



**Pontificia Universidad  
Católica del Ecuador**

Seréis mis testigos

# **PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR**

**FACULTAD DE HÁBITAT, INFRAESTRUCTURA  
Y CREATIVIDAD**

**TÍTULO DE MÁSTER EN BIOLOGÍA  
COMPUTACIONAL**

**Comparación Genómica y Evolutiva de *Trypanosoma  
cruzi* y *Trypanosoma rangeli* para evaluar su impacto  
en la patogenicidad**

Trabajo de Titulación

**Autor:** Lcda. Rueda Vera Adriana Elizabeth

**Director:** Ph.D Cervantes Pérez Sergio Alan

QUITO – ECUADOR

2025



## **DEDICATORIA**

Dedico esta tesis a las mujeres que me precedieron, mis abuelas Marina de Jesús y María Teodoza y a mi mamita Rosa, por su fortaleza, amor y ejemplo que me sostienen y acompañan siempre.

A mi hijo Thomy, inspiración y motor de mis sueños, la persona que me da una razón para continuar cada día.

A mi amor Jonathan, por su amor, confianza y aliento constante, por creer en mí aun cuando yo dudaba.

Y a mis terapeutas, Cris, Dani e Ivonne, quienes con su guía y apoyo me ayudaron a reencontrarme conmigo misma y a caminar con mayor claridad y fortaleza.

Y a mí misma, por no rendirme a pesar de los obstáculos, por continuar aun en los días más grises y creer que era posible llegar hasta aquí.

Este logro es de todos, pero también es mío.



## **AGRADECIMIENTOS**

Quiero expresar mi más profundo agradecimiento a todas las personas que hicieron posible la culminación de este proyecto.

Al amor de mi vida, Jonathan, por impulsarme siempre, creer en mí y recordarme en cada momento que soy capaz de lograr mis metas y por brindarme su apoyo emocional de manera incondicional.

A mi mamita, Rosa, por su apoyo incondicional desde siempre, siendo mi mayor ejemplo de fortaleza y dedicación.

A mi amado hijo Thomy, por la comprensión, paciencia y cariño que me brindó a lo largo de este proceso, convirtiéndose en una gran motivación para seguir adelante.

Al Dr. Alan Cervantes, mi tutor, por su valiosa guía y constante apoyo en el planteamiento y desarrollo de esta investigación.

A la Dra. Ana Poveda, por motivarme e inspirarme a emprender este posgrado y a crecer en mi formación académica y profesional.

A todas mis amigas, en especial a Paula y Nathy, por su compañía, ánimo y palabras de aliento que me recordaron que no estaba sola en este camino.

A Bryan, compañero de la maestría, por brindarme su apoyo sincero aun sin conocerme, demostrando la fuerza de la solidaridad en los momentos más importantes.

A todos ellos, mi más sincero agradecimiento.



## INDICE

DEDICATORIA.....	2
AGRADECIMIENTOS.....	3
INDICE DE FIGURAS .....	6
INDICE DE TABLAS.....	7
RESUMEN.....	8
INTRODUCCIÓN.....	12
MARCO TEORICO .....	16
1.1. Biología y Patogenicidad de <i>Trypanosoma</i> .....	16
1.2. Arquitectura Genómica en Tripanosomas .....	17
1.3. Familias Multigénicas Clave: AMASTIN Y DGF-1 .....	18
1.4. Fundamentos de Filogenia y Evolución Molecular .....	19
METODOLOGIA.....	20
2.1. Selección de genomas para análisis comparativo .....	20
2.2. Obtención de datos.....	20
2.3. Procesamiento y manejo de secuencias .....	21
2.4. Filtrado y obtención de las secuencias AMASTIN y DGF-1 .....	21
2.5. Comparación de secuencias AMASTIN y DGF-1 en cepas de <i>Trypanosoma cruzi</i> , <i>Trypanosoma rangeli</i> y <i>Trypanosoma brucei</i> .....	22
2.6. Alineamiento de secuencias.....	22
2.7. Representación gráfica del alineamiento de secuencias .....	23
2.8. Construcción de árboles filogenéticos mediante IQ-TREE2 .....	23
2.9. Visualización de árboles filogenéticos .....	23
2.10. Identificación de dominios conservados en proteínas AMASTIN y DGF-1 .....	24
RESULTADOS .....	24
3.1. Selección de genomas y características generales .....	24
3.2. Filtrado, obtención y comparación de secuencias AMASTIN y DGF-1 .....	25
3.3. Comparación de las características genómicas de <i>Trypanosoma cruzi</i> , <i>Trypanosoma rangeli</i> y <i>Trypanosoma brucei</i> .....	27
3.4. Alineamientos múltiples y conservación de Secuencias de las Familias AMASTIN y DGF-1 .....	28
3.4.1. AMASTIN.....	28
3.4.2. DGF-1.....	29
3.5. Filogenia de la familia AMASTIN .....	30



3.6. Relaciones Filogenéticas y Evolución de la Familia AMASTIN .....	31
3.7. Filogenia de la familia DGF-1 (Familia de genes dispersos-1) .....	35
3.8. Relaciones Filogenéticas y Evolución de la Familia DGF-1 .....	35
3.9. Identificación de Dominios Conservados de AMASTIN en <i>Trypanosoma cruzi</i> y <i>Trypanosoma rangeli</i> . .....	40
3.10. Análisis en <i>Trypanosoma cruzi</i> .....	40
3.11. Análisis en <i>Trypanosoma rangeli</i> .....	41
3.12. Identificación de Dominios Conservados de DGF- 1 en <i>Trypanosoma cruzi</i> y <i>Trypanosoma rangeli</i> . .....	42
3.12.1. Análisis en <i>Trypanosoma cruzi</i> .....	42
3.12.2. Análisis en <i>Trypanosoma rangeli</i> . .....	43
DISCUSIÓN.....	52
CONCLUSIONES.....	58
RECOMENDACIONES .....	59
BIBLIOGRAFIA .....	60



## INDICE DE FIGURAS

Figura 1. Comparación del número de secuencias de las familias génicas AMASTIN y DGF-1 en distintas cepas de <i>Trypanosoma cruzi</i> , <i>Trypanosoma rangeli</i> y <i>Trypanosoma brucei</i> .....	26
Figura 2. Comparación de las características genómicas de <i>Trypanosoma cruzi</i> , <i>Trypanosoma rangeli</i> y <i>Trypanosoma brucei</i> .....	27
Figura 3. Mapas de conservación de los alineamientos múltiples del gen AMASTIN. (A) <i>Trypanosoma cruzi</i> . (B) <i>Trypanosoma rangeli</i> . El grado de conservación se representa de forma binaria: Amarillo (1, conservado) y Negro (0, variable).....	29
Figura 4. Mapas de conservación de los alineamientos múltiples del gen DGF-1. (A) <i>Trypanosoma cruzi</i> . (B) <i>Trypanosoma rangeli</i> . El grado de conservación se representa de forma binaria: Amarillo (1, conservado) y Negro (0, variable).....	30
Figura 5. Árbol filogenético de la familia AMASTIN de <i>Trypanosoma cruzi</i> , obtenido con IQ-TREE v2 bajo el mejor modelo de sustitución (-m MFP), con soporte de 1,000 bootstrap ultrarrápido y 1,000 SH-aLRT.....	32
Figura 6. Árbol filogenético de la familia AMASTIN de <i>Trypanosoma rangeli</i> , obtenido con IQ-TREE v2 bajo el mejor modelo de sustitución (-m MFP), con soporte de 1,000 bootstrap ultrarrápido y 1,000 SH-aLRT.....	33
Figura 7. Filogenia de la familia AMASTIN. Árbol de máxima verosimilitud inferido con IQ-TREE2, basado en secuencias de <i>T. cruzi</i> (azul), <i>T. rangeli</i> (verde) y <i>T. brucei</i> (amarillo – outgroup) bajo el mejor modelo de sustitución (-m MFP), con soporte de 1,000 bootstrap ultrarrápido y 1,000 SH-aLRT.....	34
Figura 8. Árbol filogenético de la familia DGF-1 de <i>Trypanosoma cruzi</i> , obtenido con IQ-TREE v2 bajo el mejor modelo de sustitución (-m MFP), con soporte de 1,000 bootstrap ultrarrápido y 1,000 SH-aLRT.....	37
Figura 9. Árbol filogenético de la familia DGF-1 de <i>Trypanosoma rangeli</i> , obtenido con IQ-TREE v2 bajo el mejor modelo de sustitución (-m MFP), con soporte de 1,000 bootstrap ultrarrápido y 1,000 SH-aLRT.....	38
Figura 10. Filogenia de la familia DGF-1. Árbol de máxima verosimilitud inferido con IQ-TREE2, basado en secuencias de <i>T. cruzi</i> (rojo) y <i>T. rangeli</i> (azul) bajo el mejor modelo de sustitución (-m MFP), con soporte de 1,000 bootstrap ultrarrápido y 1,000 SH-aLRT.....	39



Figura 11. Dominio AMASTIN en Trypanosoma cruzi identificado con hmmscan (HMMER v3.4, Pfam, E-value $\leq 0.01$ ). .....	41
Figura 12. Dominio AMASTIN en Trypanosoma rangeli identificado con hmmscan (HMMER v3.4, Pfam, E-value $\leq 0.01$ ). .....	42
Figura 13. Dominio DGF-1 en Trypanosoma cruzi identificado con hmmscan (HMMER v3.4, Pfam, E-value $\leq 0.01$ ). .....	45
Figura 14. Dominio DGF-1 en Trypanosoma cruzi identificado con hmmscan (HMMER v3.4, Pfam, E-value $\leq 0.01$ ). .....	46
Figura 15. Dominio DGF-1 en Trypanosoma cruzi identificado con hmmscan (HMMER v3.4, Pfam, E-value $\leq 0.01$ ). .....	47
Figura 16. Dominio DGF-1 en Trypanosoma cruzi identificado con hmmscan (HMMER v3.4, Pfam, E-value $\leq 0.01$ ). .....	48
Figura 17. Dominio DGF-1 en Trypanosoma cruzi identificado con hmmscan (HMMER v3.4, Pfam, E-value $\leq 0.01$ ). .....	49
Figura 18. Dominio DGF-1 en Trypanosoma cruzi identificado con hmmscan (HMMER v3.4, Pfam, E-value $\leq 0.01$ ). .....	50
Figura 19. Dominio DGF-1 en Trypanosoma rangeli identificado con hmmscan (HMMER v3.4, Pfam, E-value $\leq 0.01$ ). .....	51

## **INDICE DE TABLAS**

Tabla 1. Selección de genomas para análisis comparativo.....	25
--	----



## RESUMEN

La diferencia en la patogenicidad entre *Trypanosoma cruzi* y *Trypanosoma rangeli* constituye un punto clave en parasitología ya que permite comprender por qué una de estas especies es altamente patógena para el ser humano mientras que la otra se considera no patógena. Explorar estas diferencias resulta fundamental para dilucidar los factores moleculares que determinan la capacidad de invasión y la interacción con el sistema inmune. En este contexto, las familias génicas que codifican proteínas de superficie como AMASTIN y DGF-1, adquieren un papel central, al estar directamente asociadas con los mecanismos de adhesión, invasión de células hospedadoras y evasión de la respuesta inmune. Sin embargo, su historia evolutiva y expansión en ambas especies permanecen aún poco comprendidas.

El objetivo principal de esta investigación fue analizar la expansión de las familias génicas AMASTIN y DGF-1 en *T. cruzi* y *T. rangeli* y su posible relación con la patogenicidad diferencial entre ambas especies a lo largo de su historia evolutiva. Para ello, se seleccionaron genomas completos de ambas especies, incorporando a *Trypanosoma brucei* como grupo externo. Los datos genómicos y sus anotaciones se obtuvieron de NCBI GenBank y se procesaron en un entorno bioinformático basado en Linux, utilizando Python, ClustalW, IQ-TREE2 y HMMER, entre otras herramientas, para la extracción de secuencias, alineamientos múltiples, construcción de árboles filogenéticos e identificación de dominios conservados.

Los resultados mostraron que *Trypanosoma cruzi* presentó una expansión significativa de ambas familias génicas (AMASTIN: hasta 39 copias; DGF-1: >1000 copias en algunas cepas) con alta conservación secuencial y una organización estructurada de dominios. En contraste, *T. rangeli* exhibió números reducidos de copias (AMASTIN 8; DGF-1 18) y una mayor variabilidad estructural. El análisis filogenético confirmó que las familias de proteínas AMASTIN y DGF-1 en *T. cruzi* y *T. rangeli* forman clados monofiléticos, respaldando un ancestro común dentro del subgénero *Schizotrypanum*, seguido de eventos de expansión específicos por especie. *T. brucei* careció de DGF-1 y retuvo solo una copia de AMASTIN.

Dado que los genomas analizados provienen de diferentes tecnologías de secuenciación y presentan variaciones en su calidad de ensamblaje y anotación, los números absolutos



de genes deben interpretarse con cautela. No obstante, la expansión masiva y conservación de AMASTIN y DGF-1 en *T. cruzi* sugieren una fuerte presión adaptativa relacionada con su estilo de vida intracelular. El reducido número de copias y la mayor variabilidad en *T. rangeli* reflejan una selección relajada, consistente con su ecología no patógena. Estos resultados destacan el papel de la evolución de familias génicas en la configuración de la virulencia y adaptación al hospedero en tripanosomátidos.

**Palabras clave:** *Trypanosoma cruzi*, *Trypanosoma rangeli*, AMASTIN, DGF-1, expansión de familias génicas, filogenómica, genómica comparativa.



## ABSTRACT

The difference in pathogenicity between *Trypanosoma cruzi* and *Trypanosoma rangeli* constitutes a key point in parasitology as it allows us to understand why one of these species is highly pathogenic for humans while the other is considered nonpathogenic. Exploring these differences is essential to elucidate the molecular factors that determine invasiveness and interaction with the immune system. In this context, gene families encoding surface proteins such as AMASTIN and DGF-1 play a central role, being directly associated with the mechanisms of adhesion, host cell invasion, and evasion of the immune response. However, their evolutionary history and expansion in both species remain poorly understood.

The main objective of this research was to analyze the expansion of the AMASTIN and DGF-1 gene families in *T. cruzi* and *T. rangeli* and its possible relationship with the differential pathogenicity between the two species throughout their evolutionary history. To this end, complete genomes of both species were selected, incorporating *Trypanosoma brucei* as an outgroup. Genomic data and annotations were obtained from NCBI GenBank and processed in a Linux-based bioinformatics environment using Python, ClustalW, IQ-TREE2, and HMMER, among other tools, for sequence extraction, multiple alignments, phylogenetic tree construction, and identification of conserved domains.

The results showed that *Trypanosoma cruzi* presented a significant expansion of both gene families (AMASTIN: up to 39 copies; DGF-1: >1000 copies in some strains) with high sequence conservation and a structured domain organization. In contrast, *T. rangeli* exhibited reduced copy numbers (AMASTIN 8; DGF-1 18) and greater structural variability. Phylogenetic analysis confirmed that the AMASTIN and DGF-1 protein families in *T. cruzi* and *T. rangeli* form monophyletic clades, supporting a common ancestor within the subgenus Schizotrypanum, followed by species-specific expansion events. *T. brucei* lacked DGF-1 and retained only one copy of AMASTIN.

Since the analyzed genomes originate from different sequencing technologies and exhibit variations in their assembly and annotation quality, absolute gene numbers should be interpreted with caution. Nevertheless, the massive expansion and conservation of AMASTIN and DGF-1 in *T. cruzi* suggests strong adaptive pressure related to its intracellular lifestyle. The reduced copy number and greater variability in *T. rangeli*



reflect relaxed selection, consistent with its nonpathogenic ecology. These results highlight the role of gene family evolution in shaping virulence and host adaptation in trypanosomatids.

Keywords: *Trypanosoma cruzi*, *Trypanosoma rangeli*, AMASTIN, DGF-1, gene family expansion, phylogenomics, comparative genomics.



## INTRODUCCIÓN

La enfermedad de Chagas es una de las principales parasitosis desatendidas en América Latina y representa un grave problema de salud pública con aproximadamente 6 a 7 millones de personas infectadas y una alta carga de morbilidad y mortalidad, especialmente en sus fases crónicas (Mazza, 2009; Peña-Callejas et al., 2022; World Health Organization, 2025). Se estima que cerca de 100 millones de personas están en riesgo de infección en todo el mundo y que esta enfermedad causa alrededor de 10 000 muertes anuales. En América Latina, la Organización Panamericana de la Salud señala que aproximadamente 75 millones de personas se encuentran expuestas a la enfermedad y que cada año se producen alrededor de 10 000 defunciones atribuibles al Chagas (Organización Panamericana de la Salud, 2023). En el caso de Ecuador no se dispone de una estimación exacta de la población total en riesgo ni de cifras actualizadas de mortalidad específica por la enfermedad, aunque se reconoce la existencia de zonas endémicas en varias regiones del país (Morales-Viteri et al., 2022).

Según Abad- Franch & Aguilar V, 2004 en Ecuador la prevalencia de la infección por *Trypanosoma cruzi* en 2003 se estimaba en 1,38%, lo que equivalía a cerca de 170 000 personas seropositivas. Las regiones más afectadas eran la Costa (1,99%) y la Amazonía (1,75%), en contraste con la Sierra sur, donde la prevalencia fue menor (0,65%). Dos décadas más tarde los datos de vigilancia epidemiológica muestran que la transmisión de la enfermedad continúa ya que en 2023 se confirmaron 118 casos, principalmente en las provincias de Cotopaxi, El Oro, Loja y Pichincha. En 2024 se notificaron 120 casos de los cuales el 93,7% correspondió a la fase crónica. Asimismo, hasta la semana epidemiológica 17 de 2025 se han reportado 43 casos, con mayor concentración en Cotopaxi, El Oro, Pichincha y Manabí (Ministerio de Salud Pública del Ecuador, 2025)

La enfermedad de Chagas es causada por el protozoo flagelado *Trypanosoma cruzi* y se transmite principalmente a través de insectos triatominos del género *Rhodnius*, aunque también puede propagarse por transfusiones sanguíneas, trasplantes, vía congénita y oral (World Health Organization, 2025). En contraste, *Trypanosoma rangeli* es un protozoo flagelado de la misma familia que comparte vectores, hospedadores y distribución geográfica con *T. cruzi*, pero no es patógeno para los vertebrados. Aunque puede infectar a humanos de manera asintomática, su mayor impacto ocurre en los insectos vectores a los que afecta negativamente (Guhl & Vallejo, 2003). A pesar de su aparente inocuidad



clínica, *T. rangeli* constituye un desafío diagnóstico por su similitud morfológica con *T. cruzi* y la reactividad cruzada en pruebas serológicas que pueden generar falsos positivos (Cuba, 1998; Guhl & Vallejo, 2003). No obstante, ha adquirido relevancia como modelo experimental en el estudio de las interacciones parásito-vector-hospedero, permitiendo avanzar en la comprensión de los mecanismos de transmisión, los determinantes de la patogenicidad de *T. cruzi* y el diseño de herramientas diagnósticas más específicas.

Los avances en la biología computacional y genómica comparativa han proporcionado herramientas valiosas para comprender las bases moleculares de la patogenicidad diferencial entre especies del género *Trypanosoma*. En particular, el genoma haploide de *Trypanosoma rangeli* tiene aproximadamente 24 Mb de longitud, lo que lo convierte en el genoma de tripanosomátido más pequeño y menos repetitivo secuenciado hasta ahora (Callejas-Hernández et al., 2018; Stoco et al., 2014; Vallejo et al., 2015). A pesar de su reducido tamaño, se ha reportado que *T. rangeli* comparte alrededor de 7 101 genes, equivalentes al 93 % de su contenido génico, con otros tripanosomátidos que infectan humanos, como *T. cruzi* y *Leishmania spp* (El-Sayed et al., 2005). Esta similitud genómica general sugiere que las diferencias en patogenicidad podrían atribuirse a la variación en familias génicas específicas, como las AMASTINAS, MASP, TRANS-SIALIDASAS, MUCINAS, DGF-1 Y GP63 o a la regulación diferencial de la expresión génica (Callejas-Hernández et al., 2018; Vallejo et al., 2015). En este sentido, estudios del secretoma y del contenido génico han demostrado que *T. rangeli* presenta una representación mucho menor de estas familias multigénicas asociadas con virulencia, lo que refuerza su carácter no patógeno en humanos (K. R. Bradwell et al., 2018).

