



---

# Búsqueda *in silico* de posibles blancos drogables en kinetopláستidos

---

Br. Martín Rivara Espasandín

Tutor: Pablo Smircich

**Tesina de grado**

**Licenciatura en Ciencias Biológicas, Facultad de Ciencias, Udelar**

Departamento de Genómica

Instituto de Investigaciones Biológicas Clemente Estable

Montevideo, Uruguay

2021



# Resumen

---

Los kinetoplástidos son un grupo de protozoarios flagelados, caracterizados por la presencia de una estructura compuesta por el ADN mitocondrial, llamado kinetoplasto. El grupo incluye algunos parásitos que son agentes etiológicos de enfermedades características de regiones tropicales y subtropicales, con mayor prevalencia en países subdesarrollados. Existen tres enfermedades de importancia causadas por tripanosomátidos (un subgrupo dentro de los kinetoplástidos). La tripanosomiasis americana, causada por *Trypanosoma cruzi*, la tripanosomiasis africana, causada por *Trypanosoma brucei* y la leishmaniasis, causada por distintas especies del género *Leishmania spp.* Estas forman parte de las llamadas enfermedades desatendidas, ya que históricamente se les han dedicado escasos recursos para combatirlas. Esto último, ha llevado a que los tratamientos actuales tengan efectos secundarios importantes y sean efectivos en etapas específicas de la enfermedad, haciendo que la búsqueda de nuevos blancos y fármacos sea un área de suma importancia.

En este sentido, existen diferentes alternativas para la búsqueda de nuevas drogas y todas requieren muchos recursos. Una aproximación interesante, es buscar nuevas drogas basándonos en el blanco objetivo (*target based drug discovery*), de forma de enfocar los esfuerzos en buscar moléculas que puedan afectar la actividad de un blanco particular. Para esto es necesario determinar blancos que cumplan ciertos requisitos que los hagan buenos candidatos a ser posibles blancos de droga.

En este trabajo, utilizamos una aproximación computacional para buscar candidatos a ser posibles blancos drogables comunes a *T. cruzi*, *T. brucei* y *L. major*, a nivel proteico. Realizamos agrupamientos de secuencias para determinar ortólogos entre los tres parásitos y utilizamos la base de datos Target Pathogen, una herramienta que permite la integración de información proteica como función, rol en redes metabólicas, esencialidad, similaridad con proteínas humanas, así como propiedades estructurales como drogabilidad, entre otras, con el objetivo de obtener una selección de proteínas que cumplan ciertos requisitos. Particularmente en este estudio, evaluamos todo el proteoma de los parásitos estudiando drogabilidad, similaridad con proteínas humanas, importancia en redes metabólicas, así como datos de expresión, encontrando 133 candidatos que tienen condiciones para ser posibles blancos drogables en los tres parásitos. Estos resultados representan una primera aproximación a la búsqueda de nuevos blancos drogables en tripanosomátidos, ofreciendo una base sobre la cual seguir explorando.

# Tabla de contenidos

---

<b>1. Introducción</b>	<b>6</b>
1.1. Los kinetoplastidos	6
1.2. Tripanosomiasis americana	6
1.3. Tripanosomiasis africana	7
1.4. Leishmaniasis	8
1.5. Ciclo de vida <i>Trypanosoma cruzi</i>	8
1.6. Ciclo de vida <i>Trypanosoma brucei</i>	9
1.7. Ciclo de vida <i>Leishmania spp.</i>	9
1.8. Tratamientos	10
1.8.1. Tripanosomiasis americana	11
1.8.2. Tripanosomiasis africana humana	11
1.8.3. Leishmaniasis	11
1.9. Implicancias del desarrollo de nuevas drogas	12
1.10. Target-Pathogen Database	13
1.10.1. Drogabilidad	13
1.10.2. Human off-targeting (evaluación de blancos inespecíficos)	15
1.10.3. Reacciones “cuello de botella”	15
<b>2. Objetivos</b>	<b>18</b>
2.1. Objetivo general	18
2.2. Objetivos específicos	18
<b>3. Materiales, métodos y estrategia de trabajo</b>	<b>20</b>
3.1. Genomas con los que trabajamos	20
3.2. Determinación de valores de corte, para las propiedades utilizadas	20
3.2.1. Drogabilidad	20
3.2.2. Human off-targeting	21
3.2.3. Reacciones “cuello de botella”	21
3.3. Jerarquización de las listas de candidatos por parásito	21
3.4. Agrupamiento de genes homólogos entre los tres parásitos	22
3.4.1. CD-Hit	22
3.4.2. MMseqs2	22
3.5. Jerarquización de listas con candidatos comunes a <i>T. cruzi</i> , <i>T. brucei</i> y <i>L. major</i>	23
3.6. Análisis de expresión de los candidatos	24
3.7. Análisis funcional de los candidatos	25
3.7.1. Búsqueda de términos de ontología génica	25
3.7.2. Análisis de redes de interacción	25
3.7.3. Minería de datos usando IdMiner	26
<b>4. Resultados</b>	<b>27</b>

4.1. Evaluación de la distribución de los valores para los distintos parámetros.	27
4.1.1. Drogabilidad	27
4.1.2. Human Off-Targeting	28
4.1.3. Reacciones “cuello de botella”	29
4.1.4. Otros parámetros evaluados	29
4.2. Posibles blancos drogables en cada uno de los organismos	29
4.3. Agrupamientos de genes conservados entre los tres parásitos	31
4.3.1. CD-Hit	31
4.3.2. MMseqs2	32
4.4. Posibles blancos drogables comunes entre tripanosomátidos	33
4.4.1. Análisis funcional de los candidatos comunes entre tripanosomátidos	34
4.4.2. Los candidatos encontrados son expresados en estadios en el hospedero mamífero	34
4.4.3. Relevancia biológica de los candidatos	36
<b>5. Discusión</b>	<b>43</b>
<b>6. Anexo</b>	<b>46</b>
<b>7. Bibliografía</b>	<b>51</b>



# 1. Introducción

---

## 1.1. Los kinetoplástidos

Los kinetoplástidos (Kinetoplastea) son una clase de protistas flagelados, que pertenecen al filo Euglenozoa y se caracterizan y categorizan por la presencia de una estructura compuesta por el ADN mitocondrial llamada kinetoplasto. El grupo incluye organismos con formas de vida muy diversas. Pueden ser parasíticos o de vida libre, monoxenos (un solo hospedador), heteroxenos (dos o más hospedadores), intracelulares o extracelulares. Es uno de los grupos eucariotas que contiene más cantidad y diversidad de parásitos y con un rango de huésped muy amplio (Lukeš et al., 2014). Los kinetoplástidos tienen como grupos hermanos a los Euglenida, Symbiontida y Diplonemea. Dado que estos grupos son fototróficos y fagotróficos (aunque existen algunos parásitos facultativos en Diplonemea) (Adl et al., 2012), se plantea que el parasitismo emergió en este linaje. Mediante un estudio filogenético basado en el ARN ribosomal 18S se dividió a la clase en dos subclases, Metakinetoplastina y los Prokinetoplastina (Moreira et al., 2004; Adl et al., 2012). Los primeros contienen 4 órdenes siendo uno de ellos el orden Trypanosomatida. Algunos parásitos de este orden son de importancia a nivel de salud humana y a nivel económico causando enfermedades de gran prevalencia en regiones tropicales y subtropicales. Las principales enfermedades que afectan a mamíferos, causadas por los tripanosomátidos, son la tripanosomiasis americana causada por *Trypanosoma cruzi* (*T. cruzi*), la tripanosomiasis africana causada por *Trypanosoma brucei* (*T. brucei*) y la leishmaniasis causada por distintas especies del género *Leishmania*.

Todas ellas forman parte de las *neglected diseases*, enfermedades desatendidas muy prevalentes en humanos, que afectan principalmente a países subdesarrollados, donde hay altos índices de pobreza, y que a pesar de ello no se destinan importantes recursos para combatirlas (Hotez et al., 2007).

## 1.2. Tripanosomiasis americana

La tripanosomiasis americana, mejor conocida como enfermedad de Chagas, es una enfermedad causada por el parásito *T. cruzi*, transmitida al huésped por medio de insectos hematófagos de la familia Reduviidae (Chagas, 1909), que afecta aproximadamente entre 6

y 7 millones de personas principalmente en países latinoamericanos (“Chagas disease” [https://www.who.int/news-room/fact-sheets/detail/chagas-disease-\(american-trypanosomiasis\)](https://www.who.int/news-room/fact-sheets/detail/chagas-disease-(american-trypanosomiasis))). A pesar de que la gran mayoría de casos son diagnosticados en Latinoamérica, la globalización ha llevado a que no se restrinja solo a esta región, sino a que se expanda a distintas partes del mundo (“Chagas disease” [https://www.who.int/news-room/fact-sheets/detail/chagas-disease-\(american-trypanosomiasis\)](https://www.who.int/news-room/fact-sheets/detail/chagas-disease-(american-trypanosomiasis)) ). La enfermedad presenta una fase aguda y una crónica. La fase aguda se da inmediatamente luego de la infección, y puede durar de semanas a meses. Puede ser una fase asintomática o con leves síntomas como fiebre e inflamación en la zona por la cual el parásito ingresa al organismo. En esta etapa se pueden encontrar parásitos en sangre y en casos poco frecuentes la infección puede dar lugar a inflamación del músculo cardíaco o cerebro. Luego de esta etapa las personas pueden entrar en una fase larga y asintomática llamada “crónica indeterminada”, en la cual no se detectan parásitos en sangre o se detectan muy pocos. Muchas personas quedan en esta etapa por el resto de su vida sin presentar ningún tipo de síntoma, pensando que no tienen la infección, pero muchos de ellos pueden presentar anomalías del ritmo cardíaco que puede determinar una muerte repentina, dilatación del corazón y un consecuente bombeo deficiente, o inflamación de otros tejidos como el esófago o el colon. También se observa un agravamiento de la enfermedad en personas con inmunodepresión (Pérez-Molina and Molina, 2018).

### 1.3. Tripanosomiasis africana

La tripanosomiasis africana o enfermedad del sueño, endémica del continente africano, es una enfermedad causada por el parásito *T. brucei*, transmitida al huésped por moscas tse-tse (familia Glossinidae). La enfermedad tiene principalmente dos formas que son la de progresión lenta, causada por la subespecie *Trypanosoma brucei gambiense*, endémica de del centro oeste africano, y la de progresión rápida, causada por la subespecie *Trypanosoma brucei rhodesiense*, que se presenta en su mayoría en el sur y este africano. Durante el siglo veinte afectó a millones de personas en distintas zonas del continente, con diferente prevalencia en cada una de ellas (Büscher et al., 2017). En los últimos años, han aumentado los planes de control de la enfermedad, logrando disminuir los casos significativamente, y actualmente se la considera en control (Bottieau and Clerinx, 2019). Para ser precisos, en el año 2019 se reportaron menos de 1000 casos, el valor más bajo desde que se tiene un registro sistemático (“Trypanosomiasis, human African (sleeping sickness)”

[https://www.who.int/news-room/fact-sheets/detail/trypanosomiasis-human-african-\(sleeping-sickness\)](https://www.who.int/news-room/fact-sheets/detail/trypanosomiasis-human-african-(sleeping-sickness)) ). En los últimos 10 años la gran mayoría de los casos se reportaron en la República Democrática del Congo.

## 1.4. Leishmaniasis

El término leishmaniasis, hace referencia a un conjunto de zoonosis con sintomatología variable causadas por varias especies de parásitos del género *Leishmania* y transmitida al huésped mediante insectos hembra de los géneros *Lutzomyia* y *Phlebotomus* (conocidos como *sandflies* o moscas de arena). Está ampliamente distribuida por el mundo, focalizándose en zonas pobres de países en desarrollo. Principalmente hay tres tipos de leishmaniasis, la cutánea (la más frecuente), la visceral (la más grave) y la mucocutánea. En el último informe de la OMS, se muestra que el 94% de los 22.145 casos reportados se dieron en siete países, Brasil, Etiopía, India, Kenia, Somalia, Sudán del Sur y Sudán (“Global leishmaniasis surveillance” <https://www.who.int/publications-detail-redirect/who-wer9635-401-419> ). Particularmente el 90% de los casos de la forma mucocutánea, se reportaron en el Estado Plurinacional de Bolivia, Brasil y Perú, evidenciando que varios países de nuestra región están influenciados por este tipo de enfermedades (“Global leishmaniasis surveillance” <https://www.who.int/publications-detail-redirect/who-wer9635-401-419> ).

El principal factor que asocia estas tres enfermedades, además de ser enfermedades desatendidas, es que son causadas por parásitos kinetoplástidos unicelulares primitivos muy relacionados filogenéticamente entre sí (Yazaki et al., 2017) (**Fig. 1A**), que comparten una gran parte de su genoma (El-Sayed et al., 2005) (**Fig. 1B**). Además su ciclo de vida, a pesar de sus diferencias, mantiene patrones muy similares (**Fig. 1C**). El ciclo de vida de los tres parásitos involucra dos hospederos, un artrópodo y un mamífero y todos tienen estadios infectivos, así como replicativos, tanto en el vector como en el mamífero.

## 1.5. Ciclo de vida *Trypanosoma cruzi*

*T. cruzi* ingresa al insecto triatomino en forma de tripomastigota sanguíneo, donde se transforma en epimastigota y comienza a replicar. Luego adquiere capacidad infectiva en forma de tripomastigota metacíclico y puede ingresar al hospedero mamífero. Es captado

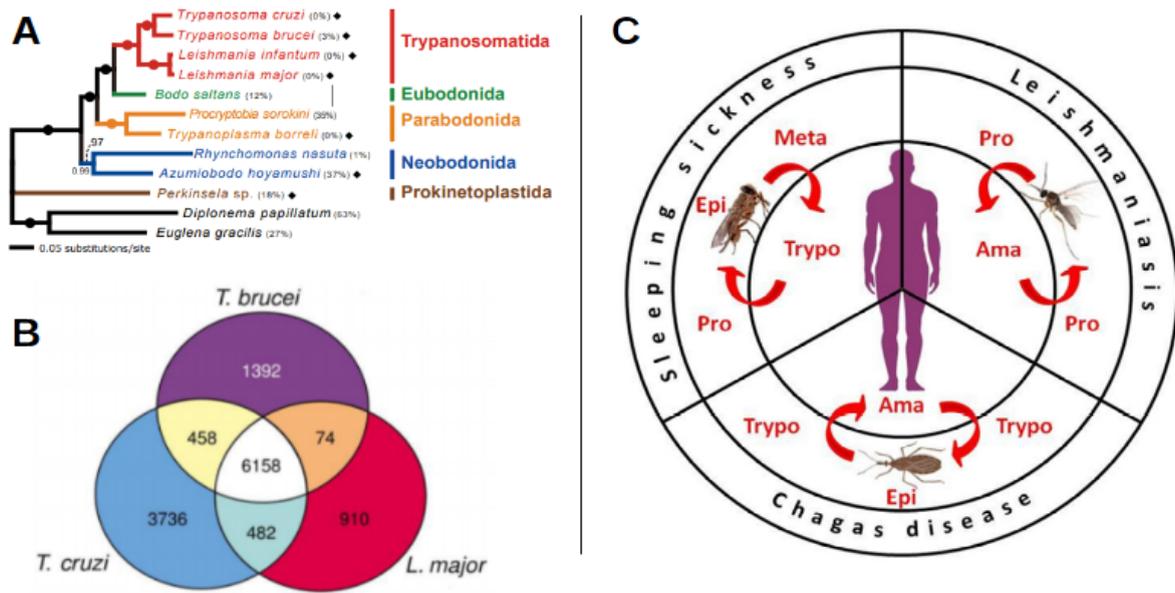
por macrófagos, donde se diferencia en amastigota, estadio replicativo intracelular, que es seguido por una diferenciación a tripomastigotas sanguíneos que son liberados al torrente sanguíneo, donde pueden pasar al vector mediante la ingesta de sangre por parte del mismo o colonizar otros tejidos (Souza, 2002; Pérez-Molina and Molina, 2018).

## 1.6. Ciclo de vida *Trypanosoma brucei*

*T. brucei*, ingresa a la mosca tse-tse como tripomastigota sanguíneo en la forma “*stumpy*” no replicativa mediante la absorción de sangre de un mamífero infectado. Rápidamente se diferencia a su forma procíclica y posteriormente al estadio replicativo epimastigota. Luego de proliferar y alcanzar las glándulas salivares, adquiere carácter infectivo y puede ingresar al hospedero mamífero como tripomastigota metacíclico. Ya en el mamífero, se diferencia a tripomastigota sanguíneo en su forma “*slender*” y prolifera en el torrente sanguíneo. Esta proliferación es seguida de una transformación a la forma “*stumpy*” cerrando el ciclo (Matthews, 2005). Vale destacar que *T. brucei* es el único de los tres parásitos en estudio que no presenta un estadio intracelular en su ciclo de vida.

## 1.7. Ciclo de vida *Leishmania spp.*

Las distintas especies del género *Leishmania*, ingresan al *sand-fly* en la forma amastigota y se diferencian en promastigotas procíclicos para luego adquirir carácter infectivo en la forma de promastigotas metacíclicos. Ingresan al hospedero en este estadio e infectan macrófagos, allí se diferencian nuevamente a amastigotas que pasan al torrente sanguíneo y al ser ingeridos por el vector se cierra el ciclo (Teixeira et al., 2013).



**Figura 1. (A)** Relación filogenética entre kinetoplástidos. Se observa como los tripanosomátidos forman un grupo monofilético. Tomado de (Yazaki et al., 2017). **(B)** Diagrama de Venn mostrando porcentajes del genoma compartidos por *T. cruzi*, *T. brucei* y *L. major*. Tomado de (El-Sayed et al., 2005). **(C)** Esquema representando el ciclo de vida de los tres parásitos, *Trypanosoma cruzi* (Enfermedad de Chagas), *Trypanosoma brucei* (Enfermedad del sueño) y *Leishmania* spp. (Leishmaniasis). Tomado de (Filardy et al., 2018).

