newly Revised Classification of the Protozoa." *The Journal of Protozoology* 27 (1): 37–58. <http://www.ncbi.nlm.nih.gov/pubmed/6989987>.
- Lin, San San, Martin Blume, Nicolas von Ahsen, Uwe Gross, and Wolfgang Bohne. 2011. "Extracellular *Toxoplasma Gondii* Tachyzoites Do Not Require Carbon Source Uptake for ATP Maintenance, Gliding Motility and Invasion in the First Hour of Their Extracellular Life." *International Journal for Parasitology* 41 (8): 835–41. <https://doi.org/10.1016/J.IJPARA.2011.03.005>.
- Lirussi, Dario, and Mariana Matrajt. 2011. "RNA Granules Present Only in Extracellular *Toxoplasma Gondii* Increase Parasite Viability." *Int. J. Biol. Sci.* Vol. 7.
- Liu, Alvin Y.C., Lex H.T. van der Ploeg, Frans A.M. Rijsewijk, Piet Borst, and P. Chambon. 1983. "The Transposition Unit of Variant Surface Glycoprotein Gene 118 of *Trypanosoma Brucei*. Presence of Repeated Elements at Its Border and Absence of Promoter-Associated Sequences." *Journal of Molecular Biology* 167 (1): 57–75. [https://doi.org/10.1016/S0022-2836\(83\)80034-5](https://doi.org/10.1016/S0022-2836(83)80034-5).

- Liu, Qian, Xuefeng Zhu, Michelle Lindström, Yonghong Shi, Ju Zheng, Xinxin Hao, Claes M. Gustafsson, and Beidong Liu. 2020. "Yeast Mismatch Repair Components Are Required for Stable Inheritance of Gene Silencing." *PLoS Genetics* 16 (5): e1008798. <https://doi.org/10.1371/JOURNAL.PGEN.1008798>.
- Liu, Wenzhong, Yubin Xie, Jiyong Ma, Xiaotong Luo, Peng Nie, Zhixiang Zuo, Urs Lahrmann, et al. 2015. "IBS: An Illustrator for the Presentation and Visualization of Biological Sequences." *Bioinformatics (Oxford, England)* 31 (20): 3359–61. <https://doi.org/10.1093/BIOINFORMATICS/BTV362>.
- Liu, Yixi, Liva Harinantenaina, Peggy J. Brodie, Jessica D. Bowman, Maria B. Cassera, Carla Slebodnick, Martin W. Callmander, et al. 2013. "Bioactive Compounds from *Stuhlmannia Moavi* from the Madagascar Dry Forest." *Bioorganic & Medicinal Chemistry* 21 (24): 7591–94. <https://doi.org/10.1016/j.bmc.2013.10.038>.
- Li, Yan, Aymeric Leveau, Qiang Zhao, Qi Feng, Hengyun Lu, Jiashun Miao, Zheyong Xue, et al. 2021. "Subtelomeric Assembly of a Multi-Gene Pathway for Antimicrobial Defense Compounds in Cereals." *Nature Communications* 12 (1). <https://doi.org/10.1038/S41467-021-22920-8>.
- Long, Haizhen, Liwei Zhang, Mengjie Lv, Zengqi Wen, Wenhao Zhang, Xiulan Chen, Peitao Zhang, et al. 2019. "H2A.Z Facilitates Licensing and Activation of Early Replication Origins." *Nature* 2019 577:7791 577 (7791): 576–81. <https://doi.org/10.1038/s41586-019-1877-9>.
- Lopez-Rubio, Jose Juan, Liliana Mancio-Silva, and Artur Scherf. 2009. "Genome-Wide Analysis of Heterochromatin Associates Clonally Variant Gene Regulation with Perinuclear Repressive Centers in Malaria Parasites." *Cell Host and Microbe* 5 (2): 179–90. <https://doi.org/10.1016/j.chom.2008.12.012>.
- Lorenzi, Hernan, Asis Khan, Michael S. Behnke, Sivaranjani Namasivayam, Lakshmiapuram S. Swapna, Michalis Hadjithomas, Svetlana Karamycheva, et al. 2016. "Local Admixture of Amplified and Diversified Secreted Pathogenesis Determinants Shapes Mosaic *Toxoplasma Gondii* Genomes." *Nature Communications* 7: 10147. <https://doi.org/10.1038/ncomms10147>.
- Lorestani, Alexander, Lilach Sheiner, Kevin Yang, Seth D. Robertson, Nivedita Sahoo, Carrie F. Brooks, David J P Ferguson, Boris Striepen, and Marc Jan Gubbels. 2010. "A *Toxoplasma* MORN1 Null Mutant Undergoes Repeated Divisions but Is Defective in Basal Assembly, Apicoplast Division and Cytokinesis." *PLoS ONE* 5 (8). <https://doi.org/10.1371/journal.pone.0012302>.
- Maas, Nancy L., Kyle M. Miller, Lisa G. DeFazio, and David P. Toczyski. 2006. "Cell Cycle and Checkpoint Regulation of Histone H3 K56 Acetylation by Hst3 and Hst4." *Molecular Cell* 23 (1): 109–19. <https://doi.org/10.1016/j.molcel.2006.06.006>.

