UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS

FACULTAD DE MEDICINA

IDENTIFICACIÓN DE LA α-L-FUCOSIDASA (ALFUC) EN *Blastocystis spp*

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTOR EN CIENCIAS

PRESENTA:

M. EN C. MARTÍNEZ OCAÑA JOEL

TUTOR PRINCIPAL DE TESIS: DR. PABLO MARAVILLA CAMPILLO
FACULTAD DE MEDICINA, UNAM

COMITÉ TUTOR: DRA. ANA FLISSER STEINBRUCH
FACULTAD DE MEDICINA, UNAM
DR. GUILLERMO AGUILAR OSORIO
FACULTAD DE QUÍMICA, UNAM

CIUDAD UNIVERSITARIA, CD. MX. 2022

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS

FACULTAD DE MEDICINA

IDENTIFICACIÓN DE LA α-L-FUCOSIDASA (ALFUC) EN *Blastocystis spp*

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTOR EN CIENCIAS

PRESENTA:

M. EN C. MARTÍNEZ OCAÑA JOEL

TUTOR PRINCIPAL DE TESIS: DR. PABLO MARAVILLA CAMPILLO
FACULTAD DE MEDICINA, UNAM

COMITÉ TUTOR: DRA. ANA FLISSER STEINBRUCH
FACULTAD DE MEDICINA, UNAM
DR. GUILLERMO AGUILAR OSORIO
FACULTAD DE QUÍMICA, UNAM

CIUDAD UNIVERSITARIA, CD. MX. 2022

COORDINACIÓN DEL POSGRADO EN CIENCIAS BIOLÓGICAS

FACULTAD DE MEDICINA

OFICIO CPCB/839/2022

ASUNTO: Oficio de Jurado

M. en C. Ivonne Ramírez Wence
Directora General de Administración Escolar, UNAM
P r e s e n t e

Me permito informar a usted qué en la reunión ordinaria del Comité Académico del Posgrado en Ciencias Biológicas, celebrada el día **25 de abril de 2022** se aprobó el siguiente jurado para el examen de grado de **DOCTOR EN CIENCIAS** del estudiante **MARTÍNEZ OCAÑA JOEL** con número de cuenta **86093844** con la tesis titulada **"IDENTIFICACIÓN DE LA α-L- FUCOSIDASA (ALFuc) EN *Blastocystis spp*",** realizada bajo la dirección del **DR. PABLO MARAVILLA CAMPILLO,** quedando integrado de la siguiente manera:

Presidente:	DR. ARTURO EDGAR ZENTENO GALINDO
Vocal:	DR. JUAN PEDRO LACLETTE SAN ROMÁN
Vocal:	DRA. CECILIA TERESA XIMÉNEZ GARCÍA
Vocal:	DR. LUIS FELIPE JIMÉNEZ GARCÍA
Secretario:	DRA. ANA FLISSER STEINBRUCH

Sin otro particular, me es grato enviarle un cordial saludo.

A T E N T A M E N T E
"POR MI RAZA HABLARÁ EL ESPÍRITU"
Ciudad Universitaria, Cd. Mx., a 13 de septiembre de 2022

COORDINADOR DEL PROGRAMA

POSGRADO EN CIENCIAS BIOLOGICAS
COORDINACIÓN

DR. ADOLFO GERARDO NAVARRO SIGÜENZA

AGRADECIMIENTOS INSTITUCIONALES

Al Posgrado en Ciencias Biológicas de la Universidad Nacional Autónoma de México UNAM, por aceptarme como alumno y permitirme continuar con mi formación profesional

Al Consejo Nacional de Ciencia y Tecnología (CONACYT), a través del Fondo FONDO SECTORIAL DE INVESTIGACIÓN EN SALUD Y SEGURIDAD SOCIAL SS/IMSS/ISSSTE-CONACYT 168619.

A mi tutor principal, Dr. Pablo Maravilla Campillo, gracias por su confianza al aceptarme como alumno de Doctorado y por todo el apoyo brindado durante el desarrollo de este trabajo.