Aunque *Trypanosoma cruzi* y *Trypanosoma rangeli* comparten una organización genómica similar, con cromosomas estructuralmente complejos, alta plasticidad y abundantes secuencias repetitivas, presentan diferencias notables en la expansión de familias génicas clave y en la capacidad de inducir patogenicidad. Ambas especies contienen familias de genes como *amastin*, *dgf-1*, *masp*, *mucinas* y *trans-sialidasas*, que codifican proteínas correspondientes (AMASTIN, DGF-1, MASP, MUCINAS y TRANS-SIALIDASAS) implicadas en procesos críticos como la invasión celular, la evasión inmunológica y la supervivencia en el hospedador. Sin embargo, *T. rangeli* posee un número significativamente menor de estas copias, así como un repertorio reducido de genes que codifican enzimas antioxidantes, lo cual podría explicar su menor capacidad



para evadir la respuesta inmune y establecer infecciones crónicas (K. R. Bradwell et al., 2018; Watanabe Costa et al., 2020). Además, aunque *T. cruzi* presenta un genoma estructurado en regiones conservadas y una región caracterizada por alta variabilidad y abundancia de familias multigénicas vinculadas con la evasión del sistema inmune y la invasión celular, una organización similar también se observa en *T. rangeli*, aunque en menor grado (Herrerros-Cabello et al., 2020; Stoco et al., 2014). Estas diferencias genómicas proporcionan una base molecular plausible para explicar la divergencia en la patogenicidad entre ambas especies de tripanosomas.

Los estudios de proteómica han aportado información clave sobre los mecanismos de patogenicidad de los tripanosomas. En *T. cruzi* se identificaron 322 proteínas secretadas potencialmente patogénicas (PSPs), frente a las 202 halladas en *T. rangeli* (Watanabe Costa et al., 2020). Entre las PSPs de *T. cruzi* destacan TRANS-SIALIDASAS, MUCINAS, proteínas asociadas a mucina (MASPs), proteínas tipo fosfolipasa A<sub>2</sub> (PLA<sub>2</sub>-like) y chaperonas Hsp70-like, todas ellas vinculadas a su virulencia (Watanabe Costa et al., 2020). Por su parte, las PSPs de *T. rangeli* están relacionadas principalmente con el metabolismo del protozooario, incluyendo carboxilasas y fosfatasas (Watanabe Costa et al., 2020).

Además, estudios recientes han mostrado que las diferencias en la tolerancia al estrés oxidativo pueden influir en la patogenicidad diferencial. *T. rangeli* presenta una menor cantidad de enzimas antioxidantes y chaperonas de choque térmico, lo que explica su alta sensibilidad al estrés oxidativo (Stoco et al., 2014). En concreto, los epimastigotes de *T. rangeli* (cepa SC-58) son cinco veces más sensibles al peróxido de hidrógeno, con un IC<sub>50</sub> de 60 µM, en comparación con los 300 µM observados en *T. cruzi* (cepa Y) (Stoco et al., 2014). Esta mayor susceptibilidad podría limitar la capacidad de *T. rangeli* para establecer infecciones persistentes en hospedadores vertebrados.

Los análisis filogenómicos realizados con 1,557 genes ortólogos permitieron construir árboles evolutivos utilizando métodos robustos como Neighbor-Joining (NJ) y Maximum Likelihood (ML), basados en una alineación concatenada de 964 591 residuos aminoacídicos. Estos árboles mostraron que *Trypanosoma cruzi* y *Trypanosoma rangeli*, ambos originarios de Sudamérica, están evolutivamente equidistantes del tripanosoma africano *T. brucei*, lo que indica que los tres comparten un ancestro común y que ni *T. cruzi* ni *T. rangeli* están más cercanamente relacionados con *T. brucei*. Este tipo de



análisis proporciona un marco evolutivo fundamental para comprender cómo las presiones selectivas han influido en la aparición de rasgos distintivos, como la capacidad patogénica en estas especies de tripanosomas (Stoco et al., 2014).

Las tecnologías de secuenciación masiva y las herramientas de análisis bioinformático han permitido avances significativos en la caracterización genómica de los tripanosomatidos (Bartholomeu et al., 2021). En *T. cruzi*, se han generado ensamblajes de alta calidad para diversas cepas, incluyendo la cepa DM25 y diferentes linajes clínicos, que han facilitado el estudio de la variabilidad genética y la expansión génica en familias vinculadas a la patogenicidad (Hoyos Sanchez et al., 2024; Wang et al., 2021). Por su parte, el genoma de *T. rangeli* cuenta con menos cepas disponibles, destacándose las cepas SC58 y AM80, cuyas secuencias han servido como referencia para estudios comparativos (K. R. Bradwell et al., 2018; Stoco et al., 2014).

La presente investigación se propone analizar detalladamente la historia evolutiva de las familias génicas AMASTIN Y DGF-1 en ambas especies, evaluando su número de copias, organización genómica y relación filogenética, con el fin de identificar patrones de expansión o contracción asociados a la patogenicidad. Para ello, se combinará la obtención de secuencias proteicas anotadas de repositorios públicos, la caracterización de familias génicas mediante herramientas como OrthoFinder y BioMart, la construcción de árboles filogenéticos con IQ-TREE y el análisis de la expansión génica a través de representaciones gráficas.

Abordar las diferencias moleculares y evolutivas entre *Trypanosoma cruzi* y *Trypanosoma rangeli* es fundamental debido a su impacto directo en la salud pública. Comprender con precisión qué factores genéticos determinan que *T. cruzi* sea patógeno mientras que *T. rangeli* no lo sea, permitirá desarrollar herramientas diagnósticas más específicas que minimicen los falsos positivos y mejoren la identificación del agente causal. Esto no solo contribuirá a un diagnóstico más preciso, sino que también facilitará el diseño de terapias más efectivas y dirigidas. Además, estudiar los procesos evolutivos que explican estas diferencias ayudará a entender cómo ha evolucionado la virulencia en los tripanosomas, lo cual es clave para anticipar posibles cambios futuros en su comportamiento patológico y para el diseño de estrategias preventivas más eficaces.



## **OBJETIVOS**

### **Objetivo principal**

Analizar la expansión de las familias génicas AMASTIN y DGF-1 en *Trypanosoma cruzi* y *Trypanosoma rangeli*, y su posible relación con la patogenicidad diferencial entre ambas especies a lo largo de su historia evolutiva.

### **Objetivos específicos**

Evaluar la historia evolutiva de las familias génicas AMASTIN y DGF-1 en *T. cruzi* y *T. rangeli* mediante árboles filogenéticos construidos con el método de máxima verosimilitud.

Identificar y comparar el número de copias de proteínas de las familias AMASTIN y DGF-1 en *T. cruzi* y *T. rangeli* a partir de archivos FASTA anotados.

Discutir la presencia o ausencia de las familias génicas AMASTIN y DGF-1 en *T. rangeli*, comparándolas con su expansión en *T. cruzi*.

## **MARCO TEORICO**

### **1.1. Biología y Patogenicidad de Trypanosoma**

El ciclo de vida de *Trypanosoma cruzi* y *Trypanosoma rangeli* incluye tres etapas principales: la forma epimastigote, que se multiplica en el intestino del insecto vector; la forma tripomastigote metacíclico, que es la fase infectante para los vertebrados, incluyendo a los humanos y la forma amastigote, que se replica dentro de las células del hospedador, especialmente en tejidos como el corazón o el sistema digestivo (Garcia et al., 2012; Guhl & Vallejo, 2003; Silva Gonçalves et al., 2018). Aunque ambos parásitos utilizan los mismos vectores e infectan a los mismos tipos de hospedadores sólo *T. cruzi* causa la enfermedad de Chagas caracterizada por inflamación del corazón (miocarditis) y agrandamiento crónico de órganos como el esófago o el colon (megavisceras). Por el contrario, *T. rangeli* coexiste sin provocar manifestaciones clínicas evidentes, lo que sugiere la ausencia o atenuación de determinantes moleculares de virulencia en esta especie (K. R. Bradwell et al., 2018).



## 1.2. Arquitectura Genómica en Tripanosomas

El genoma de los tripanosomas se compone de ADN nuclear empaquetado en 41 cromosomas y de ADN mitocondrial organizado en maxicírculos y minicírculos, que codifican ARN guía para la edición de ARN (Westenberger et al., 2006). El tamaño del genoma nuclear puede diferir de forma considerable entre especies y cepas, situándose por lo general en un rango de 30 a 70 megabases, lo que refleja diferencias en el número de cromosomas y en la extensión de regiones repetitivas (El-Sayed et al., 2005; Jackson, 2007). La presencia de repetidos en tándem y la multicopia génica contribuyen a una gran plasticidad estructural, promoviendo la expansión génica a través de duplicaciones tandem y segmentales (Jackson, 2007; Reis-Cunha et al., 2015). Estos procesos dinámicos de contracción y expansión de copias modulan la adaptación al hospedero y son determinantes en la virulencia de estos parásitos (Jackson, 2015).

La expansión del genoma en *Trypanosoma cruzi*, agente etiológico de la enfermedad de Chagas, se debe principalmente a la proliferación de familias génicas multigénicas y elementos repetitivos. Estas expansiones comprenden genes que codifican proteínas de superficie, como MUCINAS, TRANS-SIALIDASAS y PROTEÍNAS ASOCIADAS A MUCINAS (MASP), cuya diversidad funcional permite la evasión del sistema inmunológico del hospedador y la adaptación a diferentes nichos celulares. Además, los genomas de *T. cruzi* y otros tripanosomas contienen numerosos retroelementos y secuencias repetitivas dispersas que contribuyen al tamaño genómico y generan variabilidad estructural mediante duplicaciones, recombinaciones y transposiciones. Estudios comparativos han demostrado que estas expansiones no solo aumentan el número de genes, sino que también facilitan la generación de nuevas variantes génicas, lo que representa un mecanismo evolutivo clave para la supervivencia de estos parásitos frente a presiones selectivas impuestas por el hospedador y el vector (Baptista et al., 2018; Callejas-Hernández et al., 2018; Herreros-Cabello et al., 2020).

La expansión del genoma en *Trypanosoma cruzi*, se debe principalmente a la proliferación de familias génicas multigénicas y elementos repetitivos. Estas expansiones comprenden genes que codifican proteínas de superficie, como MUCINAS, TRANS-SIALIDASAS y proteínas asociadas a mucinas (MASP), cuya diversidad funcional permite la evasión del sistema inmunológico del hospedador y la adaptación a diferentes nichos celulares. Además, los genomas de *T. cruzi* y otros tripanosomas contienen



numerosos retroelementos y secuencias repetitivas dispersas que contribuyen al tamaño genómico y generan variabilidad estructural mediante duplicaciones, recombinaciones y transposiciones. Estudios comparativos han demostrado que estas expansiones no solo aumentan el número de genes, sino que también facilitan la generación de nuevas variantes génicas, lo que representa un mecanismo evolutivo clave para la supervivencia de estos parásitos frente a presiones selectivas impuestas por el hospedador y el vector (Baptista et al., 2018; Callejas-Hernández et al., 2018; Herreros-Cabello et al., 2020).

Las familias multigénicas constituyen conjuntos de genes relacionados que comparten secuencias similares y con frecuencia funciones relacionadas. Estas familias surgen principalmente por duplicaciones génicas, recombinaciones y transposiciones, procesos que generan diversidad genética y permiten la especialización funcional de los genes. En organismos eucariotas, las familias multigénicas cumplen roles fundamentales en la adaptación a diferentes condiciones ambientales y en la regulación de procesos biológicos complejos. En los tripanosomas, las familias multigénicas son particularmente abundantes y están fuertemente implicadas en la interacción con el hospedador y la supervivencia del parásito. Estas incluyen genes que codifican proteínas de superficie, enzimas implicadas en la evasión inmunológica y otras proteínas asociadas a la patogénesis, reflejando la importancia de la expansión génica como mecanismo evolutivo frente a las presiones selectivas impuestas por el hospedador y el vector. La presencia de múltiples copias y variantes génicas permite a los tripanosomas generar diversidad antigénica y funcional, constituyendo un rasgo característico de su genoma (El-Sayed et al., 2005; Weatherly et al., 2009).

### **1.3. Familias Multigénicas Clave: AMASTIN Y DGF-1**

Las AMASTINAS son glicoproteínas de membrana características de la forma amastigote de los tripanosomas, implicadas en la adaptación del parásito al ambiente intracelular hostil, particularmente por su rol en la resistencia al estrés lisosómico (Kangussu-Marcolino et al., 2013). En *Trypanosoma cruzi*, se ha identificado un repertorio amplio y diverso de genes que codifican para distintas subfamilias de AMASTINAS, con variaciones tanto en el número de copias como en sus patrones de expresión, lo que sugiere una evolución adaptativa de estas proteínas en contextos específicos del ciclo de vida del parásito (K. Bradwell, 2016). Por otro lado, la familia génica DGF-1 (Dispersed Gene Family 1), distribuida en múltiples loci cromosómicos, ha sido asociada con



funciones clave en la invasión celular y la evasión del sistema inmunitario del hospedero, aunque su función completa aún no ha sido completamente dilucidada (Herrerros-Cabello et al., 2025). Comparativamente, el genoma de *T. rangeli* presenta un número considerablemente menor de copias tanto de genes de *amastina* como de la familia *dgf-1*, lo que se ha propuesto como un factor que podría contribuir a su falta de patogenicidad en humanos (Watanabe Costa et al., 2020).

#### **1.4. Fundamentos de Filogenia y Evolución Molecular**

La reconstrucción filogenética distingue ortólogos (genes divergidos tras un evento de especiación) de parálogos (genes originados por duplicación) lo cual es fundamental para interpretar relaciones evolutivas y funcionales en genomas complejos (Nevers et al., 2022). Los métodos de máxima verosimilitud, implementados en IQ-TREE 2, emplean modelos de sustitución de aminoácidos específicos para cada sitio y permiten inferir la topología de los árboles filogenéticos con alta eficiencia computacional (Minh et al., 2020). El soporte estadístico se evalúa mediante bootstraps ultrarrápidos (UFBoot2) que ofrecen estimaciones de confianza comparables a los métodos clásicos, pero en una fracción del tiempo de cálculo (Hoang et al., 2018). Al mapear eventos de duplicación y pérdida sobre la filogenia resultante, este enfoque posibilita rastrear la expansión y contracción de familias génicas multigénicas, como la subfamilia  $\beta$ -amastina, cuya regulación específica en distintas fases de desarrollo de *T. cruzi* ha sido demostrada mediante proteómica cuantitativa (Avila et al., 2018), y la familia DGF-1, cuyo patrón de dispersión genómica sugiere un papel clave en las interacciones huésped-parásito (Ramírez, 2023).

La evolución de las familias génicas en Trypanosoma ha sido fundamental para la adaptación y patogenicidad de estos parásitos. La expansión y diversificación de genes, como las *mucinas*, *trans-sialidasas*, proteínas asociadas a mucinas (MASP) y proteasas de superficie mayor (MSP), han permitido a los tripanosomas evadir el sistema inmunológico del hospedador y colonizar diversos tejidos. Estas familias multigénicas, altamente variables y sometidas a recombinación, generan una diversidad antigénica que facilita la persistencia crónica de la infección y la propagación del parásito. Además, la coevolución de estas familias con regiones teloméricas y subteloméricas ha influido en la estructura y dinámica del genoma de *T. cruzi*, contribuyendo a su plasticidad genética y



capacidad de adaptación (Baptista et al., 2018; El-Sayed et al., 2005; Herreros-Cabello et al., 2020; Hoyos Sanchez et al., 2024; Weatherly et al., 2009).

## **METODOLOGIA**

### **2.1. Selección de genomas para análisis comparativo**

Se seleccionaron los genomas completos de *Trypanosoma cruzi* y *Trypanosoma rangeli*, incorporando a *Trypanosoma brucei* como outgroup para los análisis filogenéticos. Estos genomas fueron obtenidos mediante diferentes tecnologías de secuenciación (Sanger, Illumina, PacBio, Nanopore) y presentan variaciones en el grado de ensamblaje y anotación.

La selección de los genomas se realizó utilizando un *DataFrame* de Python que contenía información sobre cada cepa, incluyendo el tamaño del genoma, el porcentaje de GC, el número total de genes y el número de genes codificantes.

Se aplicaron dos criterios principales para la inclusión de un genoma en el análisis: (i) disponibilidad de la anotación genómica completa y (ii) presencia de genes codificadores de proteínas correctamente identificados. Los genomas que cumplían con ambos criterios fueron conservados para los análisis posteriores, mientras que aquellos con datos incompletos o con anotaciones insuficientes fueron descartados.

A partir de las anotaciones disponibles, se identificaron las familias génicas AMASTIN y DGF-1, cuyos repertorios se analizaron mediante búsquedas de homología, caracterización estructural y reconstrucciones filogenéticas. Debido a la heterogeneidad técnica de los ensamblajes, los números absolutos de genes deben interpretarse con cautela, priorizando la comparación de patrones generales de expansión y conservación.

Esta estrategia permitió garantizar que los datos utilizados fueran de alta calidad, facilitando comparaciones precisas de las familias génicas AMASTIN y DGF-1 entre las distintas cepas de *Trypanosoma*.

### **2.2. Obtención de datos**

Los genomas completos y sus anotaciones fueron obtenidos de la base de datos pública NCBI GenBank (<https://www.ncbi.nlm.nih.gov/>) bajo los siguientes números de



accesión: *Trypanosoma cruzi* CL Brener (PRJNA11755), *T. cruzi* Y clone C6 (PRJNA554625), *T. cruzi* Brazil clone A4 (PRJNA512864), *T. cruzi* Dm28c05.1 (PRJNA433042), *T. cruzi* TCC (PRJNA432753), *T. cruzi* Berenice (PRJNA498808), *T. cruzi* Sylvio X10/1 (PRJNA40815), *T. cruzi* Dm28c95.1 (PRJNA225928), *T. cruzi* Dm28c05.2 (PRJNA330977), *T. cruzi* B7 (PRJNA77843), *T. cruzi* G (PRJNA315397), *T. cruzi* CL (PRJNA315397), *Trypanosoma rangeli* AM80 (PRJNA315397), *T. rangeli* SC58 (PRJNA18945) y *Trypanosoma brucei* 927/4 GUTat10.1 (PRJNA11756). Estos recursos constituyeron la base para los análisis genómicos y comparativos realizados en este estudio.

### **2.3. Procesamiento y manejo de secuencias**

Las secuencias correspondientes a cada familia fueron obtenidas y procesadas en un entorno bioinformático implementado en Linux. Todo el flujo de trabajo, que incluyó el procesamiento, organización y manejo de los archivos, se desarrolló en un cuaderno de Jupyter Notebook ejecutado sobre una máquina virtual configurada en VirtualBox Manager, con sistema operativo Debian GNU/Linux 12 (Debian 12-server tty, kernel 6.1.0, arquitectura AMD64)

### **2.4. Filtrado y obtención de las secuencias AMASTIN y DGF-1**

El procedimiento inició con la filtración de las anotaciones de los genomas de referencia de *T. cruzi*, *T. rangeli* y *T. brucei*, empleando los comandos *grep*, *cut* y *sed* para localizar las cabeceras que contenían las palabras clave “AMASTIN” o “DGF-1” y depurar los identificadores correspondientes. Posteriormente, mediante la herramienta *SeqKit* se extrajeron las secuencias proteicas completas asociadas a dichos identificadores, generando archivos FASTA específicos para cada especie y familia génica. ’

Con el fin de garantizar la correcta clasificación de las secuencias y evitar falsos positivos derivados de anotaciones incompletas o erróneas, las proteínas seleccionadas fueron validadas mediante la identificación de dominios conservados utilizando HMMER v3.4 contra la base de datos Pfam . De esta manera, se confirmó que las secuencias obtenidas pertenecían a las familias génicas de interés.



## **2.5. Comparación de secuencias AMASTIN y DGF-1 en cepas de *Trypanosoma cruzi*, *Trypanosoma rangeli* y *Trypanosoma brucei*.**

Para representar gráficamente la cantidad de secuencias pertenecientes a las familias génicas AMASTIN y DGF-1 en distintas cepas de *Trypanosoma*, se empleó Python (v.3.x) junto con las librerías Pandas (v.1.x), Matplotlib (v.3.x) y Seaborn (v.0.11.x).

Primero, la información se organizó en un diccionario de Python con las columnas: Organismo, Cepa, Número de Secuencias AMASTIN y Número de Secuencias DGF-1, donde cada fila correspondió a una cepa específica. Posteriormente, los datos se convirtieron en un *DataFrame* mediante Pandas para su procesamiento.

Con el objetivo de facilitar la comparación, el *Data Frame* se transformó al formato largo (*long format*) utilizando la función `melt()`, agrupando las columnas de las proteínas AMASTIN y DGF-1 bajo una sola variable denominada Parámetro, con sus valores asociados en la columna Valor.

Para la visualización, se generó un gráfico de barras horizontales con la función `catplot()` de Seaborn, separando los resultados de AMASTIN y DGF-1 en paneles independientes (`col="Parámetro"`). En el eje Y se ubicaron las cepas analizadas y en el eje X el número de secuencias detectadas. Se aplicó una paleta de colores categórica para diferenciar los organismos y resaltar las variaciones entre especies.