## 1.8. Tratamientos

Los tratamientos actuales para estas enfermedades no son satisfactorios, en parte debido a la falta de recursos brindados para estudiarlas. Los fármacos utilizados tienen grandes efectos secundarios y son poco eficaces, además de que existen problemas generados por parásitos resistentes. A su vez muchos de estos compuestos se utilizan hace más de 50 años lo que evidencia esta desatención.

### 1.8.1. Tripanosomiasis americana

La tripanosomiasis americana, actualmente se trata con quimioterapias basadas principalmente en el uso de dos medicamentos, Nifurtimox y Benznidazol (“Chagas disease” [https://www.who.int/news-room/fact-sheets/detail/chagas-disease-\(american-trypanosomiasis\)](https://www.who.int/news-room/fact-sheets/detail/chagas-disease-(american-trypanosomiasis))). Ambos fármacos tienen efectos secundarios importantes (Rodrigues Coura and de Castro, 2002), y solamente son efectivos en etapas tempranas de la enfermedad debido, principalmente, a resistencias generadas por el parásito (Nozaki et al., 1996; Wilkinson et al., 2008).

### 1.8.2. Tripanosomiasis africana humana

Actualmente cuatro fármacos son utilizados para el tratamiento de la tripanosomiasis africana humana, pentamidina, suramina, melarsoprol y eflornitina (“Trypanosomiasis, human African (sleeping sickness)” [https://www.who.int/news-room/fact-sheets/detail/trypanosomiasis-human-african-\(sleeping-sickness\)](https://www.who.int/news-room/fact-sheets/detail/trypanosomiasis-human-african-(sleeping-sickness))). Muchos de ellos son efectivos en etapas específicas de la enfermedad, por lo cual se utilizan combinaciones de los mismos. A su vez, el Nifurtimox, utilizado para el tratamiento de la tripanosomiasis americana, es utilizado en situaciones específicas en combinación con algunos de los fármacos mencionados anteriormente. Todos ellos pueden tener efectos secundarios, que varían en intensidad y muchas veces son dosis dependientes (Babokhov et al., 2013). Recientemente el fármaco oral Fexinidazol se aprobó para combatir ambas etapas de la enfermedad (Mesu et al., 2018). Su principal relevancia es que al ser oral, es sencillo de llevar a zonas de difícil acceso, como muchas zonas del continente africano donde se desarrolla esta enfermedad.

### 1.8.3. Leishmaniasis

Con respecto a la leishmaniasis, los tratamientos varían según el tipo de enfermedad. Se utiliza la anfotericina B, la miltefosina, pentamidina, entre otros y su forma más grave, la visceral, se trata mediante antimoniales pentavalentes (Aronson et al., 2017). Todos estos medicamentos tienen efectos secundarios importantes (Oliveira et al., 2011). Particularmente, los antimoniales pentavalentes se dan al paciente mediante inyecciones

por el curso de 28 a 30 días y pueden ser cardiotóxicos y causar arritmias (Sundar and Chakravarty, 2010).

Todos estos problemas respecto a los tratamientos actuales, resaltan la importancia de invertir tiempo y recursos, en investigación y desarrollo de nuevos fármacos y tratamientos.

## 1.9. Implicancias del desarrollo de nuevas drogas

La búsqueda y el desarrollo de nuevas drogas es un proceso largo y muy costoso (DiMasi et al., 2016), lo que hace realmente importante enfocar los esfuerzos iniciales en un grupo pequeño de blancos potenciales. Muchos de los proyectos de investigación y desarrollo en este área fallan, porque ya avanzada la investigación encuentran que el blanco tiene características que dificultan su drogabilidad o el compuesto que utilizan no es adecuado para el mismo, bajando su efectividad (Brown and Superti-Furga, 2003; Cook et al., 2014).

Existen diferentes aproximaciones para la búsqueda de nuevas drogas. Una aproximación clásica es el cribado fenotípico (*screening* fenotípico), donde se prueban muchos compuestos y se evalúa el efecto que tienen sobre células, modelos animales, etc. Esta aproximación tiene dos grandes desventajas. La primera es que se obtienen muy pocos candidatos haciendo un gran esfuerzo, y la segunda es que una vez encontrado un candidato hay que estudiar qué blanco está siendo afectado, lo que requiere esfuerzos importantes y muchos recursos. La principal ventaja de este enfoque es que no tiene ningún sesgo sobre qué vías metabólicas o qué procesos se busca atacar (Croston, 2017). También existen aproximaciones computacionales para la búsqueda de nuevas drogas (Pereira et al., 2020). En este sentido, el objetivo es evaluar compuestos que cumplen ciertas características. Un enfoque muy utilizado en el *screening* computacional, es la búsqueda de nuevas drogas basándose en el blanco objetivo (*target-based drug discovery*) y no en características generales (Croston, 2017). De esta forma, se buscan compuestos que puedan unirse y modificar la actividad de un blanco específico, ahorrando tiempo y recursos. Esta aproximación tiene algunas desventajas, como por ejemplo el posible sesgo hacia ciertas vías metabólicas (Croston, 2017). También puede suceder que lo estudiado de forma computacional, finalmente no es funcional en el organismo, ya que la complejidad de la biología excede las representaciones computacionales que podamos realizar (Swinney, 2013). La búsqueda de nuevas drogas basándonos en el blanco, igualmente es muy interesante y tiene grandes ventajas a nivel de recursos y tiempo, pero requiere determinar algunos blancos a estudiar.

Para esto, un enfoque ómico, integrando información estructural, funcional, datos de expresión, así como datos de secuencia permitiría, a partir del proteoma del patógeno, obtener un primer grupo de posibles blancos drogables con características interesantes, restringiendo la búsqueda y el desarrollo de fármacos a un número mucho más reducido y manejable de proteínas. En este sentido Fernández Do Porto y cols (Sosa et al., 2018), desarrollaron una base de datos para obtener posibles blancos drogables.

## 1.10. Target-Pathogen Database

Target-Pathogen-Database es una base de datos *online*, especializada en patógenos humanos, que permite la integración de información proteica como, función, rol metabólico, *off-targeting* (conservación a nivel de secuencia aminoacídica con su ortólogo en humanos), información estructural como drogabilidad, etc (Sosa et al., 2018). Mediante el uso de filtros y puntajes, se puede obtener una lista jerarquizada de proteínas del patógeno de interés que cumplen con los requisitos buscados. Este tipo de herramientas, permite tener un primer conjunto de proteínas que en principio tienen potencial para ser posibles blancos drogables, de forma de enfocar los esfuerzos de investigación y desarrollo de nuevas drogas, en estos candidatos.

En una primera aproximación y a modo general, un potencial blanco debería cumplir tres grandes requisitos. Debería ser drogable, osea tener la capacidad de unir una molécula tipo droga (ver sección Drogabilidad), debería estar ausente o ser muy diferente a su homólogo en humanos y debería ser esencial o muy importante para la biología del patógeno (Hughes et al., 2011; Wyatt et al., 2011).

En este trabajo buscamos obtener una lista jerarquizada de posibles candidatos a ser blancos drogables en kinetoplástidos, particularmente en *T. cruzi*, *T. brucei* y *L. major*. Para eso evaluamos la relevancia de varios de los parámetros disponibles en la base de datos Target Pathogen Database, para finalmente trabajar con tres de ellas: drogabilidad, *human off-targeting* (identidad de secuencia con proteínas humanas) y *chokepoint* (reacciones “cuello de botella”).

### 1.10.1. Drogabilidad

Es un parámetro que indica la probabilidad de una estructura tipo bolsillo (o sea una cavidad en una proteína) de unir moléculas tipo droga (*drug-like*). Existen una serie de propiedades, compartidas por la gran mayoría de las moléculas aprobadas como drogas, que se utilizan para definir si una molécula es *drug-like* o no (por revisiones ver (Lipinski, 2004; Schneider,

2013)). Vale destacar que, en general si una molécula no cumple con estos requisitos, es probable que sea problemática en términos de absorción y permeabilidad, pero si los cumple no necesariamente significa que sea un gran candidato a droga. Volviendo a la drogabilidad, una misma proteína puede tener más de un bolsillo, y cada uno tiene un valor de drogabilidad asignado. Los valores varían entre 0 y 1, siendo 0 un bolsillo “no drogable” y 1 un bolsillo con propiedades que en principio cumple los requisitos para la unión de una molécula *drug-like*, de forma excelente. El procedimiento para calcular estos valores, comienza con el estudio de estructuras proteicas basadas en modelos cristalográficos (si los hay para esa proteína) o en su defecto un modelo estructural de la misma creado con el software MODELLER (Šali and Blundell, 1993; Fiser et al., 2000). Una vez obtenida la estructura, se hace uso del software fpocket (Schmidtke et al., 2010) y se calcula el valor de drogabilidad (*Drug Score (DS)*) (Schmidtke and Barril, 2010). fpocket se basa en el concepto de esferas alfa, introducido por Woodward y cols (Liang et al., 1998) que son esferas que contactan con cuatro átomos en su límite y no contienen ningún átomo en su interior. El *software*, busca estas esferas en la estructura proteica y luego analiza las propiedades de las mismas. Dada la densidad de átomos, en el interior de la proteína las esferas son más chicas que en el exterior. fpocket usa el radio de esfera para determinar cuáles son aquellas que tienen mayor probabilidad de ser cavidades. De forma general, las cavidades y hendiduras suelen ser esferas de tamaño intermedio. Una vez identificados los distintos bolsillos, se calcula el *DS* para cada bolsillo utilizando un modelo que evalúa las propiedades de los residuos aminoacídicos y las características de la zona en general para determinar que tan “drogable” es la cavidad (Schmidtke and Barril, 2010). A modo de resumen, se vió que la drogabilidad correlaciona de buena forma con la densidad hidrofóbica de la cavidad, es decir cómo están distribuidos los grupos hidrofóbicos, la polaridad de los residuos aminoacídicos (el promedio de los residuos involucrados), y la hidrofobicidad de los residuos aminoacídicos (el promedio de los residuos involucrados). De esta forma, se obtiene para cada bolsillo identificado, un valor entre 0 y 1 que indica su drogabilidad. Target-Pathogen Database presenta el bolsillo con mayor *DS* para cada proteína que esté ingresada a la base de datos, pero también se puede acceder a los valores de las demás cavidades. En (Sosa et al., 2018), definen cuatro categorías en las cuales colocar las proteínas, en base a este parámetro. “*Non druggable*” (ND;  $DS \leq 0.2$ ), “*poor druggable*” (PD;  $0.2 < DS \leq 0.5$ ), “*druggable*” (D;  $0.5 < DS \leq 0.7$ ) y “*highly druggable*” (HD;  $DS > 0.7$ ).

### 1.10.2. *Human off-targeting* (evaluación de blancos inespecíficos)

Otro de los parámetros que utilizamos es el *human off-targeting*. Este parámetro hace referencia a la similitud de secuencia entre proteínas del patógeno y del hospedero (en este caso humano). El objetivo de este parámetro es ver que nivel de identidad a nivel de secuencia primaria tienen las proteínas del patógeno con las proteínas humanas, de forma de evitar blancos de alta similitud. Esto es importante ya que son conocidos los efectos fuera del blanco objetivo (*off-target effects*) por parte de distintos fármacos, es decir que la droga afecta un blanco que no era el deseado. Proteínas con alto grado de identidad, tienen mayor probabilidad de unir la misma molécula, por lo tanto, evaluando los niveles de identidad se puede tener una aproximación a qué proteínas blanco son más adecuadas para no generar este tipo de efectos. De esta forma, Target-Pathogen Database mediante una búsqueda utilizando NCBI-BLASTp (Altschul et al., 1990), que tiene como *query* a todas las proteínas del patógeno y como *subject* a todas las proteínas humanas, obtiene valores de identidad que son usados para calcular el *human off-targeting*. En (Sosa et al., 2018), describen que el NCBI-BLASTp se filtró por  $evaluate < 10^{-7}$ . Este parámetro es calculado como  $1 - \text{máxima identidad de alineamiento}$ . Entonces los valores de *human off-targeting* varían entre 0 y 1, siendo 1 una proteína que no tuvo reporte contra el proteoma humano, y 0 una proteína que tuvo un reporte perfecto. Finalmente si una proteína del patógeno tiene múltiples reportes en la base de datos humana, el que tenga la máxima identidad de secuencia va a ser usado para el cálculo.

### 1.10.3. Reacciones “cuello de botella”

Mientras que los dos parámetros anteriores tienen su base en la secuencia y estructura de las proteínas, este parámetro busca darle una visión más funcional a la selección de posibles blancos drogables. Una reacción “cuello de botella (de estrangulamiento)” o *chokepoint* es aquella que consume de forma única un sustrato o genera de forma única un producto, por lo cual es de vital importancia para cierta vía metabólica. Cada una de las proteínas del parásito tiene dos posibles valores para este parámetro, Verdadero o Falso (*True* o *False*), Verdadero si participa en una reacción *chokepoint* o Falso si no es así. De esta manera, nos permite seleccionar aquellas proteínas que en principio son de gran importancia para el metabolismo del parásito, dándole un aspecto funcional a la selección.

Trabajando con estas propiedades logramos obtener una lista de posibles blancos drogables jerarquizados para cada uno de los parásitos de estudio, y dado que comparten un gran porcentaje del genoma, buscamos cuáles de estos blancos son comunes a *T. cruzi*, *T. brucei*

y *L. major*. Este último aspecto, creemos que es interesante dada la cantidad de recursos y tiempo que requiere el desarrollo de una droga. Con una aproximación como esta se podría enfocar el trabajo en blancos potenciales que puedan tener efecto sobre los tres parásitos, implicando un posible ahorro de recursos.



## 2. Objetivos

---

### 2.1. Objetivo general

1. Colaborar con la búsqueda de nuevos blancos drogables, para tratar enfermedades desatendidas causadas por kinetoplástidos, como son la tripanosomiasis americana, la tripanosomiasis africana y los distintos tipos de leishmaniasis.

### 2.2. Objetivos específicos

1. Realizar una búsqueda de posibles blancos drogables para *T. cruzi*, *T. brucei* y *L. major*, utilizando la base de datos Target-Pathogen.
2. Obtener genes homólogos entre los tres parásitos, mediante el uso de algoritmos de agrupamiento.
3. Generar una lista jerarquizada de genes que puedan ser posibles blancos drogables comunes a los tres parásitos.
4. Realizar un análisis de las funciones conocidas de la lista de genes obtenidos explorando su relevancia como posibles blancos drogables.



## 3. Materiales, métodos y estrategia de trabajo

---

### 3.1. Genomas con los que trabajamos

La base de datos Target-Pathogen, ofrece genomas de varios patógenos humanos para poder trabajar con ellos. Para *T. cruzi* se ofrece el genoma de la cepa *CL Brener*, para *T. brucei* la cepa *DAL972* y para *L. major* la cepa *Friedlin*. Por lo tanto, este estudio se llevó adelante utilizando estos genomas.

### 3.2. Determinación de valores de corte, para las propiedades utilizadas

Una primera selección de proteínas para la determinación de candidatos se obtuvo mediante el filtrado por dos parámetros que son la drogabilidad y *human off-targeting*. En primera instancia para decidir qué valores usar como punto de corte, fuimos variando los mismos y observando qué selección obteníamos. Esto permitió determinar valores que capten una cantidad suficiente de posibles blancos para continuar el filtrado con otros parámetros.

#### 3.2.1. Drogabilidad

Como mencionan en el artículo (Sosa et al., 2018), los valores de drogabilidad varían en un rango de 0 a 1 y representan la capacidad de una proteína para unir un compuesto tipo droga. Recordamos que son calculados a partir de la búsqueda estructural de bolsillos proteicos, en los cuales se analizan las propiedades fisicoquímicas de los aminoácidos que lo componen, para determinar un valor de drogabilidad para ese bolsillo. Las propiedades más relevantes para la determinación de la drogabilidad, fueron mencionadas anteriormente (sección Target-Pathogen Database). Según Fernández Do Porto y cols (Sosa et al., 2018), proteínas con  $DS \geq 0.5$ , son consideradas drogables. Evaluando puntos de corte más exigentes, decidimos que  $DS \geq 0.5$  era un buen punto de corte, donde obteníamos un número de proteínas drogables que nos permitía continuar trabajando y agregando filtros para afinar la selección.

### 3.2.2. Human off-targeting

El *human off-targeting* varía en un rango de 0 a 1 y es una representación de posibles efectos no deseados sobre proteínas del huésped humano. El parámetro se determina de acuerdo a la identidad a nivel de secuencia primaria de las proteínas humanas con las proteínas del parásito. Es calculado a partir de alineamientos por NCBI-BLASTp del proteoma del patógeno contra el proteoma humano. Evaluando los resultados de distintos puntos de corte, decidimos que  $HOT \geq 0.7$ , es un valor razonable con el cual somos exigentes en que haya bajos niveles de similaridad de secuencia y obtenemos un número de proteínas razonable que son buenos candidatos a ser blancos para drogas.

### 3.2.3. Reacciones “cuello de botella”

Este parámetro da un aspecto biológico y funcional a la selección de blancos, complementando el enfoque estructural que le brindan los otros dos parámetros utilizados. No utilizamos esta propiedad de las proteínas para realizar un filtrado, sino que fue usada para jerarquizar las listas obtenidas, de forma que los primeros lugares del *ranking* estén ocupados por posibles blancos drogables que en principio son importantes para la biología del parásito.

Con estos parámetros filtramos para obtener las listas de posibles blancos drogables para *T. cruzi*, *T. brucei* y *L. major*. Estas listas las jerarquizamos y posteriormente buscamos cuáles de estos candidatos eran comunes a los tres parásitos de estudio.

