

5. Análisis genómico comparativo con diferentes especies y subgenotipos del apicomplexa *Cryptosporidium*

Laura Marcela Arias-Agudelo¹, Gisela M. García-Montoya^{1,2}, Felipe Cabarcas^{1,3}, Juan F. Alzate^{1,2} Ana Luz Galván-Díaz⁴

INTRODUCCIÓN Y PLANTEAMIENTO DEL PROBLEMA

Cryptosporidium es un apicomplexa considerado una importante causa de diarrea en vertebrados, incluyendo el hombre. Estudios multicéntricos en países en desarrollo lo clasifican como uno de los cinco patógenos asociados a diarrea en niños menores de cinco años. En la actualidad se describen 38 especies que difieren en cuanto a la especificidad de hospedero, hábitat, distribución geográfica y virulencia; siendo *C.hominis*, *C.parvum* y *C.meleagridis* las más frecuentes en humanos. Además, se han descrito familias de subgenotipos que representan variaciones intra-especie basadas en el polimorfismo del gen gp60. Aunque se tienen disponibles genomas de diferentes especies y subgenotipos, son escasos los análisis comparativos que exploran las bases genéticas que expliquen las diferencias fenotípicas descritas para este protozoo. Estudios previos enfocados en *C.hominis* y *C.parvum*, han descrito una similitud a nivel de nucleótidos de 95-97%(6) postulándose que las diferencias se deben a polimorfismos sobre regiones codificantes y cambios en la regulación génica.

Por lo anterior en este trabajo se pretende realizar análisis de genómica comparada con diferentes genomas de las especies intestinales más frecuentes en humanos y explorar la relación entre las variaciones genéticas identificadas con las diferencias en cuanto al rango de hospederos y potencial de transmisión.

¹ Centro Nacional de Secuenciación Genómica CNSG. SIU Universidad de Antioquia.

² Facultad de Medicina. Universidad de Antioquia

³ Grupo SISTEMIC, Ingeniería electrónica, Universidad de Antioquia

⁴ Grupo de Microbiología ambiental, Escuela de Microbiología. Universidad de Antioquia

Correspondencia: Ana Luz Galván-Díaz; ana.galvan@udea.edu.co

Macroproyecto: "Estudio de prevalencia, diversidad genética y genómica de *Cryptosporidium* en población VIH positiva de Antioquia" COLCIENCIAS 115-777 57608

OBJETIVO GENERAL

Realizar análisis comparativo estructural y funcional con diferentes genomas de especies y subgenotipos de *Cryptosporidium* descritos en humanos.

MATERIALES Y MÉTODOS:

Se descargaron del SRA los reads de 24 aislados de *Cryptosporidium*: *C.hominis*, *C.parvum* y *C.meleagridis*; incluyendo para cada una de las especies diferentes subgenotipos. Los reads fueron ensamblados *de novo* con SPAdes. Para los análisis comparativos se utilizó como referencia el genoma de *C.parvum* Iowa II. Se buscaron las variantes de un sólo nucleótido (SNVs) con MUMmer y dnaDIFF y los polimorfismos de secuencias largas (LSPs) con MUMmer y Assemblytics(9). Para la detección de variantes en CDSs se mapearon los reads con BWA y se anotaron con SIFT4G(11). Los genes con cambios no sinónimos deletéreos fueron anotados realizando asignación de términos GO y análisis de ortología con KEGG.

RESULTADOS PRELIMINARES Y DISCUSIÓN:

El tamaño de los genomas de los 24 aislados fue aproximadamente de 9.0 Mb; contenido medio de GC de 30% y una cobertura del mapeo por encima de 100X en la mayoría de los aislados. El porcentaje de similitud frente a *C.parvum* fue de 96,8 % en los aislados de *C.hominis* y 91,5 % en *C.meleagridis*. En *C.parvum* se identificaron diferencias intra-genotípicas, con porcentajes de identidad entre un 99,7 y 99,9 % para aislados zoonóticos y del 99,5 % para antroponóticos.