Finalmente, la figura fue exportada en alta resolución (300 dpi) en formato PNG mediante `plt.savefig()`.

## **2.6. Alineamiento de secuencias**

Para el análisis comparativo de las familias génicas AMASTIN y DGF-1 en *Trypanosoma cruzi*, *T. rangeli* y *T. brucei*, se efectuó un alineamiento múltiple de las secuencias proteicas utilizando el software ClustalW (versión 2.1). Como datos de entrada se utilizaron los archivos en formato FASTA con las secuencias proteicas completas previamente obtenidas, a partir de los cuales se generaron alineamientos en formato CLUSTAL. Este procedimiento permitió obtener alineamientos consistentes, esenciales para la identificación de regiones conservadas y el desarrollo de análisis filogenéticos comparativos entre las especies estudiadas.



## 2.7. Representación gráfica del alineamiento de secuencias

Para visualizar de manera clara el grado de conservación en el alineamiento múltiple de las secuencias de la familia DGF-1 en *Trypanosoma cruzi*, se empleó la librería Biopython para cargar el archivo en formato Clustal. Las secuencias fueron convertidas a una matriz de caracteres mediante pandas, donde cada fila representó una secuencia y cada columna una posición en el alineamiento.

A partir de esta matriz, se generó un perfil de consenso que permitió identificar el residuo más frecuente en cada posición. Con esta información, se construyó una matriz binaria (1 = posición conservada, 0 = posición variable), la cual fue representada gráficamente como un mapa de calor mediante seaborn y matplotlib.

## 2.8. Construcción de árboles filogenéticos mediante IQ-TREE2

Para inferir las relaciones filogenéticas de las familias génicas AMASTIN y DGF-1 en *Trypanosoma cruzi* y *T. rangeli*, utilizando a *T. brucei* como grupo externo, se empleó el software IQ-TREE2. Los alineamientos múltiples de aminoácidos obtenidos para cada familia génica sirvieron como archivos de entrada. La selección del modelo evolutivo más adecuado se realizó automáticamente mediante la función ModelFinder Plus (MFP). La solidez de las topologías resultantes se evaluó con 1.000 réplicas de bootstrap ultrarrápido (-bb 1000) y mediante la prueba de razón de verosimilitud aproximada aLRT (-alrt 1000), proporcionando valores de soporte estadístico complementarios para cada rama. Los análisis se ejecutaron con asignación automática de núcleos de procesamiento (-nt AUTO), optimizando el rendimiento computacional y generando árboles filogenéticos robustos para ambas familias génicas en las especies analizadas.

## 2.9. Visualización de árboles filogenéticos

Los árboles filogenéticos de las familias AMASTIN y DGF-1 en *T. cruzi*, *T. rangeli* y *T. brucei* (como grupo externo) fueron visualizados utilizando la plataforma web iTOL (Letunic & Bork, 2024). Para ello, se cargaron los archivos .treefile generados con IQ-TREE y se personalizó su apariencia ajustando los colores de ramas y etiquetas, además de incorporar los valores de soporte estadístico de bootstrap y UFBoot. Finalmente, los árboles fueron exportados en formato PNG, facilitando una representación clara y



comparativa de las relaciones filogenéticas entre las especies y las familias génicas analizadas.

## **2.10. Identificación de dominios conservados en proteínas AMASTIN y DGF-1**

Para la familia DGF-1 de *Trypanosoma cruzi*, que contenía un total de 429 secuencias, se realizó una reducción del conjunto para obtener representantes de cada clado y subclado identificados en el árbol filogenético. Para ello, se utilizó el paquete ETE3 en Python, seleccionando, dentro de cada agrupamiento definido por color, la secuencia con menor distancia filogenética promedio respecto a las demás del mismo grupo. Esta estrategia permitió conservar la diversidad filogenética del conjunto original, minimizando la redundancia de secuencias en los análisis posteriores.

Las secuencias proteicas de las familias AMASTIN y DGF-1 de *Trypanosoma cruzi* y *T. rangeli* fueron analizadas mediante HMMER v3.4 en su versión web (<https://www.ebi.ac.uk/Tools/hmmer/>). Para cada conjunto de secuencias se utilizó la herramienta hmmscan, comparándolas contra la base de datos Pfam con el objetivo de identificar dominios conservados asociados a cada familia proteica. Las coincidencias se evaluaron mediante los valores estadísticos de significancia E-value (independiente y condicional), estableciendo como criterio de detección un E-value  $\leq 0.01$ . Para cada proteína se registraron los dominios identificados, su ubicación dentro de la secuencia (posición inicial y final) y la anotación funcional correspondiente. Finalmente, las posiciones de los dominios AMASTIN y DGF-1 identificados mediante HMMER fueron integradas en sus respectivos árboles filogenéticos, indicando las coordenadas correspondientes dentro de cada proteína.

## **RESULTADOS**

### **3.1. Selección de genomas y características generales**

De las 45 cepas secuenciadas de *Trypanosoma cruzi* inicialmente consideradas, únicamente 12 cumplieron con los criterios de inclusión establecidos, los cuales contemplaban (i) la disponibilidad de una anotación genómica completa y (ii) la presencia de genes codificadores de proteínas correctamente identificados. Los genomas que no cumplieron estas condiciones fueron descartados por presentar información incompleta o anotaciones insuficientes. En *Trypanosoma rangeli*, se identificaron únicamente dos



genomas disponibles, ambos conformes a los requisitos, por lo tanto, incorporados al análisis. En cuanto a *Trypanosoma brucei*, se empleó exclusivamente su genoma de referencia, utilizado como outgroup para los análisis comparativos y filogenéticos (Tabla 1).

Tabla 1. Selección de genomas para análisis comparativo

Numero de accesión	Organismo	Cepa	Tamaño (Mb)	% GC	Anotación Genes Totales	Anotación Genes codificador de proteínas
PRJNA11755	<i>Trypanosoma cruzi</i>	<b>CL Brener</b>	89,9	51,5	25210	19607
PRJNA554625	<i>Trypanosoma cruzi</i>	Y clone C6	47,2	51,5	17651	14336
PRJNA512864	<i>Trypanosoma cruzi</i>	Brazil clone A4	45,6	51,5	18708	14287
PRJNA433042	<i>Trypanosoma cruzi</i>	Dm28c05.1	53,3	51,5	19115	15319
PRJNA432753	<i>Trypanosoma cruzi</i>	TCC	87,1	51,5	29302	24191
PRJNA498808	<i>Trypanosoma cruzi</i>	Berenice	40,8	51	14352	13307
PRJNA40815	<i>Trypanosoma cruzi</i>	Sylvio X10/1	38,6	51	10876	10847
PRJNA225928	<i>Trypanosoma cruzi</i>	Dm28c95.1	27,3	50,5	11398	11348
PRJNA330977	<i>Trypanosoma cruzi</i>	Dm28c05.2	50,9	51,5	22870	13643
PRJNA77843	<i>Trypanosoma cruzi</i>	B7	34,2	51	10127	10104
PRJNA315397	<i>Trypanosoma cruzi</i>	G	25,2	50	13503	12708
PRJNA315397	<i>Trypanosoma cruzi</i>	CL	26,8	50,5	34344	32278
PRJNA315397	<i>Trypanosoma rangeli</i>	<b>AM80</b>	21,2	52	10589	10104
PRJNA18945	<i>Trypanosoma rangeli</i>	SC58	15,1	53	7475	7475
PRJNA11756	<i>Trypanosoma brucei</i>	<b>927/4 GUTat10.1</b>	26,1	46,5	10221	8758

### 3.2. Filtrado, obtención y comparación de secuencias AMASTIN y DGF-1

El análisis comparativo del número de secuencias pertenecientes a las familias génicas AMASTIN y DGF-1 en *Trypanosoma cruzi*, *Trypanosoma rangeli* y *Trypanosoma brucei* evidencia diferencias marcadas en la abundancia de estas familias multigénicas entre especies y cepas.

En la familia AMASTIN, las cepas de *T. cruzi* muestran una gran variabilidad, con valores que van desde tan solo 2 secuencias (cepa Berenice) hasta un máximo de 39 secuencias



(cepa TCC). Cepas representativas como Y clone C6 y Dm28c05.2 presentan 38 y 21 secuencias respectivamente, mientras que otras como Sylvio X10/1 y G poseen un número reducido (2 y 6, respectivamente). En contraste, *T. rangeli* exhibe una menor abundancia, con 8 secuencias en la cepa AM80 y ninguna en SC58, mientras que *T. brucei* presenta únicamente una copia anotada (cepa 927/4 GUTat10.1).

En cuanto a la familia DGF-1, se observa una expansión mucho más notable en *T. cruzi*. Algunas cepas como CL y B7 alcanzan más de 1,000 secuencias (1021 y 958, respectivamente), mientras que otras como Y clone C6 y TCC presentan valores intermedios (429 y 190, respectivamente). Por el contrario, *T. rangeli* muestra una cantidad considerablemente menor, con 18 secuencias en la cepa AM80 y 9 en SC58. Finalmente, *T. brucei* carece de anotaciones para esta familia génica en la cepa analizada.

En conjunto, estos resultados destacan una diferencia sustancial en la expansión génica entre especies, donde *T. cruzi* presenta una marcada proliferación tanto en AMASTIN como en DGF-1, en contraste con la menor abundancia observada en *T. rangeli* y la casi ausencia de estas familias en *T. brucei*. Este patrón sugiere procesos de duplicación génica diferenciales, posiblemente relacionados con las estrategias adaptativas particulares de cada especie de tripanosomátido (Figura 1).

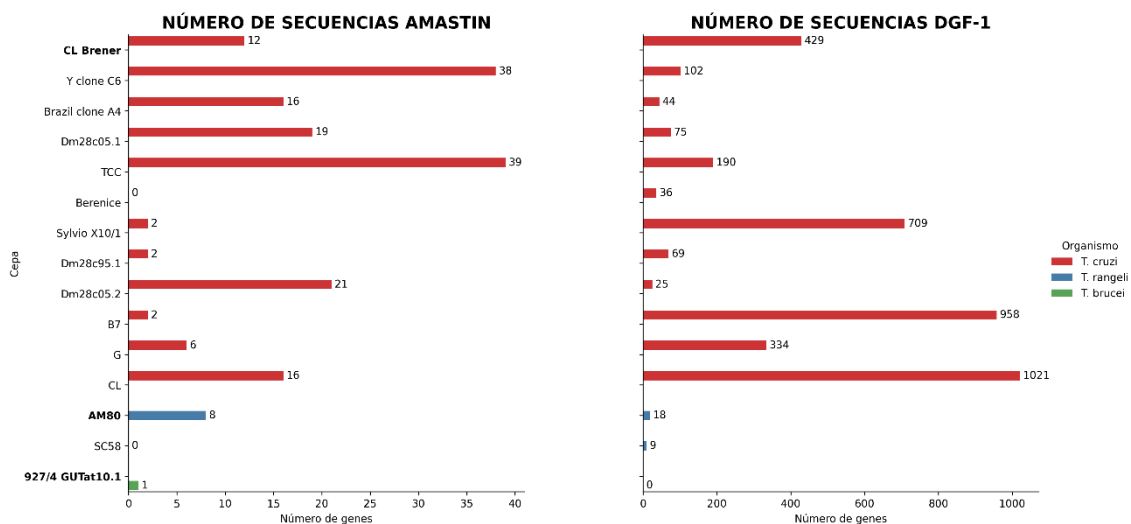


Figura 1. Comparación del número de secuencias de las familias génicas AMASTIN y DGF-1 en distintas cepas de *Trypanosoma cruzi*, *Trypanosoma rangeli* y *Trypanosoma brucei*.



### 3.3. Comparación de las características genómicas de *Trypanosoma cruzi*, *Trypanosoma rangeli* y *Trypanosoma brucei*.

El análisis comparativo de las características genómicas de *Trypanosoma cruzi*, *Trypanosoma rangeli* y *Trypanosoma brucei* revela patrones diferenciados en cuanto a tamaño del genoma, composición de bases y carga génica (Figura 2).

El tamaño del genoma varía significativamente entre especies y cepas. *T. cruzi* presenta la mayor amplitud (30–55 Mb), con la cepa CL Brener como la más extensa y Sylvio X10/1 y B7 en rangos intermedios. *T. rangeli* posee un genoma más pequeño (~20 Mb) y *T. brucei* uno intermedio (~25 Mb).

El contenido de GC es similar en las tres especies, con valores cercanos al 50 %, lo que indica una composición nucleotídica conservada evolutivamente.

En la anotación génica, *T. cruzi* destaca por su abundancia y variabilidad (15,000–30,000+ genes), con CL Brener y CL como las cepas más ricas. *T. rangeli* presenta valores intermedios (10,000–15,000) y *T. brucei* el menor número (~9,000).

La anotación de genes codificantes sigue el mismo patrón, *T. cruzi* muestra la mayor complejidad, *T. rangeli* ocupa una posición intermedia y *T. brucei* posee el repertorio más reducido.

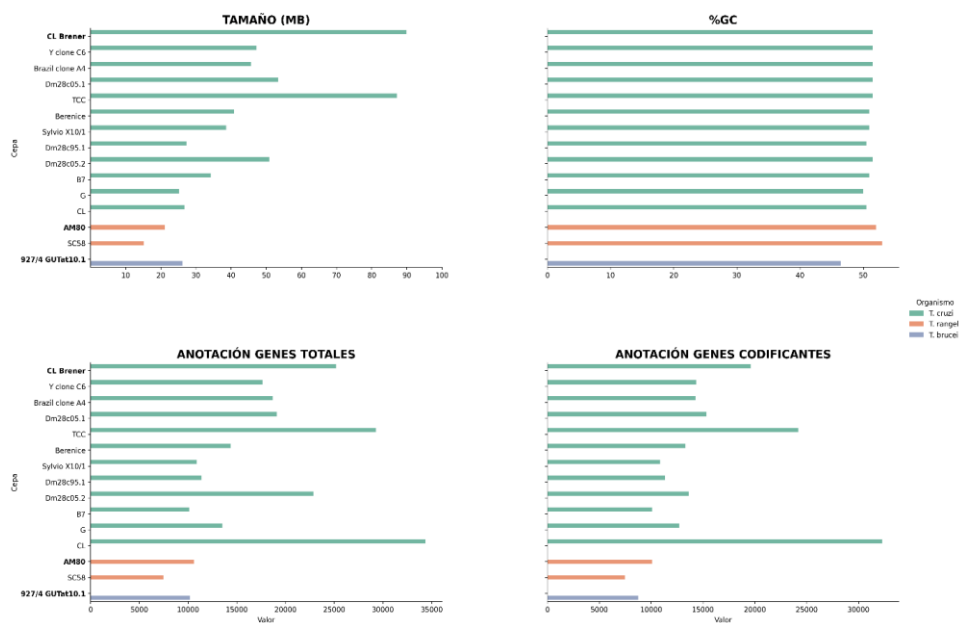


Figura 2. Comparación de las características genómicas de *Trypanosoma cruzi*, *Trypanosoma rangeli* y *Trypanosoma brucei*.



### **3.4. Alineamientos múltiples y conservación de Secuencias de las Familias AMASTIN y DGF-1**

Para evaluar el grado de conservación de las familias génicas de interés, se realizó un análisis de los alineamientos múltiples de las secuencias de AMASTIN y DGF-1 obtenidas de *Trypanosoma cruzi* y *Trypanosoma rangeli*. Como se describió en la metodología, los alineamientos generados con ClustalW fueron procesados mediante un pipeline bioinformático en Python utilizando las librerías Biopython, pandas, seaborn y matplotlib. Este proceso permitió calcular el residuo consenso para cada posición y generar una matriz binaria donde 1 representa una posición conservada y 0 una posición variable, la cual fue visualizada como un mapa de calor para su interpretación.

#### **3.4.1. AMASTIN**

El análisis comparativo del perfil de conservación del gen AMASTIN entre las dos especies revela estrategias evolutivas distintas:

En *Trypanosoma cruzi*, el mapa de conservación del alineamiento de AMASTIN (Figura 3A) confirma un alto grado de conservación a lo largo de la secuencia analizada, con la gran mayoría de las posiciones categorizadas como conservadas (valor 1). Esta conservación extrema refuerza el papel crucial de AMASTIN en la supervivencia intracelular del parásito.

En contraste, el análisis del gen AMASTIN en *Trypanosoma rangeli* (Figura 3B), revela un perfil de conservación significativamente más variable. El mapa de calor muestra un patrón mosaico, con regiones cortas conservadas intercaladas con extensos tramos variables. Está marcada diferencia en el grado de conservación de la secuencia de AMASTIN entre *T. cruzi* (altamente conservado) y *T. rangeli* (variable) sugiere diferencias fundamentales en las presiones selectivas que actúan sobre este gen. Estas diferencias podrían estar directamente relacionadas con la distinta patogenicidad y estrategias de vida de ambos parásitos, donde AMASTIN es crucial para la supervivencia intracelular de *T. cruzi*, pero podría tener un rol diferente o estar sujeto a una evolución más relajada en *T. rangeli*.

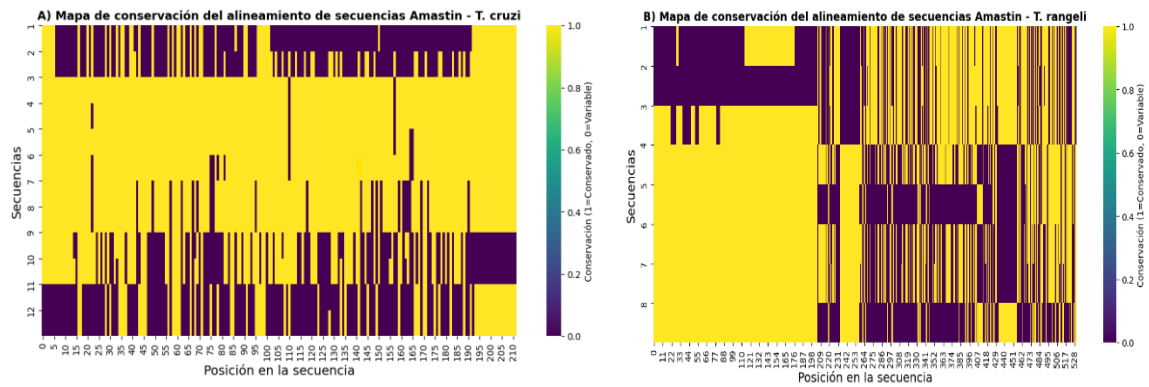


Figura 3. Mapas de conservación de los alineamientos múltiples del gen AMASTIN. (A) *Trypanosoma cruzi*. (B) *Trypanosoma rangeli*. El grado de conservación se representa de forma binaria: Amarillo (1, conservado) y Negro (0, variable).

### 3.4.2. DGF-1

El análisis del mapa de conservación del alineamiento múltiple de la familia DGF-1 mostró diferencias notables entre *Trypanosoma cruzi* y *Trypanosoma rangeli*.

En *T. cruzi*, se identificaron regiones extensas altamente conservadas en los extremos N- y C-terminales de las proteínas, mientras que el centro del alineamiento presentó elevada variabilidad, lo que indica la presencia de dominios funcionales conservados flanqueados por regiones divergentes posiblemente asociadas a diversificación funcional (Figura 4A). Por su parte, en *T. rangeli*, la conservación se distribuyó en segmentos alternados de alta y baja conservación a lo largo de la secuencia, sugiriendo una reorganización estructural con dominios funcionales intercalados por regiones variables (Figura 4B). Estos resultados reflejan que, aunque ambas especies conservan zonas esenciales para la función de las proteínas DGF-1, difieren en la distribución y extensión de las regiones variables, lo que podría estar relacionado con procesos evolutivos específicos y distintos mecanismos de interacción con el hospedero.

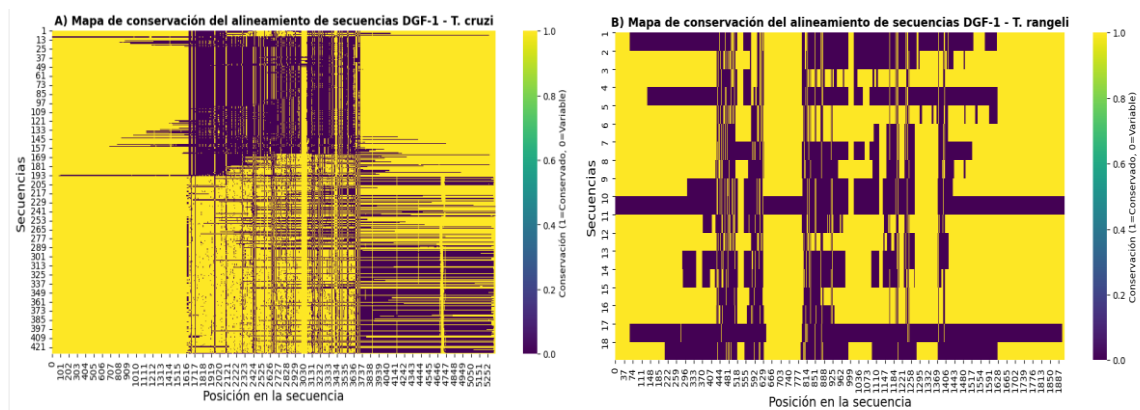


Figura 4. Mapas de conservación de los alineamientos múltiples del gen DGF-1. (A) *Trypanosoma cruzi*. (B) *Trypanosoma rangeli*. El grado de conservación se representa de forma binaria: Amarillo (1, conservado) y Negro (0, variable).