A los miembros de mi Comité Tutor, Dra. Ana Flisser Steinbruch, y Dr. Guillermo Aguilar Osorio, por su apoyo durante el desarrollo de este proyecto de tesis y sus valiosos comentarios y sugerencias para mejorarlo

AGRADECIMIENTOS A TÍTULO PERSONAL

A la Universidad Nacional Autónoma de México UNAM, que ha sido mi segunda casa y parte fundamental de mi formación profesional

Al Hospital General “Dr Manuel Gea González” por permitirme formar parte de su personal, como miembro del Área de Investigación

Al Dr. Edgar Zenteno Galindo, Dr. Juan Pedro Laclette San Román, Dra. Cecilia Teresa Ximénez García, Dr. Luis Felipe Jiménez García y Dra. Ana Flisser Steinbruch por sus comentarios y valiosa contribución para mejorar la elaboración de la presente tesis

Al Dr. Williams Arony Martínez Flores por sus valiosas contribuciones para desarrollar el presente trabajo

A la Dr. Angélica Olivo Díaz por sus acertados comentarios y su apoyo para el desarrollo de este trabajo

DEDICATORIA

A mi Madre María Eugenia Ocaña Salazar† y a mi Padre Paulino Martínez Tapia†; a donde quiera que ellos se encuentren ahora…

A mis hermanos: Pablo, Rosa†, Santos†, Lilia, Blanca y Laura

Índice

Lista de figuras

Lista de tablas

I.- Resumen

Blastocystis es un microorganismo intestinal que parasita a seres humanos y a diversos animales, es el eucarionte intestinal más común a nivel mundial y en los últimos años, se ha convertido en el parásito más identificado en heces de humanos en diferentes países. Se calcula que podrían existir más de mil millones de portadores humanos de *Blastocystis* en todo el mundo. No obstante, su alta frecuencia, su papel en la patogenia humana intestinal sigue siendo ambigua, ya que algunos estudios lo asocian a sintomatología intestinal, mientras que otros estudios lo exoneran. *Blastocystis* presenta una gran variabilidad morfológica con 4 estadios principales: vacuolar (o cuerpo central), granular, ameboide y quiste. *Blastocystis* es un parásito con gran diversidad genética. Basándose en el análisis de la secuencia del gen de la subunidad pequeña ribosomal (*SSUrDNA* por sus siglas en inglés), se han identificado varios subtipos (STs). Los STs 1-3 son los más frecuentes en la población humana, sin embargo, también se han encontrado otros STs en humanos y animales (ST4-10, ST12, ST14 y ST16). Entre los STs 1-4 colectivamente representan el 90% que tienen como hospedero al humano, mientras que el 10% restante lo constituyen ST5-ST17, que se han observado en otros mamíferos y aves.

En países en vías de desarrollo en zonas tropicales, se presentan altas frecuencias de *Blastocystis*, asociado a hábitos higiénicos deficientes, carencia de agua potable o ambos, mientras que en países desarrollados se encuentra un comportamiento estacional y una prevalencia baja. Con respecto a México, se han publicado varios estudios que muestran una frecuencia variable del 7 al 80%, dependiendo de la población estudiada, su condición de rural o urbana y la técnica de identificación. Estudios sobre epidemiología molecular en nuestro país en poblaciones urbanas y rurales, tanto en adultos como en niños de Guerreo, Michoacán. Morelos y Sonora han mostrado que los principales STs que circulan son el ST1, ST2 y ST3, identificándose además en muy baja frecuencia (<5%) el ST7.

Los microorganismos intestinales pueden expresar/secretar diferentes moléculas que interactúan con la mucosa intestinal del huésped, entre ellas α-L-fucosidasa (ALFuc), una glucósido hidrolasa (GH). Según la Carbohydrate-Active EnZYmes Database" n.d. (CAZy), ALFuc es una enzima que cataliza la eliminación hidrolítica de residuos de L-fucosa que se unen al extremo no reductor de las cadenas de glicanos, como las mucinas. Según su secuencia de aminoácidos, las α-L-fucosidasas se clasifican en dos familias: GH29 y GH95. Las enzimas GH29 son una amplia familia de fucosidasas de retención con actividad en

enlaces α(1,2)-, α(1,3)-, α(1,4)- y α(1,6)-L-fucosilo. la GH29 se ha dividido en subfamilias A y B, con GH29A activa en una amplia gama de ligandos, mientras que GH29B es específica para α(1,3)- y α(1,4)-L-fucosilo ligandos. Mientras que las enzimas GH95 son una pequeña familia de fucosidasas invertidasas con actividad sobre α(1,2)-fucosil galactosa.

