

Identificación de las características genómicas del kDNA de *Trypanosoma cruzi* I (ADN mitocondrial) procedente de zonas endémicas de la enfermedad de Chagas en el Tolima y sus implicaciones en la biología de la tripanosomiasis americana.

Autores:

Hader Sebastián Ospina Zapata¹; *María Camila Hoyos Sánchez*²; *Brayhan Darío Suárez*¹; *Aura María Rodríguez*¹; *Valentina Herrera Sánchez*¹; *Carlos Mario Ospina*¹; *Hamilton Julián Barbosa*¹; *Julio César Carranza Martínez*¹; *Gustavo Adolfo Vallejo*¹; *María Clara Echeverry*³; *Jorge Duitama*²; *Daniel Alfonso Urrea*¹

Correo: hsospinaz@ut.edu.co

Afiliación:

¹ Laboratorio de Investigaciones en Parasitología Tropical (LIPT), Universidad del Tolima.

² TICSw: Tecnologías de Información y Construcción de Software, Universidad de los Andes.

³ Infecciones y Salud en el Trópico, Universidad Nacional de Colombia, Sede Bogotá.

Resumen:

La enfermedad de Chagas o tripanosomiasis americana es causada por el parásito protozoo *Trypanosoma cruzi*, clase Kinetoplastea, familia Trypanosomatidae. En Colombia se estima que existen entre 700.000 y 1.200.000 habitantes infectados con *T. cruzi* y alrededor de 8.000.000 de personas en riesgo de adquirir la infección. Los kinetoplastidos se caracterizan por presentar una única mitocondria ramificada con una intrincada organización de su propio ADN, molécula que puede representar hasta un 20-25% de la cantidad total de ADN por célula. Conocer las características genómicas del cinetoplasto es fundamental, pues este tiene un impacto significativo en la de transmisión, patogenicidad, desarrollo y evolución de los tripanosomátidos.

El objetivo general del presente trabajo fue identificar las características genómicas de los maxicírculos y minicírculos (kDNA) de *T. cruzi* que circula en Colombia, para lo cual se realizó el ensamblaje, anotación y comparación del maxicírculo de *T. cruzi* con otras moléculas reportadas anteriormente. Para ello, se utilizó una cepa procedente del municipio de Coyaima - Tolima, la cual se secuenció por tecnología PacBio Hifi, obteniendo un total de 4337 Mb aproximadamente, distribuidas en 206.530 reads y una media de calidad de Q26. El ensamblaje se realizó empleando NGSEP, Canu, Flye y HifiAsm con el fin de comparar los resultados. Para la búsqueda de los maxicírculos se utilizó la herramienta Blast, comparando los contigs ensamblados, con una base de datos local construida con 40 genomas de referencia de maxicírculos de tripanosomátidos (NCBI, TriTrypDB). Adicionalmente, con ayuda del paquete bioinformático EMBOSS, se filtraron los contigs con el menor porcentaje de Guanina/Citocina, con el fin de confirmar la presencia del maxicírculo. Se logró obtener una molécula circular con un tamaño de 47.166 pb, la cual se anotó manualmente obteniendo un total de 2 genes ribosomales y 18 genes estructurales. Para el análisis comparativo se utilizó el software MUMmer, realizando una alineación global del maxicírculo obtenido con sus referencias homologas reportadas previamente, adicional a esto, para conocer la variación intraespecífica de la molécula, se evaluó la presencia de SNV mediante la

estrategia de mapeo con 15 lecturas de Illumina de cepas de *T. cruzi* pertenecientes a diferentes haplotipos reportadas en la base de datos SRA, utilizando el software NGSEP.

Los resultados muestran porcentajes de identidad nucleotídica del 98,51, 98,20, 92,75, 92,53, 92,50, 90,02, 86,40% con *T. cruzi* Y, *T. cruzi* Silvio, *T. cruzi* CL Brener, *T. cruzi* 231, *T. cruzi* MT3663, *T. cruzi* Esmeraldo y *T. cruzi* Marinkellei, respectivamente. Por otro lado, la comparación con otras especies de tripanosomátidos mostraron porcentajes de 80,30, 80,29, 79,92, 79,27 y 78,92% con *T. brucei*, *T. equiperdum*, *T. rangeli*, *T. vivax* y *T. lewisi* respectivamente. Además, con especies de Leishmanias del subgénero Viannia presentaron un porcentaje de identidad entre 77,98% - 78,07%. Estos resultados muestran la utilidad de la secuenciación de tercera generación para la obtención de genomas complejos como los mitocondriales con variación en tamaños y secuencias en la región control y bajos porcentajes de GC.