### 3.5. Filogenia de la familia AMASTIN

El análisis filogenético de la familia de genes AMASTIN mediante IQ-TREE2 reveló una distribución diferencial altamente significativa entre las tres especies de tripanosomátidos analizadas:

En *Trypanosoma cruzi* se identificó un conjunto de doce (12) secuencias de AMASTIN (EAN80890.1, EAN82364.1, EAN82366.1, EAN84309.1, EAN84343.1, EAN89651.1, EAN90538.1, EAN90540.1, EAN92151.1, EAN96073.1, EAN98402.1, EAN98408.1). Este número, significativamente mayor, sugiere una expansión génica sustancial de esta familia en *T. cruzi* (Figura 5).

En *Trypanosoma rangeli* se identificaron ocho (8) secuencias (RNE95532.1, RNE95533.1, RNE97308.1, RNE97310.1, RNE97311.1, RNE99637.1, RNE99638.1, RNF04654.1). La presencia de un número considerable de secuencias refuerza la idea de que la familia AMASTIN es un componente ancestral del subgénero *Schizotrypanum* (Figura 6).

En *Trypanosoma brucei* solo se identificó una 1 secuencia homóloga (EAN79808.1). Esta marcada diferencia cuantitativa corrobora hallazgos previos de que la familia AMASTIN verdadera está prácticamente ausente en este linaje, consistente con su estrategia de vida extracelular.



### **3.6. Relaciones Filogenéticas y Evolución de la Familia AMASTIN**

El árbol filogenético, reconstruido bajo el mejor modelo evolutivo seleccionado por ModelFinder Plus y con soporte estadístico robusto (-bb 1000, -alrt 1000), mostró una topología clara y altamente informativa (Figura 7).

**Agrupamiento Específico y Soporte Estadístico:** El árbol filogenético de la familia Amastin muestra que las secuencias de *T. cruzi* (azul) y *T. rangeli* (verde) tienden a formar grupos separados, aunque no de manera completamente cerrada. Esto significa que, aunque existe una clara diferenciación entre especies, algunas copias de genes comparten relaciones que sugieren duplicaciones y divergencias ocurridas antes de que las especies se separaran. Aunque en la figura no se muestran los valores de soporte, los análisis con 1,000 réplicas de bootstrap ultrarrápido y SH-aLRT indicaron niveles de confianza moderados a altos (>70%), lo que confirma la solidez de los agrupamientos principales. El análisis filogenético de las secuencias Amastin mostró que las copias de *T. cruzi* y *T. rangeli* tienden a agruparse con mayor cercanía entre sí que con las de *T. brucei*. Sin embargo, al haberse considerado únicamente tres especies y sin un outgroup definido, la topología obtenida no permite establecer con certeza relaciones filogenéticas concluyentes, aunque sí evidencia una diferenciación clara entre las especies analizadas.

**Expansión Diferencial Post-Especiación:** Se detectó una diferencia en el número de copias de Amastin entre *T. cruzi* (12 secuencias) y *T. rangeli* (8 secuencias) lo que refleja una expansión más amplia en *T. cruzi*. Este patrón podría indicar un proceso de duplicación y diversificación génica más marcado en esta especie, en concordancia con su adaptación como parásito intracelular obligado en mamíferos. Sin embargo, también es necesario considerar que parte de la variación en el número de genes puede estar influida por diferencias en las estrategias de secuenciación, ensamblaje y anotación de los genomas de referencia disponibles para cada especie.

La secuencia homóloga de *T. brucei* se ubicó separada de las copias de *T. cruzi* y *T. rangeli*, sin agruparse dentro de los clados principales de estas especies. Esta posición confirma su mayor divergencia respecto a los miembros del subgénero *Schizotrypanum* y la utilidad de *T. brucei* como referencia comparativa externa.



### *Amastin Trypanosoma cruzi*



*Figura 5. Árbol filogenético de la familia AMASTIN de Trypanosoma cruzi, obtenido con IQ-TREE v2 bajo el mejor modelo de sustitución (-m MFP), con soporte de 1,000 bootstrap ultrarrápido y 1,000 SH-aLRT. Las ramas se muestran coloreadas de manera diferencial para resaltar clados internos y facilitar la visualización; los colores no corresponden a especies ni categorías taxonómicas específicas.*

### AMASTIN *Trypanosoma rangeli*

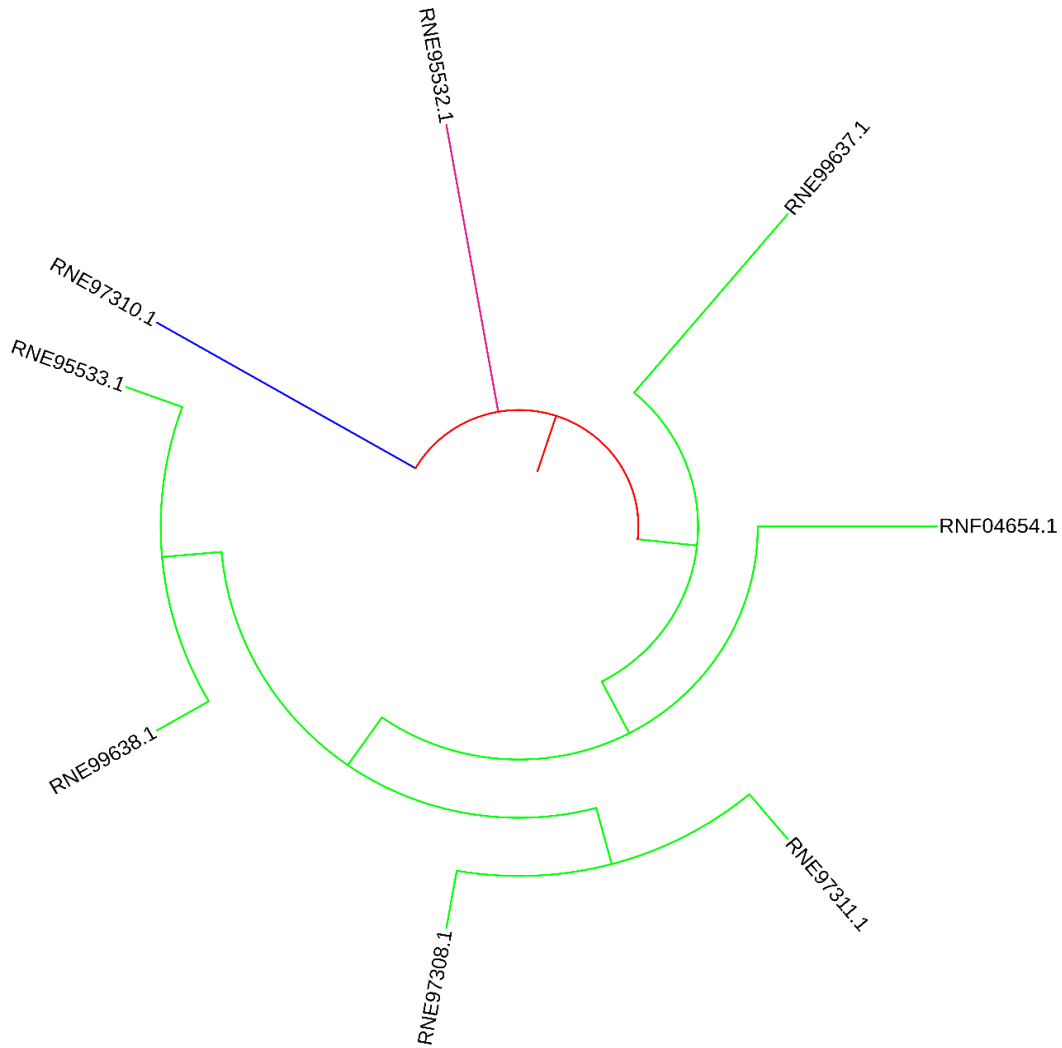


Figura 6. Árbol filogenético de la familia AMASTIN de *Trypanosoma rangeli*, obtenido con IQ-TREE v2 bajo el mejor modelo de sustitución (-m MFP), con soporte de 1,000 bootstrap ultrarrápido y 1,000 SH-aLRT. Las ramas se muestran coloreadas de manera diferencial para resaltar clados internos y facilitar la visualización; los colores no corresponden a especies ni categorías taxonómicas específicas.

**ARBOL FILOGENETICO DE LA FAMILIA AMASTIN EN *Trypanosoma cruzi*, *Trypanosoma rangeli* y *Trypanosoma brucei***

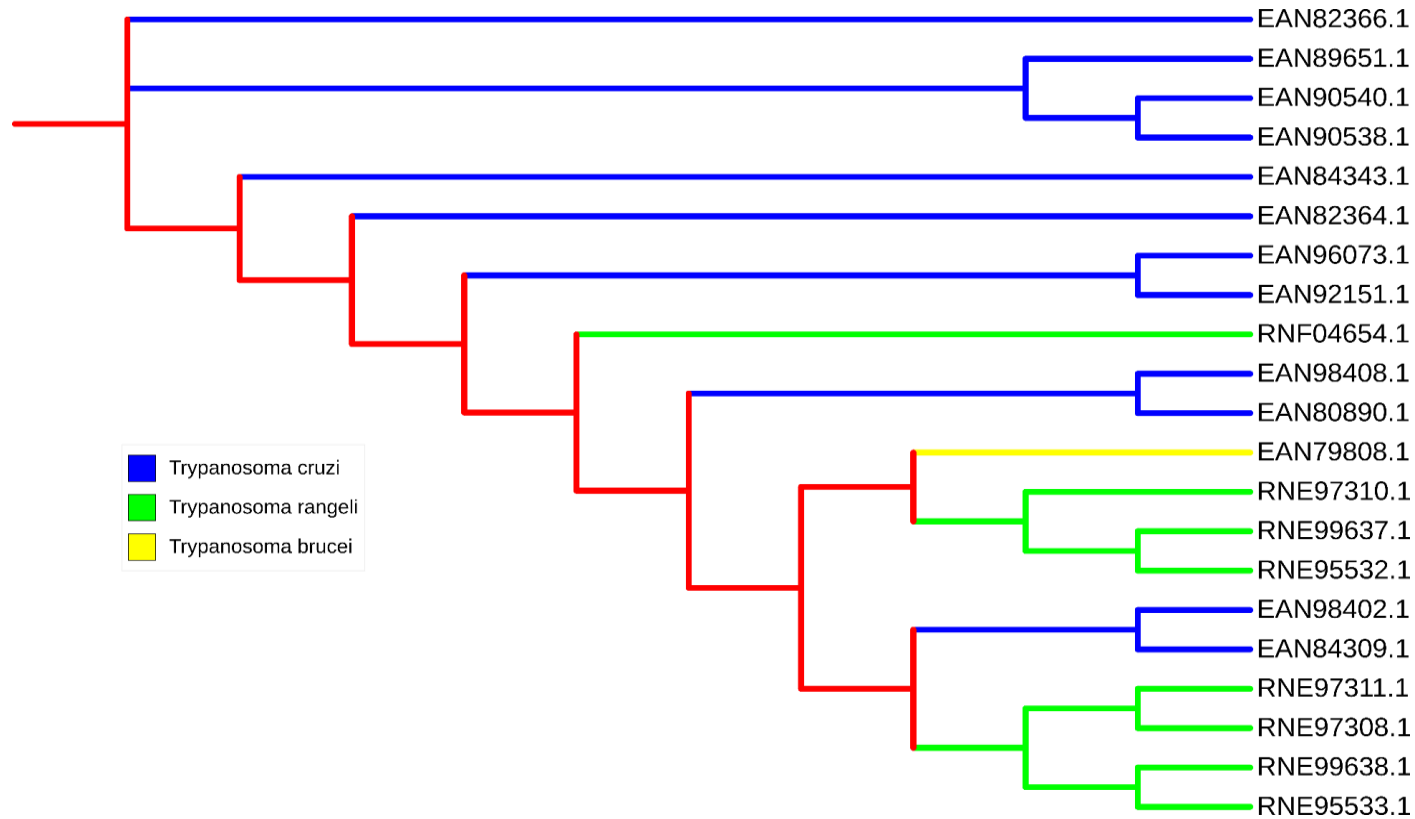


Figura 7. Filogenia de la familia AMASTIN. Árbol de máxima verosimilitud inferido con IQ-TREE2, basado en secuencias de *T. cruzi* (azul), *T. rangeli* (verde) y *T. brucei* (amarillo – outgroup) bajo el mejor modelo de sustitución (-m MFP), con soporte de 1,000 bootstrap ultrarrápido y 1,000 SH-aLRT. Las ramas se muestran coloreadas de manera diferencial para resaltar clados internos y facilitar la visualización; los colores no corresponden a especies ni categorías taxonómicas específicas.



### **3.7. Filogenia de la familia DGF-1 (Familia de genes dispersos-1)**

Inicialmente se construyeron árboles filogenéticos independientes para cada especie, *T. cruzi* (Figura 8) y *T. rangeli* (Figura 9). Estos análisis permitieron evaluar la organización interna de la familia DGF-1 sin la influencia de la comparación interespecífica. En ambos casos se evidenciaron expansiones génicas notables y la conformación de múltiples subclados, lo que demuestra que la diversificación de DGF-1 también ocurrió a nivel intraespecífico, reforzando su carácter de familia multigénica altamente dinámica.

En el caso del árbol combinado, el análisis filogenético de la familia génica DGF-1 realizado con IQ-TREE2 reveló un patrón atípico de segregación entre las especies de tripanosomátidos analizadas:

- En *Trypanosoma cruzi* se identificaron múltiples secuencias de la familia DGF-1, en total 429 secuencias, las cuales no conformaron un clado monofilético exclusivo. En cambio, se distribuyeron en diferentes ramas del árbol junto con secuencias de *T. rangeli*.
- En *Trypanosoma rangeli* también se detectaron 18 secuencias de DGF-1, igualmente intercaladas con las de *T. cruzi*, lo que refleja un alto grado de similitud entre genes de ambas especies.
- De *Trypanosoma brucei* no se incluyeron secuencias homólogas de DGF-1, debido a que en su genoma de referencia no se encuentra ninguna anotación correspondiente a esta familia génica. En consecuencia, esta especie no actuó como grupo externo en el análisis.

### **3.8. Relaciones Filogenéticas y Evolución de la Familia DGF-1**

El árbol filogenético, reconstruido bajo el mejor modelo evolutivo seleccionado por ModelFinder Plus y con soporte estadístico robusto (-bb 1000, -alrt 1000), mostró una topología compleja y no congruente con la especiación (Figura 10).

**Clados Mixtos e Intercalación de Especies:** Las secuencias de *T. cruzi* y *T. rangeli* aparecieron intercaladas en múltiples ramas, sin formar agrupamientos monofiléticos por especie. Este patrón puede explicarse por dos escenarios no excluyentes (i) procesos de evolución concertada como por ejemplo, recombinación o conversión génica, que habrían homogenizado parcialmente las copias dentro y entre linajes o (ii) duplicaciones génicas



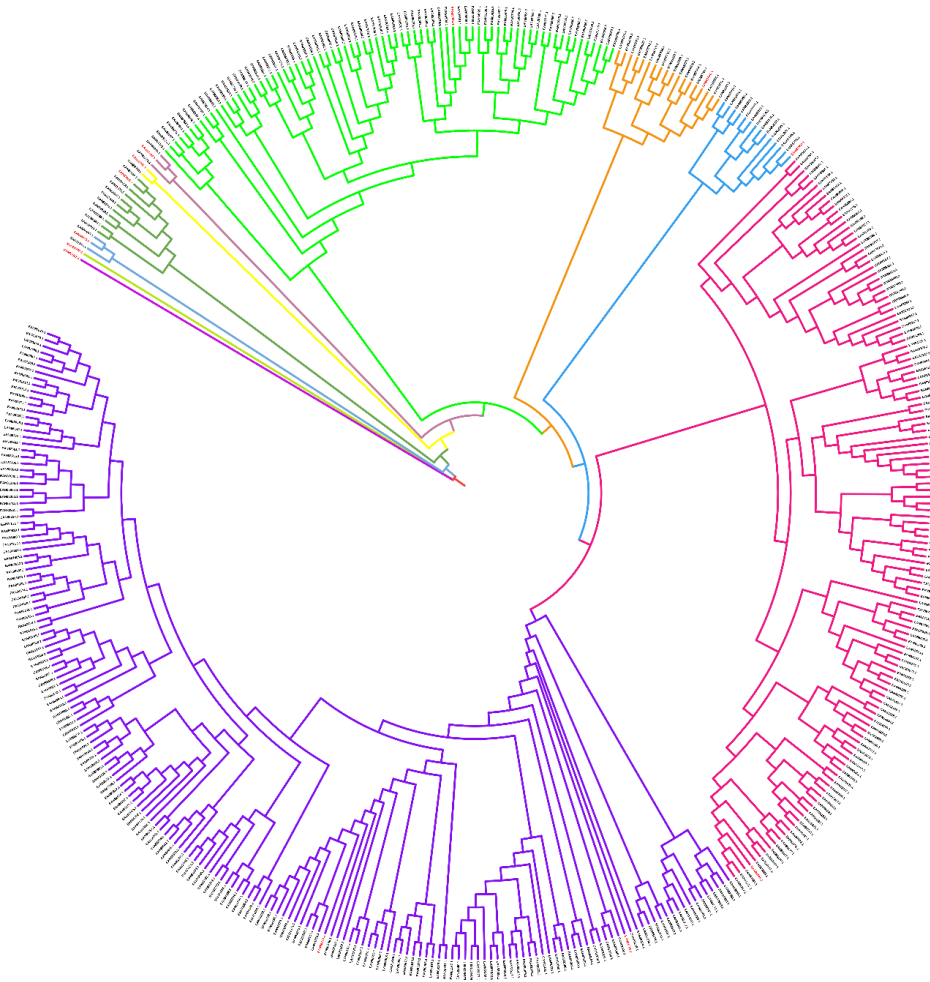
que ocurrieron antes de la especiación, de modo que ambos linajes heredaron conjuntos de copias estrechamente relacionadas. La topología observada no permite, por sí sola, distinguir claramente entre estas alternativas.

**Implicaciones Evolutivas:** El análisis mostró que las secuencias de *T. cruzi* y *T. rangeli* no se agruparon en clados monofiléticos por especie, sino que aparecieron intercaladas en distintas ramas del árbol. Este patrón indica que la similitud entre ciertas copias de ambos tripanosomas es mayor que la observada dentro de cada especie, lo que sugiere que parte de la diversificación de la familia DGF-1 ocurrió antes de la separación de los linajes y que posteriormente se mantuvo una alta conservación de secuencias en ambas especies.

**Limitaciones Filogenéticas:** La ausencia de *T. brucei* como grupo externo impidió enraizar el árbol con certeza absoluta, pero no afectó la interpretación principal: la familia DGF-1 conserva un patrón de homogeneización que enmascara la divergencia evolutiva entre especies.