## 3.3. Jerarquización de las listas de candidatos por parásito

A partir de estas listas generamos un *ranking* de proteínas, donde el *Score (S)* para ser ubicadas en el mismo viene dado por la fórmula  $S = 1 * DS + 1 * HOT + 5 * CP$ , donde *DS* es *DrugScore*, *HOT* es *Human Off-Targeting* y *CP* es *Chokepoint reaction* (reacción “cuello de botella”). Para este último parámetro los valores *True* o *False*, se transforman numéricamente en 1 y 0 respectivamente. Filtrando con los valores de corte mencionados anteriormente y asignando estos pesos a cada uno de estos parámetros, obtenemos un *ranking* de candidatos a ser posibles blancos, con proteínas drogables, ausentes en humanos o con baja similaridad respecto de su homólogo, con las primeras posiciones ocupadas por proteínas que en principio son de gran importancia para la biología de los parásitos. Las listas jerarquizadas para cada uno de los parásitos se encuentran detalladas en la carpeta de Google Drive

[https://drive.google.com/drive/folders/1VW\\_IJfMpvK8\\_O4EGsfw7QSaERFAGK9DG?usp=sh  
aring](https://drive.google.com/drive/folders/1VW_IJfMpvK8_O4EGsfw7QSaERFAGK9DG?usp=sharing) (**T. cruzi:** 'candidatos-cruzi-jerarquizados.csv', **T. brucei:** 'candidatos-brucei-jerarquizados.csv', **L. major:** 'candidatos-major-jerarquizados.csv').

### 3.4. Agrupamiento de genes homólogos entre los tres parásitos

Para poder evaluar posibles blancos comunes a los tres parásitos, probamos con diferentes algoritmos de agrupamiento con el objetivo de obtener grupos de genes homólogos y sus respectivos parálogos y así evaluarlos en conjunto.

#### 3.4.1. CD-Hit

Primero utilizamos CD-Hit (Li and Godzik, 2006; Fu et al., 2012), un software que permite armar agrupamientos de proteínas basadas en su identidad de secuencia. Corrimos cdhit-est para formar agrupamientos de proteínas a partir de un archivo .fasta que contiene todas las proteínas con  $DS \geq 0.5$  para los tres parásitos. Para evaluar el punto de corte de identidad de secuencia a utilizar, realizamos varias corridas donde en cada una de ellas los agrupamientos de proteínas se forman si su identidad de secuencia pasa cierto umbral. Probamos con (0.5, 0.6, 0.7, 0.8 y 0.9) y evaluamos los resultados obtenidos. Los comandos ejecutados con sus respectivos parámetros y el *script* que escribimos para parsear la salida se encuentra en la carpeta de Google Drive [https://drive.google.com/drive/folders/1VW\\_IJfMpvK8\\_O4EGsfw7QSaERFAGK9DG?usp=sh  
aring](https://drive.google.com/drive/folders/1VW_IJfMpvK8_O4EGsfw7QSaERFAGK9DG?usp=sharing) (**comandos:** 'comandos-cd-hit', **script:** 'parsear-cluster.py').

#### 3.4.2. MMseqs2

Por otro lado, probamos el software MMseqs2 (Steinegger and Söding, 2017). Este software tiene tres módulos que actúan de forma secuencial, un módulo de pre-filtrado, uno de alineamiento y otro de agrupamiento. El primero evalúa la similitud de secuencia de un grupo de proteínas en una base de datos de consulta con una base de datos blanco y filtra por un cierto *cut-off*. El segundo toma estos alineamientos que pasaron el primer filtro y los somete a alineamientos tipo Smith-Waterman vectorizado (con *gaps*). Finalmente son agrupados mediante el tercer módulo. Los métodos de agrupamiento forman grafos no dirigidos a partir de los alineamientos. En estos grafos, cada vértice es una proteína y un conector se agrega si esas dos proteínas cumplen los requisitos planteados. En este caso, utilizamos un agrupamiento realizado junto con un integrante de nuestro departamento en el

marco de su tesis de maestría. En este agrupamiento se usó una cobertura mínima de 70%, una identidad de secuencia mayor a 50% y un *e-value* menor a 0.001 para determinar un *hit*. Se usó el método de alineamiento número 3 que es el más preciso y el modo de agrupamiento 0, que usa el algoritmo *Greedy Set Cover* para agrupar. Este algoritmo toma el nodo con mayor cantidad de conexiones y todos los nodos conectados a él, forma un agrupamiento y se eliminan del grafo. Esto se repite hasta que todos los nodos queden asignados a un agrupamiento. Si quedan nodos sin conexiones se le asigna su propio agrupamiento. Finalmente, luego de *Greedy Set cover* hay un paso donde ciertos nodos se reasignan a un centroide diferente si tiene un mejor valor de alineamiento.

Una vez obtenidos los agrupamientos, observamos que muchos contenían más de un gen por parásito. Decidimos que cada agrupamiento quedara compuesto por una única proteína de cada uno de los parásitos. Para decidir qué trío es el definitivo de ese agrupamiento, calculamos la desviación estándar de *DS* (*stdDS*) y *HOT* (*stdHOT*) para todas las combinaciones que involucran una proteína de cada uno de los parásitos. Finalmente eliminamos aquellas combinaciones con *stdDS* > 0.1 y *stdHOT* > 0.1, quedándonos con tríos con *DS* y *HOT* similares en cada uno de los agrupamientos. De esta forma cada agrupamiento queda compuesto por una proteína de cada uno de los parásitos con valores de *DS*  $\geq 0.5$ , y *HOT*  $\geq 0.7$ , similares entre sí. Finalmente si en un mismo agrupamiento existen dos combinaciones que pasan estos filtros, nos quedamos con la que tenga un mayor *DS* en promedio. Para este filtrado escribimos un *script* disponible en la carpeta de Google Drive [https://drive.google.com/drive/folders/1VW\\_IJfMpvK8\\_O4EGsfw7QSaERFAGK9DG?usp=sh](https://drive.google.com/drive/folders/1VW_IJfMpvK8_O4EGsfw7QSaERFAGK9DG?usp=sharing) [aring](#) (*'obtener-candidatos.py'*).

Luego de obtener los agrupamientos filtrados, se realizan los pasos de jerarquización de blancos. Así obtendremos un grupo de proteínas comunes con potencial para ser posibles blancos drogables (O3).

### 3.5. Jerarquización de listas con candidatos comunes a *T. cruzi*, *T. brucei* y *L. major*

Una vez obtenida la lista con los candidatos comunes a los tres parásitos, realizamos una nueva jerarquización basada en el *Score* obtenido en cada uno de los parásitos. Para esto evaluamos el *Score* de cada una de las tres proteínas que conforman un agrupamiento, calculamos la media y la desviación estándar, y creamos un nuevo *ranking*, ahora de

agrupamientos, usando esos valores. El *ranking* fue determinado usando de forma descendente la media y como segundo factor la desviación estándar de forma ascendente. La tabla con los candidatos comunes jerarquizados se encuentra en la carpeta de Google Drive

[https://drive.google.com/drive/folders/1VW\\_IJfMpvK8\\_O4EGsfw7QSaERFAGK9DG?usp=sharing](https://drive.google.com/drive/folders/1VW_IJfMpvK8_O4EGsfw7QSaERFAGK9DG?usp=sharing) ('*candidatos-comunes-jerarquizados.csv*').

### 3.6. Análisis de expresión de los candidatos

El uso de fármacos afecta a los parásitos en los estadios que se encuentran en el hospedero mamífero, por lo tanto es importante analizar la expresión de los candidatos en estas etapas del ciclo de vida. En el caso de *T. cruzi*, serían los estadios tripomastigota y amastigota, en el caso de *T. brucei* los estadios tripomastigota metacíclico y tripomastigota sanguíneo, y en *L. major* amastigota. Para el análisis utilizamos los datos generados por Burleigh y cols (Li et al., 2016), donde estudian el transcriptoma de los estadios Epimastigota, Tripomastigota y Amastigota (en diferentes horas post-infección) para *T. cruzi*. Trabajamos con estos datos ya que son de excelente calidad y describen varios estadios pudiéndose comparar la expresión génica entre ellos sin los problemas que se generan al intentar cruzar datos de transcriptómica provenientes de diferentes experimentos. Como los candidatos que estamos estudiando son comunes a los tres parásitos, asumimos que los ortólogos se van a comportar de forma similar en los estadios correspondientes.

Realizamos un mapa de calor con los valores de expresión normalizados de los candidatos en los tres estadios evaluados (para amastigotas usamos datos de 48h post-infección), sin agrupar y sin escalar por fila ya que lo que nos interesa evaluar es si su expresión es similar en los diferentes estadios y no necesariamente formar grupos de proteínas. Los genes se ordenaron por expresión en amastigotas de forma descendente.

Por otro lado, quisimos evaluar si los genes obtenidos son de baja, media o alta expresión. Para eso comparamos los valores de expresión del total de los genes detectados en el transcriptoma con los valores de expresión de los candidatos obtenidos. Realizamos un *violin plot* solapado con un *box plot* y analizamos si las diferencias eran estadísticamente significativas mediante un *test de Student (t-test)*.

## 3.7. Análisis funcional de los candidatos

### 3.7.1. Búsqueda de términos de ontología génica

Finalmente, una vez obtenidas las listas de candidatos, realizamos un análisis funcional de los mismos, de manera de entender mejor los roles de esas proteínas en la biología de los parásitos. Para esto primero evaluamos los términos de ontología génica (términos GO del inglés *Gene Ontology*) asociados a esos genes candidatos y buscamos si existe sobrerrepresentación de alguno de ellos, usando una herramienta disponible en la base de datos de tripanosomátidos TritrypDB (Aslett et al., 2010). Para este análisis utilizamos los identificadores de *L. major*, que es el organismo más estudiado de los tres y por lo tanto es esperable que haya una mejor anotación funcional de sus genes. Corrimos el análisis buscando términos de función molecular con los parámetros por defecto, que busca tanto términos curados como computados y usa un punto de corte de p-valor de 0.05 para determinar un enriquecimiento. Ese resultado fue analizado en Revigo (Supek et al., 2011), para sintetizar y visualizar los resultados de este tipo de análisis. Como parámetros marcamos que use una reducción importante de los términos que se grafican, marcando la opción “*Tiny (0.4)*”, para ver los términos más representados de forma de tener una visión general de lo que contiene la lista. El resto de los parámetros los dejamos por defecto, de forma que descarta términos GO obsoletos y usa la medida SimRel para armar el espacio semántico.

### 3.7.2. Análisis de redes de interacción

Luego realizamos un análisis de redes de interacción entre los genes candidatos de forma de ver si se encuentra alguna vía metabólica bien representada, utilizando la herramienta STRING (Szklarczyk et al., 2019). Para este análisis también utilizamos los identificadores de *L. major*, por el motivo ya comentado. Para el análisis dejamos los parámetros por defecto.

### 3.7.3. Minería de datos usando IdMiner

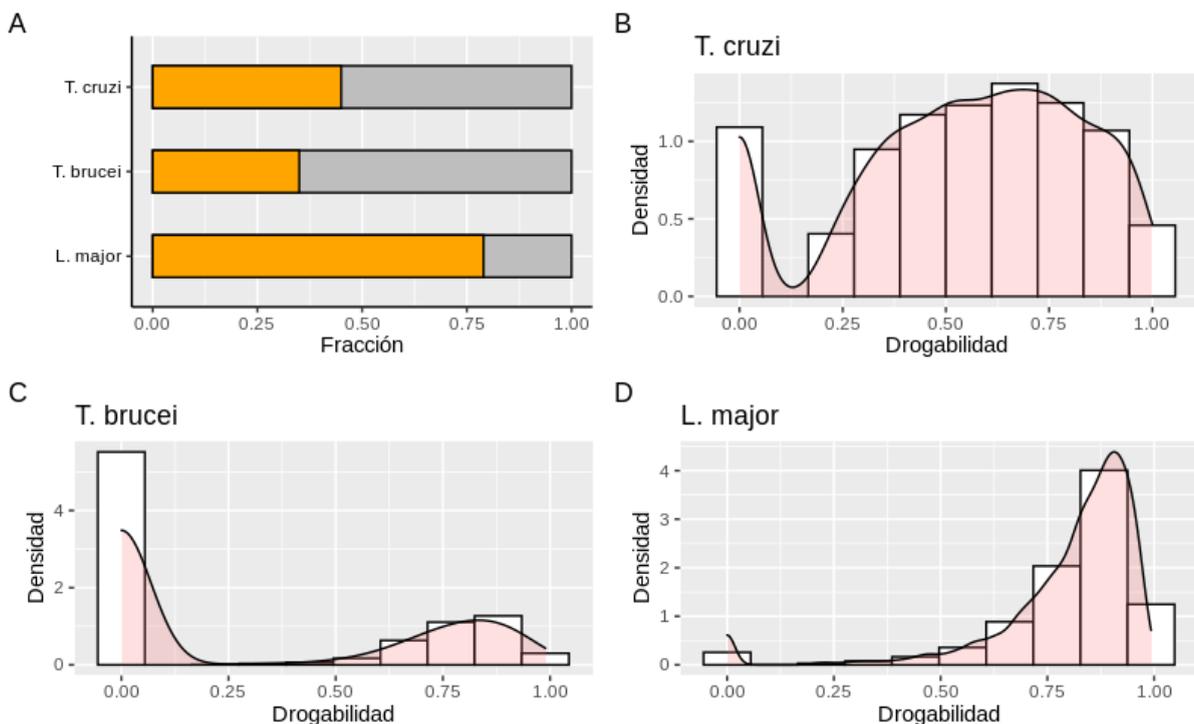
Finalmente buscamos analizar la literatura, para explorar su potencial biológico como posibles blancos drogables. Para esto utilizamos el software IdMiner (Radio, 2020), desarrollado por nuestro laboratorio, que permite mediante minería de texto, identificar términos sobrerrepresentados en artículos que mencionan alguno de los genes en estudio. Con esta aproximación podemos encontrar artículos que mencionan los genes candidatos y que tengan algún término de interés asociado. En este caso nos enfocamos en los artículos que contienen el término “*Drug*”, para buscar posibles asociaciones de estos genes con la búsqueda de drogas.

## 4. Resultados

### 4.1. Evaluación de la distribución de los valores para los distintos parámetros.

#### 4.1.1. Drogabilidad

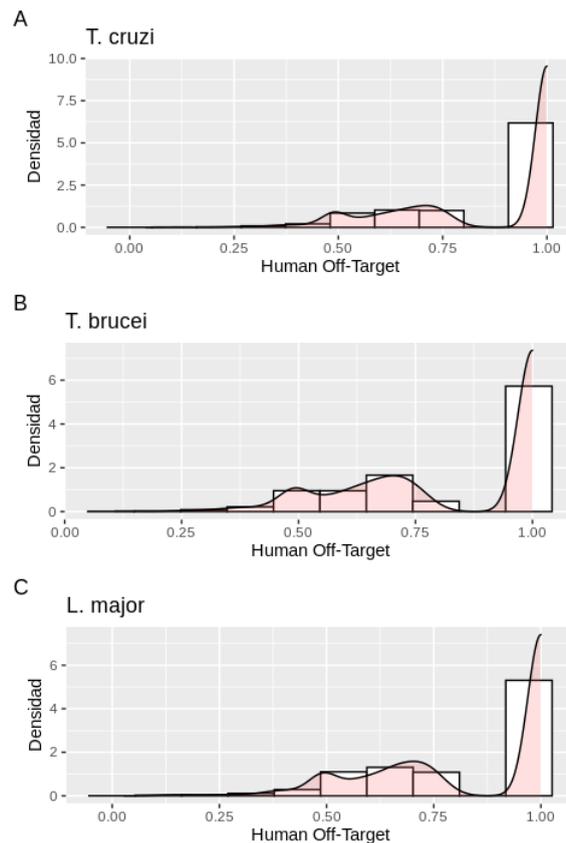
En primer lugar, para trabajar con la propiedad drogabilidad, buscamos tener una visión global de la distribución de estos valores en el proteoma de los parásitos. Encontramos que para *T. cruzi*, de los aprox. 10.300 genes anotados, 4592 tienen asignado un valor de drogabilidad. Para *T. brucei* de los aprox. 9900 genes anotados, 3399 tienen asignado un valor de drogabilidad, mientras que para *L. major*, de los aprox. 8400 genes anotados 6656 tienen asignado un valor de drogabilidad (**Fig. 2A**). Es llamativa la gran diferencia que existe en el número de proteínas con drogabilidad asignada para *L. major*. Esto podría explicar parcialmente las diferencias observadas en las distribuciones de los valores de drogabilidad entre los tres parásitos (**Fig 2B, C y D**). De todas formas para los 3 organismos se puede detectar un gran número de proteínas drogables que nos permite continuar con el trabajo.



**Figura 2.** Distribución de valores de drogabilidad en cada uno de los parásitos. **(A)** Se muestra la fracción de genes que tiene asignado al menos un valor de drogabilidad. *T. cruzi* tiene un 45% de genes con drogabilidad asignada, *T. brucei* 35% y *L. major* 79%. **(B), (C), (D)** son gráficos de densidad superpuestos con un histograma, donde se muestra la distribución de los valores de drogabilidad para *T. cruzi*, *T. brucei* y *L. major* respectivamente.

#### 4.1.2. Human Off-Targeting

Para esta propiedad también evaluamos la distribución de valores en el proteoma de los parásitos. Encontramos una distribución de valores muy similar en los tres organismos, enriquecida en valores altos (**Fig. 3**). Este enriquecimiento es esperable dado que los tripanosomátidos son de los eucariotas más primitivos lo cual hace que su similitud con proteínas humanas sea en general baja (Burki et al., 2020).



**Figura 3.** Distribución de valores de Human Off-Targeting para cada uno de los parásitos. **(A), (B)** y **(C)** son gráficos de densidad superpuestos con un histograma, donde se muestra la distribución de los valores de human off-targeting para *T. cruzi*, *T. brucei* y *L. major* respectivamente. Se observa que en los tres organismos la distribución sigue un patrón similar enriquecido en valores altos, cercanos a 1.