Se ha demostrado una importante vínculo entre la α-L-fucosidasa humana (denominada FUCA2) y la adhesión, crecimiento y patogenicidad de *Helycobacter pylori*. Además, se ha documentado que el crecimiento y la invasión de las cepas *Campylobacter jejuni fuc*+, 129,108 y NCTC 11168 aumenta en presencia de L-fucosidasas activas liberadas por *Bacteriodes fragilis*, lo que demuestra que *C. jejuni* depende de las fucosidasas externas para su posterior crecimiento e invasión. También se ha demostrado que la capacidad de las bifidobacterias para metabolizar compuestos L-fucosilados (presentes en la leche materna) por medio de fucosidasas es un mecanismo indispensable para dar forma al microbioma intestinal en los primeros meses de vida en humanos.

Existe mucha información sobre ALFuc en bacterias; sin embargo, ALFuc no ha sido identificada experimentalmente en *Blastocystis*. Por lo tanto, el objetivo de la presente tesis fue identificar ALFuc en sobrenadantes de cultivos axénicos comerciales de *Blastocystis* ST1. Además de secuenciar el gen completo de *alfuc* a partir de *Blastocystis* STs 1-3 y comparar la secuencia de aminoácidos de ALFuc en STs 1-3.

En esta tesis se realizaron cultivos de *Blastocystis* ATCC ST1 50610 y 50177 para obtener El extracto de proteínas de *Blastocystis* (EpB) y proteínas de excreción/secreción (Pe/s), posteriormente se sometieron a SDS-PAGE y *Western blot*. Además, se analizaron 18 muestras fecales de portadores humanos de *Blastocystis* mediante PCR-secuenciación para subtipificación. Así mismo se realizó la secuenciación completa del gen *alfuc*, de estas 18 muestras, posteriormente se tuvo la traducción a proteína. Estas secuencias se analizaron para identificar dominios conservados, además de realizar modelado de proteínas en 3D y finalmente se realizó análisis filogenético del gen *alfuc*, comparándolo con la filogenia del gen *SSUrDNA*.

los resultados mostraron la identificación de ALFuc de *Blastocystis* (~51 kDa), principalmente en las proteínas de excreción/secreción. Además, el análisis de las secuencias de aminoácidos de ALFc STs 1-3 mostró la conservación del dominio α-L-fucosidasa y los sitios catalíticos de la familia de la glicósido hidrolasas GH29 en los 3 subtipos; curiosamente, el dominio de unión a la galactosa se reconoció solo en ST1 y ST2.

Los resultados del análisis filogenético del gen *alfuc*, confirmaron una topología idéntica conservada con la topología del gen *SSUrDNA*.

Nuestros resultados muestran, por primera vez , que ALFuc es un producto de secreción de *Blastocystis*. En todas las secuencias de ALFuc se identificó el dominio IPR000933 de la familia GH29, de clan CL0058 de las α-L-fucosidasas. y un dominio tipo C 5/8 (IPR000421) de unión a carbohidratos conservado en ST1 y ST2. Además, las ALFuc de *Blastocystis* STs 1-3 tienen conservado el sitio catalítico para la familia de las glucósido hidrolasas-29 (GH29), con el residuo nucleófilo asp (D) y el residuo general ácido/base Glu (E). Además, el gen *alfuc* es un candidato prometedor de un marcador filogenético adicional al gen *SSUrDNA,* que es el clásicamente usado para analizar la filogenia de *Blastocystis*.