### DGF-1 *Trypanosoma cruzi*



*Figura 8. Árbol filogenético de la familia DGF-1 de Trypanosoma cruzi, obtenido con IQ-TREE v2 bajo el mejor modelo de sustitución (-m MFP), con soporte de 1,000 bootstrap ultrarrápido y 1,000 SH-aLRT. Las ramas se muestran coloreadas de manera diferencial para resaltar clados internos y facilitar la visualización; los colores no corresponden a especies ni categorías taxonómicas específicas.*



### DGF-1 *Trypanosoma rangeli*

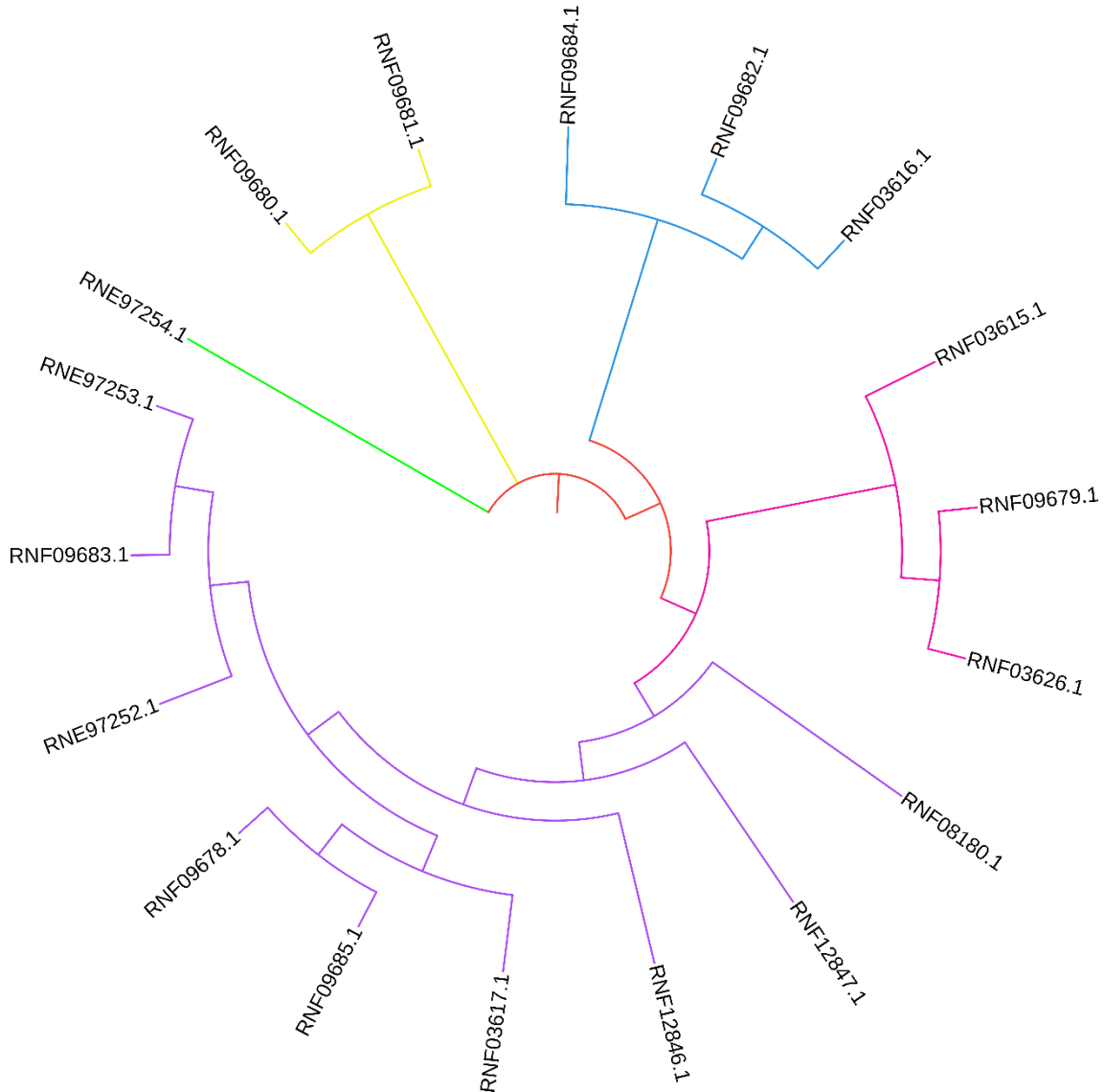


Figura 9. Árbol filogenético de la familia DGF-1 de *Trypanosoma rangeli*, obtenido con IQ-TREE v2 bajo el mejor modelo de sustitución (-m MFP), con soporte de 1,000 bootstrap ultrarrápido y 1,000 SH-aLRT. Las ramas se muestran coloreadas de manera diferencial para resaltar clados internos y facilitar la visualización; los colores no corresponden a especies ni categorías taxonómicas específicas.



ARBOL FILOGENÉTICO DE LA FAMILIA DGF-1 EN *Trypanosoma cruzi* y  
*Trypanosoma rangeli*

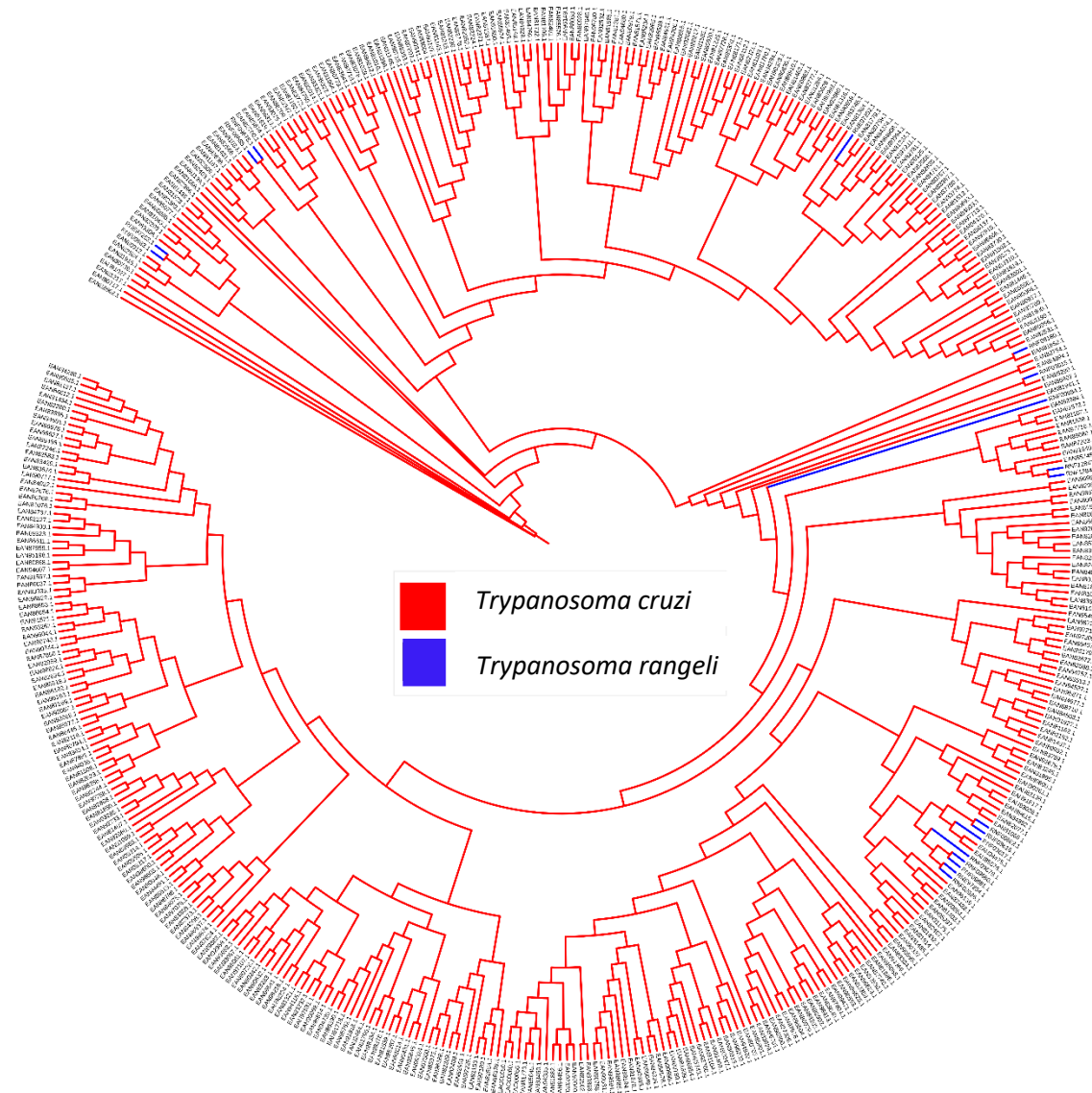


Figura 10. Filogenia de la familia DGF-1. Árbol de máxima verosimilitud inferido con IQ-TREE2, basado en secuencias de *T. cruzi* (rojo) y *T. rangeli* (azul) bajo el mejor modelo de sustitución (-m MFP), con soporte de 1,000 bootstrap ultrarrápido y 1,000 SH-aLRT. Las ramas se muestran coloreadas de manera diferencial para resaltar clados internos y facilitar la visualización; los colores no corresponden a especies ni categorías taxonómicas específicas.



### **3.9. Identificación de Dominios Conservados de AMASTIN en *Trypanosoma cruzi* y *Trypanosoma rangeli*.**

Como parte del análisis bioinformático para caracterizar las familias de proteínas de interés, se identificaron los dominios conservados en las secuencias de AMASTIN de *Trypanosoma cruzi* y *Trypanosoma rangeli* mediante el uso de la herramienta hmmscan de HMMER v3.4 contra la base de datos Pfam. El criterio de significancia estadística se estableció en un valor E-value  $\leq 0.01$ .

Las coordenadas precisas de los dominios AMASTIN identificadas en este análisis para cada proteína individual fueron integradas en sus respectivos árboles filogenéticos. Esta superposición de información permite correlacionar directamente la relación evolutiva (filogenia) con la variación estructural (posición del dominio conservado), facilitando la identificación de clados que comparten organizaciones de dominios específicas y permitiendo inferencias más robustas sobre la relación estructura-función-evolución en esta familia de proteínas.

### **3.10. Análisis en *Trypanosoma cruzi***

El análisis de dominios conservados en las secuencias de AMASTIN de *T. cruzi* revela un alto grado de conservación estructural entre los diferentes isotipos proteicos. Se identificó de manera consistente el dominio característico AMASTIN en todas las secuencias analizadas.

**Localización y Conservación:** Las coordenadas del dominio AMASTIN son notablemente uniformes en la mayoría de las secuencias, agrupándose en dos intervalos principales:

- Grupo 1: La mayoría de las proteínas presentan el dominio conservado entre las posiciones de aminoácido 14 y 171 de la secuencia primaria.
- Grupo 2: Un conjunto menor de secuencias muestra el dominio en posiciones ligeramente variables (20-182 y 20-96, respectivamente), lo que podría indicar la presencia de isoformas con extensiones N- o C-terminales diferenciales.

Este patrón de conservación sugiere una fuerte presión evolutiva para mantener la integridad estructural y funcional del dominio AMASTIN en *T. cruzi*, lo que es



consistente con su papel crucial en la supervivencia intracelular del parásito durante su etapa amastigote (Figura 11).

### Identificación y localización del dominio AMASTIN en secuencias de *Trypanosoma cruzi*.

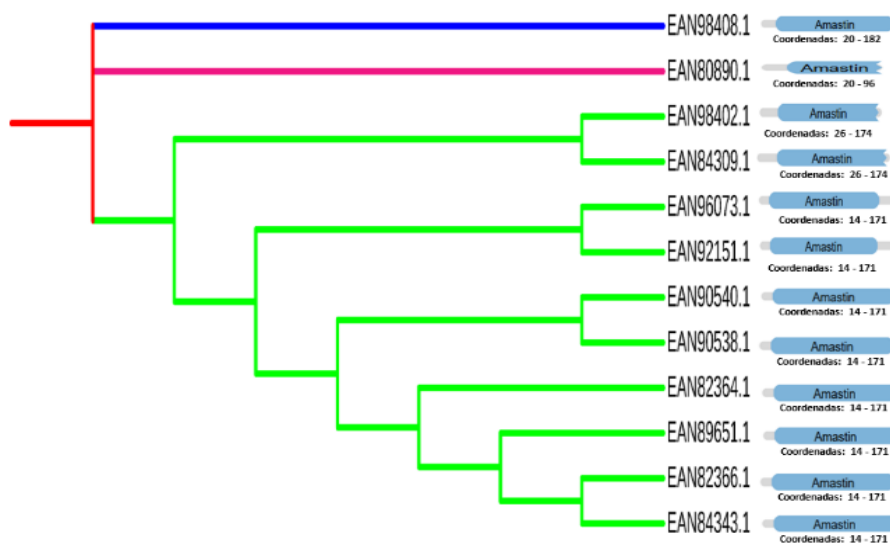


Figura 11. Dominio AMASTIN en *Trypanosoma cruzi* identificado con hmmscan (HMMER v3.4, Pfam, E-value  $\leq 0.01$ ). Las ramas se muestran coloreadas de manera diferencial para resaltar clados internos y facilitar la visualización; los colores no corresponden a especies ni categorías taxonómicas específicas.

### 3.11. Análisis en *Trypanosoma rangeli*

El perfil de dominios conservados de AMASTIN en *T. rangeli* presenta una mayor heterogeneidad en comparación con *T. cruzi*, tanto en la longitud como en la posición del dominio dentro de la secuencia proteica (Figura 12).

**Variabilidad Estructural:** Las coordenadas del dominio AMASTIN varían significativamente entre las diferentes proteínas:

Se observan dominios de longitud completa en posiciones como 167-311 (RNE97310.1), 280-425 (RNE95532.1) y 346-491 (RNE99637.1).

Otras entradas (ej. RNF04654.1 y RNE99638.1) están anotadas con el dominio en la región 3-120, lo que sugiere la posible identificación de dominios truncados o la presencia de genes que codifican para isoformas más cortas de la proteína.



La mayor diversidad en la ubicación del dominio AMASTIN en *T. rangeli* podría reflejar un proceso evolutivo divergente o una especialización funcional diferente a la de *T. cruzi*, posiblemente asociada a las diferencias en la patogenicidad y el ciclo de vida entre ambas especies.

#### Identificación y localización del dominio Amastin en secuencias de *Trypanosoma rangeli*.

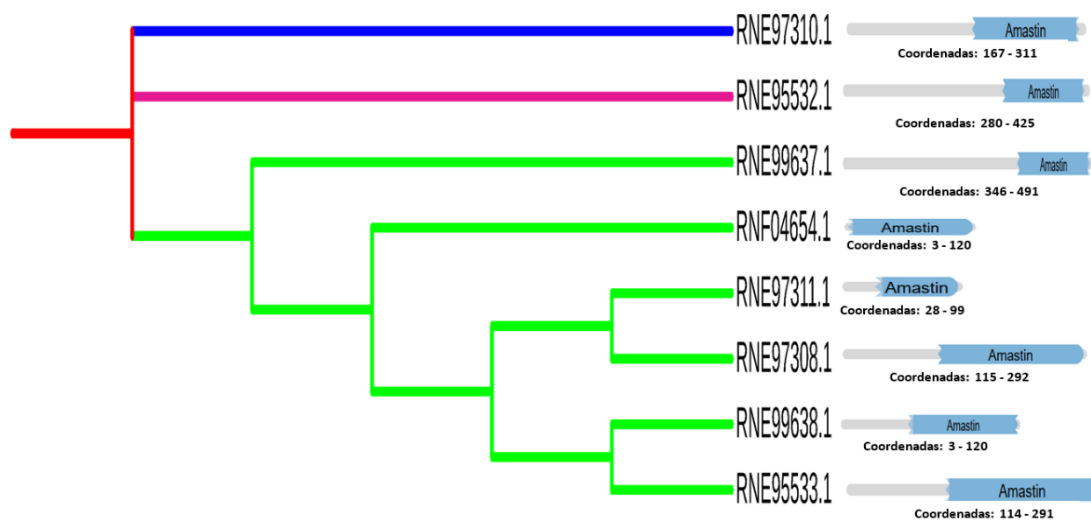


Figura 12. Dominio AMASTIN en *Trypanosoma rangeli* identificado con hmmscan (HMMER v3.4, Pfam,  $E\text{-value} \leq 0.01$ ). Las ramas se muestran coloreadas de manera diferencial para resaltar clados internos y facilitar la visualización; los colores no corresponden a especies ni categorías taxonómicas específicas.

### 3.12. Identificación de Dominios Conservados de DGF- 1 en *Trypanosoma cruzi* y *Trypanosoma rangeli*.

El análisis comparativo de la arquitectura de dominios de la familia DGF-1 entre *Trypanosoma cruzi* y *Trypanosoma rangeli* revela patrones distintivos que reflejan posibles divergencias evolutivas y adaptativas.

#### 3.12.1. Análisis en *Trypanosoma cruzi*

El análisis mediante HMMER reveló que las proteínas de la familia DGF-1 en *Trypanosoma cruzi* están formadas por múltiples dominios repetidos en tándem, distribuidos a lo largo de la secuencia proteica. Esta arquitectura modular, típica de proteínas de superficie en tripanosomátidos, está relacionada con su capacidad de evasión inmune (Figura 13).



La comparación entre secuencias de diferentes clados mostró que, aunque ciertos dominios son altamente conservados en posiciones específicas (particularmente en los extremos N- y C-terminales), existe una marcada variabilidad en el número y ubicación de las repeticiones internas. Esta variabilidad refleja procesos evolutivos dinámicos, como duplicaciones y recombinaciones, que favorecen la diversificación antigénica del parásito.

Al integrar estos datos con el árbol filogenético, se evidenció que los patrones de organización de dominios están asociados a linajes específicos, lo que sugiere que la evolución de la familia DGF-1 está impulsada por la expansión y contracción de sus repeticiones. Además, no se detectaron otros dominios conservados distintos a DGF-1, indicando que estas proteínas están conformadas casi exclusivamente por su dominio característico (Figura 14, 15, 16, 17, 18).

En conjunto, estos resultados confirman que la diversidad estructural de la familia DGF-1 refleja adaptaciones evolutivas que podrían tener un papel clave en la interacción parásito-hospedero.

### **3.12.2. Análisis en *Trypanosoma rangeli*.**

El análisis de las secuencias de la familia DGF-1 en *Trypanosoma rangeli* mostró que estas proteínas poseen una arquitectura modular basada en repeticiones de su dominio característico, aunque de forma menos extensa y compleja que en *T. cruzi*. Algunas secuencias presentan múltiples dominios, pero con longitudes totales menores y variabilidad notable en el número y la disposición de estas repeticiones.

Se identificaron dominios bien conservados en posiciones clave (inicio y final de las proteínas) pero también se observaron intervalos que podrían corresponder a variantes truncadas o posibles dominios adicionales, indicando mayor diversidad estructural y funcional. La integración de estas coordenadas con el árbol filogenético permitió detectar clados con patrones de dominios compartidos, sugiriendo que la evolución de DGF-1 en *T. rangeli* ha sido dinámica, probablemente mediada por duplicaciones y recombinaciones génicas (Figura 19).

En conclusión, la familia DGF-1 en *T. rangeli* presenta repetición modular de dominios con menor complejidad que *T. cruzi*, pero con considerable variabilidad filogenética y



estructural. Este enfoque combinando análisis de dominios y filogenia resultó eficaz para capturar la diversidad de la familia y explorar posibles correlaciones entre estructura y evolución funcional.

### Identificación y localización del dominio DGF-1 en secuencias de

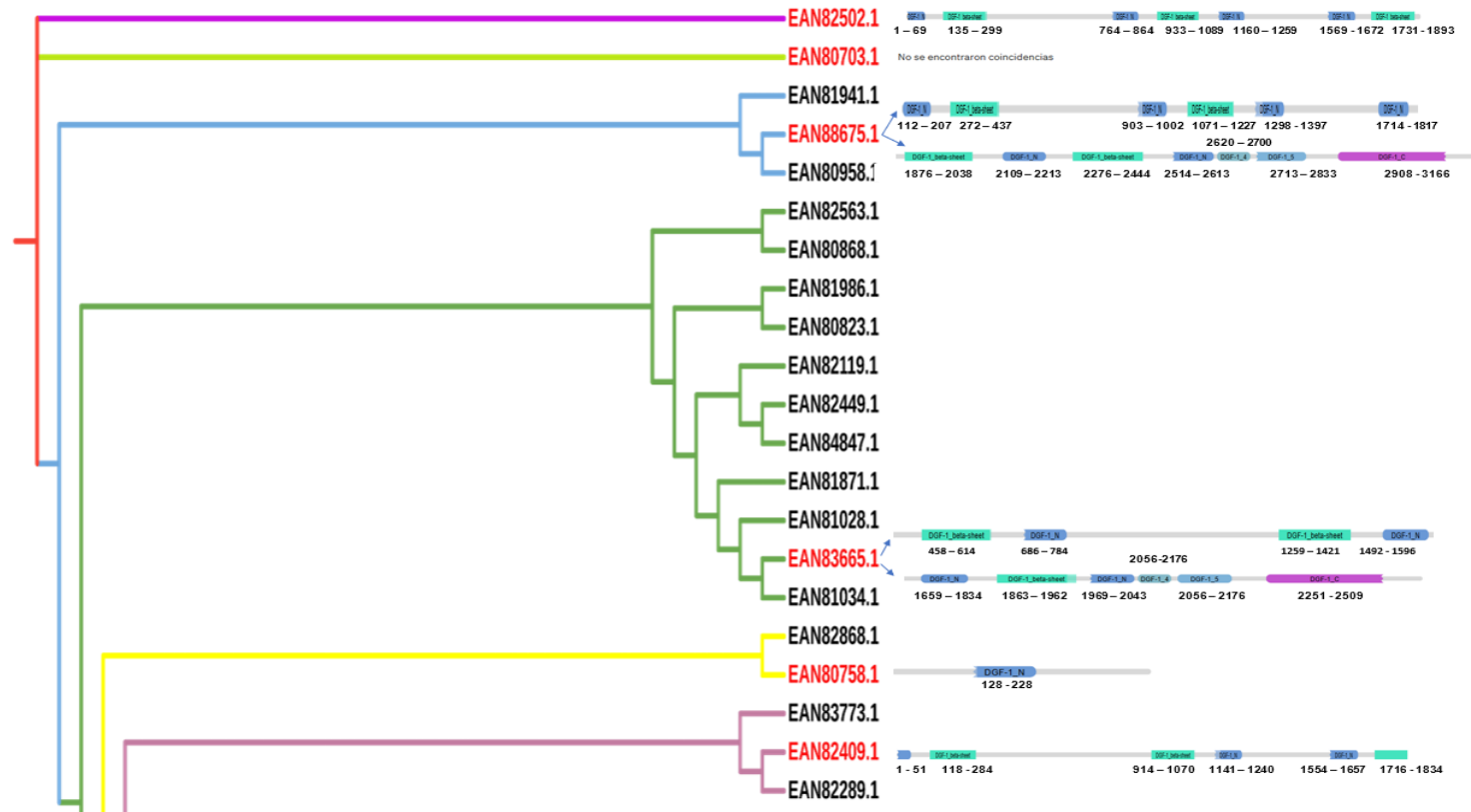


Figura 13. Dominio DGF-1 en *Trypanosoma cruzi* identificado con hmmscan (HMMER v3.4, Pfam, E-value  $\leq 0.01$ ). Las ramas se muestran coloreadas de manera diferencial para resaltar clados internos y facilitar la visualización; los colores no corresponden a especies ni categorías taxonómicas específicas.