### 4.1.3. Reacciones “cuello de botella”

Evaluando la distribución de valores en el proteoma de los parásitos vimos que *T. cruzi*, tiene 575 proteínas anotadas, de las cuales 466 participan en reacciones *chokepoint* y 109 no lo hacen. *L. major* tiene 348 proteínas con anotación para esta propiedad, de las cuales 135 participan en reacciones *chokepoint* y 213 no. Finalmente para *T. brucei* hay 386 proteínas anotadas, de las cuales 156 participan en reacciones *chokepoint* y 230 no. La anotación para esta propiedad es muy pobre en los tres parásitos dada la falta de estudios funcionales en estos parásitos.

### 4.1.4. Otros parámetros evaluados

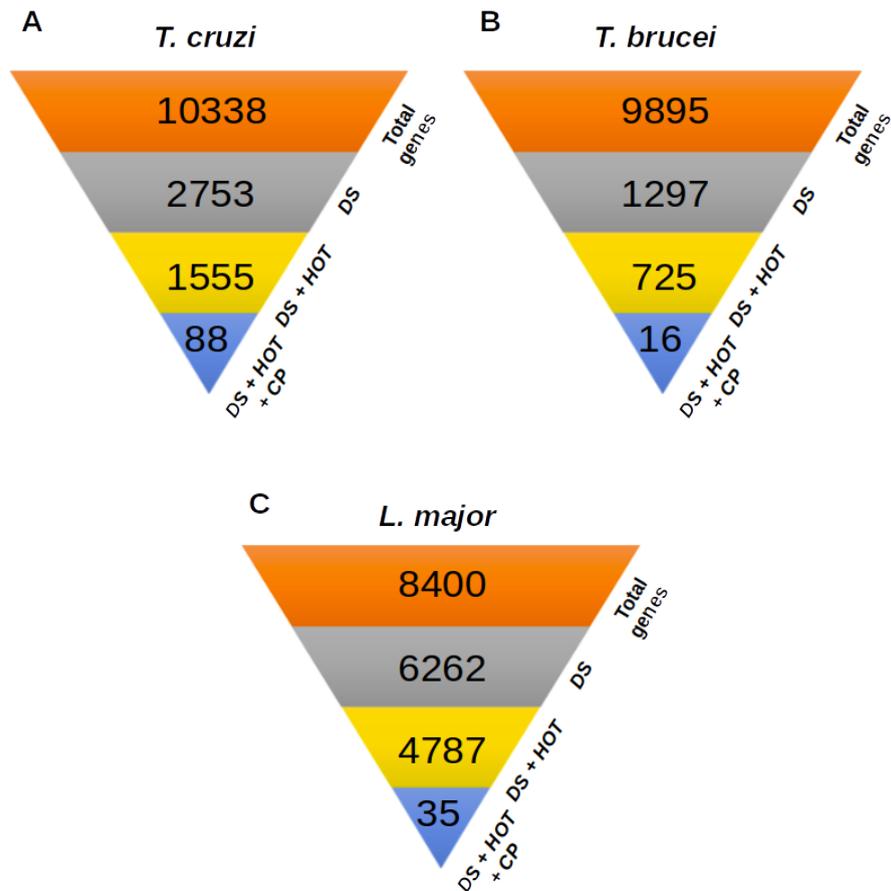
Existe mucha más información interesante y útil a la hora de buscar posibles blancos drogables. Target-Pathogen Database ofrece más parámetros para realizar un filtrado más específico, pero trabajar con parásitos poco estudiados restringe las posibilidades y buscar blancos comunes a los tres parásitos también. Evaluamos tener en cuenta proteínas que tengan reportes en la base de datos de genes esenciales DEG (*Database of Essential Genes*) (Zhang et al., 2004), pero como los reportes eran bajos no lo consideramos para este trabajo. También evaluamos la centralidad de las proteínas, un parámetro que tiene en cuenta el rol de esa proteína en redes metabólicas. Si una proteína tiene alta centralidad, significa que gran parte de las vías metabólicas pasan por ese punto y por lo tanto esa proteína es importante en el metabolismo. En este caso, como las redes metabólicas de tripanosomátidos no están bien comprendidas, la gran mayoría de las proteínas no tienen anotación y por esa razón descartamos su uso.

## 4.2. Posibles blancos drogables en cada uno de los organismos

Usando los filtros mencionados anteriormente ( $DS \geq 0.5$  y  $HOT \geq 0.7$ ), obtuvimos una lista de proteínas para cada uno de los parásitos, que tienen potencial para ser blancos drogables.

Para *T. cruzi*, de las 4592 proteínas con anotación de drogabilidad, 2753 tienen al menos un bolsillo drogable ( $DS \geq 0.5$ ). De estas 2753, 1555 además tienen un  $HOT \geq 0.7$ , y 88 de estas están anotadas como participantes de una reacción *chokepoint* (**Fig. 4A**). Para *T. brucei*, de las 3399 proteínas con anotación de drogabilidad, 1297 tienen al menos un bolsillo drogable ( $DS \geq 0.5$ ), de las cuales 725 tienen un  $HOT \geq 0.7$ . De estas 725 solo 16 tienen una anotación que indique que participan en una reacción *chokepoint* (**Fig. 4B**). Para

*L. major*, de las 6656 proteínas con anotación de drogabilidad, 6262 tienen al menos un bolsillo drogable ( $DS \geq 0.5$ ), de las cuales 4787 tienen un  $HOT \geq 0.7$ . De este grupo solamente 35 están anotadas como participantes de una reacción *chokepoint* (**Fig. 4C**). Una apreciación interesante, es notar la baja anotación de proteínas involucradas en reacciones *chokepoint*. Esto seguramente esté asociado a la falta de estudios detallados sobre las vías metabólicas en estos parásitos.



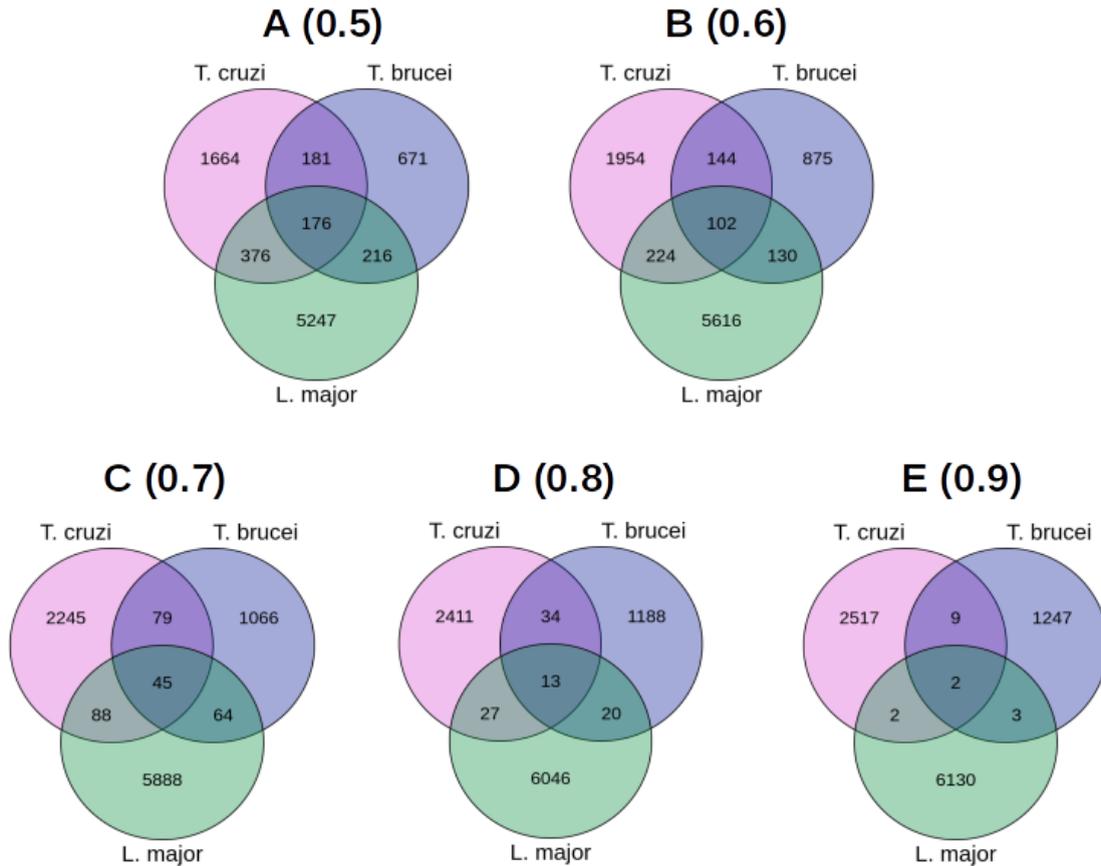
**Figura 4.** Posibles blancos drogables para *T. cruzi*, *T. brucei* y *L. major*. La figura muestra tres gráficos de pirámide invertida donde cada escalón marca la cantidad de genes que cumplieron con el filtro detallado a la derecha (A) *T. cruzi*, (B) *T. brucei*, (C) *L. major*. DS refiere a  $DS \geq 0.5$ , HOT refiere a  $HOT \geq 0.7$  y CP refiere a que están involucrados en una reacción *chokepoint* (“cuello de botella”).

### 4.3. Agrupamientos de genes conservados entre los tres parásitos

Con el objetivo de encontrar posibles blancos drogables comunes a los tres parásitos, buscamos generar agrupamientos de genes ortólogos, sobre los cuales seguir agregando características interesantes que nos den buenos candidatos. Esta estrategia, permite finalmente obtener posibles blancos drogables que cumplan las características buscadas tanto en *T. cruzi*, *T. brucei* como *L. major*, como una primera aproximación a candidatos interesantes para combatir las tres enfermedades en simultáneo.

#### 4.3.1. CD-Hit

La agrupación de secuencias con *CD-Hit* dió como resultado agrupamientos muy chicos compartidos entre los tres parásitos, para cualquiera de los puntos de corte (**Fig. 5**). Cuando usamos un porcentaje de identidad de 0.5 obtuvimos 176 agrupamientos comunes a los tres parásitos (**Fig. 5A**), con 0.6 obtuvimos 102 (**Fig. 5B**), con 0.7 obtuvimos 45 (**Fig. 5C**), con 0.8 obtuvimos 13 (**Fig. 5D**) y con 0.9 obtuvimos solamente 2 (**Fig. 5E**). Esperábamos obtener más agrupamientos compartidos por los tres organismos dado el porcentaje del genoma que comparten mencionado anteriormente (El-Sayed et al., 2005) (**Fig. 1B**) y la cantidad de proteínas con  $DS \geq 0.5$  que observamos (*T. cruzi* (2753), *T. brucei* (1297) y *L. major* (6262)). Observando estos resultados decidimos que la cantidad de agrupamientos obtenidos, para cualquiera de los puntos de corte, era insuficiente para poder seguir con los otros filtros. Para evaluar si con otras metodologías de agrupamiento obteníamos los mismos resultados, probamos utilizando MMseqs2.



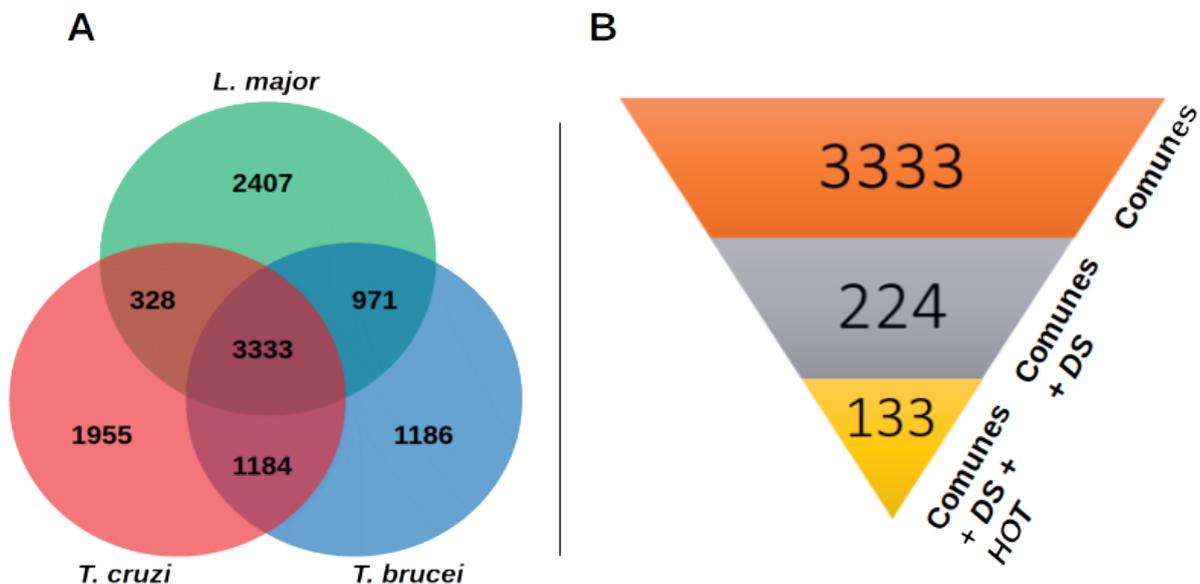
**Figura 5.** Diagrama de Venn, mostrando el número de *clusters* compartidos entre los *T. cruzi*, *T. brucei* y *L. major*, para proteínas con  $DS > 0.5$ . Cada uno de los gráficos corresponde a uno de los distintos *cut-off* de identidad de secuencia evaluados. **(A)** Se muestra el diagrama de Venn resultante de un *cut-off* de 0.5, **(B)** *cut-off* de 0.6, **(C)** *cut-off* de 0.7, **(D)** *cut-off* de 0.8 y **(E)** *cut-off* de 0.9.

#### 4.3.2. MMseqs2

La clusterización con MMseqs2 dió como resultado 3333 agrupamientos que contienen proteínas de *T. cruzi*, *T. brucei* y *L. major* (**Fig. 6A**). De estos, 224 tienen un  $DS \geq 0.5$ . Este número es un poco más grande que el obtenido por CD-Hit (175, **Fig. 5A**) para el mismo umbral por lo que decidimos quedarnos con este *software* que nos permitió trabajar con otros filtros, con el objetivo de determinar blancos más interesantes.

#### 4.4. Posibles blancos drogables comunes entre tripanosomátidos

Mediante el uso de la base de datos Target-Pathogen y de algoritmos de agrupamiento de secuencias, logramos obtener una lista de genes candidatos a ser posibles blancos drogables comunes a *T. cruzi*, *T. brucei* y *L. major*. Filtrando por  $DS \geq 0.5$  y  $HOT \geq 0.7$ , obtuvimos 133 proteínas que en principio tienen características interesantes como bolsillos drogables y poca similitud de secuencia con proteínas humanas, lo que los hace buenos candidatos (**Fig. 6B**). De los 133 grupos filtrados, 24 tienen anotación de proteína hipotética en los 3 organismos y 29 más tienen al menos esta anotación en alguno de los organismos. El resto de agrupamientos tienen una anotación descriptiva para los tres organismos. Para entender más profundamente la relevancia biológica de estos posibles blancos, fue que realizamos los análisis de términos de ontología génica, redes de interacción y minería de texto.



**Figura 6.** (A) Agrupamientos obtenidos mediante el uso de MMseqs2. Se muestra un diagrama de Venn representando la cantidad de agrupamientos comunes a los tres parásitos, compartidos entre dos de ellos y exclusivos de cada uno. (B) Posibles blancos drogables comunes a *T. cruzi*, *T. brucei* y *L. major*. La figura muestra un gráfico de pirámide invertida donde cada escalón marca la cantidad de proteínas que cumplieron con el filtro detallado a la derecha. Con la metodología utilizada encontramos 3333 agrupamientos que contienen proteínas de los tres parásitos, de los cuales 224 cumplen con que en los tres organismos las proteínas del agrupamiento tienen al menos un bolsillo

drogable ( $DS \geq 0.5$ ) y 133 cumplen con el requisito de que los tres representantes tengan poca similitud de secuencia con su homólogo en humanos ( $HOT \geq 0.7$ ).