Abstract

Blastocystis is an intestinal microorganism that parasitizes humans and various animals, it is the most common intestinal eukaryote worldwide, in recent years, it has become the most identified parasite in human feces in different countries. It is estimated that there could be more than one billion human carriers of *Blastocystis* worldwide. Despite its high frequency, its role in human intestinal pathogenesis remains ambiguous, since some studies associate it with intestinal symptoms, while other studies exonerate it. *Blastocystis* presents a great morphological variability with 4 main stages: vacuolar (or central body), granular, amoeboid and cyst. *Blastocystis* is a parasite with great genetic diversity. Based on sequence analysis of the small ribosomal subunit (*SSUrDNA*) gene, several subtypes (STs) have been identified. STs 1-3 are the most prevalent in the human population, however, other STs have also been found in humans and animals (ST4-10, ST12, ST14 and ST16). Among the STs 1-4 collectively they represent 90% that have the human as host, while the remaining 10% are ST5-ST17, which have been observed in other mammals and birds.

In underdeveloped countries, in tropical areas, there are high frequencies of *Blastocystis*, associated with poor hygienic habits, lack of drinking water or both, while in developed countries there is a seasonal behavior and a low prevalence. With respect to Mexico, several studies have been published that show a variable frequency of 7 to 80%, depending on the population studied, its rural or urban condition and the identification technique. Studies on molecular epidemiology in our country in urban and rural populations, both in adults and children of Guerreo, Michoacán. Morelos and Sonora have shown that the main STs that circulate are ST1, ST2 and ST3, also identifying in very low frequency (<5%) the ST7.

Intestinal microorganisms can express/secrete different molecules that interact with the host's intestinal mucosa, including α-L-fucosidase (ALFuc), a glycoside hydrolase (GH). According to the Carbohydrate-Active EnZYmes Database"n.d. (CAZy), ALFuc is an enzyme that catalyzes the hydrolytic removal of L-fucose residues that bind to the non-reducing end of glycan chains, such as mucins. According to their amino acid sequence, α-L-fucosidases are classified into two families: GH29 and GH95. GH29 enzymes are a broad family of retention fucosidases with activity in α(1,2)-, α(1,3)-, α(1,4)- and α(1,6)-L-fucosyl bonds. GH29 has been divided into subfamilies A and B, with GH29A active in a wide range of ligands, while GH29B is specific for α(1,3)- and α(1,4)-L-fucosyl ligands. While GH95 enzymes are a small family of inverted fucosidases with activity on α(1,2)-fucosyl galactose.

An important connection between the human α-L-fucosidase (called FUCA2) and the adhesion, growth and pathogenicity of *Helicobacter pylori* has been demonstrated. In addition, the growth and invasion of *Campylobacter jejuni* fuc+, 129,108 and NCTC 11168 strains has been documented to increase in the presence of active L-fucosidases released by *Bacteriodes fragilis*, demonstrating that *C. jejuni* depends on external fucosidases for subsequent growth and invasion. The ability of bifidobacteria to metabolize L-fucosylated compounds (present in breast milk) by fucosidases has also been shown to be an indispensable mechanism for shaping the gut microbiome in the first months of life in humans.

There is a lot of information about ALFuc in bacteria; however, ALFuc has not been experimentally identified in *Blastocystis*. Therefore, the objective of this thesis was to identify ALFuc in supernatants of commercial axenic cultures of *Blastocystis* ST1. In addition to sequencing the entire *alfuc* gene from *Blastocystis* STs 1-3 and comparing the amino acid sequence of ALFuc in STs 1-3.

In this thesis, *Blastocystis* ATCC ST1 50610 and 50177 cultures were processed to obtain *Blastocystis* protein extract (EpB) and excretion/secretion proteins (Pe/s), then subjected to SDS-PAGE and Western blot. In addition, 18 fecal samples from human *Blastocystis* carriers were analyzed by PCR-sequencing for subtyping. Likewise, the complete sequencing of the *alfuc* gene was carried out, of these 18 samples, later the translation to protein was had. These sequences were analyzed to identify conserved domains, in addition to performing 3D protein modeling and finally phylogenetic analysis of the *alfuc* gene was performed, comparing it with the phylogeny of the *SSUrDNA* gene.

The results showed the identification of AlFuc of *Blastocystis* (~51 kDa), mainly in the excretion/secretion proteins. In addition, the analysis of the amino acid sequences of ALFc STs 1-3 showed the conservation of the α-L-fucosidase domain and the catalytic sites of the glycoside hydrolase family GH29 in the 3 subtypes; interestingly, the galactose-binding domain was recognized only in ST1 and ST2.