La especie que presentó el mayor número de SNVs fue *C.meleagridis* seguida de *C.hominis*. En *C.parvum* se identificó una acumulación de SNVs superior en los aislados antroponóticos comparado con los zoonóticos. La mayoría de los cambios de un solo nucleótido se encontraron en las regiones codificantes; en cuanto a su ubicación en los cromosomas, se observó una distribución homogénea en *C.parvum*, mientras que para *C.hominis* y *C.meleagridis*, una distribución heterogénea, identificándose una mayor cantidad en los cromosomas 1 y 3. No se evidenciaron diferencias significativas en la acumulación de SNVs sinónimos y no sinónimos por especie.

Los genes con cambios no sinónimos deletéreos se encontraron principalmente en *C.meleagridis* con 1017, *C.hominis* 715, y *C.parvum* 288. De estos genes, 183 son compartidos por las tres especies y en su mayoría codifican proteínas hipotéticas. El análisis de ortología indicó que se encuentran implicados en el procesamiento de la información genética y procesos metabólicos.

La mayoría de los indels se presentaron en regiones codificantes, con un rango entre 50 a 500pb de longitud. El mayor número de inserciones y deleciones se detectaron en *C.hominis* y *C.meleagridis* respectivamente.

CONCLUSIÓN:

Se identificaron eventos de variabilidad genética inter-especie e intra-genotípica, siendo los *SNVs* el principal polimorfismo detectado en el genoma de las especies de *Cryptosporidium* analizadas. *C.hominis* y *C.meleagridis* presentaron el mayor número tanto de *SNVs* como de indels. Entre los aislados de cada especie no se observaron diferencias en los parámetros evaluados. Se requieren estudios que mejoren la anotación de los genomas disponibles, para así poder explorar el efecto de las variantes estructurales genómicas sobre las características fenotípicas descritas en las especies del género.

BIBLIOGRAFÍA

1. Kotloff KL, Blackwelder WC, Nasrin D, Nataro JP, Farag TH, Van Eijk A, et al. The Global Enteric Multi-center Study (GEMS) of diarrheal disease in infants and young children in developing countries: Epidemiologic and clinical methods of the case/control study. *Clin Infect Dis*. 2012 Dec;55 Suppl 4(Suppl 4):S232-45.
2. Platts-mills JA, Babji S, Bodhidatta L, Gratz J, Haque R, Havt A, et al. Pathogen-specific burdens of community diarrhoea in developing countries: a multisite birth cohort study (MALED). *Lancet Glob Heal*. 2015;3:564-75.
3. Yaoyu Feng, Una M Ryan, Lihua Xiao. Genetic Diversity and Population Structure of *Cryptosporidium*. *Trends Parasitol*. *Trends Parasitol*. 2018 Nov;34(11):997-1011.
4. Šlapeta J. Cryptosporidiosis and *Cryptosporidium* species in animals and humans: a thirty colour rainbow?. *Int J Parasitol*. 2013;43(12-13):957-970.
5. Xiao L, Feng Y. Molecular epidemiologic tools for waterborne pathogens *Cryptosporidium* spp. and *Giardia duodenalis*. *Food Waterborne Parasitol*. 2017;8-9:14-32.
6. Mazurie AJ, Alves JM, Ozaki LS, Zhou S, Schwartz DC, Buck GA. Comparative genomics of *cryptosporidium*. *Int J Genomics*. 2013;2013:832756.
7. Leinonen R, Sugawara H, Shumway M; International Nucleotide Sequence Database Collaboration. The sequence read archive. *Nucleic Acids Res*. 2011;39(Database issue):D19-D21.
8. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J Comput Biol*. 2012;19(5):455-77.
9. Nattestad M, Schatz MC. Assemblytics: A web analytics tool for the detection of variants from an assembly. *Bioinformatics*. 2016;32(19):3021-3.
10. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010 Mar 1;26(5):589-95.
11. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. *Nat Protoc*. 2016;11(1):1-9.
12. Finke W, Rachimow C, Pfützner B. Gene Ontology: tool for the unification of biology. *Hydrolog und Wasserbewirtschaftung*. 2004;48(1):2-11.
13. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*. 1999;27(1):29-34.