#### 4.4.1. Análisis funcional de los candidatos comunes entre tripanosomátidos

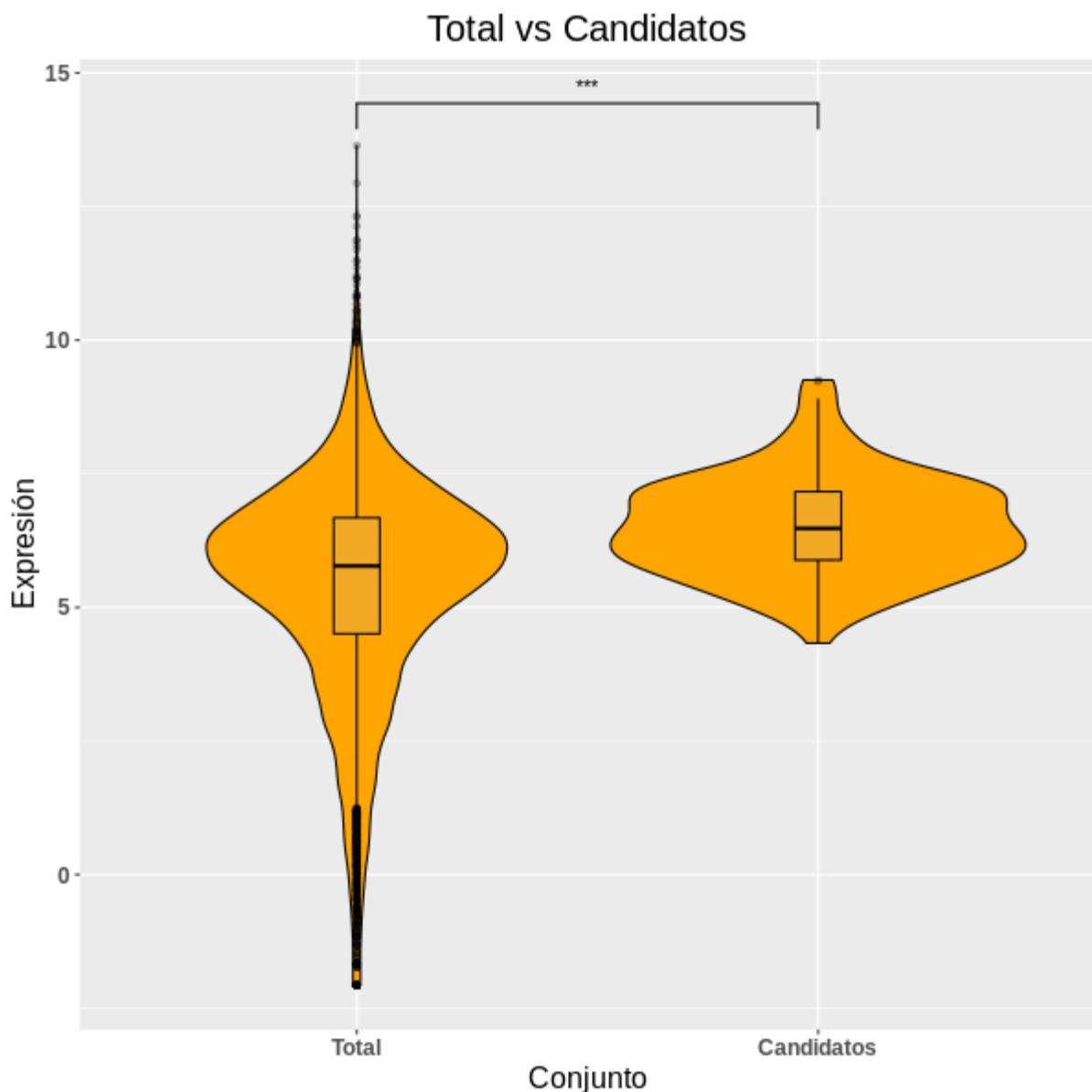
El análisis de términos de ontología génica, mostró algunos términos interesantes enriquecidos en las listas (**Fig. Sup. 1**). Cabe destacar que no necesariamente deberíamos encontrar una sobrerrepresentación en un estudio exploratorio como este, ya que los parámetros utilizados para el filtrado no están asociados a alguna función particular, sino que son meramente estructurales. Igualmente realizamos este análisis como una primera aproximación para entender la composición de las listas de genes obtenidas. Los términos encontrados se muestran en la (**Fig. Sup. 1**). Se observa un enriquecimiento en *serine* proteasas y términos relacionados a este, que son enzimas que están relacionadas con procesos infectivos en tripanosomátidos (Vermelho et al., 2010) (ver más adelante). También se ve la presencia de términos como actividad ascorbato-peroxidasa y actividad ribosa 5-fosfato isomerasa.

Respecto al análisis de redes de interacción, no encontramos redes sobrerrepresentadas en nuestra lista (**Fig. Sup. 2**). La gran mayoría de los candidatos quedan aislados y solamente se ven algunas pocas interacciones. Este análisis también fue exploratorio, con el objetivo de tener una visión general del contenido de las listas. Tampoco es esperable que en un estudio de este tipo se vean vías bien representadas en las listas.

#### 4.4.2. Los candidatos encontrados son expresados en estadios en el hospedero mamífero

Analizando la expresión de los candidatos en datos para *T. cruzi*, observamos que son genes expresados de forma similar tanto en epimastigotas, tripomastigotas como en amastigotas (**Fig Sup 3**). Por otro lado, quisimos evaluar si son genes con baja, media o alta expresión en el estadio intracelular amastigota (en el hospedero mamífero). Para eso comparamos los valores de expresión de los candidatos con los valores de expresión del total de los genes del parásito para el estadio amastigota, observando que nuestros candidatos tienen una media de expresión levemente mayor, pero estadísticamente significativa (**Fig. 7**). Esto nos muestra que los candidatos obtenidos son genes de expresión media y media alta en el estadio amastigota, con alguna excepción, por lo cual

son interesantes como posibles blancos drogables. También ordenamos el total de los genes del parásito según su expresión en amastigotas y obtuvimos el lugar que ocupan nuestros candidatos en el *ranking* de expresión. Como estamos trabajando con agrupamientos formados por genes ortólogos, asumimos que la expresión de los mismos en *T. brucei* y *L. major* va a ser similar en sus respectivos estadios en el hospedero. Los valores de expresión de los candidatos en los tres estadios mencionados, así como el lugar que ocupan en el *ranking* de expresión del total de los genes en amastigotas, se encuentran en la tabla en la carpeta de Google Drive [https://drive.google.com/drive/folders/1VW\\_IJfMpvK8\\_O4EGsfw7QSaERFAGK9DG?usp=sharing](https://drive.google.com/drive/folders/1VW_IJfMpvK8_O4EGsfw7QSaERFAGK9DG?usp=sharing) ('*candidatos-comunes-jerarquizados.csv*').



**Figura 7.** Los genes candidatos tienen una expresión media en el estadio intracelular amastigota en *T. cruzi*. Se muestra *violin plot* solapado con *box plot*, comparando los valores de expresión del total de genes del parásito (“Total”) con los candidatos a posibles blancos drogables (“Candidatos”). Se realizó *test* de *Student* para comparar la media de expresión de ambos conjuntos de genes, obteniendo una diferencia significativa ( $pvalor < 0.001 (***)$ ), con medias de 5.41 para “Total” y 6.52 para “Candidatos”.

#### 4.4.3. Relevancia biológica de los candidatos

Finalmente la minería de texto, nos permitió enfocar la búsqueda bibliográfica relacionada a estos candidatos y su posible relevancia como blanco drogable. Utilizando el IdMiner logramos encontrar varios artículos que nos permitieron entender de mejor forma el rol que cumplen estos candidatos en los organismos y analizar las consecuencias de alterar su normal funcionamiento.

El primer y el cuarto lugar de la lista están ocupados por tRNA pseudouridina sintasas, proteínas encargadas de realizar modificaciones en los ARN de transferencia. Las pseudouridina sintasas no han sido muy estudiadas en tripanosomátidos, salvo por la CBF5 en *T. brucei* (Barth et al., 2005; Chikne et al., 2016; Rajan et al., 2019). Michaeli y cols (Barth et al., 2005) demostraron que se inhibe el crecimiento del parásito al silenciar la expresión de CBF5, evidenciando la esencialidad de esta proteína para el mismo. Los genes candidatos no son los mismos que en este estudio, pero es una evidencia de la importancia de este tipo de proteínas en los tripanosomátidos. Estos genes ocupan los lugares 4901 y 4985 de un total de 10.463 genes detectados por Burleigh y cols (Li et al., 2016), ubicándose en el tercer quintil de expresión.

En el segundo lugar del *ranking* está la dihidroorotato deshidrogenasa, una enzima encargada de catalizar la conversión de L-dihidroorotato a orotato, la única reacción de reducción-oxidación en la vía de síntesis *de novo* de nucleótidos de pirimidina. Esta enzima está cristalizada en tripanosomátidos (Pinheiro et al., 2008; Arakaki et al., 2008; Cordeiro et al., 2012), y se viene estudiando como posible blanco drogable. Algunos estudios demostraron la esencialidad de esta vía metabólica para el crecimiento de los parásitos (Arakaki et al., 2008; Hashimoto et al., 2012; Wilson et al., 2012; Tiwari and Dubey, 2018), y particularmente se vió que en *T. cruzi* es aún más importante en el estadio amastigota, al cual apuntan generalmente los diferentes tratamientos. Estos estudios también resaltan la importancia de atacar tanto la vía de síntesis *de novo*, como la vía de reciclaje de nucleótidos, al diseñar un nuevo tratamiento, ya que si bien no es la principal vía, también colabora. Se han utilizado las estructuras cristalográficas para buscar regiones de la

proteína con potencial para unir una droga (Pinheiro et al., 2008), y se han realizado estudios para encontrar inhibidores de esta enzima (Yoshino et al., 2015; Inaoka et al., 2017; Chibli et al., 2018). Este gen ocupa el lugar 2870 en el *ranking* de expresión mencionado, perteneciendo al segundo quintil de expresión en amastigotas.

El tercer puesto, lo ocupa un blanco interesante que es la ribosa 5-fosfato isomerasa (Rpi) B (RpiB), una enzima que participa en la rama no oxidativa de la vía de las pentosas fosfato. Particularmente está involucrada en la interconversión de ribosa 5-fosfato y ribulosa 5-fosfato. Existen dos grandes grupos de Rpi, las RpiA presentes en la gran mayoría de los organismos y las RpiB que se encuentran reducidas a bacterias y protozoarios principalmente. Los tripanosomátidos tienen RpiB y carecen de RpiA, inversamente de lo que sucede en humanos (Stern et al., 2007, 2011; Loureiro et al., 2015; Faria et al., 2016a). Por otro lado, se ha observado que una disminución de la expresión de RpiB, afecta negativamente el crecimiento de tripanosomátidos *in vitro* y su infectividad *in vivo* (Loureiro et al., 2015; Faria et al., 2016a). Afectar esta vía metabólica podría tener consecuencias a nivel de síntesis de ácidos nucleicos, así como desde el punto de vista energético en amastigotas (Burchmore and Barrett, 2001). Se han realizado análisis para la búsqueda de inhibidores de esta enzima en *T. cruzi* (de V. C. Sinatti et al., 2017) y en *Leishmania infantum* (Dickie et al., 2021), encontrando compuestos interesantes. Todos estos puntos ubican a RpiB en una posición interesante como posible blanco drogable en tripanosomátidos. Además esta proteína se encuentra en el segundo quintil de expresión en amastigotas ocupando el lugar 2690 en el *ranking*.

El siguiente lugar del *ranking*, está ocupado por un agrupamiento anotado como Enoil-CoA isomerasa/hidratasa en *T. brucei* y *L. major*, pero como proteína ribosomal mitocondrial S47 en *T. cruzi*. Esta discordancia hace que sea difícil evaluarla como un posible blanco común a los tres organismos. Igualmente analizamos la relevancia de la Enoil-CoA isomerasa/hidratasa como posible blanco drogable. Esta enzima participa en B-Oxidación de ácidos grasos, una vía muy importante para el normal crecimiento intracelular en *Leishmania spp* (Semini et al., 2020). Por otro lado, también en *Leishmania* se han encontrado inhibidores de esta vía (Herman et al., 1987; Parreira de Aquino et al., 2021), por lo que quizá la inhibición de esta enzima pueda tener resultados similares. Particularmente este gen es de los más expresados en amastigotas de *T. cruzi* ocupando el lugar 599 del *ranking*, ubicándose en el primer quintil de expresión.

El puesto número 6 del *ranking* está ocupado por la fosfono-piruvato descarboxilasa, una enzima que forma parte de la vía de biosíntesis del amino-etil-fosfonato (AEP). AEP es un componente importante de varias macromoléculas, muy presente en fosfono-lípidos de

invertebrados marinos, bacterias y algunos protozoarios (Mukhamedova and Glushenkova, 2000). AEP forma parte de una gran proporción de las moléculas de membrana de *T. cruzi* (Carreira et al., 1996; Almeida and Gazzinelli, 2001; Acosta-Serrano, 2001), pero no está presente en *L. major* ni *T. brucei* (Ferguson et al., 1982; Turco et al., 2001; Guan and Mäser, 2017), es más estos últimos carecen de algunas enzimas de la vía de síntesis del AEP. Es probable que los ortólogos en *T. brucei* y *L. major*, estén cumpliendo otra función, que sería interesante determinar. Respecto a la importancia de este fosfonato para *T. cruzi*, Smith y Booth (Booth and Smith, 2020), comentan que en resultados no mostrados ven una disminución del crecimiento de los parásitos al realizar un *knockout* simple de la fosfoenolpiruvato mutasa, una enzima clave en esta vía de síntesis del AEP. Ya se ha propuesto esta vía de síntesis para buscar posibles blancos drogables (Lopes, 2015; Guan and Mäser, 2017), pero en principio no afectaría de igual forma a los tres parásitos, ya que es funcional solamente en *T. cruzi*. De todas formas atacar este grupo de ortólogos encontrado como candidato, podría ser interesante en los tres parásitos, ya que quizá en *T. brucei* y *L. major* también cumplen funciones importantes. Este gen se encuentra en el segundo quintil de expresión, ocupando el lugar 3089.

La siguiente es una diadenosina tetra-fosfatasa, esta enzima en específico no ha sido muy estudiada, pero en la base de datos TritypDB, se encuentra que su ortólogo en otras subespecies de *T. brucei* es la ApaH-like fosfatasa, enzima responsable de la degradación de ARN mensajeros (ARNm) mediante el mecanismo de *decapping* (Kramer, 2017). Kramer (Kramer, 2017), demostró su esencialidad, mediante el uso de ARN interferente, mostrando que puede ser un blanco drogable interesante. Se encuentra en el tercer quintil de expresión, ocupando el lugar 5872.

El siguiente lugar del *ranking* está integrado por fosfatasa ácidas, enzimas que tiene su óptimo a pH bajo. Se ha visto que las distintas fosfatasa ácidas están relacionadas a variaciones en los niveles de virulencia en *L. major* (Baghaei and Mesripour, 2015), dándole una importancia biológica interesante a estos posibles blancos. Por otro lado, se reportó que en *T. brucei*, la fosfatasa ácida de membrana MBAP-1 (del inglés), es esencial para la sobrevivencia de los parásitos (Engstler et al., 2005), pero no así en *Leishmania mexicana*, donde se observó que esta enzima no es necesaria para ningún proceso esencial en el hospedero (Benzel et al., 2000). Estos datos no dejan tan claro si este podría ser un blanco drogable interesante que sea esencial para los organismos en cuestión. Es interesante destacar su expresión en amastigotas, donde se ubican en la posición 1356, entrando en el primer quintil de expresión.

La ascorbato peroxidasa (APX), ocupa el lugar 9 en la lista. Es un blanco sumamente interesante, es una de las enzimas encargadas de contrarrestar el estrés oxidativo generado por el hospedero en tripanosomátidos (Krauth-Siegel et al., 2007; Castro and Tomás, 2008), y está ausente en humanos. El ácido ascórbico es un gran antioxidante, actuando mediante la reducción de radicales libres, secuestrando especies reactivas no radicales o funcionando como dador de electrones para las distintas APXs (Halliwell, 2001). La función de esta enzima ha sido descrita en *T. cruzi* así como en *L. major*, pero no así en *T. brucei*, por lo cual es probable que este ortólogo esté cumpliendo otra función que en principio desconocemos. Algo muy interesante de esta enzima, es que se ha visto que cepas de *Leishmania spp.* resistentes a Anfotericina B, así como cepas de *T. cruzi* resistentes al Benznidazol tienen niveles elevados de APX (Nogueira et al., 2012; Kumar et al., 2014), llevando a una mejor resolución del estrés oxidativo (Dolai et al., 2009; Taylor et al., 2015). En *L. major* la ausencia de APX lleva a un aumento de la apoptosis (Pal et al., 2010), demostrando que tiene una gran importancia en la biología del parásito, pero en *T. cruzi* hay posturas encontradas respecto a la relevancia como posible blanco drogable. Por un lado, se plantea que no es esencial para la sobrevivencia ni la virulencia de los parásitos, aunque afecta un poco su infectividad (Taylor et al., 2015), y por otro lado, se plantea que al ser tan importante para combatir el estrés oxidativo y estar sobreexpresada en cepas resistentes es un gran blanco terapéutico (Nogueira et al., 2012). APX, ya ha estado en el foco de la búsqueda de nuevos blancos drogables y se han buscado y encontrado posibles inhibidores de la misma en *Leishmania* (Mansuri et al., 2017; Moreira et al., 2018; Kashif et al., 2020). Además de estas interesantes propiedades, APX ocupa el lugar 2441 del *ranking* de expresión, estando en el segundo quintil, pero cerca del primero.

Otra proteína involucrada en la resistencia a medicamentos que aparece como candidato (lugar 110) es el transportador de membrana ABCG1 que pertenece a la familia de transportadores ABC. En *T. cruzi* se demostró su participación en la resistencia a Benznidazol (Zingales et al., 2015). Por otro lado, en *L. major*, se vió que parásitos con doble knock-out para este gen son menos infectivos y virulentos (Manzano et al., 2017). Este gen ocupa el lugar 4747 en expresión, perteneciendo al tercer quintil.

Las metacaspasas, otras de las candidatas (lugar 15), son un grupo de proteasas, presentes en hongos, plantas y protozoarios, vinculadas a procesos de muerte celular programada y regulación del estrés, entre otros (Vandana et al., 2019). En nuestra lista encontramos una metacaspasa como posible blanco drogable y es interesante ya que no existe en humanos. Se han encontrado diferentes vías por las cuales se llega a la apoptosis mediante metacaspasas en *L. major* (Basmacıyan et al., 2018), algunas de las cuales involucran la

activación de una metacaspasa, y otras la inactivación. Esto abre el debate acerca de la búsqueda de inhibidores o de potenciadores de la actividad de esta enzima, como posibles drogas, algo muy interesante ya que es una forma diferente de pensar el accionar de las drogas. Con respecto a su nivel de expresión, se encuentra en el tercer quintil, en la posición 4683.