- Makarova, Kira S., L. Aravind, and Eugene v. Koonin. 2000. "A Novel Superfamily of Predicted Cysteine Proteases from Eukaryotes, Viruses and *Chlamydia Pneumoniae*." *Trends in Biochemical Sciences* 25 (2): 50–52. [https://doi.org/10.1016/S0968-0004\(99\)01530-3](https://doi.org/10.1016/S0968-0004(99)01530-3).
- Mancio-Silva, Liliana, Jose Juan Lopez-Rubio, Aurélie Claes, and Artur Scherf. 2013. "Sir2a Regulates RDNA Transcription and Multiplication Rate in the Human Malaria Parasite *Plasmodium Falciparum*." *Nature Communications* 4: 1530. <https://doi.org/10.1038/ncomms2539>.
- Mancio-Silva, Liliana, Ana Paola Rojas-Meza, Miguel Vargas, Artur Scherf, and Rosaura Hernandez-Rivas. 2008. "Differential Association of Orc1 and Sir2 Proteins to Telomeric Domains in *Plasmodium Falciparum*." *Journal of Cell Science* 121 (Pt 12): 2046–53. <https://doi.org/10.1242/jcs.026427>.
- Mann, T, and C Beckers. 2001. "Characterization of the Subpellicular Network, a Filamentous Membrane Skeletal Component in the Parasite *Toxoplasma Gondii*." *Molecular and Biochemical Parasitology* 115 (2): 257–68. <http://www.ncbi.nlm.nih.gov/pubmed/11420112>.
- Martin, Cyrus, and Yi Zhang. 2005. "The Diverse Functions of Histone Lysine Methylation." *Nature Reviews Molecular Cell Biology* 6 (11): 838–49. <https://doi.org/10.1038/nrm1761>.
- Matrajt, Mariana, Sergio O. Angel, Viviana Pszenny, Eduardo Guarnera, David S. Roos, and Juan C. Garberi. 1999. "Arrays of Repetitive DNA Elements in the Largest Chromosomes of *Toxoplasma Gondii*." *Genome* 42 (2): 265–69. <https://doi.org/10.1139/G98-120>.
- McClintock, Barbara. 1929. "CHROMOSOME MORPHOLOGY IN ZEA MAYS." *Science (New York, N.Y.)* 69 (1798): 629. <https://doi.org/10.1126/SCIENCE.69.1798.629>.
- Mefford, Heather C., and Barbara J. Trask. 2002. "The Complex Structure and Dynamic Evolution of Human Subtelomeres." *Nature Reviews. Genetics* 3 (2): 91–102. <https://doi.org/10.1038/NRG727>.
- Mei, Shiyue, Yuanhang Liu, Yue Bao, Yuan Zhang, Siping Min, Yifei Liu, Yun Huang, et al. 2014. "Dendritic Cell-Associated MiRNAs Are Modulated via Chromatin Remodeling in Response to Different Environments." *PLOS ONE* 9 (4): e90231. <https://doi.org/10.1371/JOURNAL.PONE.0090231>.
- Melo, E J de, R O Mayerhoffer, and W de Souza. 2000. "Hydroxyurea Inhibits Intracellular *Toxoplasma Gondii* Multiplication." *FEMS Microbiology Letters* 185 (1): 79–82. <http://www.ncbi.nlm.nih.gov/pubmed/10731610>.
- Mendes, T. A.O., F. P. Lobo, T. S. Rodrigues, G. F. Rodrigues-Luiz, W. D. Darocha, R. T. Fujiwara, S. M.R. Teixeira, and D. C. Bartholomeu. 2013. "Repeat-Enriched Proteins Are Related to Host Cell Invasion and

- Immune Evasion in Parasitic Protozoa.” *Molecular Biology and Evolution* 30 (4): 951–63. <https://doi.org/10.1093/MOLBEV/MST001>.
- Meneghini, Marc D., Michelle Wu, and Hiten D. Madhani. 2003. “Conserved Histone Variant H2A.Z Protects Euchromatin from the Ectopic Spread of Silent Heterochromatin.” *Cell* 112 (5): 725–36. [https://doi.org/10.1016/S0092-8674\(03\)00123-5](https://doi.org/10.1016/S0092-8674(03)00123-5).
- Merrick, Catherine J, and Manoj T Duraisingh. 2007. “*Plasmodium Falciparum* Sir2: An Unusual Sirtuin with Dual Histone Deacetylase and ADP-Ribosyltransferase Activity.” *Eukaryotic Cell* 6 (11): 2081–91. <https://doi.org/10.1128/EC.00114-07>.
- Miao, Jun, Qi Fan, Long Cui, Junsuo Li, Jianyong Li, and Liwang Cui. 2006. “The Malaria Parasite *Plasmodium Falciparum* Histones: Organization, Expression, and Acetylation.” <https://doi.org/10.1016/j.gene.2005.10.022>.
- Michishita, Eriko, Ronald A. McCord, Elisabeth Berber, Mitomu Kioi, Hesus Padilla-Nash, Mara Damian, Peggie Cheung, et al. 2008. “SIRT6 Is a Histone H3 Lysine 9 Deacetylase That Modulates Telomeric Chromatin.” *Nature* 452 (7186): 492–96. <https://doi.org/10.1038/NATURE06736>.
- Miman, Ozlem, Elif Aktan Mutlu, Ozlem Ozcan, Metin Atambay, Rifat Karlidag, and Suheyla Unal. 2010. “Is There Any Role of *Toxoplasma Gondii* in the Etiology of Obsessive–Compulsive Disorder?” *Psychiatry Research* 177 (1–2): 263–65. <https://doi.org/10.1016/j.psychres.2009.12.013>.
- Mishra, Pragyan, Shibangini Beura, Sweta Sikder, Ajit Ku Dhal, Madavan Vasudevan, Manjima Roy, Joydeep Rakshit, Roli Budhwar, Tapas K Kundu, and Rahul Modak. 2022. “Vp1524, a *Vibrio Parahaemolyticus* NAD<sup>+</sup> Dependent Deacetylase, Regulates Host Response during Infection by Induction of Host Histone Deacetylation.” *Journal of Biochemistry*, March. <https://doi.org/10.1093/JB/MVAC027>.
- Moraes Barros, Roberto R, Marjorie M Marini, Cristiane Antônio, Danielle R Cortez, Andrea M Miyake, Fábio M Lima, Jeronimo C Ruiz, et al. 2012. “Anatomy and Evolution of Telomeric and Subtelomeric Regions in the Human Protozoan Parasite *Trypanosoma Cruzi*.” *BMC Genomics* 13 (1): 229. <https://doi.org/10.1186/1471-2164-13-229>.
- Morrisette, Naomi. 2015. “Targeting Toxoplasma Tubules: Tubulin, Microtubules, and Associated Proteins in a Human Pathogen.” <https://doi.org/10.1128/EC.00225-14>.
- Morrisette, Naomi S, and L David Sibley. 2002. “Cytoskeleton of Apicomplexan Parasites.” *Microbiology and Molecular Biology Reviews: MMBR* 66 (1): 21–38; table of contents. <http://www.ncbi.nlm.nih.gov/pubmed/11875126>.