### Identificación y localización del dominio DGF-1 en secuencias de

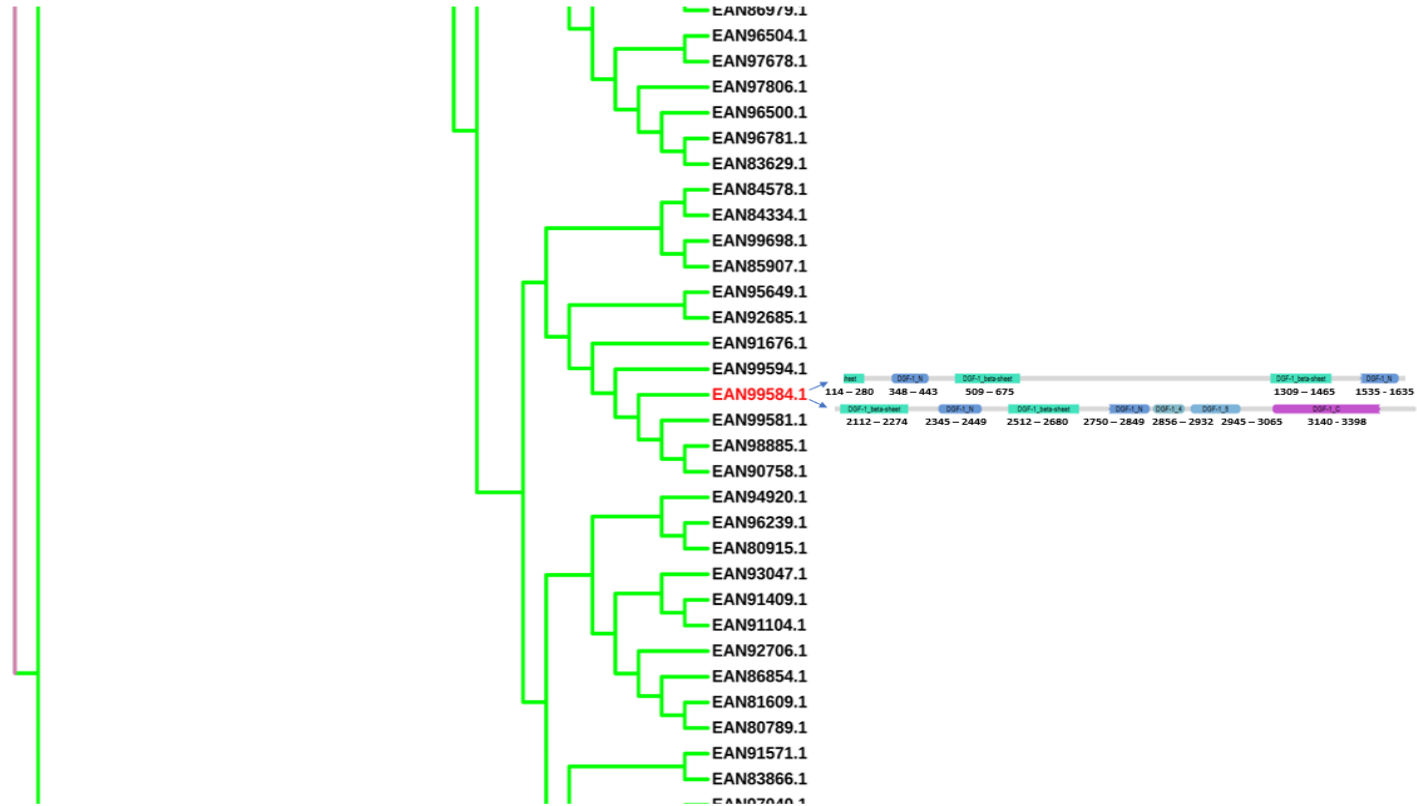


Figura 14. Dominio DGF-1 en *Trypanosoma cruzi* identificado con hmmscan (HMMER v3.4, Pfam,  $E\text{-value} \leq 0.01$ ). Las ramas se muestran coloreadas de manera diferencial para resaltar clados internos y facilitar la visualización; los colores no corresponden a especies ni categorías taxonómicas específicas.

### Identificación y localización del dominio DGF-1 en secuencias de *Trypanosoma*

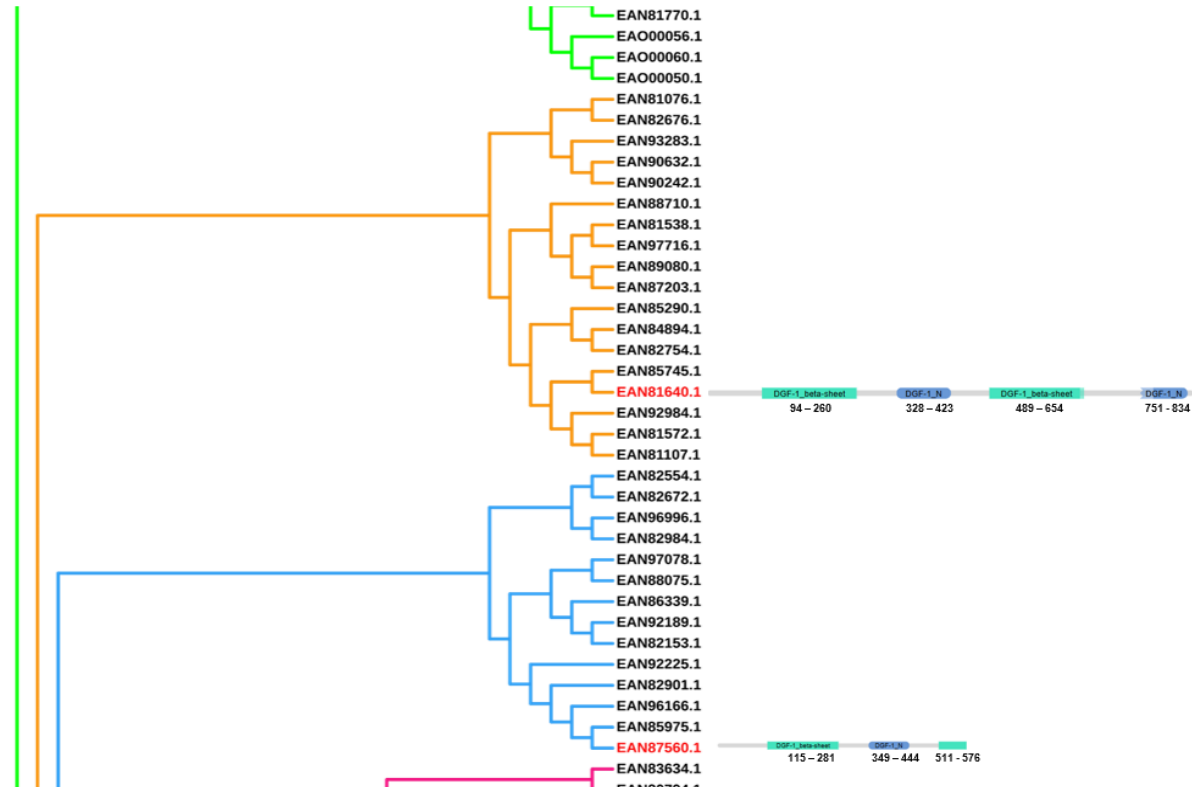


Figura 15. Dominio DGF-1 en *Trypanosoma cruzi* identificado con hmmscan (HMMER v3.4, Pfam,  $E\text{-value} \leq 0.01$ ). Las ramas se muestran coloreadas de manera diferencial para resaltar clados internos y facilitar la visualización; los colores no corresponden a especies ni categorías taxonómicas específicas.

### Identificación y localización del dominio DGF-1 en secuencias de *Trypanosoma*

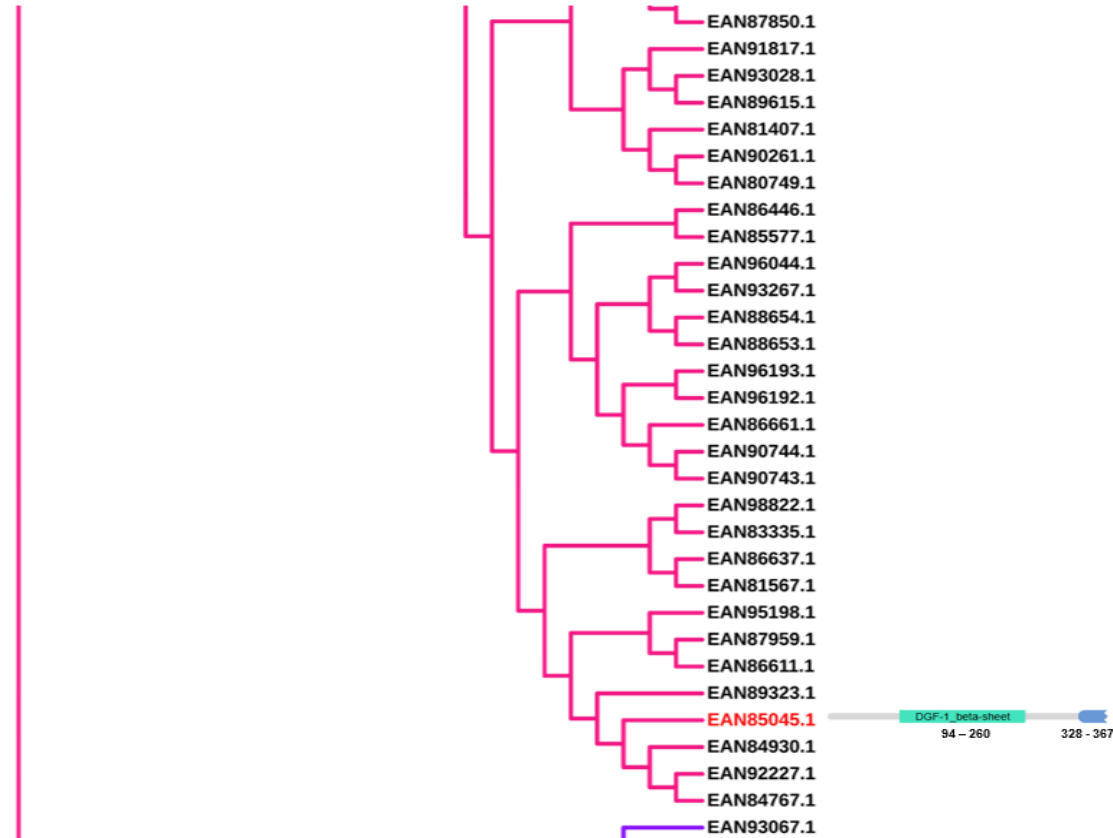


Figura 16. Dominio DGF-1 en *Trypanosoma cruzi* identificado con hmmscan (HMMER v3.4, Pfam,  $E\text{-value} \leq 0.01$ ). Las ramas se muestran coloreadas de manera diferencial para resaltar clados internos y facilitar la visualización; los colores no corresponden a especies ni categorías taxonómicas específicas.

### Identificación y localización del dominio DGF-1 en secuencias de *Trypanosoma*

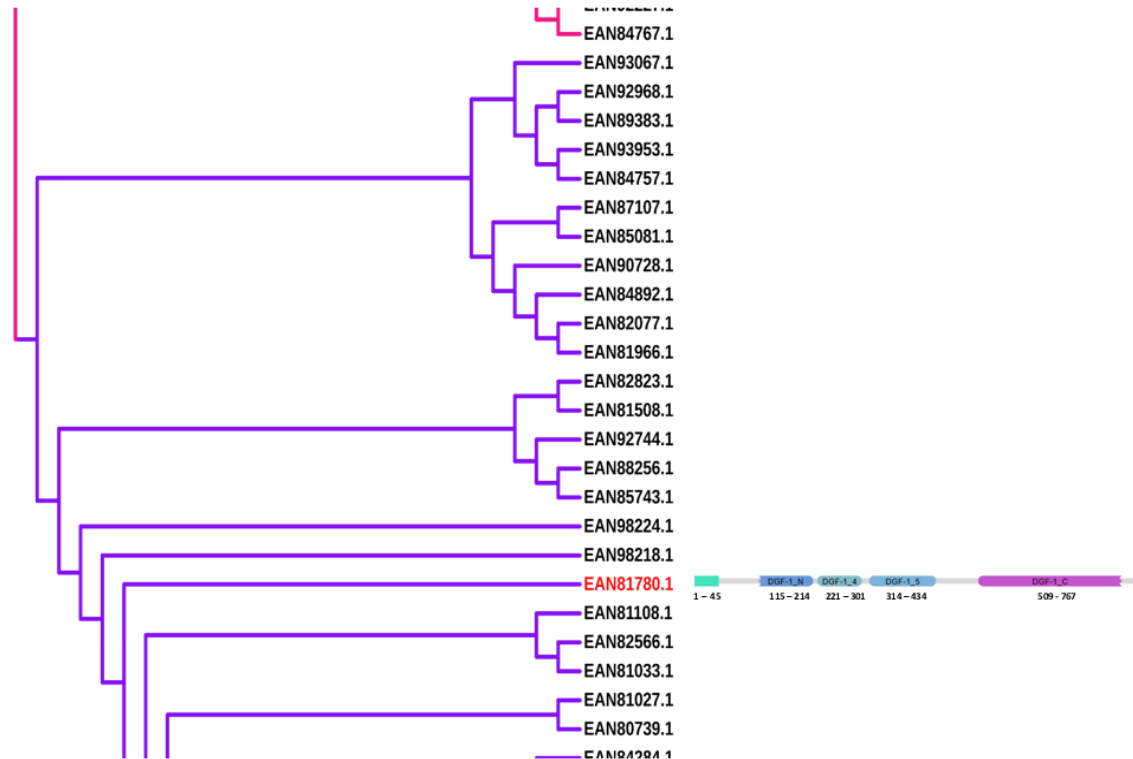


Figura 17. Dominio DGF-1 en *Trypanosoma cruzi* identificado con hmmscan (HMMER v3.4, Pfam,  $E\text{-value} \leq 0.01$ ). Las ramas se muestran coloreadas de manera diferencial para resaltar clados internos y facilitar la visualización; los colores no corresponden a especies ni categorías taxonómicas específicas.

### Identificación y localización del dominio DGF-1 en secuencias de

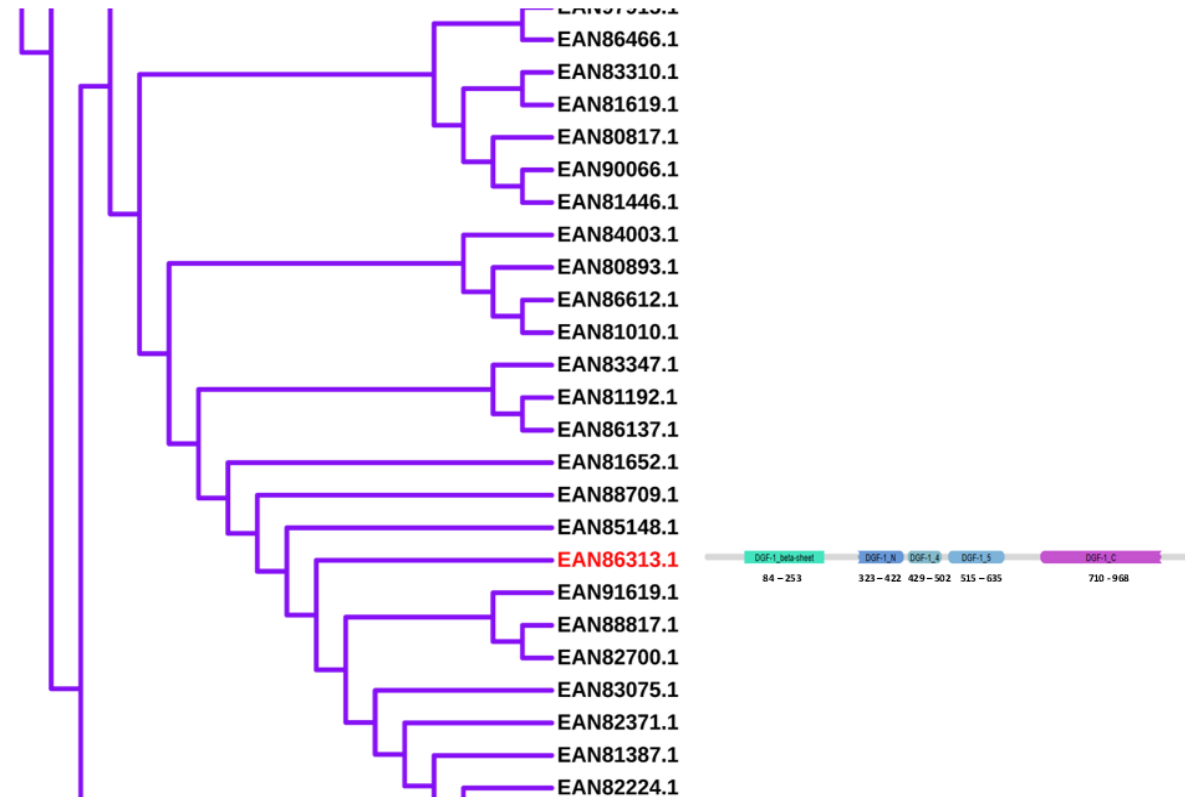


Figura 18. Dominio DGF-1 en *Trypanosoma cruzi* identificado con hmmscan (HMMER v3.4, Pfam,  $E\text{-value} \leq 0.01$ ). Las ramas se muestran coloreadas de manera diferencial para resaltar clados internos y facilitar la visualización; los colores no corresponden a especies ni categorías taxonómicas específicas.