Las asparagina sintetasas (AS) son enzimas encargadas de la síntesis del aminoácido Asparagina (Asn) a partir del aspartato. Existen dos grandes tipos de AS, las de tipo B presentes en procariotas y eucariotas, y las de tipo A que están presentes en procariotas y llamativamente también en kinetoplastidos y otros protozoarios (Faria et al., 2016b). Nuestra proteína candidata (lugar 46), es de tipo A por lo cual en principio es muy interesante dada su ausencia en humanos (Manhas et al., 2014). Respecto a su esencialidad en la biología de los parásitos, en *T. brucei* se encontró que en presencia de Asn es totalmente funcional y ni su sobrevivencia, ni la infectividad, ni la virulencia se vieron afectados (Loureiro et al., 2013). Por otro lado, en *Leishmania donovani*, se observó mediante estudios de reemplazo de genes que este gen es esencial para la viabilidad de los parásitos (Manhas et al., 2014), mientras que en *Leishmania infantum* se demostró lo mismo que en *T. brucei*, que solamente es esencial en ausencia de Asn en el medio (Faria et al., 2016b). Estos estudios demuestran que no es clara la esencialidad de este posible blanco en tripanosomátidos. Cordeiro da Silva y cols (Loureiro et al., 2013) plantean que la única combinación posible sería una terapia que debería involucrar un inhibidor de esta enzima así como una depleción de Asn extracelular (ej. L- asparaginasa) o un bloqueo de los transportadores de Asn. Según estas investigadoras, esta combinación es muy dificultosa, dado que demanda muchos recursos de desarrollo y logística, y el uso de dos drogas diferentes puede dar lugar al desarrollo de resistencias. Además agregan que se ha visto que en cáncer, el tratamiento con L-asparaginasa genera efectos adversos importantes (Appel et al., 2008; Cohen et al., 2010). Este gen es altamente expresado a nivel transcriptómico en el hospedero, ocupando el lugar 716, entrando en el primer quintil de expresión.

Otro de los candidatos interesantes es la oligopeptidasa B (OpB) (lugar 99). Esta enzima es una serina proteasa, presente en procariotas (Polgár, 1997; Fenno et al., 2001; Morty et al., 2002; Yan et al., 2006), tripanosomátidos (Morty et al., 2001, 2005; de Matos Guedes et al., 2007; Swenerton et al., 2011), algunas plantas (Matheson et al., 1995; Guo et al., 1998; Bagarozzi et al., 1998) y está predicha en genomas del reino Fungi. Una característica que hace a las OpB un blanco interesante es que hasta el momento no están descritas en mamíferos. A su vez, en *T. cruzi* se vio que está involucrada en la invasión a células no fagocíticas por parte del parásito (Burleigh et al., 1997). *T. cruzi*, ingresa a las células no

fagocíticas a través de lisosomas. Genera una liberación de  $\text{Ca}^{2+}$  intracelular en las células del hospedero, lo que determina un aumento de la presencia de lisosomas en la membrana, a través de los cuáles ingresa para posteriormente diferenciarse al estadio amastigota. OpB genera un agonista que se une a un receptor de membrana que determina la señalización por  $\text{Ca}^{2+}$  y su posterior invasión. En este estudio se vió que un tratamiento con anticuerpos anti-OpB disminuye de forma importante la capacidad invasiva de *T. cruzi*, algo que luego fue confirmado en experimentos *in vivo* en ratones. Además de tener todas estas cualidades OpB es muy expresada en el hospedero, ocupando el lugar 1137 en el ranking de expresión, perteneciendo al primer quintil.

Estos son los posibles blancos drogables con mejor puntaje utilizando nuestra metodología y algunos otros interesantes que cumplen los requisitos básicos y que mediante minería de texto usando IdMiner los asociamos a la búsqueda de drogas. Es interesante destacar que estos candidatos son expresados en los estadios en el hospedero, característica imprescindible para ser un posible blanco drogable. Esta lista incluye muchos otros candidatos que pueden tener una gran importancia para la biología de estos parásitos y que debemos seguir estudiando.



## 5. Discusión

---

Las enfermedades desatendidas (“*neglected diseases*”) tienen una gran incidencia en países de bajos recursos económicos y en vías de desarrollo. Dentro de este grupo se encuentran aquellas causadas por tripanosomátidos, como la tripanosomiasis americana, tripanosomiasis africana humana y los distintos tipos de leishmaniasis. Estas enfermedades han sido poco estudiadas por muchos años, teniendo como principal consecuencia, tratamientos poco efectivos, con efectos secundarios serios. Con el objetivo de comenzar a revertir esta situación, recientemente comenzó una búsqueda exhaustiva de posibles candidatos a drogas para enfermedades desatendidas, enmarcada en la organización sin fines de lucro DNDi (*Drugs for Neglected Diseases*, <https://dndi.org>), donde están incluidas las enfermedades causadas por tripanosomátidos.

La búsqueda de nuevos fármacos puede ser abordada desde distintas aproximaciones mencionadas anteriormente (ver sección “Implicancias del desarrollo de nuevas drogas”), que tienen sus ventajas y desventajas. En este trabajo buscamos tener una primera visualización de proteínas comunes a *T. cruzi*, *T. brucei* y *L. major*, que tienen propiedades interesantes para ser posibles blancos drogables, de forma de definir ciertos candidatos sobre los cuales enfocar estudios posteriores. Para esto, buscamos ortólogos compartidos entre los tres parásitos, utilizando metodologías de agrupamiento de secuencias y evaluamos su drogabilidad, *human off-targeting*, así como su importancia para la biología de los parásitos, usando la base de datos Target-Pathogen, y búsquedas bibliográficas. De esta forma logramos obtener una lista jerarquizada de 133 candidatos comunes a los tripanosomátidos, que en principio tienen cualidades para ser estudiados como posibles blancos drogables. Es importante aclarar que esta aproximación es un primer estudio general sobre todo el proteoma de estos parásitos, que permite reducir el espectro de búsqueda, pero las proteínas candidatas no necesariamente cumplen con todos los requisitos para ser blancos drogables.

El análisis bibliográfico de los candidatos, nos permitió comprobar que algunos de los candidatos obtenidos ya habían sido evaluados como posibles blancos de droga, validando esta metodología como una primera aproximación. También logramos obtener otros

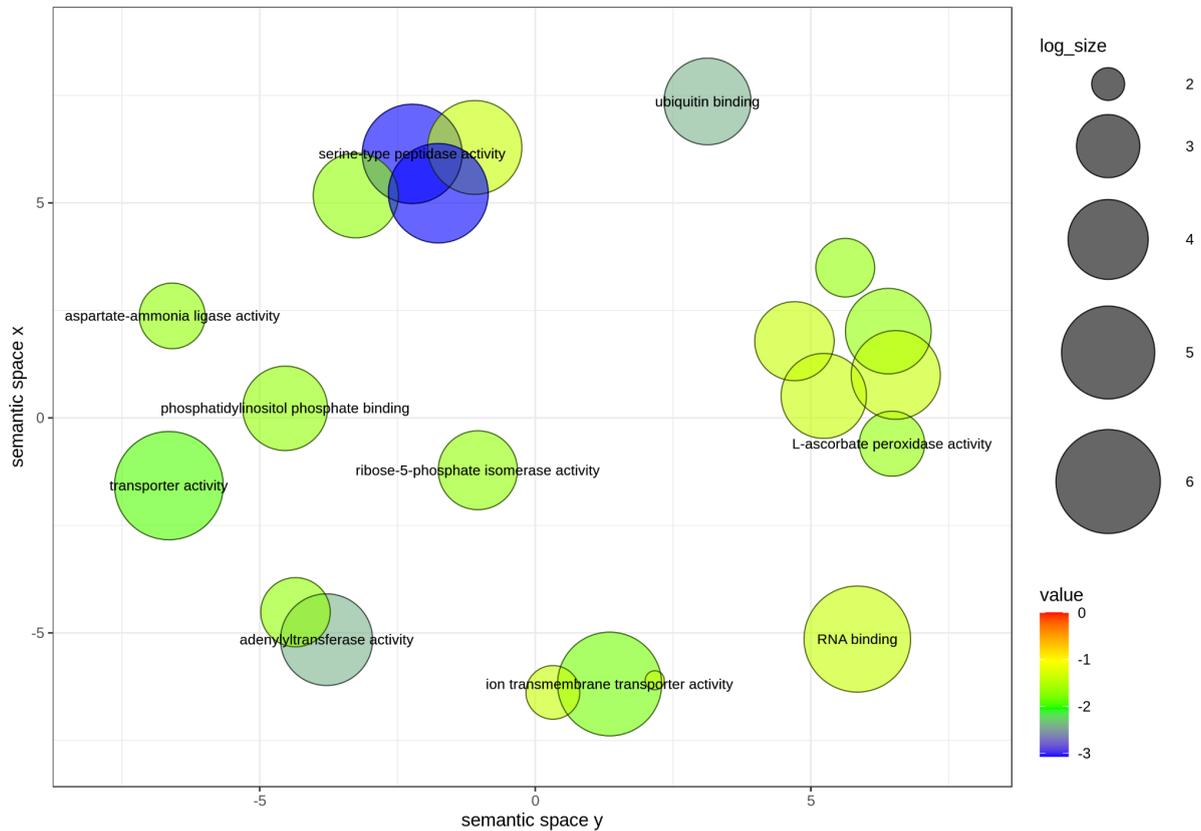
candidatos que no han sido evaluados como posibles blancos de droga que podrían ser interesantes de investigar.

Por otro lado, es interesante destacar que la búsqueda de candidatos a partir de agrupamientos de proteínas, permite reducir el espectro a blancos interesantes comunes a los organismos de interés, lo que aumenta las posibilidades de encontrar tratamientos que puedan combatir distintas enfermedades en simultáneo. Trabajar a partir de agrupamientos reduce la posibilidad de encontrar blancos individuales para cada uno de los parásitos, por lo cual entendemos que es una aproximación complementaria a la búsqueda individual, que no requiere grandes esfuerzos y aporta información que puede ser de importancia a la hora de decidir dónde enfocar estudios posteriores.

Finalmente, comentar que la base de datos utilizada permite la integración de todo tipo de información, por lo tanto, a medida que exista más investigación en estos parásitos, en cualquier aspecto, va servir para obtener candidatos mucho más interesantes y poder enfocar de forma más precisa los esfuerzos y recursos.

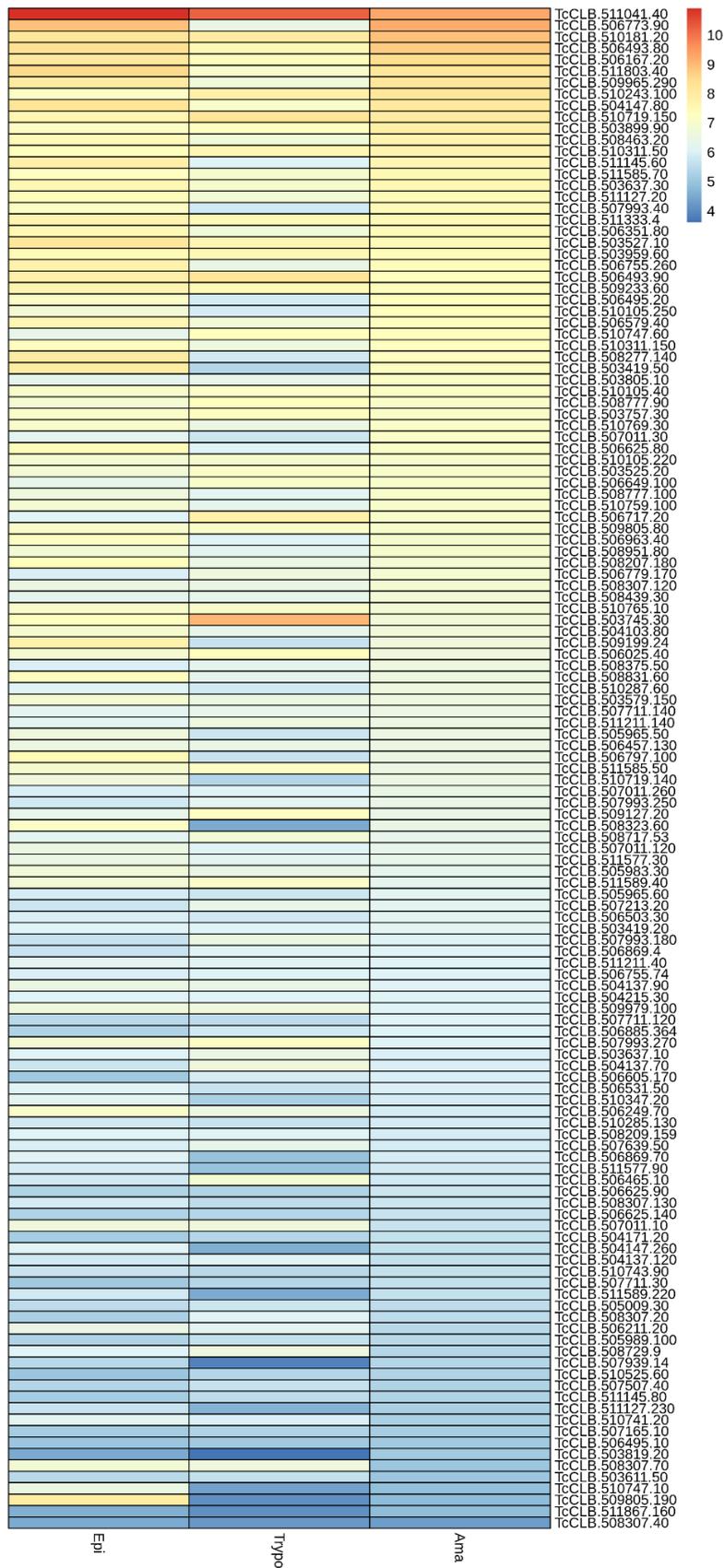


## 6. Anexo



**Figura Suplementaria 1.** Términos de ontología génica enriquecidos en la lista obtenida de *L. major*. Se observa un *scatter-plot* donde cada círculo representa un término de ontología génica, su color representa el p-valor asociado a ese enriquecimiento y el tamaño a la frecuencia en la que está ese término en las listas. Los ejes x e y representan un espacio semántico dónde términos más relacionados se encuentran más cercanos entre sí.





**Figura Suplementaria 3.** Mapa de calor con valores de expresión (transcriptoma) de los genes candidatos, para los estadios epimastigota (**Epi**), tripomastigota (**Trypo**) y amastigota (48h)

post-infección)(**Ama**) de *T. cruzi*. La escala representa los valores de expresión normalizados (cpm). Los genes fueron ordenados según su expresión en amastigotas de forma descendente. Se observa que no hay grandes diferencias de expresión entre estadios.



## 7. Bibliografía

---

- Acosta-Serrano, A., 2001. The mucin-like glycoprotein super-family of *Trypanosoma cruzi*: structure and biological roles. *Mol. Biochem. Parasitol.* 114, 143–150. [https://doi.org/10.1016/S0166-6851\(01\)00245-6](https://doi.org/10.1016/S0166-6851(01)00245-6)
- Adl, S.M., Simpson, A.G.B., Lane, C.E., Lukeš, J., Bass, D., Bowser, S.S., Brown, M.W., Burki, F., Dunthorn, M., Hampl, V., Heiss, A., Hoppenrath, M., Lara, E., le Gall, L., Lynn, D.H., McManus, H., Mitchell, E.A.D., Mozley-Stanridge, S.E., Parfrey, L.W., Pawlowski, J., Rueckert, S., Shadwick, L., Schoch, C.L., Smirnov, A., Spiegel, F.W., 2012. The Revised Classification of Eukaryotes. *J. Eukaryot. Microbiol.* 59, 429–514. <https://doi.org/10.1111/j.1550-7408.2012.00644.x>
- Almeida, I.C., Gazzinelli, R.T., 2001. Proinflammatory activity of glycosylphosphatidylinositol anchors derived from *Trypanosoma cruzi*: structural and functional analyses. *J. Leukoc. Biol.* 70, 467–477. <https://doi.org/10.1189/jlb.70.4.467>
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Appel, I.M., Hop, W.C.J., van Kessel-Bakvis, C., Stigter, R., Pieters, R., 2008. L-Asparaginase and the effect of age on coagulation and fibrinolysis in childhood acute lymphoblastic leukemia. *Thromb. Haemost.* 100, 330–337. <https://doi.org/10.1160/TH07-10-0620>
- Arakaki, T.L., Buckner, F.S., Gillespie, J.R., Malmquist, N.A., Phillips, M.A., Kalyuzhnyi, O., Luft, J.R., DeTitta, G.T., Verlinde, C.L.M.J., Van Voorhis, W.C., Hol, W.G.J., Merritt, E.A., 2008. Characterization of *Trypanosoma brucei* dihydroorotate dehydrogenase as a possible drug target; structural, kinetic and RNAi studies. *Mol. Microbiol.* 68, 37–50. <https://doi.org/10.1111/j.1365-2958.2008.06131.x>
- Aronson, N., Herwaldt, B.L., Libman, M., Pearson, R., Lopez-Velez, R., Weina, P., Carvalho, E., Ephros, M., Jeronimo, S., Magill, A., 2017. Diagnosis and Treatment of Leishmaniasis: Clinical Practice Guidelines by the Infectious Diseases Society of America (IDSA) and the American Society of Tropical Medicine and Hygiene (ASTMH). *Am. J. Trop. Med. Hyg.* 96, 24–45. <https://doi.org/10.4269/ajtmh.16-84256>
- Aslett, M., Aurrecochea, C., Berriman, M., Brestelli, J., Brunk, B.P., Carrington, M., Depledge, D.P., Fischer, S., Gajria, B., Gao, X., Gardner, M.J., Gingle, A., Grant, G., Harb, O.S., Heiges, M., Hertz-Fowler, C., Houston, R., Innamorato, F., Iodice, J., Kissinger, J.C., Kraemer, E., Li, W., Logan, F.J., Miller, J.A., Mitra, S., Myler, P.J., Nayak, V., Pennington, C., Phan, I., Pinney, D.F., Ramasamy, G., Rogers, M.B., Roos, D.S., Ross, C., Sivam, D., Smith, D.F., Srinivasamoorthy, G., Stoeckert, C.J., Subramanian, S., Thibodeau, R., Tivey, A., Treatman, C., Velarde, G., Wang, H., 2010. TriTrypDB: a functional genomic resource for the Trypanosomatidae. *Nucleic Acids Res.* 38, D457–D462. <https://doi.org/10.1093/nar/gkp851>
- Babokhov, P., Sanyaolu, A.O., Oyibo, W.A., Fagbenro-Beyioku, A.F., Iriemenam, N.C., 2013. A current analysis of chemotherapy strategies for the treatment of human African trypanosomiasis. *Pathog. Glob. Health* 107, 242–252. <https://doi.org/10.1179/2047773213Y.0000000105>
- Bagarozzi, D.A., Potempa, J., Travis, J., 1998. Purification and Characterization of an Arginine-specific Peptidase from Ragweed (*Ambrosia artemisiifolia*) Pollen. *Am. J. Respir. Cell Mol. Biol.* 18, 363–369. <https://doi.org/10.1165/ajrcmb.18.3.2825>
- Baghaei, M., Mesripour, M., 2015. Characterization of Acid Phosphatase in the Promastigotes of Three Isolates of *Leishmania major*. *Iran. J. Med. Sci.* 28, 1–8.
- Barth, S., Hury, A., Liang, X., Michaeli, S., 2005. Elucidating the Role of H/ACA-like RNAs in trans-Splicing and rRNA Processing via RNA Interference Silencing of the