- Mostoslavsky, Raul, Katrin F. Chua, David B. Lombard, Wendy W. Pang, Miriam R. Fischer, Lionel Gellon, Pingfang Liu, et al. 2006. "Genomic Instability and Aging-like Phenotype in the Absence of Mammalian SIRT6." *Cell* 124 (2): 315–29. <https://doi.org/10.1016/J.CELL.2005.11.044>.
- Munera López, Jonathan, Agustina Ganuza, Silvina S Bogado, Daniela Muñoz, Diego M Ruiz, William J Sullivan, Laura Vanagas, and Sergio O Angel. 2019. "Evaluation of ATM Kinase Inhibitor KU-55933 as Potential Anti-*Toxoplasma Gondii* Agent." *Frontiers in Cellular and Infection Microbiology* 9 (February): 26. <https://doi.org/10.3389/fcimb.2019.00026>.
- Nardelli, Sheila C., Natalie C. Silmon de Monerri, Laura Vanagas, Xiaonan Wang, Zoi Tampaki, William J. Sullivan, Sergio O. Angel, and Kami Kim. 2022. "Genome-Wide Localization of Histone Variants in *Toxoplasma Gondii* Implicates Variant Exchange in Stage-Specific Gene Expression." *BMC Genomics* 23 (1). <https://doi.org/10.1186/S12864-022-08338-6>.
- North, Brian J., and Eric Verdin. 2004. "Sirtuins: Sir2-Related NAD-Dependent Protein Deacetylases." *Genome Biology* 5 (5). <https://doi.org/10.1186/GB-2004-5-5-224>.
- Ohme-Takagi, M., and H Shinshi. 1995. "Ethylene-Inducible DNA Binding Proteins That Interact with an Ethylene-Responsive Element." *THE PLANT CELL ONLINE* 7 (2): 173–82. <https://doi.org/10.1105/tpc.7.2.173>.
- Olins, Ada L., Aurélie Ernst, Monika Zwerger, Harald Herrmann, and Donald E. Olins. 2010. "An in Vitro Model for Pelger-Huët Anomaly." *Nucleus* 1 (6): 506–12. <https://doi.org/10.4161/nucl.1.6.13271>.
- Oliveira, Tatiane R, Carmen Fernandez-Becerra, Maria Carolina S Jimenez, Hernando A del Portillo, and Irene S Soares. 2006. "Evaluation of the Acquired Immune Responses to *Plasmodium Vivax* VIR Variant Antigens in Individuals Living in Malaria-Endemic Areas of Brazil." *Malaria Journal* 5 (1): 83. <https://doi.org/10.1186/1475-2875-5-83>.
- Omelianczyk, Radoslaw Igor, Han Ping Loh, Marvin Chew, Regina Hoo, Sebastian Baumgarten, Laurent Renia, Jianzhu Chen, and Peter R Preiser. 2020. "Rapid Activation of Distinct Members of Multigene Families in *Plasmodium* Spp." <https://doi.org/10.1038/s42003-020-1081-3>.
- Ota, H, E Tokunaga, K Chang, M Hikasa, K Iijima, M Eto, K Kozaki, M Akishita, Y Ouchi, and M Kaneki. 2006. "Sirt1 Inhibitor, Sirtinol, Induces Senescence-like Growth Arrest with Attenuated Ras-MAPK Signaling in Human Cancer Cells." *Oncogene* 25 (2): 176–85. <https://doi.org/10.1038/sj.onc.1209049>.