### Identificación y localización del dominio DGF-1 en secuencias de *Trypanosoma*

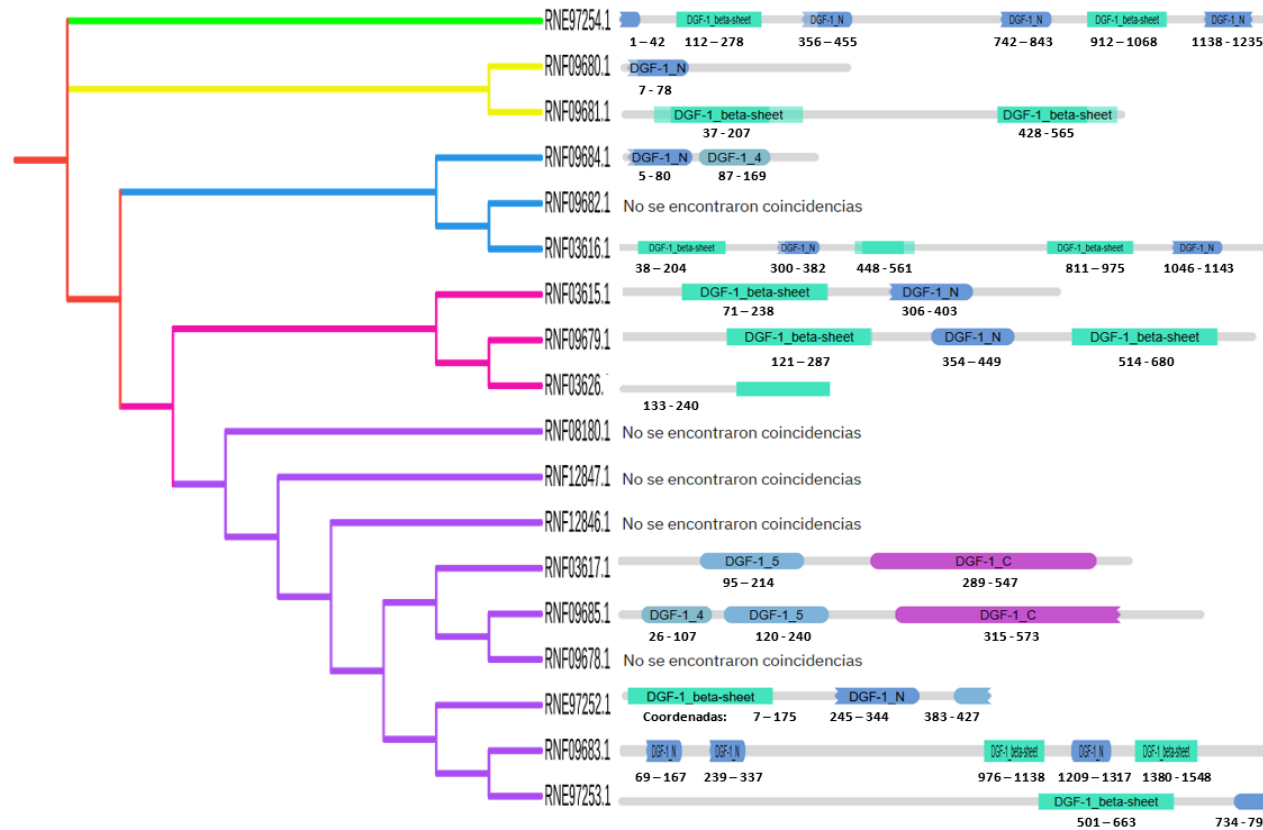


Figura 19. Dominio DGF-1 en *Trypanosoma rangeli* identificado con hmmscan (HMMER v3.4, Pfam,  $E$ -value  $\leq 0.01$ ). Las ramas se muestran coloreadas de manera diferencial para resaltar clados internos y facilitar la visualización; los colores no corresponden a especies ni categorías taxonómicas específicas.



## DISCUSIÓN

La selección rigurosa de genomas, como la descrita en este estudio, constituye un paso metodológico fundamental que impacta directamente la validez de los análisis genómicos comparativos y filogenéticos. El estricto criterio de inclusión, que priorizó la integridad de la anotación genómica y la correcta identificación de genes codificantes de proteínas, aunque resultó en la exclusión del 73% de las cepas iniciales de *Trypanosoma cruzi*, asegura la alta calidad y confiabilidad de los datos utilizados. No obstante, como se muestra en la Tabla 1, los genomas seleccionados presentan una amplia variación tanto en el tamaño total como en el número de genes anotados. Estas diferencias pueden reflejar, en parte, particularidades biológicas entre cepas, pero también son consecuencia de los métodos de secuenciación, ensamblaje y anotación empleados en cada proyecto genómico. Por ello, aunque los criterios de selección garantizan la confiabilidad de las secuencias analizadas, los resultados comparativos deben interpretarse considerando estas fuentes potenciales de variabilidad.

Este enfoque de calidad sobre cantidad es crucial para minimizar artefactos analíticos y conclusiones erróneas, ya que genomas mal anotados introducen un alto nivel de ruido e incertidumbre en estudios de genómica comparada (Prieto-Baños et al., 2025). Sin embargo, esta estricta selección tiene una limitación pues al excluir muchas cepas, podría no reflejar completamente la diversidad genética dentro de *T. cruzi*, lo que a su vez podría afectar la interpretación de las relaciones filogenéticas entre sus linajes (Omeranovic et al., 2025). Esta disparidad en la disponibilidad de recursos genómicos de calidad se hace aún más evidente al comparar con *Trypanosoma rangeli*, donde solo dos genomas cumplieron los requisitos, la cepa SC-58, un representante del linaje TrD aislado de roedores y nunca reportado en humanos (Machado et al., 2000) y la cepa AM80, proveniente de un aislamiento humano en la región amazónica (Caballero et al., 2015). El origen contrastante de estas cepas tiene implicaciones directas para la interpretación de los resultados, ya que mientras AM80 ofrece un marco de comparación más cercano con *T. cruzi*, SC-58 refleja adaptaciones propias de reservorios animales, lo que limita la generalización de los hallazgos a nivel poblacional. Finalmente, la elección del genoma de referencia de *Trypanosoma brucei* como grupo externo está bien fundamentada, ya que su posición filogenética claramente distante dentro del género permite enraizar de manera efectiva el árbol evolutivo y polarizar las características genéticas bajo estudio,



una práctica estándar y robusta en análisis filogenéticos (K. R. Bradwell et al., 2018). En conjunto, la estrategia de selección adoptada, a pesar de las limitaciones logísticas, establece una base sólida y reproducible para los análisis subsiguientes.

Respecto a la marcada variabilidad en el número de copias de las familias génicas AMASTIN y DGF-1 entre *Trypanosoma cruzi*, *T. rangeli* y *T. brucei* sugiere fuertemente que su expansión diferencial está intrínsecamente ligada a sus distintos modos de vida y estrategias de parasitismo. La expansión masiva de DGF-1, superando las 1000 copias en algunas cepas de *T. cruzi*, junto con la alta variabilidad en el número de AMASTINas (de 2 a 39 copias) refleja la acción de procesos evolutivos como la duplicación génica y la diversificación dirigida por presiones selectivas específicas (K. Bradwell, 2016; K. R. Bradwell et al., 2018; El-Sayed et al., 2005). Se plantea que la expansión de estas familias génicas en *T. cruzi* está vinculada a su capacidad de adaptación a nuevos hospederos y ambientes, en el marco de su complejo ciclo de vida intracelular en mamíferos, así como a la necesidad de evadir la respuesta inmune del hospedador. Estas funciones han sido propuestas previamente para las proteínas de superficie codificadas por las familias AMASTIN y DGF-1 (Cardoso et al., 2015). En contraste, la marcada reducción o ausencia de estas familias génicas en *T. rangeli* (no patógeno) y *T. brucei* (parásito extracelular) respalda esta hipótesis, sugiriendo que la presión selectiva para mantener un amplio repertorio de genes AMASTIN y DGF-1 es menor en parásitos que no invaden activamente células del hospedador o que ocupan nichos ecológicos distintos (K. R. Bradwell et al., 2018). La variabilidad intraespecífica observada entre diferentes cepas de *T. cruzi* (por ejemplo, Berenice vs. TCC) podría estar asociada con diferencias en su tropismo tisular o en su virulencia. Esto sugiere que el tamaño de las familias multigénicas AMASTIN y DGF-1 podría constituir un factor adaptativo clave, influyendo en la especificidad y el éxito de la infección (Silvestrini et al., 2024).

La marcada diferencia en el tamaño del genoma y la carga génica entre *Trypanosoma cruzi*, *T. rangeli* y *T. brucei* refleja trayectorias evolutivas divergentes, moldeadas por sus distintas estrategias de vida y presiones selectivas. En particular, la expansión genómica masiva y la alta variabilidad en el número de genes observadas en *T. cruzi* (30–55 Mb; 15,000 –30,000 genes) son coherentes con su estilo de vida intracelular obligado y su notable diversidad intraespecífica. Su historia evolutiva incluye duplicaciones génicas, expansión de familias multigénicas de superficie y acumulación de elementos repetitivos,



procesos que han contribuido a la plasticidad de su genoma y a su capacidad de adaptarse a una amplia variedad de hospederos intracelulares (K. R. Bradwell et al., 2018; Herreros-Cabello et al., 2025)

En contraste, el genoma de *Trypanosoma brucei* presenta una estructura más compacta (26 Mb; 9,000 genes), lo que sugiere una evolución orientada hacia una mayor eficiencia replicativa y una regulación génica más precisa. Esta organización genómica está estrechamente asociada a su ciclo de vida extracelular en mamíferos y a su sofisticado mecanismo de variación antigénica (Berriman et al., 2005). A diferencia de *T. cruzi*, que depende de la expansión de familias multigénicas para adaptarse a diversos nichos intracelulares, *T. brucei* ha optimizado su genoma para expresar una única glicoproteína de superficie variante (VSG) en cada célula, mediante un sistema de conmutación antigénica que involucra la recombinación homóloga de genes VSG silenciosos en sitios de expresión específicos. Este mecanismo permite al parásito evadir eficazmente la respuesta inmune del hospedador sin necesidad de mantener un repertorio masivo de genes de superficie, reflejando así una estrategia evolutiva de especialización y eficiencia (Staneva et al., 2022).

El tamaño genómico de *Trypanosoma rangeli* (21 Mb; 10,000–15,000 genes) se ubica en un rango intermedio respecto al de otros tripanosomátidos. Este tamaño genómico intermedio sugiere una presión selectiva menos intensa para la retención de genes redundantes o la expansión de familias génicas asociadas a la evasión inmune compleja, características presentes en especies como *T. cruzi* (Stoco et al., 2014). En contraste, *T. rangeli* ha conservado un repertorio más reducido de genes multigénicos, como las mucinas y las trans-sialidasas, lo que indica una estrategia evolutiva adaptada a su ciclo de vida extracelular y a su interacción con hospedadores menos complejos. Esta simplificación genética también sugiere un estilo de vida más simple, con menor gasto energético que podría facilitar una replicación más eficiente y una menor carga metabólica, permitiendo al parásito mantener su viabilidad en condiciones ecológicas específicas sin la necesidad de mecanismos complejos de evasión inmune (Watanabe Costa et al., 2020).

La notable conservación del contenido de GC (~50%) entre estas especies sugiere la existencia de restricciones evolutivas significativas en la composición nucleotídica basal,



posiblemente asociadas con la maquinaria transcripcional y replicativa común en los tripanosomátidos (El-Sayed et al., 2005).

El análisis de conservación de secuencias revela patrones evolutivos divergentes en las familias AMASTIN y DGF-1 entre *Trypanosoma cruzi* y *Trypanosoma rangeli*, que reflejan directamente sus diferentes estrategias de vida y relaciones con el hospedero. En *T. cruzi*, la conservación del dominio AMASTIN es notoriamente alta, con coordenadas muy similares en la mayoría de las copias (aproximadamente entre los aminoácidos 140–175), lo que indica una fuerte presión selectiva purificadora para mantener su integridad estructural y funcional. Este grado de conservación sugiere que la proteína cumple un rol esencial en la supervivencia intracelular del parásito en el hospedero mamífero, posiblemente vinculada a procesos de invasión celular y modulación del fagosoma (Cruz et al., 2012; Jackson, 2010). En contraste, *T. rangeli* presenta un patrón más variable, con un rango amplio en la localización del dominio (entre los aminoácidos 3–491) y variaciones significativas en su extensión. Esta heterogeneidad refleja una presión selectiva más relajada, congruente con la biología de *T. rangeli* como parásito no patógeno incapaz de establecer un ciclo intracelular en mamíferos, lo que sugiere que la función de AMASTIN en esta especie podría ser redundante o haber experimentado una divergencia funcional significativa (Stoco et al., 2014).

El análisis comparativo de la familia DGF-1 revela estrategias evolutivas contrastantes entre *T. cruzi* y *T. rangeli*. En *T. cruzi*, la arquitectura con dominios terminales conservados y una región central hipervariable constituye un rasgo característico de familias multigénicas implicadas en la evasión inmune, donde los dominios estables cumplen funciones estructurales o de anclaje esenciales, mientras que la variabilidad central facilita la diversificación antigénica y la adhesión a distintos ligandos del hospedero. En contraste, la organización más fragmentada y menos variable de DGF-1 en *T. rangeli* sugiere una presión selectiva reducida para mantener un repertorio antigénico amplio, coherente con su ciclo de vida no patógeno y su interacción menos estrecha con las defensas del hospedero vertebrado. En conjunto, estos patrones de conservación y variabilidad evidencian cómo las presiones selectivas impuestas por modos de parasitismo divergentes han moldeado la evolución funcional de la familia DGF-1 en ambos tripanosomátidos (Pablos & Osuna, 2012; Ramírez, 2023).



El análisis filogenético y estructural de la familia AMASTIN revela patrones evolutivos profundamente divergentes entre los tripanosomátidos estudiados, consistentes con sus distintas estrategias biológicas. La expansión significativa de AMASTINas en *T. cruzi* (12 secuencias) en comparación con *T. rangeli* (8) y especialmente con *T. brucei* (1), sugiere una fuerte presión selectiva para la diversificación de esta familia en linajes adaptados a la vida intracelular. La topología del árbol filogenético muestra que las secuencias de *T. cruzi* y *T. rangeli* tienden a formar agrupamientos diferenciados, aunque no completamente monofiléticos, lo que sugiere que parte de la diversificación de la familia AMASTIN ocurrió antes de la separación de estas especies. Este patrón es consistente con la hipótesis de un origen ancestral común de la familia en el subgénero *Schizotrypanum*, seguido de expansiones diferenciales en cada linaje (Hamilton et al., 2004). La notable expansión en *T. cruzi* probablemente refleja una adaptación a su nicho intracelular obligado en mamíferos, donde las AMASTIN, glicoproteínas de superficie, desempeñan roles cruciales en la invasión y supervivencia dentro de las células hospedadoras (Cruz et al., 2012; Jackson, 2010).

La conservación estructural del dominio AMASTIN en *T. cruzi*, con coordenadas notablemente uniformes (posiciones 14-171), indica una fuerte restricción evolutiva para mantener la integridad funcional de este dominio, probablemente esencial para su función en la infectividad (Avila et al., 2018). En *T. rangeli*, la heterogeneidad en la longitud y localización del dominio AMASTIN incluyendo isoformas truncadas (por ejemplo, 3-120) y variantes de longitud completa en posiciones variables (por ejemplo, 167-311) sugiere una relajación de las presiones selectivas o una diversificación funcional alternativa. Esta variabilidad estructural podría estar directamente vinculada con la incapacidad de este parásito para replicarse eficientemente en células mamíferas y con su reducida patogenicidad, en contraste con la fuerte restricción evolutiva observada en *T. cruzi* (Stoco et al., 2014).

La secuencia homóloga única de *T. brucei* se ubicó separada de los clados principales de *T. cruzi* y *T. rangeli*, lo que refleja la reducción extrema de esta familia en un parásito extracelular. Su posición confirma la utilidad de *T. brucei* como referencia comparativa externa y respalda la idea de que la presión selectiva para retener genes involucrados en la invasión intracelular es mínima en linajes con ciclo extracelular (Jackson, 2015). En conjunto, estos hallazgos subrayan cómo las presiones selectivas impuestas por los



distintos modos de parasitismo han tallado la evolución de la familia AMASTIN, con una conservación estructural que refleja restricción funcional en *T. cruzi* y una mayor variabilidad que sugiere una evolución más neutra en *T. rangeli*.

La arquitectura de dominios de la familia DGF-1, caracterizada por repeticiones en tándem de su dominio característico, revela una clara firma evolutiva de adaptación a la interacción con el hospedero en los tripanosomátidos. La expansión masiva y la elevada variabilidad en el número y disposición de estos dominios observada en *Trypanosoma cruzi* son consistentes con su papel como una familia multigénica central en la evasión de la respuesta inmune del hospedero mamífero. Esta estrategia de variación antigénica estructural, basada en la generación de un vasto repertorio de variantes de superficie mediante duplicaciones, recombinaciones y contracciones génicas, es un mecanismo bien documentado en *T. cruzi* para persistir en ambientes inmunológicamente hostiles (Jackson, 2015; Pablos & Osuna, 2012; Ramírez, 2023). La conservación de los dominios en los extremos N- y C-terminales sugiere que estas regiones cumplen una función estructural o de anclaje esencial que se mantiene bajo fuerte presión selectiva, mientras que la hipervariabilidad de las repeticiones internas proporciona la plasticidad necesaria para la diversificación antigénica (Herreros-Cabello et al., 2025; Kawashita et al., 2009).

En marcado contraste, la arquitectura menos extensa y compleja de DGF-1 en *Trypanosoma rangeli* refleja una presión selectiva radicalmente diferente. La notable reducción en el número de repeticiones y la presencia de posibles variantes truncadas son coherentes con el estilo de vida no patógeno de este parásito y su incapacidad para establecer infecciones intracelulares crónicas en mamíferos (K. R. Bradwell et al., 2018; Stoco et al., 2014). La relajación de la presión selectiva para mantener un amplio arsenal de genes de superficie permitiría una evolución más neutra o divergente de esta familia en *T. rangeli*, donde su función podría ser vestigial o haber sido reconvertida para un rol diferente en su ciclo de vida, que involucra principalmente la glándula salival de insectos vectores del género *Rhodnius* (Dos-Santos et al., 2012). La correlación observada entre los patrones de dominios y los clados filogenéticos en ambas especies demuestra que la evolución de las familias multigénicas no es aleatoria, sino que está guiada por procesos de duplicación y recombinación que son moldeados por las presiones adaptativas específicas de cada linaje (Jackson, 2015; Pablos & Osuna, 2012). En conjunto, la disparidad en la complejidad estructural de DGF-1 entre *T. cruzi* y *T. rangeli* sirve como



un ejemplo paradigmático de cómo los modos de parasitismo divergentes esculpen la arquitectura genómica y la diversificación de las familias de genes de superficie.

## CONCLUSIONES

Los resultados de este estudio sugieren y complementan estudios anteriores sobre la existencia de una expansión diferencial significativa de las familias génicas AMASTIN y DGF-1 entre *Trypanosoma cruzi* y *Trypanosoma rangeli*, la cual está directamente correlacionada con su distinta capacidad patogénica y su estrategia de vida. La expansión masiva y la alta conservación secuencial y estructural en *T. cruzi* reflejan una fuerte presión selectiva para mantener genes cruciales para la invasión y supervivencia intracelular en el hospedero mamífero. Por el contrario, la reducción, variabilidad y relajación evolutiva observadas en *T. rangeli* son consistentes con su estilo de vida no patógeno y extracelular. Así, la patogenicidad de *T. cruzi* está sustentada, al menos en parte, por la expansión y diversificación adaptativa de estas familias multigénicas de superficie.

El análisis filogenético confirmó la historia evolutiva divergente de ambas familias génicas. Para la familia AMASTIN, la topología del árbol mostró agrupamientos diferenciados de *T. cruzi* y *T. rangeli*, aunque no completamente monofiléticos, lo que sugiere que parte de la diversificación de esta familia ocurrió antes de la separación de ambas especies. La secuencia homóloga de *T. brucei* se ubicó apartada de los clados principales, consistente con la reducción extrema de esta familia en un parásito extracelular y con la hipótesis de que la radiación de AMASTIN tuvo lugar en el ancestro común del subgénero *Schizotrypanum*, tras su divergencia del linaje de *T. brucei*.

De manera crucial, el análisis de la familia DGF-1 reveló un patrón evolutivo relevante y claramente diferenciado respecto al observado en AMASTIN. La topología filogenética también mostró una agrupación de secuencias por especie, con *T. cruzi* y *T. rangeli* formando clados separados y bien soportados. Esto demuestra que a pesar de la expansión masiva e independiente de DGF-1 en *T. cruzi*, las secuencias de ambas especies comparten un ancestro común exclusivo, corroborando que la familia también estaba presente en el antepasado del subgénero *Schizotrypanum*. La evolución de DGF-1



ha estado marcada por una dinámica de expansión y contracción de repeticiones en tándem de su dominio característico, la cual ha sido radicalmente más intensa en el linaje de *T. cruzi* tras la especiación. Este patrón de evolución convergente en ambas familias expansión en *T. cruzi* y contracción/variabilidad en *T. rangeli* sustenta firmemente que su diversificación ha sido moldeada por las presiones selectivas asociadas al nicho intracelular.

Se logró identificar y cuantificar de manera robusta el número de copias de ambas familias. *T. cruzi* exhibió una expansión génica extrema, particularmente en DGF-1 (hasta >1000 copias), con una alta variabilidad intraespecífica. *T. rangeli* mostró un número considerablemente menor de copias para ambas familias (AMASTIN: max. 8; DGF-1: max. 18), mientras que *T. brucei* prácticamente carece de ellas (1 copia de AMASTIN, 0 de DGF-1). Esta cuantificación sustenta la base del argumento de la expansión diferencial.

Se pudo discutir no solo la presencia, sino también las notables diferencias cualitativas entre las especies. *T. rangeli* posee las familias génicas, confirmando su origen ancestral compartido con *T. cruzi*. Sin embargo, la discusión va más allá de la mera presencia pues la arquitectura de dominios menos compleja, la mayor variabilidad secuencial y la heterogeneidad en la posición de los dominios en *T. rangeli* sugieren una evolución neutra o una potencial neofuncionalización, coherente con su reducida presión adaptativa para mantener genes de virulencia intracelular.