- Trypanosoma brucei CBF5 Pseudouridine Synthase. *J. Biol. Chem.* 280, 34558–34568. <https://doi.org/10.1074/jbc.M503465200>
- Basmaciyani, L., Azas, N., Casanova, M., 2018. Different apoptosis pathways in Leishmania parasites. *Cell Death Discov.* 4, 90. <https://doi.org/10.1038/s41420-018-0092-z>
- Benzel, I., Weise, F., Wiese, M., 2000. Deletion of the gene for the membrane-bound acid phosphatase of Leishmania mexicana. *Mol. Biochem. Parasitol.* 111, 77–86. [https://doi.org/10.1016/S0166-6851\(00\)00306-6](https://doi.org/10.1016/S0166-6851(00)00306-6)
- Booth, L.-A., Smith, T.K., 2020. Lipid metabolism in Trypanosoma cruzi: A review. *Mol. Biochem. Parasitol.* 240, 111324. <https://doi.org/10.1016/j.molbiopara.2020.111324>
- Bottieau, E., Clerinx, J., 2019. Human African Trypanosomiasis: Progress and Stagnation. *Infect. Dis. Clin. North Am.* 33, 61–77. <https://doi.org/10.1016/j.idc.2018.10.003>
- Brown, D., Superti-Furga, G., 2003. Rediscovering the sweet spot in drug discovery. *Drug Discov. Today* 8, 1067–1077. [https://doi.org/10.1016/s1359-6446\(03\)02902-7](https://doi.org/10.1016/s1359-6446(03)02902-7)
- Burchmore, R.J.S., Barrett, M.P., 2001. Life in vacuoles – nutrient acquisition by Leishmania amastigotes. *Int. J. Parasitol.* 31, 1311–1320. [https://doi.org/10.1016/S0020-7519\(01\)00259-4](https://doi.org/10.1016/S0020-7519(01)00259-4)
- Burki, F., Roger, A.J., Brown, M.W., Simpson, A.G.B., 2020. The New Tree of Eukaryotes. *Trends Ecol. Evol.* 35, 43–55. <https://doi.org/10.1016/j.tree.2019.08.008>
- Burleigh, B.A., Caler, E.V., Webster, P., Andrews, N.W., 1997. A Cytosolic Serine Endopeptidase from Trypanosoma cruzi Is Required for the Generation of Ca<sup>2+</sup> Signaling in Mammalian Cells. *J. Cell Biol.* 136, 609–620. <https://doi.org/10.1083/jcb.136.3.609>
- Büscher, P., Cecchi, G., Jamonneau, V., Priotto, G., 2017. Human African trypanosomiasis. *The Lancet* 390, 2397–2409. [https://doi.org/10.1016/S0140-6736\(17\)31510-6](https://doi.org/10.1016/S0140-6736(17)31510-6)
- Carreira, J.C., Jones, C., Wait, R., Previato, J.O., Mendonça-Previato, L., 1996. Structural variation in the glycoinositolphospholipids of different strains of Trypanosoma cruzi. *Glycoconj. J.* 13, 955–966. <https://doi.org/10.1007/BF01053191>
- Castro, H., Tomás, A.M., 2008. Peroxidases of Trypanosomatids. *Antioxid. Redox Signal.* 10, 1593–1606. <https://doi.org/10.1089/ars.2008.2050>
- Chagas, C., 1909. Nova tripanozomíaze humana: estudos sobre a morfologia e o ciclo evolutivo do Schizotrypanum cruzi n. gen., n. sp., agente etiológico de nova entidade morbida do homem. *Mem. Inst. Oswaldo Cruz* 1, 159–218. <https://doi.org/10.1590/S0074-02761909000200008>
- Chagas disease [WWW Document], n.d. URL [https://www.who.int/news-room/fact-sheets/detail/chagas-disease-\(american-trypanosomiasis\)](https://www.who.int/news-room/fact-sheets/detail/chagas-disease-(american-trypanosomiasis)) (accessed 8.6.21).
- Chibli, L.A., Schmidt, T.J., Nonato, M.C., Calil, F.A., Da Costa, F.B., 2018. Natural products as inhibitors of Leishmania major dihydroorotate dehydrogenase. *Eur. J. Med. Chem.* 157, 852–866. <https://doi.org/10.1016/j.ejmech.2018.08.033>
- Chikne, V., Doniger, T., Rajan, K.S., Bartok, O., Eliaz, D., Cohen-Chalamish, S., Tschudi, C., Unger, R., Hashem, Y., Kadener, S., Michaeli, S., 2016. A pseudouridylation switch in rRNA is implicated in ribosome function during the life cycle of Trypanosoma brucei. *Sci. Rep.* 6, 25296. <https://doi.org/10.1038/srep25296>
- Cohen, H., Bielora, B., Harats, D., Toren, A., Pinhas-Hamiel, O., 2010. Conservative treatment of L-asparaginase-associated lipid abnormalities in children with acute lymphoblastic leukemia: Asparaginase Induced Hypertriglyceridemia. *Pediatr. Blood Cancer* 54, 703–706. <https://doi.org/10.1002/pbc.22305>
- Cook, D., Brown, D., Alexander, R., March, R., Morgan, P., Satterthwaite, G., Pangalos, M.N., 2014. Lessons learned from the fate of AstraZeneca’s drug pipeline: a five-dimensional framework. *Nat. Rev. Drug Discov.* 13, 419–431. <https://doi.org/10.1038/nrd4309>
- Cordeiro, A.T., Feliciano, P.R., Pinheiro, M.P., Nonato, M.C., 2012. Crystal structure of dihydroorotate dehydrogenase from Leishmania major. *Biochimie* 94, 1739–1748. <https://doi.org/10.1016/j.biochi.2012.04.003>

- Croston, G.E., 2017. The utility of target-based discovery. *Expert Opin. Drug Discov.* 12, 427–429. <https://doi.org/10.1080/17460441.2017.1308351>
- de Matos Guedes, H.L., Carneiro, M.P.D., de Oliveira Gomes, D.C., Rossi-Bergmann, B., De-Simone, S.G., 2007. Oligopeptidase B from *Leishmania amazonensis*: molecular cloning, gene expression analysis and molecular model. *Parasitol. Res.* 101, 865–875. <https://doi.org/10.1007/s00436-007-0630-8>
- de V. C. Sinatti, V., R. Baptista, L.P., Alves-Ferreira, M., Dardenne, L., Herminio Martins da Silva, J., Guimarães, A.C., 2017. In silico identification of inhibitors of ribose 5-phosphate isomerase from *Trypanosoma cruzi* using ligand and structure based approaches. *J. Mol. Graph. Model.* 77, 168–180. <https://doi.org/10.1016/j.jmgm.2017.08.007>
- Dickie, E.A., Ronin, C., Sá, M., Ciesielski, F., Trouche, N., Tavares, J., Santarem, N., Major, L.L., Pemberton, I.K., MacDougall, J., Smith, T.K., Cordeiro-da-Silva, A., Ciapetti, P., 2021. Towards chemical validation of *Leishmania infantum* ribose 5-phosphate isomerase as a drug target. <https://doi.org/10.1128/AAC.01892-20>
- DiMasi, J.A., Grabowski, H.G., Hansen, R.W., 2016. Innovation in the pharmaceutical industry: New estimates of R&D costs. *J. Health Econ.* 47, 20–33. <https://doi.org/10.1016/j.jhealeco.2016.01.012>
- Dolai, S., Yadav, R.K., Pal, S., Adak, S., 2009. Overexpression of Mitochondrial *Leishmania major* Ascorbate Peroxidase Enhances Tolerance to Oxidative Stress-Induced Programmed Cell Death and Protein Damage. *Eukaryot. Cell* 8, 1721–1731. <https://doi.org/10.1128/EC.00198-09>
- El-Sayed, N.M., Myler, P.J., Blandin, G., Berriman, M., Crabtree, J., Aggarwal, G., Caler, E., Renaud, H., Worthey, E.A., Hertz-Fowler, C., Ghedin, E., Peacock, C., Bartholomeu, D.C., Haas, B.J., Tran, A.-N., Wortman, J.R., Alsmark, U.C.M., Angiuoli, S., Anupama, A., Badger, J., Bringaud, F., Cadag, E., Carlton, J.M., Cerqueira, G.C., Creasy, T., Delcher, A.L., Djikeng, A., Embley, T.M., Hauser, C., Ivens, A.C., Kummerfeld, S.K., Pereira-Leal, J.B., Nilsson, D., Peterson, J., Salzberg, S.L., Shallom, J., Silva, J.C., Sundaram, J., Westenberger, S., White, O., Melville, S.E., Donelson, J.E., Andersson, B., Stuart, K.D., Hall, N., 2005. Comparative Genomics of Trypanosomatid Parasitic Protozoa. *Science* 309, 404–409. <https://doi.org/10.1126/science.1112181>
- Engstler, M., Weise, F., Bopp, K., Grünfelder, C.G., Günzel, M., Heddergott, N., Overath, P., 2005. The membrane-bound histidine acid phosphatase TbMBAP1 is essential for endocytosis and membrane recycling in *Trypanosoma brucei*. *J. Cell Sci.* 118, 2105–2118. <https://doi.org/10.1242/jcs.02327>
- Faria, J., Loureiro, I., Santarém, N., Cecílio, P., Macedo-Ribeiro, S., Tavares, J., Cordeiro-da-Silva, A., 2016a. Disclosing the essentiality of ribose-5-phosphate isomerase B in Trypanosomatids. *Sci. Rep.* 6, 26937. <https://doi.org/10.1038/srep26937>
- Faria, J., Loureiro, I., Santarém, N., Macedo-Ribeiro, S., Tavares, J., Cordeiro-da-Silva, A., 2016b. *Leishmania infantum* Asparagine Synthetase A Is Dispensable for Parasites Survival and Infectivity. *PLoS Negl. Trop. Dis.* 10, e0004365. <https://doi.org/10.1371/journal.pntd.0004365>
- Fenno, J.C., Lee, S.Y., Bayer, C.H., Ning, Y., 2001. The *opdB* Locus Encodes the Trypsin-Like Peptidase Activity of *Treponema denticola*. *Infect. Immun.* 69, 6193–6200. <https://doi.org/10.1128/IAI.69.10.6193-6200.2001>
- Ferguson, M.A.J., Allen, A.K., Snary, D., 1982. The detection of phosphonolipids in the protozoan *Trypanosoma cruzi*. *Biochem. J.* 207, 171–174.
- Filardy, A.A., Guimarães-Pinto, K., Nunes, M.P., Zukeram, K., Fliess, L., Pereira, L., Oliveira Nascimento, D., Conde, L., Morrot, A., 2018. Human Kinetoplastid Protozoan Infections: Where Are We Going Next? *Front. Immunol.* 9, 1493. <https://doi.org/10.3389/fimmu.2018.01493>
- Fiser, A., Do, R.K., Sali, A., 2000. Modeling of loops in protein structures. *Protein Sci. Publ.*

- Protein Soc. 9, 1753–1773.
- Fu, L., Niu, B., Zhu, Z., Wu, S., Li, W., 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. <https://doi.org/10.1093/bioinformatics/bts565>
- Global leishmaniasis surveillance: 2019–2020, a baseline for the 2030 roadmap [WWW Document], n.d. URL <https://www.who.int/publications-detail-redirect/who-wer9635-401-419> (accessed 11.10.21).
- Guan, X.L., Mäser, P., 2017. Comparative sphingolipidomics of disease-causing trypanosomatids reveal unique lifecycle- and taxonomy-specific lipid chemistries. *Sci. Rep.* 7, 13617. <https://doi.org/10.1038/s41598-017-13931-x>
- Guo, Z.-J., Lamb, C., Dixon, R.A., 1998. A serine protease from suspension-cultured soybean cells. *Phytochemistry* 47, 547–553. [https://doi.org/10.1016/S0031-9422\(97\)00441-X](https://doi.org/10.1016/S0031-9422(97)00441-X)
- Halliwell, B., 2001. Vitamin C and genomic stability. *Mutat. Res.* 475, 29–35. [https://doi.org/10.1016/s0027-5107\(01\)00072-0](https://doi.org/10.1016/s0027-5107(01)00072-0)
- Hashimoto, M., Morales, J., Fukai, Y., Suzuki, S., Takamiya, S., Tsubouchi, A., Inoue, S., Inoue, M., Kita, K., Harada, S., Tanaka, A., Aoki, T., Nara, T., 2012. Critical importance of the de novo pyrimidine biosynthesis pathway for *Trypanosoma cruzi* growth in the mammalian host cell cytoplasm. *Biochem. Biophys. Res. Commun.* 417, 1002–1006. <https://doi.org/10.1016/j.bbrc.2011.12.073>
- Herman, J.D., Gallalee, J.V., Best, J.M., 1987. Sodium stibogluconate (pentostam) inhibition of glucose catabolism via the glycolytic pathway, and fatty acid  $\beta$ -oxidation in *leishmania mexicana* amastigotes. *Biochem. Pharmacol.* 36, 197–201. [https://doi.org/10.1016/0006-2952\(87\)90689-7](https://doi.org/10.1016/0006-2952(87)90689-7)
- Hotez, P.J., Molyneux, D.H., Fenwick, A., Kumaresan, J., Sachs, S.E., Sachs, J.D., Savioli, L., 2007. Control of Neglected Tropical Diseases. *N. Engl. J. Med.* 357, 1018–1027. <https://doi.org/10.1056/NEJMra064142>
- Hughes, J., Rees, S., Kalindjian, S., Philpott, K., 2011. Principles of early drug discovery. *Br. J. Pharmacol.* 162, 1239–1249. <https://doi.org/10.1111/j.1476-5381.2010.01127.x>
- Inaoka, D.K., Iida, M., Hashimoto, S., Tabuchi, T., Kuranaga, T., Balogun, E.O., Honma, T., Tanaka, A., Harada, S., Nara, T., Kita, K., Inoue, M., 2017. Design and synthesis of potent substrate-based inhibitors of the *Trypanosoma cruzi* dihydroorotate dehydrogenase. *Bioorg. Med. Chem.* 25, 1465–1470. <https://doi.org/10.1016/j.bmc.2017.01.009>
- Kashif, M., Paladhi, A., Singh, R., Bhattacharyya, S., Hira, S.K., Manna, P.P., 2020. Leishmanicidal Activity of an In Silico-Screened Novel Inhibitor against Ascorbate Peroxidase of *Leishmania donovani*. *Antimicrob. Agents Chemother.* 64, e01766-19. <https://doi.org/10.1128/AAC.01766-19>
- Kramer, S., 2017. The ApaH-like phosphatase TbALPH1 is the major mRNA decapping enzyme of trypanosomes. *PLOS Pathog.* 13, e1006456. <https://doi.org/10.1371/journal.ppat.1006456>
- Krauth-Siegel, L.R., Comini, M.A., Schlecker, T., 2007. The Trypanothione System, in: Flohé, L., Harris, J.R. (Eds.), *Peroxisredoxin Systems, Subcellular Biochemistry*. Springer Netherlands, Dordrecht, pp. 231–251. [https://doi.org/10.1007/978-1-4020-6051-9\\_11](https://doi.org/10.1007/978-1-4020-6051-9_11)
- Kumar, A., Das, S., Purkait, B., Sardar, A.H., Ghosh, A.K., Dikhit, M.R., Abhishek, K., Das, P., 2014. Ascorbate Peroxidase, a Key Molecule Regulating Amphotericin B Resistance in Clinical Isolates of *Leishmania donovani*. *Antimicrob. Agents Chemother.* 58, 6172–6184. <https://doi.org/10.1128/AAC.02834-14>
- Li, W., Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>
- Li, Y., Shah-Simpson, S., Okrah, K., Belew, A.T., Choi, J., Caradonna, K.L., Padmanabhan, P., Ndegwa, D.M., Temanni, M.R., Bravo, H.C., El-Sayed, N.M., Burleigh, B.A., 2016.