- Ottaviani, Alexandre, Eric Gilson, and Frédérique Magdinier. 2008. "Telomeric Position Effect: ¿From the Yeast Paradigm to Human Pathologies?" *Biochimie* 90 (1): 93–107. <https://doi.org/10.1016/j.biochi.2007.07.022>.
- Painter, Heather J., Tracey L. Campbell, and Manuel Llinás. 2011. "The Apicomplexan AP2 Family: Integral Factors Regulating Plasmodium Development." *Molecular and Biochemical Parasitology* 176 (1): 1–7. <https://doi.org/10.1016/j.molbiopara.2010.11.014>.
- Palm, Wilhelm, and Titia de Lange. 2008. "How Shelterin Protects Mammalian Telomeres." *Annual Review of Genetics* 42 (1): 301–34. <https://doi.org/10.1146/annurev.genet.41.110306.130350>.
- Pays, Etienne, Marie France Delauw, Suzanne van Assel, Monique Laurent, Tony Vervoort, Nestor van Meirvenne, and Maurice Steinert. 1983. "Modifications of a *Trypanosoma Brucei* Antigen Gene Repertoire by Different DNA Recombinational Mechanisms." *Cell* 35 (3 Pt 2): 721–31. [https://doi.org/10.1016/0092-8674\(83\)90105-8](https://doi.org/10.1016/0092-8674(83)90105-8).
- Pays, Etienne, Luc Vanhamme, and Magali Berberof. 1994. "Genetic Controls for the Expression of Surface Antigens in African Trypanosomes." *Annual Review of Microbiology* 48: 25–52. <https://doi.org/10.1146/ANNUREV.MI.48.100194.000325>.
- Peck, Barrie, Chun-Yuan Chen, Ka-Kei Ho, Paolo di Fruscia, Stephen S Myatt, R Charles Coombes, Matthew J Fuchter, Chwan-Deng Hsiao, and Eric W-F Lam. 2010. "SIRT Inhibitors Induce Cell Death and P53 Acetylation through Targeting Both SIRT1 and SIRT2." *Molecular Cancer Therapeutics* 9 (4): 844–55. <https://doi.org/10.1158/1535-7163.MCT-09-0971>.
- Pérez-Toledo, Karla, Ana Paola Rojas-Meza, Liliana Mancio-Silva, Nora Adriana Hernández-Cuevas, Dulce Maria Delgadillo, Miguel Vargas, Santiago Martínez-Calvillo, Artur Scherf, and Rosaura Hernandez-Rivas. 2009. "*Plasmodium Falciparum* Heterochromatin Protein 1 Binds to Tri-Methylated Histone 3 Lysine 9 and Is Linked to Mutually Exclusive Expression of Var Genes." *Nucleic Acids Research* 37 (8): 2596–2606. <https://doi.org/10.1093/nar/gkp115>.
- Petter, Michaela, Malin Haeggström, Ayman Khattab, Victor Fernandez, Mo-Quen Klinkert, and Mats Wahlgren. 2007. "Variant Proteins of the *Plasmodium Falciparum* RIFIN Family Show Distinct Subcellular Localization and Developmental Expression Patterns." *Molecular and Biochemical Parasitology* 156 (1): 51–61. <https://doi.org/10.1016/j.molbiopara.2007.07.011>.
- Petter, Michaela, Chin Chin Lee, Timothy J Byrne, Katja E Boysen, Jennifer Volz, Stuart A Ralph, Alan F Cowman, Graham v Brown, and Michael F Duffy. 2011. "Expression of *P. Falciparum* Var Genes Involves

Exchange of the Histone Variant H2A.Z at the Promoter.” *PLoS Pathogens* 7 (2): e1001292. <https://doi.org/10.1371/journal.ppat.1001292>.

- Petter, Michaela, Shamista A. Selvarajah, Chin Chin Lee, Wai Hoe Chin, Archana P. Gupta, Zbynek Bozdech, Graham v. Brown, and Michael F. Duffy. 2013. “H2A.Z and H2B.Z Double-Variant Nucleosomes Define Intergenic Regions and Dynamically Occupy *Var* Gene Promoters in the Malaria Parasite *Plasmodium Falciparum*.” *Molecular Microbiology* 87 (6): 1167–82. <https://doi.org/10.1111/mmi.12154>.
- Pomel, Sebastien, Flora C. Y. Luk, and Con J. M. Beckers. 2008. “Host Cell Egress and Invasion Induce Marked Relocations of Glycolytic Enzymes in *Toxoplasma Gondii* Tachyzoites.” Edited by Dominique Soldati-Favre. *PLoS Pathogens* 4 (10): e1000188. <https://doi.org/10.1371/journal.ppat.1000188>.
- Portillo, Hernando A. del, Michael Lanzer, Sergio Rodriguez-Malaga, Fidel Zavala, and Carmen Fernandez-Becerra. 2004. “Variant Genes and the Spleen in *Plasmodium Vivax* Malaria.” *International Journal for Parasitology* 34 (13–14): 1547–54. <https://doi.org/10.1016/j.ijpara.2004.10.012>.
- Price, Robert Jordan, Esther Weindling, Judith Berman, and Alessia Buscaino. 2019. “Chromatin Profiling of the Repetitive and Nonrepetitive Genomes of the Human Fungal Pathogen *Candida Albicans*.” *MBio* 10 (4). <https://doi.org/10.1128/MBIO.01376-19>.
- Probst, Aline v., Elaine Dunleavy, and Geneviève Almouzni. 2009. “Epigenetic Inheritance during the Cell Cycle.” *Nature Reviews Molecular Cell Biology* 10 (3): 192–206. <https://doi.org/10.1038/nrm2640>.
- Radke, Joshua B, Olivier Lucas, Erandi K de Silva, YanFen Ma, William J Sullivan, Louis M Weiss, Manuel Llinas, and Michael W White. 2013. “ApiAP2 Transcription Factor Restricts Development of the *Toxoplasma* Tissue Cyst.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (17): 6871–76. <https://doi.org/10.1073/pnas.1300059110>.
- Radke, JayR, MichaelS Behnke, AaronJ Mackey, JoshB Radke, DavidS Roos, and MichaelW White. 2005. “The Transcriptome of *Toxoplasma Gondii*.” *BMC Biology* 3 (1): 26. <https://doi.org/10.1186/1741-7007-3-26>.
- Radke, Jay R., Michael N. Guerini, Maria Jerome, and Michael W. White. 2003. “A Change in the Premitotic Period of the Cell Cycle Is Associated with Bradyzoite Differentiation in *Toxoplasma Gondii*.” *Molecular and Biochemical Parasitology* 131 (2): 119–27. [https://doi.org/10.1016/S0166-6851\(03\)00198-1](https://doi.org/10.1016/S0166-6851(03)00198-1).
- Radke, Jay R, Boris Striepen, Michael N Guerini, Maria E Jerome, David S Roos, and Michael W White. 2001. “Defining the Cell Cycle for the Tachyzoite Stage of *Toxoplasma Gondii*.” *Molecular and Biochemical Parasitology* 115 (2): 165–75. [https://doi.org/10.1016/S0166-6851\(01\)00284-5](https://doi.org/10.1016/S0166-6851(01)00284-5).