Este estudio sentó las bases con un conjunto de datos de alta calidad, pero la exclusión de muchas cepas de *T. cruzi* por falta de anotaciones completas limita la captura de toda la diversidad intraespecífica.

## RECOMENDACIONES

Se recomienda impulsar proyectos de secuenciación genómica de *Trypanosoma cruzi* y *Trypanosoma rangeli* en Ecuador, con el fin de disponer de un repertorio representativo de cepas locales. Contar con genomas secuenciados a nivel nacional permitiría realizar análisis comparativos entre linajes endémicos y cepas de otros países, lo que aportaría información clave sobre la diversidad genética, la evolución de familias multigénicas y los factores asociados a la patogenicidad. Este enfoque fortalecería la investigación en



genómica comparativa y contribuiría a la vigilancia epidemiológica y al desarrollo de estrategias de control más específicas para la realidad epidemiológica del país.

Se recomienda desarrollar estudios comparativos entre *Trypanosoma cruzi* y *Trypanosoma rangeli* aislados tanto de vectores triatomínicos como de muestras humanas. Este enfoque permitiría identificar posibles diferencias genómicas y funcionales asociadas al ciclo de transmisión y a la interacción con distintos hospedadores, generando información clave para comprender los mecanismos de infectividad, la dinámica de coinfecciones y la relevancia epidemiológica de ambas especies en contextos endémicos.

Se recomienda que en futuras investigaciones se priorice la secuenciación y anotación de nuevos genomas, particularmente de *Trypanosoma rangeli*, con el fin de ampliar la base de datos comparativa disponible.

La expansión de las familias AMASTIN y DGF-1 sugiere un papel clave en la supervivencia intracelular y la evasión inmune, por lo que se recomienda complementar el análisis genómico con estudios de transcriptómica y proteómica.

## **BIBLIOGRAFIA**

- Abad- Franch, F., & Aguilar V, H. (2004). Control de la enfermedad de Chagas en el Ecuador. OPS – OMS – Ministerio de Salud Pública del Ecuador. En *Control de la enfermedad de Chagas en el Ecuador: Vol. Sección I, II, III y VI*.
- Avila, C. C., Mule, S. N., Rosa-Fernandes, L., Viner, R., Barisón, M. J., Costa-Martins, A. G., De Oliveira, G. S., Teixeira, M. M. G., Marinho, C. R. F., Silber, A. M., & Palmisano, G. (2018). Proteome-Wide Analysis of *Trypanosoma cruzi* Exponential and Stationary Growth Phases Reveals a Subcellular Compartment-Specific Regulation. *Genes*, 9(8), 413. <https://doi.org/10.3390/genes9080413>
- Baptista, R. P., Reis-Cunha, J. L., DeBarry, J. D., Chiari, E., Kissinger, J. C., Bartholomeu, D. C., & Macedo, A. M. (2018). Assembly of highly repetitive genomes using short reads: The genome of discrete typing unit III *Trypanosoma cruzi* strain 231. *Microbial Genomics*, 4(4), e000156. <https://doi.org/10.1099/mgen.0.000156>



- Bartholomeu, D. C., Teixeira, S. M. R., & Cruz, A. K. (2021). Genomics and functional genomics in *Leishmania* and *Trypanosoma cruzi*: Statuses, challenges and perspectives. *Memórias Do Instituto Oswaldo Cruz*, *116*, e200634. <https://doi.org/10.1590/0074-02760200634>
- Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., Lennard, N. J., Caler, E., Hamlin, N. E., Haas, B., Böhme, U., Hannick, L., Aslett, M. A., Shallom, J., Marcello, L., Hou, L., Wickstead, B., Alsmark, U. C. M., Arrowsmith, C., ... El-Sayed, N. M. (2005). The genome of the African trypanosome *Trypanosoma brucei*. *Science (New York, N.Y.)*, *309*(5733), 416–422. <https://doi.org/10.1126/science.1112642>
- Bradwell, K. (2016). Genomic comparisons and genome architecture of divergent *Trypanosoma* species. *Theses and Dissertations*. <https://doi.org/10.25772/MYWV-2K59>
- Bradwell, K. R., Koparde, V. N., Matveyev, A. V., Serrano, M. G., Alves, J. M. P., Parikh, H., Huang, B., Lee, V., Espinosa-Alvarez, O., Ortiz, P. A., Costa-Martins, A. G., Teixeira, M. M. G., & Buck, G. A. (2018). Genomic comparison of *Trypanosoma conorhini* and *Trypanosoma rangeli* to *Trypanosoma cruzi* strains of high and low virulence. *BMC Genomics*, *19*(1), 770. <https://doi.org/10.1186/s12864-018-5112-0>
- Caballero, Z. C., Costa-Martins, A. G., Ferreira, R. C., Alves, J. M. P., Serrano, M. G., Camargo, E. P., Buck, G. A., Minoprio, P., & Teixeira, M. M. G. (2015). Phylogenetic and syntenic data support a single horizontal transference to a *Trypanosoma* ancestor of a prokaryotic proline racemase implicated in parasite evasion from host defences. *Parasites & Vectors*, *8*(1). <https://doi.org/10.1186/s13071-015-0829-y>
- Callejas-Hernández, F., Rastrojo, A., Poveda, C., Gironès, N., & Fresno, M. (2018). Genomic assemblies of newly sequenced *Trypanosoma cruzi* strains reveal new genomic expansion and greater complexity. *Scientific Reports*, *8*(1), 14631. <https://doi.org/10.1038/s41598-018-32877-2>
- Cardoso, M. S., Reis-Cunha, J. L., & Bartholomeu, D. C. (2015). Evasion of the Immune Response by *Trypanosoma cruzi* during Acute Infection. *Frontiers in Immunology*, *6*, 659. <https://doi.org/10.3389/fimmu.2015.00659>



- Cruz, M. C., Souza-Melo, N., da Silva, C. V., DaRocha, W. D., Bahia, D., Araújo, P. R., Teixeira, S. R., & Mortara, R. A. (2012). Trypanosoma cruzi: Role of  $\delta$ -Amastin on Extracellular Amastigote Cell Invasion and Differentiation. *PLoS ONE*, 7(12), e51804. <https://doi.org/10.1371/journal.pone.0051804>
- Cuba, C. A. (1998). Revisión de los aspectos biológicos y diagnósticos del Trypanosoma (Herpetosoma) rangeli. *Revista da Sociedade Brasileira de Medicina Tropical*, 31, 207–220. <https://doi.org/10.1590/S0037-86821998000200007>
- Dos-Santos, A. L. A., Dick, C. F., Alves-Bezerra, M., Silveira, T. S., Paes, L. S., Gondim, K. C., & Meyer-Fernandes, J. R. (2012). Interaction between Trypanosoma rangeli and the Rhodnius prolixus salivary gland depends on the phosphotyrosine ecto-phosphatase activity of the parasite. *International Journal for Parasitology*, 42(9), 819–827. <https://doi.org/10.1016/j.ijpara.2012.05.011>
- El-Sayed, N. M., Myler, P. J., Blandin, G., Berriman, M., Crabtree, J., Aggarwal, G., Caler, E., Renauld, 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., ... Hall, N. (2005). Comparative genomics of trypanosomatid parasitic protozoa. *Science (New York, N.Y.)*, 309(5733), 404–409. <https://doi.org/10.1126/science.1112181>
- Garcia, E. S., Castro, D. P., Figueiredo, M. B., & Azambuja, P. (2012). Parasite-mediated interactions within the insect vector: Trypanosoma rangeli strategies. *Parasites & Vectors*, 5(1), 105. <https://doi.org/10.1186/1756-3305-5-105>
- Guhl, F., & Vallejo, G. A. (2003). Trypanosoma (Herpetosoma) rangeli Tejera, 1920: An updated review. *Memorias Do Instituto Oswaldo Cruz*, 98(4), 435–442. <https://doi.org/10.1590/s0074-02762003000400001>
- Hamilton, P. B., Stevens, J. R., Gaunt, M. W., Gidley, J., & Gibson, W. C. (2004). Trypanosomes are monophyletic: Evidence from genes for glyceraldehyde phosphate dehydrogenase and small subunit ribosomal RNA. *International Journal for Parasitology*, 34(12), 1393–1404. <https://doi.org/10.1016/j.ijpara.2004.08.011>
- Herreros-Cabello, A., Callejas-Hernández, F., Gironès, N., & Fresno, M. (2020). Trypanosoma Cruzi Genome: Organization, Multi-Gene Families, Transcription, and Biological Implications. *Genes*, 11(10), 1196. <https://doi.org/10.3390/genes11101196>



- Herreros-Cabello, A., Callejas-Hernández, F., Gironès, N., & Fresno, M. (2025). Trypanosoma cruzi: Genomic Diversity and Structure. *Pathogens*, *14*(1), Article 1. <https://doi.org/10.3390/pathogens14010061>
- Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, *35*(2), 518–522. <https://doi.org/10.1093/molbev/msx281>
- Hoyos Sanchez, M. C., Ospina Zapata, H. S., Suarez, B. D., Ospina, C., Barbosa, H. J., Carranza Martinez, J. C., Vallejo, G. A., Urrea Montes, D., & Duitama, J. (2024). A phased genome assembly of a Colombian Trypanosoma cruzi TcI strain and the evolution of gene families. *Scientific Reports*, *14*, 2054. <https://doi.org/10.1038/s41598-024-52449-x>
- Jackson, A. P. (2007). Tandem gene arrays in Trypanosoma brucei: Comparative phylogenomic analysis of duplicate sequence variation. *BMC Evolutionary Biology*, *7*(1), 54. <https://doi.org/10.1186/1471-2148-7-54>
- Jackson, A. P. (2010). The Evolution of Amastin Surface Glycoproteins in Trypanosomatid Parasites. *Molecular Biology and Evolution*, *27*(1), 33–45. <https://doi.org/10.1093/molbev/msp214>
- Jackson, A. P. (2015). Genome evolution in trypanosomatid parasites. *Parasitology*, *142 Suppl 1*(Suppl 1), S40-56. <https://doi.org/10.1017/S0031182014000894>
- Kangussu-Marcolino, M. M., de Paiva, R. M. C., Araújo, P. R., de Mendonça-Neto, R. P., Lemos, L., Bartholomeu, D. C., Mortara, R. A., daRocha, W. D., & Teixeira, S. M. R. (2013). Distinct genomic organization, mRNA expression and cellular localization of members of two amastin sub-families present in Trypanosoma cruzi. *BMC Microbiology*, *13*, 10. <https://doi.org/10.1186/1471-2180-13-10>
- Kawashita, S. Y., da Silva, C. V., Mortara, R. A., Burleigh, B. A., & Briones, M. R. S. (2009). Homology, paralogy and function of DGF-1, a highly dispersed Trypanosoma cruzi specific gene family and its implications for information entropy of its encoded proteins. *Molecular and Biochemical Parasitology*, *165*(1), 19–31. <https://doi.org/10.1016/j.molbiopara.2008.12.010>
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0.



- Bioinformatics*, 23(21), 2947–2948.  
<https://doi.org/10.1093/bioinformatics/btm404>
- Letunic, I., & Bork, P. (2024). Interactive Tree of Life (iTOL) v6: Recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research*, 52(W1), W78–W82. <https://doi.org/10.1093/nar/gkae268>
- Machado, E. M., Alvarenga, N. J., Romanha, A. J., & Grisard, E. C. (2000). A simplified method for sample collection and DNA isolation for polymerase chain reaction detection of *Trypanosoma rangeli* and *Trypanosoma cruzi* in triatomine vectors. *Memórias Do Instituto Oswaldo Cruz*, 95, 863–866. <https://doi.org/10.1590/S0074-02762000000600021>
- Mazza, S. (2009). La enfermedad de Chagas en la República Argentina. En J. d. R. Carneiro, N. Azevedo, J. Araújo, J. Lannes-Vieira, M. d. N. C. Soeiro, & L. Klein, *Clássicos em Doença de Chagas: Histórias e perspectivas no centenário da descoberta* (pp. 361–384). FIOCRUZ.
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, 37(5), 1530–1534. <https://doi.org/10.1093/molbev/msaa015>
- Ministerio de Salud Pública del Ecuador. (2025). *Boletín epidemiológico: Enfermedad de Chagas, SE 17/2025*. Dirección Nacional de Vigilancia Epidemiológica. MSP.
- Morales-Viteri, D., Quinatoa Tutillo, P. A., Sánchez Mackenzie, D. D., Cagua Ordoñez, J. C., & Veloz Pérez, H. R. (2022). Enfermedad de Chagas en el Ecuador: Una revisión sistemática de los aspectos epidemiológicos y entomológicos. *INSPILIP*. <https://doi.org/10.31790/inspilip.v5i1.2>
- Nevers, Y., Jones, T. E. M., Jyothi, D., Yates, B., Ferret, M., Portell-Silva, L., Codo, L., Cosentino, S., Marcet-Houben, M., Vlasova, A., Poidevin, L., Kress, A., Hickman, M., Persson, E., Piližota, I., Guijarro-Clarke, C., the OpenEBench team the Quest for Orthologs Consortium, Iwasaki, W., Lecompte, O., ... Altenhoff, A. (2022). The Quest for Orthologs orthology benchmark service in 2022. *Nucleic Acids Research*, 50(W1), W623–W632. <https://doi.org/10.1093/nar/gkac330>
- Omeranovic, A., Van Long, F. N., Boubaker, A., Turgeon, A., & Nabi, H. (2025). Underrepresented populations in genomic research: A qualitative study of



- researchers' perspectives. *BMC Medical Genomics*, 18(1), 72.  
<https://doi.org/10.1186/s12920-025-02140-5>
- Organización Panamericana de la Salud. (2023, abril 13). *Menos del 10% de las personas con Chagas han sido diagnosticadas—OPS/OMS | Organización Panamericana de la Salud*. <https://www.paho.org/es/noticias/13-4-2023-menos-10-personas-con-chagas-han-sido-diagnosticadas>
- Pablos, L. M. D., & Osuna, A. (2012). Multigene Families in *Trypanosoma cruzi* and Their Role in Infectivity. *Infection and Immunity*.  
<https://doi.org/10.1128/iai.06225-11>
- Peña-Callejas, G., González, J., Jiménez-Cortés, J. G., Fuentes-Vicente, J. A. de, Salazar-Schettino, P. M., Bucio-Torres, M. I., Cabrera-Bravo, M., Flores-Villegas, A. L., Peña-Callejas, G., González, J., Jiménez-Cortés, J. G., Fuentes-Vicente, J. A. de, Salazar-Schettino, P. M., Bucio-Torres, M. I., Cabrera-Bravo, M., & Flores-Villegas, A. L. (2022). Enfermedad de Chagas: Biología y transmisión de *Trypanosoma cruzi*. *TIP. Revista especializada en ciencias químico-biológicas*, 25. <https://doi.org/10.22201/fesz.23958723e.2022.449>
- Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., & Finn, R. D. (2018). HMMER web server: 2018 update. *Nucleic Acids Research*, 46(W1), W200–W204.  
<https://doi.org/10.1093/nar/gky448>
- Prieto-Baños, S., Nevers, Y., Altenhoff, A., Warwick Vesztröcy, A., Dessimoz, C., & Glover, N. M. (2025). Annotation matters: The effect of structural gene annotation on orthology inference. *Bioinformatics*, 41(7), btaf365.  
<https://doi.org/10.1093/bioinformatics/btaf365>
- Python Software Foundation. (2024). Python.Org. <https://www.python.org/>
- Ramírez, J. L. (2023). The Elusive *Trypanosoma cruzi* Disperse Gene Protein Family (DGF-1). *Pathogens (Basel, Switzerland)*, 12(2), 292.  
<https://doi.org/10.3390/pathogens12020292>
- Reis-Cunha, J. L., Rodrigues-Luiz, G. F., Valdivia, H. O., Baptista, R. P., Mendes, T. A. O., de Moraes, G. L., Guedes, R., Macedo, A. M., Bern, C., Gilman, R. H., Lopez, C. T., Andersson, B., Vasconcelos, A. T., & Bartholomeu, D. C. (2015). Chromosomal copy number variation reveals differential levels of genomic plasticity in distinct *Trypanosoma cruzi* strains. *BMC Genomics*, 16(1), 499.  
<https://doi.org/10.1186/s12864-015-1680-4>



- Shen, W., Le, S., Li, Y., & Hu, F. (2016). SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. *PLOS ONE*, *11*(10), e0163962. <https://doi.org/10.1371/journal.pone.0163962>
- Silva Gonçalves, C. S., Ávila, A. R., de Souza, W., Motta, M. C. M., & Cavalcanti, D. P. (2018). Revisiting the *Trypanosoma cruzi* metacyclogenesis: Morphological and ultrastructural analyses during cell differentiation. *Parasites & Vectors*, *11*(1), 83. <https://doi.org/10.1186/s13071-018-2664-4>
- Silvestrini, M. M. A., Alessio, G. D., Frias, B. E. D., Sales Júnior, P. A., Araújo, M. S. S., Silvestrini, C. M. A., Brito Alvim de Melo, G. E., Martins-Filho, O. A., Teixeira-Carvalho, A., & Martins, H. R. (2024). New insights into *Trypanosoma cruzi* genetic diversity, and its influence on parasite biology and clinical outcomes. *Frontiers in Immunology*, *15*, 1342431. <https://doi.org/10.3389/fimmu.2024.1342431>
- Staneva, D. P., Bresson, S., Auchynnikava, T., Spanos, C., Rappsilber, J., Jeyaprakash, A. A., Tollervey, D., Matthews, K. R., & Allshire, R. C. (2022). The SPARC complex defines RNAPII promoters in *Trypanosoma brucei*. *eLife*, *11*, e83135. <https://doi.org/10.7554/eLife.83135>
- Stoco, P. H., Wagner, G., Talavera-Lopez, C., Gerber, A., Zaha, A., Thompson, C. E., Bartholomeu, D. C., Lückemeyer, D. D., Bahia, D., Loreto, E., Prestes, E. B., Lima, F. M., Rodrigues-Luiz, G., Vallejo, G. A., Filho, J. F. da S., Schenkman, S., Monteiro, K. M., Tyler, K. M., de Almeida, L. G. P., ... Grisard, E. C. (2014). Genome of the avirulent human-infective trypanosome—*Trypanosoma rangeli*. *PLoS Neglected Tropical Diseases*, *8*(9), e3176. <https://doi.org/10.1371/journal.pntd.0003176>
- Vallejo, G. A., Suárez, J., Olaya, J. L., Gutierrez, S. A., & Carranza, J. C. (2015). *Trypanosoma rangeli*: Un protozoo infectivo y no patógeno para el humano que contribuye al entendimiento de la transmisión vectorial y la infección por *Trypanosoma cruzi*, agente causal de la enfermedad de Chagas. *Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales*, 111–122. <https://doi.org/10.18257/raccefyn.143>
- Wang, W., Peng, D., Baptista, R. P., Li, Y., Kissinger, J. C., & Tarleton, R. L. (2021). Strain-specific genome evolution in *Trypanosoma cruzi*, the agent of Chagas



disease. *PLoS Pathogens*, 17(1), e1009254.  
<https://doi.org/10.1371/journal.ppat.1009254>

Watanabe Costa, R., Batista, M. F., Meneghelli, I., Vidal, R. O., Nájera, C. A., Mendes, A. C., Andrade-Lima, I. A., da Silveira, J. F., Lopes, L. R., Ferreira, L. R. P., Antoneli, F., & Bahia, D. (2020). Comparative Analysis of the Secretome and Interactome of *Trypanosoma cruzi* and *Trypanosoma rangeli* Reveals Species Specific Immune Response Modulating Proteins. *Frontiers in Immunology*, 11, 1774. <https://doi.org/10.3389/fimmu.2020.01774>

Weatherly, D. B., Boehlke, C., & Tarleton, R. L. (2009). Chromosome level assembly of the hybrid *Trypanosoma cruzi* genome. *BMC Genomics*, 10(1), 255. <https://doi.org/10.1186/1471-2164-10-255>

Westenberger, S. J., Cerqueira, G. C., El-Sayed, N. M., Zingales, B., Campbell, D. A., & Sturm, N. R. (2006). *Trypanosoma cruzi* mitochondrial maxicircles display species- and strain-specific variation and a conserved element in the non-coding region. *BMC Genomics*, 7(1), 60. <https://doi.org/10.1186/1471-2164-7-60>

World Health Organization. (2025). *Chagas disease (American trypanosomiasis)*. [https://www.who.int/es/news-room/fact-sheets/detail/chagas-disease-\(american-trypanosomiasis\)](https://www.who.int/es/news-room/fact-sheets/detail/chagas-disease-(american-trypanosomiasis))