- Transcriptome Remodeling in *Trypanosoma cruzi* and Human Cells during Intracellular Infection. *PLoS Pathog.* 12, e1005511. <https://doi.org/10.1371/journal.ppat.1005511>
- Liang, J., Edelsbrunner, H., Woodward, C., 1998. Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. *Protein Sci. Publ. Protein Soc.* 7, 1884–1897.
- Lipinski, C.A., 2004. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov. Today Technol.* 1, 337–341. <https://doi.org/10.1016/j.ddtec.2004.11.007>
- Lopes, F., 2015. Expression And Functional Analysis Of Lipids And Glycolipids From The Mammal-Dwelling Stages Of *Trypanosoma Cruzi*. Open Access Theses Diss.
- Loureiro, I., Faria, J., Clayton, C., Macedo-Ribeiro, S., Santarém, N., Roy, N., Cordeiro-da-Siva, A., Tavares, J., 2015. Ribose 5-Phosphate Isomerase B Knockdown Compromises *Trypanosoma brucei* Bloodstream Form Infectivity. *PLoS Negl. Trop. Dis.* 9, e3430. <https://doi.org/10.1371/journal.pntd.0003430>
- Loureiro, I., Faria, J., Clayton, C., Ribeiro, S.M., Roy, N., Santarém, N., Tavares, J., Cordeiro-da-Silva, A., 2013. Knockdown of Asparagine Synthetase A Renders *Trypanosoma brucei* Auxotrophic to Asparagine. *PLoS Negl. Trop. Dis.* 7, e2578. <https://doi.org/10.1371/journal.pntd.0002578>
- Lukeš, J., Skalický, T., Týč, J., Votýpka, J., Yurchenko, V., 2014. Evolution of parasitism in kinetoplastid flagellates. *Mol. Biochem. Parasitol.* 195, 115–122. <https://doi.org/10.1016/j.molbiopara.2014.05.007>
- Manhas, R., Tripathi, P., Khan, S., Sethu Lakshmi, B., Lal, S.K., Gowri, V.S., Sharma, A., Madhubala, R., 2014. Identification and Functional Characterization of a Novel Bacterial Type Asparagine Synthetase A. *J. Biol. Chem.* 289, 12096–12108. <https://doi.org/10.1074/jbc.M114.554642>
- Mansuri, R., Kumar, A., Rana, S., Panthi, B., Ansari, M.Y., Das, S., Dikhit, M.R., Sahoo, G.C., Das, P., 2017. In Vitro Evaluation of Antileishmanial Activity of Computationally Screened Compounds against Ascorbate Peroxidase To Combat Amphotericin B Drug Resistance. *Antimicrob. Agents Chemother.* 61, e02429-16. <https://doi.org/10.1128/AAC.02429-16>
- Manzano, J.I., Perea, A., León-Guerrero, D., Campos-Salinas, J., Piacenza, L., Castanys, S., Gamarro, F., 2017. Leishmania LABCG1 and LABCG2 transporters are involved in virulence and oxidative stress: functional linkage with autophagy. *Parasit. Vectors* 10, 267. <https://doi.org/10.1186/s13071-017-2198-1>
- Matheson, N., Schmidt, J., Travis, J., 1995. Isolation and properties of an angiotensin II-cleaving peptidase from mesquite pollen. *Am. J. Respir. Cell Mol. Biol.* 12, 441–448. <https://doi.org/10.1165/ajrcmb.12.4.7695924>
- Matthews, K.R., 2005. The developmental cell biology of *Trypanosoma brucei*. *J. Cell Sci.* 118, 283–290. <https://doi.org/10.1242/jcs.01649>
- Mesu, V.K.B.K., Kalonji, W.M., Bardonneau, C., Mordt, O.V., Blesson, S., Simon, F., Delhomme, S., Bernhard, S., Kuziena, W., Lubaki, J.-P.F., Vuvu, S.L., Ngima, P.N., Mbembo, H.M., Ilunga, M., Bonama, A.K., Heradi, J.A., Solomo, J.L.L., Mandula, G., Badibabi, L.K., Dama, F.R., Lukula, P.K., Tete, D.N., Lumbala, C., Scherrer, B., Strub-Wourgaft, N., Tarral, A., 2018. Oral fexinidazole for late-stage African *Trypanosoma brucei gambiense* trypanosomiasis: a pivotal multicentre, randomised, non-inferiority trial. *The Lancet* 391, 144–154. [https://doi.org/10.1016/S0140-6736\(17\)32758-7](https://doi.org/10.1016/S0140-6736(17)32758-7)
- Moreira, D. de S., Xavier, M.V., Murta, S.M.F., 2018. Ascorbate peroxidase overexpression protects *Leishmania braziliensis* against trivalent antimony effects. *Mem. Inst. Oswaldo Cruz* 113, e180377. <https://doi.org/10.1590/0074-02760180377>
- Moreira, D., López-García, P., Vickerman, K., 2004. An updated view of kinetoplastid phylogeny using environmental sequences and a closer outgroup: proposal for a new classification of the class Kinetoplastea. *Int. J. Syst. Evol. Microbiol.* 54, 1861–1875. <https://doi.org/10.1099/ijs.0.63081-0>

- Morty, R.E., Fülöp, V., Andrews, N.W., 2002. Substrate Recognition Properties of Oligopeptidase B from *Salmonella enterica* Serovar Typhimurium. *J. Bacteriol.* 184, 3329–3337. <https://doi.org/10.1128/JB.184.12.3329-3337.2002>
- Morty, R.E., Lonsdale-Eccles, J.D., Mentele, R., Auerswald, E.A., Coetzer, T.H.T., 2001. Trypanosome-Derived Oligopeptidase B Is Released into the Plasma of Infected Rodents, Where It Persists and Retains Full Catalytic Activity. *Infect. Immun.* 69, 2757–2761. <https://doi.org/10.1128/IAI.69.4.2757-2761.2001>
- Morty, R.E., Shih, A.Y., Fülöp, V., Andrews, N.W., 2005. Identification of the reactive cysteine residues in oligopeptidase B from *Trypanosoma brucei*. *FEBS Lett.* 579, 2191–2196. <https://doi.org/10.1016/j.febslet.2005.03.014>
- Mukhamedova, Kh.S., Glushenkova, A.I., 2000. Natural Phosphonolipids. *Chem. Nat. Compd.* 36, 329–341. <https://doi.org/10.1023/A:1002804409503>
- Nogueira, F.B., Rodrigues, J.F.A., Correa, M.M.S., Ruiz, J.C., Romanha, A.J., Murta, S.M.F., 2012. The level of ascorbate peroxidase is enhanced in benzimidazole-resistant populations of *Trypanosoma cruzi* and its expression is modulated by stress generated by hydrogen peroxide. *Mem. Inst. Oswaldo Cruz* 107, 494–502. <https://doi.org/10.1590/S0074-02762012000400009>
- Nozaki, T., Engel, J.C., Dvorak, J.A., 1996. Cellular and molecular biological analyses of nifurtimox resistance in *Trypanosoma cruzi*. *Am. J. Trop. Med. Hyg.* 55, 111–117. <https://doi.org/10.4269/ajtmh.1996.55.111>
- Oliveira, L.F., Schubach, A.O., Martins, M.M., Passos, S.L., Oliveira, R.V., Marzochi, M.C., Andrade, C.A., 2011. Systematic review of the adverse effects of cutaneous leishmaniasis treatment in the New World. *Acta Trop.* 118, 87–96. <https://doi.org/10.1016/j.actatropica.2011.02.007>
- Pal, S., Dolai, S., Yadav, R.K., Adak, S., 2010. Ascorbate Peroxidase from *Leishmania major* Controls the Virulence of Infective Stage of Promastigotes by Regulating Oxidative Stress. *PLoS ONE* 5, e11271. <https://doi.org/10.1371/journal.pone.0011271>
- Parreira de Aquino, G., Mendes Gomes, M.A., Köpke Salinas, R., Laranjeira-Silva, M.F., 2021. Lipid and fatty acid metabolism in trypanosomatids. *Microb. Cell* 8, 262–275. <https://doi.org/10.15698/mic2021.11.764>
- Pereira, C.A., Sayé, M., Reigada, C., Silber, A.M., Labadie, G.R., Miranda, M.R., Valera-Vera, E., 2020. Computational approaches for drug discovery against trypanosomatid-caused diseases. *Parasitology* 147, 611–633. <https://doi.org/10.1017/S0031182020000207>
- Pérez-Molina, J.A., Molina, I., 2018. Chagas disease. *The Lancet* 391, 82–94. [https://doi.org/10.1016/S0140-6736\(17\)31612-4](https://doi.org/10.1016/S0140-6736(17)31612-4)
- Pinheiro, M.P., Iulek, J., Cristina Nonato, M., 2008. Crystal structure of *Trypanosoma cruzi* dihydroorotate dehydrogenase from Y strain. *Biochem. Biophys. Res. Commun.* 369, 812–817. <https://doi.org/10.1016/j.bbrc.2008.02.074>
- Polgár, L., 1997. A potential processing enzyme in prokaryotes: Oligopeptidase B, a new type of serine peptidase. *Proteins Struct. Funct. Genet.* 28, 375–379. [https://doi.org/10.1002/\(SICI\)1097-0134\(199707\)28:3<375::AID-PROT7>3.0.CO;2-B](https://doi.org/10.1002/(SICI)1097-0134(199707)28:3<375::AID-PROT7>3.0.CO;2-B)
- Radio, S., 2020. *sradiouy/ldMiner*.
- Rajan, K.S., Doniger, T., Cohen-Chalamish, S., Chen, D., Semo, O., Aryal, S., Glick Saar, E., Chikne, V., Gerber, D., Unger, R., Tschudi, C., Michaeli, S., 2019. Pseudouridines on *Trypanosoma brucei* spliceosomal small nuclear RNAs and their implication for RNA and protein interactions. *Nucleic Acids Res.* 47, 7633–7647. <https://doi.org/10.1093/nar/gkz477>
- Rodrigues Coura, J., de Castro, S.L., 2002. A critical review on Chagas disease chemotherapy. *Mem. Inst. Oswaldo Cruz* 97, 3–24. <https://doi.org/10.1590/s0074-02762002000100001>
- Šali, A., Blundell, T.L., 1993. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* 234, 779–815. <https://doi.org/10.1006/jmbi.1993.1626>
- Schmidtke, P., Barril, X., 2010. Understanding and Predicting Druggability. A

- High-Throughput Method for Detection of Drug Binding Sites. *J. Med. Chem.* 53, 5858–5867. <https://doi.org/10.1021/jm100574m>
- Schmidtke, P., Le Guilloux, V., Maupetit, J., Tuffery, P., 2010. fpocket: online tools for protein ensemble pocket detection and tracking. *Nucleic Acids Res.* 38, W582–W589. <https://doi.org/10.1093/nar/gkq383>
- Schneider, G., 2013. Prediction of Drug-Like Properties, Madame Curie Bioscience Database [Internet]. Landes Bioscience.
- Semini, G., Paape, D., Blume, M., Sernee, M.F., Peres-Alonso, D., Calvignac-Spencer, S., Döllinger, J., Jehle, S., Saunders, E., McConville, M.J., Aebischer, T., 2020. Leishmania Encodes a Bacterium-like 2,4-Dienoyl-Coenzyme A Reductase That Is Required for Fatty Acid  $\beta$ -Oxidation and Intracellular Parasite Survival. *mBio* 11, e01057-20. <https://doi.org/10.1128/mBio.01057-20>
- Sosa, E.J., Burguener, G., Lanzarotti, E., Defelipe, L., Radusky, L., Pardo, A.M., Marti, M., Turjanski, A.G., Fernández Do Porto, D., 2018. Target-Pathogen: a structural bioinformatic approach to prioritize drug targets in pathogens. *Nucleic Acids Res.* 46, D413–D418. <https://doi.org/10.1093/nar/gkx1015>
- Souza, W., 2002. Basic Cell Biology of *Trypanosoma cruzi*. *Curr. Pharm. Des.* 8, 269–285. <https://doi.org/10.2174/1381612023396276>
- Steinegger, M., Söding, J., 2017. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* 35, 1026–1028. <https://doi.org/10.1038/nbt.3988>
- Stern, A.L., Burgos, E., Salmon, L., Cazzulo, J.J., 2007. Ribose 5-phosphate isomerase type B from *Trypanosoma cruzi*: kinetic properties and site-directed mutagenesis reveal information about the reaction mechanism. *Biochem. J.* 401, 279–285. <https://doi.org/10.1042/BJ20061049>
- Stern, A.L., Naworyta, A., Cazzulo, J.J., Mowbray, S.L., 2011. Structures of type B ribose 5-phosphate isomerase from *Trypanosoma cruzi* shed light on the determinants of sugar specificity in the structural family: T. cruzi ribose-5-P isomerase structure/activity. *FEBS J.* 278, 793–808. <https://doi.org/10.1111/j.1742-4658.2010.07999.x>
- Sundar, S., Chakravarty, J., 2010. Antimony toxicity. *Int. J. Environ. Res. Public Health* 7, 4267–4277. <https://doi.org/10.3390/ijerph7124267>
- Supek, F., Bošnjak, M., Škunca, N., Šmuc, T., 2011. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLOS ONE* 6, e21800. <https://doi.org/10.1371/journal.pone.0021800>
- Swenerton, R.K., Zhang, S., Sajid, M., Medzihradzsky, K.F., Craik, C.S., Kelly, B.L., McKerrow, J.H., 2011. The Oligopeptidase B of *Leishmania* Regulates Parasite Enolase and Immune Evasion \*. *J. Biol. Chem.* 286, 429–440. <https://doi.org/10.1074/jbc.M110.138313>
- Swinney, D.C., 2013. Phenotypic vs. Target-Based Drug Discovery for First-in-Class Medicines. *Clin. Pharmacol. Ther.* 93, 299–301. <https://doi.org/10.1038/clpt.2012.236>
- Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., Jensen, L.J., Mering, C. von, 2019. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47, D607–D613. <https://doi.org/10.1093/nar/gky1131>
- Taylor, M.C., Lewis, M.D., Francisco, A.F., Wilkinson, S.R., Kelly, J.M., 2015. The *Trypanosoma cruzi* Vitamin C Dependent Peroxidase Confers Protection against Oxidative Stress but Is Not a Determinant of Virulence. *PLoS Negl. Trop. Dis.* 9, e0003707. <https://doi.org/10.1371/journal.pntd.0003707>
- Teixeira, D.E., Benchimol, M., Rodrigues, J.C.F., Crepaldi, P.H., Pimenta, P.F.P., de Souza, W., 2013. The Cell Biology of *Leishmania*: How to Teach Using Animations. *PLoS Pathog.* 9, e1003594. <https://doi.org/10.1371/journal.ppat.1003594>
- Tiwari, K., Dubey, V.K., 2018. Fresh insights into the pyrimidine metabolism in the

- trypanosomatids. *Parasit. Vectors* 11, 87. <https://doi.org/10.1186/s13071-018-2660-8>
- Trypanosomiasis, human African (sleeping sickness) [WWW Document], n.d. URL [https://www.who.int/news-room/fact-sheets/detail/trypanosomiasis-human-african-\(sleeping-sickness\)](https://www.who.int/news-room/fact-sheets/detail/trypanosomiasis-human-african-(sleeping-sickness)) (accessed 7.29.21).
- Turco, S.J., Späth, G.F., Beverley, S.M., 2001. Is lipophosphoglycan a virulence factor? A surprising diversity between *Leishmania* species. *Trends Parasitol.* 17, 223–226. [https://doi.org/10.1016/S1471-4922\(01\)01895-5](https://doi.org/10.1016/S1471-4922(01)01895-5)
- Vandana, Dixit, R., Tiwari, R., Katyal, A., Pandey, K.C., 2019. Metacaspases: Potential Drug Target Against Protozoan Parasites. *Front. Pharmacol.* 10, 790. <https://doi.org/10.3389/fphar.2019.00790>
- Vermelho, A.B., Branquinha, M.H., D'Avila-Levy, C.M., Souza dos Santos, A.L., Dias, E.P. de S., Nogueira de Melo, A.C., 2010. Biological Roles of Peptidases in Trypanosomatids~!2009-11-26~!2010-02-15~!2010-03-18~! *Open Parasitol. J.* 4, 5–23. <https://doi.org/10.2174/1874421401004010005>
- Wilkinson, S.R., Taylor, M.C., Horn, D., Kelly, J.M., Cheeseman, I., 2008. A mechanism for cross-resistance to nifurtimox and benznidazole in trypanosomes. *Proc. Natl. Acad. Sci. U. S. A.* 105, 5022–5027. <https://doi.org/10.1073/pnas.0711014105>
- Wilson, Z.N., Gilroy, C.A., Boitz, J.M., Ullman, B., Yates, P.A., 2012. Genetic Dissection of Pyrimidine Biosynthesis and Salvage in *Leishmania donovani*. *J. Biol. Chem.* 287, 12759–12770. <https://doi.org/10.1074/jbc.M112.346502>
- Wyatt, P.G., Gilbert, I.H., Read, K.D., Fairlamb, A.H., 2011. Target Validation: Linking Target and Chemical Properties to Desired Product Profile. *Curr. Top. Med. Chem.* 11, 1275–1283. <https://doi.org/10.2174/156802611795429185>
- Yan, J.-B., Wang, G.-Q., Du, P., Zhu, D.-X., Wang, M.-W., Jiang, X.-Y., 2006. High-level expression and purification of *Escherichia coli* oligopeptidase B. *Protein Expr. Purif.* 47, 645–650. <https://doi.org/10.1016/j.pep.2006.01.018>
- Yazaki, E., Ishikawa, S.A., Kume, K., Kumagai, A., Kamaishi, T., Tanifuji, G., Hashimoto, T., Inagaki, Y., 2017. Global Kinetoplastea phylogeny inferred from a large-scale multigene alignment including parasitic species for better understanding transitions from a free-living to a parasitic lifestyle. *Genes Genet. Syst.* 92, 35–42. <https://doi.org/10.1266/ggs.16-00056>
- Yoshino, R., Yasuo, N., Inaoka, D.K., Hagiwara, Y., Ohno, K., Orita, M., Inoue, M., Shiba, T., Harada, S., Honma, T., Balogun, E.O., da Rocha, J.R., Montanari, C.A., Kita, K., Sekijima, M., 2015. Pharmacophore Modeling for Anti-Chagas Drug Design Using the Fragment Molecular Orbital Method. *PLoS ONE* 10, e0125829. <https://doi.org/10.1371/journal.pone.0125829>
- Zhang, R., Ou, H.-Y., Zhang, C.-T., 2004. DEG: a database of essential genes. *Nucleic Acids Res.* 32, D271–D272. <https://doi.org/10.1093/nar/gkh024>
- Zingales, B., Araujo, R.G.A., Moreno, M., Franco, J., Aguiar, P.H.N., Nunes, S.L., Silva, M.N., lenne, S., Machado, C.R., Brandão, A., 2015. A novel ABCG-like transporter of *Trypanosoma cruzi* is involved in natural resistance to benznidazole. *Mem. Inst. Oswaldo Cruz* 110, 433–444. <https://doi.org/10.1590/0074-02760140407>