- Ragunathan, Kaushik, Gloria Jih, and Danesh Moazed. 2015. "Epigenetic Inheritance Uncoupled from Sequence-Specific Recruitment." *Science* 348 (6230): 1258699. <https://doi.org/10.1126/science.1258699>.
- Ramirez, Jose Luis. 2019. "An Evolutionary View of *Trypanosoma Cruzi* Telomeres." *Frontiers in Cellular and Infection Microbiology* 9 (January): 439. <https://doi.org/10.3389/FCIMB.2019.00439>.
- Razakantoanina, Valérie, Isabelle Florent, and Ginette Jaureguiberry. 2008. "Plasmodium Falciparum: Functional Mitochondrial ADP/ATP Transporter in *Escherichia Coli* Plasmic Membrane as a Tool for Selective Drug Screening." *Experimental Parasitology* 118 (2): 181–87. <https://doi.org/10.1016/J.EXPPARA.2007.07.015>.
- Reddy, Bharat D, Yu Wang, Lifang Niu, Emily C Higuchi, Samuel B Marguerat, Jürg Bähler, Gerald R Smith, and Songtao Jia. 2011. "Elimination of a Specific Histone H3K14 Acetyltransferase Complex Bypasses the RNAi Pathway to Regulate Pericentric Heterochromatin Functions." *Genes & Development* 25 (3): 214–19. <https://doi.org/10.1101/gad.1993611>.
- Reid, Adam James, Sarah J Vermont, James a Cotton, David Harris, Grant a Hill-Cawthorne, Stephanie Könen-Waisman, Sophia M Latham, et al. 2012. "Comparative Genomics of the Apicomplexan Parasites *Toxoplasma Gondii* and *Neospora Caninum*: Coccidia Differing in Host Range and Transmission Strategy." *PLoS Pathogens* 8 (3): e1002567. <https://doi.org/10.1371/journal.ppat.1002567>.
- Reynolds, T. B., and G. R. Fink. 2001. "Bakers' Yeast, a Model for Fungal Biofilm Formation." *Science (New York, N.Y.)* 291 (5505): 878–81. <https://doi.org/10.1126/SCIENCE.291.5505.878>.
- Robinson, James T., Helga Thorvaldsdóttir, Wendy Winckler, Mitchell Guttman, Eric S. Lander, Gad Getz, and Jill P. Mesirov. 2011. "Integrative Genomics Viewer." *Nature Biotechnology* 29 (1): 24–26. <https://doi.org/10.1038/nbt.1754>.
- Rogakou, Emmy P., Duane R. Pilch, Ann H. Orr, Vessela S. Ivanova, and William M. Bonner. 1998. "DNA Double-Stranded Breaks Induce Histone H2AX Phosphorylation on Serine 139." *Journal of Biological Chemistry* 273 (10): 5858–68. <https://doi.org/10.1074/jbc.273.10.5858>.
- Rusche, Laura N., Ann L. Kirchmaier, and Jasper Rine. 2003. "The Establishment, Inheritance, and Function of Silenced Chromatin in *Saccharomyces Cerevisiae*." *Annual Review of Biochemistry* 72 (1): 481–516. <https://doi.org/10.1146/annurev.biochem.72.121801.161547>.
- Salcedo-Amaya, A. M., M. A. van Driel, B. T. Alako, M. B. Trelle, A. M. G. van den Elzen, A. M. Cohen, E. M. Janssen-Megens, et al. 2009. "Dynamic Histone H3 Epigenome Marking during the Intraerythrocytic Cycle

- of *Plasmodium Falciparum*." *Proceedings of the National Academy of Sciences* 106 (24): 9655–60. <https://doi.org/10.1073/pnas.0902515106>.
- Sato, Shigeharu. 2011. "The Apicomplexan Plastid and Its Evolution." *Cellular and Molecular Life Sciences* 68 (8): 1285–96. <https://doi.org/10.1007/s00018-011-0646-1>.
- Sautel, Céline F., Dominique Cannella, Olivier Bastien, Sylvie Kieffer, Delphine Aldebert, Jérôme Garin, Isabelle Tardieux, Hassan Belrhali, and Mohamed-Ali Hakimi. 2007. "SET8-Mediated Methylations of Histone H4 Lysine 20 Mark Silent Heterochromatic Domains in Apicomplexan Genomes." *Molecular and Cellular Biology* 27 (16): 5711–24. <https://doi.org/10.1128/MCB.00482-07>.
- Scherf, Artur, Luisa M Figueiredo, and Lúcio H Freitas-junior. 2001. "Plasmodium Telomeres: A Pathogen's Perspective," 409–14.
- Schmittgen, Thomas D, and Kenneth J Livak. 2008. "Analyzing Real-Time PCR Data by the Comparative C(T) Method." *Nature Protocols* 3 (6): 1101–8. <http://www.ncbi.nlm.nih.gov/pubmed/18546601>.
- Schotta, G., Monika Lachner, Kavitha Sarma, Anja Ebert, Roopsha Sengupta, Gunter Reuter, Danny Reinberg, and Thomas Jenuwein. 2004. "A Silencing Pathway to Induce H3-K9 and H4-K20 Trimethylation at Constitutive Heterochromatin." *Genes & Development* 18 (11): 1251–62. <https://doi.org/10.1101/gad.300704>.
- Serrano, Lourdes, Paloma Martínez-Redondo, Anna Marazuela-Duque, Berta N Vazquez, Scott J Dooley, Philipp Voigt, David B Beck, et al. 2012. "The Tumor Suppressor SirT2 Regulates Cell Cycle Progression and Genome Stability by Modulating the Mitotic Deposition of H4K20 Methylation." <https://doi.org/10.1101/gad.211342.112>.
- Sgambato, A, R Ardito, B Faraglia, A Boninsegna, F I Wolf, and A Cittadini. 2001. "Resveratrol, a Natural Phenolic Compound, Inhibits Cell Proliferation and Prevents Oxidative DNA Damage." *Mutation Research* 496 (1–2): 171–80. <http://www.ncbi.nlm.nih.gov/pubmed/11551493>.
- Shankaranarayana, Gurumurthy D., Mohammad R. Motamedi, Danesh Moazed, and Shiv I.S. Grewal. 2003. "Sir2 Regulates Histone H3 Lysine 9 Methylation and Heterochromatin Assembly in Fission Yeast." *Current Biology : CB* 13 (14): 1240–46. [https://doi.org/10.1016/S0960-9822\(03\)00489-5](https://doi.org/10.1016/S0960-9822(03)00489-5).
- Sheffield, H G, and M L Melton. 1968. "The Fine Structure and Reproduction of *Toxoplasma Gondii*." *The Journal of Parasitology* 54 (2): 209–26. <http://www.ncbi.nlm.nih.gov/pubmed/5647101>.

- Shia, W.-J., Bing Li, and Jerry L Workman. 2006. "SAS-Mediated Acetylation of Histone H4 Lys 16 Is Required for H2A.Z Incorporation at Subtelomeric Regions in *Saccharomyces Cerevisiae*." *Genes & Development* 20 (18): 2507–12. <https://doi.org/10.1101/gad.1439206>.
- Sibley, L David, Asis Khan, James W Ajioka, and Benjamin M Rosenthal. 2009. "Genetic Diversity of *Toxoplasma Gondii* in Animals and Humans." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 364 (1530): 2749–61. <https://doi.org/10.1098/rstb.2009.0087>.
- Sidik, Saima M., Diego Huet, Suresh M. Ganesan, My Hang Huynh, Tim Wang, Armiyaw S. Nasamu, Prathapan Thiru, et al. 2016. "A Genome-Wide CRISPR Screen in *Toxoplasma* Identifies Essential Apicomplexan Genes." *Cell* 166 (6): 1423-1435.e12. <https://doi.org/10.1016/j.cell.2016.08.019>.
- Sindikubwabo, Fabien, Shuai Ding, Tahir Hussain, Philippe Ortet, Mohamed Barakat, Sebastian Baumgarten, Dominique Cannella, et al. 2017. "Modifications at K31 on the Lateral Surface of Histone H4 Contribute to Genome Structure and Expression in Apicomplexan Parasites." *ELife* 6 (November). <https://doi.org/10.7554/ELIFE.29391>.
- Singh, Santosh Kumar, Saswati Banerjee, Edward P Acosta, James W Lillard, and Rajesh Singh. 2017. "Resveratrol Induces Cell Cycle Arrest and Apoptosis with Docetaxel in Prostate Cancer Cells via a P53/P21WAF1/CIP1 and P27KIP1 Pathway." *Oncotarget* 8 (10): 17216–28. <https://doi.org/10.18632/oncotarget.15303>.
- Striepen, Boris, Carly N Jordan, Sarah Reiff, and Giel G van Dooren. 2007. "Building the Perfect Parasite: Cell Division in Apicomplexa." *PLoS Pathogens* 3 (6): e78. <https://doi.org/10.1371/journal.ppat.0030078>.
- Strobl, Jeannine S, Meredith Cassell, Sheila M Mitchell, Christopher M Reilly, and David S Lindsay. 2007. "Scriptaid and Suberoylanilide Hydroxamic Acid Are Histone Deacetylase Inhibitors with Potent Anti-*Toxoplasma Gondii* Activity in Vitro." *J. Parasitol.* 93 (3): 694–700.
- Suka, Noriyuki, Kunheng Luo, and Michael Grunstein. 2002. "Sir2p and Sas2p Opposingly Regulate Acetylation of Yeast Histone H4 Lysine16 and Spreading of Heterochromatin." *Nature Genetics* 32 (3): 378–83. <https://doi.org/10.1038/ng1017>.
- Talbert, Paul B, Kami Ahmad, Geneviève Almouzni, Juan Ausió, Frederic Berger, Prem L Bhalla, William M Bonner, et al. 2012. "A Unified Phylogeny-Based Nomenclature for Histone Variants." *Epigenetics & Chromatin* 5 (1): 7. <https://doi.org/10.1186/1756-8935-5-7>.
- Talbert, Paul B, and Steven Henikoff. 2006. "Spreading of Silent Chromatin: Inaction at a Distance." *Nature Reviews. Genetics* 7 (10): 793–803. <https://doi.org/10.1038/nrg1920>.

- Tanny, Jason C., Gustave J. Dowd, Julie Huang, Helmuth Hilz, and Danesh Moazed. 1999. "An Enzymatic Activity in the Yeast Sir2 Protein That Is Essential for Gene Silencing." *Cell* 99 (7): 735–45. [https://doi.org/10.1016/S0092-8674\(00\)81671-2](https://doi.org/10.1016/S0092-8674(00)81671-2).
- Tenter, a M, a R Heckerth, and L M Weiss. 2000. "Toxoplasma Gondii: From Animals to Humans." *International Journal for Parasitology* 30 (12–13): 1217–58. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3109627&tool=pmcentrez&rendertype=abstract>.
- Tenter, A M, and A M Johnson. 1997. "Phylogeny of the Tissue Cyst-Forming Coccidia." *Advances in Parasitology* 39: 69–139. <http://www.ncbi.nlm.nih.gov/pubmed/9241815>.
- Teunissen, A. W. R. H., and H. Y. Steensma. 1995. "The Dominant Flocculation Genes of *Saccharomyces Cerevisiae* Constitute a New Subtelomeric Gene Family." *Yeast* 11 (11): 1001–13. <https://doi.org/10.1002/YEA.320111102>.
- Thomazeau K, Curien G, Dumas R, and Biou V. 2001. "Crystal Structure of Threonine Synthase from *Arabidopsis Thaliana*." *Protein Science: A Publication of the Protein Society* 10 (3): 638–48. <https://doi.org/10.1110/PS.44301>.
- Thurtle, Deborah M, and Jasper Rine. 2013. "The Molecular Topography of Silenced Chromatin in *Saccharomyces Cerevisiae*." <https://doi.org/10.1101/gad.230532.113>.
- Tonkin, Christopher J., Céline K. Carret, Manoj T. Duraisingh, Till S. Voss, Stuart A. Ralph, Mirja Hommel, Michael F. Duffy, et al. 2009. "Sir2 Paralogues Cooperate to Regulate Virulence Genes and Antigenic Variation in *Plasmodium Falciparum*." *PLoS Biology* 7 (4): 0771–88. <https://doi.org/10.1371/journal.pbio.1000084>.
- Trojer, Patrick, and Danny Reinberg. 2007. "Facultative Heterochromatin: ¿Is There a Distinctive Molecular Signature?" *Molecular Cell* 28 (1): 1–13. <https://doi.org/10.1016/j.molcel.2007.09.011>.
- Tuteja, Renu, and Narendra Tuteja. 2000. "Ku Autoantigen: A Multifunctional DNA-Binding Protein." *Critical Reviews in Biochemistry and Molecular Biology* 35 (1): 1–33. <https://doi.org/10.1080/10409230091169177>.
- Tyagi, Alpna, Mallikarjuna Gu, Takenori Takahata, Barbara Frederick, Chapla Agarwal, Sunitha Siriwardana, Rajesh Agarwal, and Robert A. Sclafani. 2011. "Resveratrol Selectively Induces DNA Damage, Independent of Smad4 Expression, in Its Efficacy against Human Head and Neck Squamous Cell

- Carcinoma." *Clinical Cancer Research* 17 (16): 5402–11. <https://doi.org/10.1158/1078-0432.CCR-11-1072>.
- Vanagas, Laura, Susana Marisol Contreras, and Sergio Oscar Angel. 2019. *Apicomplexa and Histone Variants: What's New ?*. InTechOpen. DOI: 10.5772/intechopen.81409.
- Vanagas, Laura, Maria C Dalmasso, Jean F Dubremetz, Enrique L Portiansky, Donald E Olins, and Sergio O Angel. 2013. "Epichromatin Is Conserved in *Toxoplasma Gondii* and Labels the Exterior Parasite Chromatin throughout the Cell Cycle NIH Public Access." *Parasitology* 140 (9): 1104–10. <https://doi.org/10.1017/S0031182013000504>.
- Vanagas, Laura, Victoria Jeffers, and Ss Bogado. 2012. "Toxoplasma Histone Acetylation Remodelers as Novel Drug Targets." *Expert Review of Anti-Infective Therapy* 10 (10): 1189–1201. <https://doi.org/10.1586/eri.12.100>.
- Vaquero, Alejandro, Michael Scher, Donghoon Lee, Hediye Erdjument-Bromage, Paul Tempst, and Danny Reinberg. 2004. "Human SirT1 Interacts with Histone H1 and Promotes Formation of Facultative Heterochromatin." *Molecular Cell* 16 (1): 93–105. <https://doi.org/10.1016/j.molcel.2004.08.031>.
- Vaquero, A, R Sternglanz, and D Reinberg. 2007. "NAD P-Dependent Deacetylation of H4 Lysine 16 by Class III HDACs." *Oncogene* 26: 5505–20. <https://doi.org/10.1038/sj.onc.1210617>.
- Varunan, Shalu M., Jaishree Tripathi, Sunanda Bhattacharyya, Tanvi Suhane, and Mrinal Kanti Bhattacharyya. 2013. "Plasmodium Falciparum Origin Recognition Complex Subunit 1 (PfOrc1) Functionally Complements  $\Delta$ sir3 Mutant of *Saccharomyces Cerevisiae*." *Molecular and Biochemical Parasitology* 191 (1): 28–35. <https://doi.org/10.1016/J.MOLBIOPARA.2013.08.004>.
- Vempati, Rahul Kumar. 2012. "DNA Damage in the Presence of Chemical Genotoxic Agents Induce Acetylation of H3K56 and H4K16 but Not H3K9 in Mammalian Cells." *Molecular Biology Reports* 39 (1): 303–8. <https://doi.org/10.1007/s11033-011-0739-9>.
- Verdaasdonk, Jolien S., and Kerry Bloom. 2011. "Centromeres: Unique Chromatin Structures That Drive Chromosome Segregation." *Nature Reviews. Molecular Cell Biology* 12 (5): 320–32. <https://doi.org/10.1038/NRM3107>.
- Vergnes, Baptiste, Laurent Vanhille, Ali Ouaisi, and Denis Sereno. 2005. "Stage-Specific Antileishmanial Activity of an Inhibitor of SIR2 Histone Deacetylase." *Acta Tropica* 94: 107–15. <https://doi.org/10.1016/j.actatropica.2005.03.004>.

- Vittecoq, Marion, Eric Elguero, Kevin D. Lafferty, Benjamin Roche, Jacques Brodeur, Michel Gauthier-Clerc, Dorothée Missé, and Frédéric Thomas. 2012. "Brain Cancer Mortality Rates Increase with *Toxoplasma Gondii* Seroprevalence in France." *Infection, Genetics and Evolution* 12 (2): 496–98. <https://doi.org/10.1016/j.meegid.2012.01.013>.
- Volpe, T. A., Catherine Kidner, Ira M Hall, Grace Teng, Shiv I S Grewal, and Robert A Martienssen. 2002. "Regulation of Heterochromatic Silencing and Histone H3 Lysine-9 Methylation by RNAi." *Science* 297 (5588): 1833–37. <https://doi.org/10.1126/science.1074973>.
- Vonlaufen, Nathalie, Arunasalam Naguleswaran, Isabelle Coppens, and William J Sullivan. 2010. "MYST Family Lysine Acetyltransferase Facilitates Ataxia Telangiectasia Mutated (ATM) Kinase-Mediated DNA Damage Response in *Toxoplasma Gondii*." *The Journal of Biological Chemistry* 285 (15): 11154–61. <https://doi.org/10.1074/jbc.M109.066134>.
- Walker, Robert, Mathieu Gissot, Matthew M. Croken, Ludovic Huot, David Hot, Kami Kim, and Stanislas Tomavo. 2013. "The *Toxoplasma* Nuclear Factor TgAP2XI-4 Controls Bradyzoite Gene Expression and Cyst Formation." *Molecular Microbiology* 87 (December 2012): 641–55. <https://doi.org/10.1111/mmi.12121>.
- Walker, Robert, Mathieu Gissot, Ludovic Huot, Tchilabalo Dilezitoko Alayi, David Hot, Guillemette Marot, Christine Schaeffer-Reiss, Alain van Dorsselaer, Kami Kim, and Stanislas Tomavo. 2013. "*Toxoplasma* Transcription Factor TgAP2XI-5 Regulates the Expression of Genes Involved in Parasite Virulence and Host Invasion." *Journal of Biological Chemistry* 288 (43): 31127–38. <https://doi.org/10.1074/jbc.M113.486589>.
- Wang, Jiyong, Sharon T. Jia, and Songtao Jia. 2016. "New Insights into the Regulation of Heterochromatin." *Trends in Genetics* 32 (5): 284–94. <https://doi.org/10.1016/j.tig.2016.02.005>.
- Wang, Jiyong, Stephanie T. Lawry, Allison L. Cohen, and Songtao Jia. 2014. "Chromosome Boundary Elements and Regulation of Heterochromatin Spreading." *Cellular and Molecular Life Sciences* 71 (24): 4841–52. <https://doi.org/10.1007/s00018-014-1725-x>.
- Weiss, Louis M., and Jitender. P. Dubey. 2009. "Toxoplasmosis: A History of Clinical Observations." *International Journal for Parasitology* 39 (8): 895–901. <https://doi.org/10.1016/j.ijpara.2009.02.004>.
- Xia, Jing, Aarthi Venkat, Rachel E. Bainbridge, Michael L. Reese, Karine G. le Roch, Ferhat Ay, and Jon P. Boyle. 2021. "Third-Generation Sequencing Revises the Molecular Karyotype for *Toxoplasma Gondii* and Identifies Emerging Copy Number Variants in Sexual Recombinants." *Genome Research* 31 (5): 834–51. <https://doi.org/10.1101/GR.262816.120>.

- Xu, W S, R B Parmigiani, and P a Marks. 2007. "Histone Deacetylase Inhibitors: Molecular Mechanisms of Action." *Oncogene* 26: 5541–52. <https://doi.org/10.1038/sj.onc.1210620>.
- Yolken, R. H., F. B Dickerson, and E. Fuller Torrey. 2009. "Toxoplasma and Schizophrenia." *Parasite Immunology* 31 (11): 706–15. <https://doi.org/10.1111/j.1365-3024.2009.01131.x>.
- Zhang, Haiying, Douglas N. Roberts, and Bradley R. Cairns. 2005. "Genome-Wide Dynamics of Htz1, a Histone H2A Variant That Poises Repressed/Basal Promoters for Activation through Histone Loss." *Cell* 123 (2): 219–31. <https://doi.org/10.1016/j.cell.2005.08.036>.
- Zhang, Ke, Kerstin Mosch, Wolfgang Fischle, and Shiv I S Grewal. 2008. "Roles of the Clr4 Methyltransferase Complex in Nucleation, Spreading and Maintenance of Heterochromatin." *Nature Structural & Molecular Biology* 15 (4): 381–88. <https://doi.org/10.1038/nsmb.1406>.
- Zhou, Wei, Juan-Hua Quan, Young-Ha Lee, Dae-Whan Shin, and Guang-Ho Cha. 2013. "Toxoplasma Gondii Proliferation Require Down-Regulation of Host Nox4 Expression via Activation of PI3 Kinase/Akt Signaling Pathway." Edited by Ira J. Blader. *PloS One* 8 (6): e66306. <https://doi.org/10.1371/journal.pone.0066306>.

Esta versión corresponde a la última revisión con correcciones sugeridas por la Dra. Marcela Cucher, jurada de esta tesis.

Lic. Susana Marisol Contreras

A handwritten signature in blue ink, consisting of stylized initials and a surname.

Dr. Sergio O. Angel

A handwritten signature in blue ink, featuring a cursive style with a prominent horizontal stroke at the end.