Phylogenetic analysis of the *alfuc* gene confirmed an identical topology conserved with the topology of the *SSUrDNA* gene. Our results show, experimentally for the first time, that ALFuc is a product of *Blastocystis* secretion. In all ALFuc sequences, the IPR000933 domain of the GH29 family, clan CL0058 of the α-L-fucosidases, was identified. A carbohydrate-binding C-5/8 (IPR000421) type-C domain was identified as conserved in ST1 and ST2. In addition, the ALFuc of *Blastocystis* STs 1-3 have preserved the catalytic site for the family

of glycoside hydrolases-29 (GH29), with the nucleophilic residue asp (D) and the general acid/base residue Glu (E). In addition, the *alfuc* gene is a promising candidate for an additional phylogenetic marker to the *SSUrDNA* gene, which is the one classically used to analyze the phylogeny of *Blastocystis*.

II.- Introducción

1.- Antecedentes históricos de *Blastocystis* y taxonomía

Blastocystis es un protista Stramenopile intestinal que parasita a seres humanos y a diversos animales, presentado una amplia distribución mundial. Recientemente se ha incrementado su frecuencia en portadores que son diagnosticados por medio de estudios coproparasitoscópicos, y métodos de biología molecular, llegándose a encontrar una frecuencia de más del 60% en países en vías de desarrollo (1–3). A pesar de que *Blastocystis* fue descrito a principios del siglo XX, sólo se había estudiado su morfología y ha sido en la última década en la que se han obtenido avances significativos en el conocimiento de la biología de este microorganismo.

Las primeras descripciones detalladas que se hicieron de *Blastocystis*, fueron realizadas por Alexieff en 1911, en aquel entonces nombró al microorganismo como *Blastocystis enterocola*, considerándolo una levadura. Alexieff aplicó el mismo nombre a las células que observó en las heces de las ratas, cobayos, pollos, reptiles y sanguijuelas (4). En 1912 Brumpt acuñó el nombre *Blastocystis hominis*, cambiando el género específico de *enterocola* a *hominis*, debido a que Brump trabajaba con heces humanas (5). En aquel momento, sus observaciones tuvieron mucha relevancia e influencia tempranas y el nombre de *B. hominis* fue ampliamente aceptado y reconocido en todo el mundo. Brumpt realizó una descripción detallada de *B. hominis*, ubicándolo dentro de los géneros de levadura, *Schizosaccharomyces* y *Saccharomyces* y así permaneció hasta mediados del siglo XX. como una “inofensiva levadura intestinal” (6). Eventualmente, comenzaron a aparecer los informes de *B. hominis* como un “agente de los intestinos”, presente en heces de individuos con o sin enfermedad. No obstante, sin tener evidencia clara de su patogenicidad, *B. hominis* fue incluido en los estudios de parásitos intestinales, junto con amebas y flagelados, y a menudo sin referencia a su clasificación como levadura (4,7)

Estudios posteriores de fisiología y microscopía electrónica indicaron una probabilidad mayor de que *Blastocystis* perteneciera a los protozoos. Sin embargo, más adelante estudios filogenéticos moleculares basados en el análisis de la secuencia DNA de la subunidad pequeña del ARN ribosomal (*SSUrDNA*) (8), demostraron que *B. hominis* (una cepa denominada Nand) y un aislado de *Blastocystis* de cobayo estaban claramente ubicados dentro de un complejo grupo heterogéneo, denominado "Stramenopiles", que incluyen protistas unicelulares y multicelulares cuyos flagelos llevan pelos tubulares tripartitos, Tanto

los heterótrofos como los autótrofos (fotosintéticos) pertenecen a este grupo, incluyendo algas marrones, algas marrones doradas y diatomeas (8).

Con el advenimiento de las tecnologías de secuenciación, fue posible secuenciar y analizar por completo, diferentes genes, como el de la proteína de choque térmico de tipo citosólico de 70 kDa, el factor de elongación de la traducción 2, y la subunidad β no catalítica de la ATPasa vacuolar de *Blastocystis*. Todos estos análisis de filogenia molecular confirmaron la ubicaron de *Blastocystis* dentro de los Stramenopiles, cuyos parientes más cercanos son los alveolata (9–12). (figura 1)

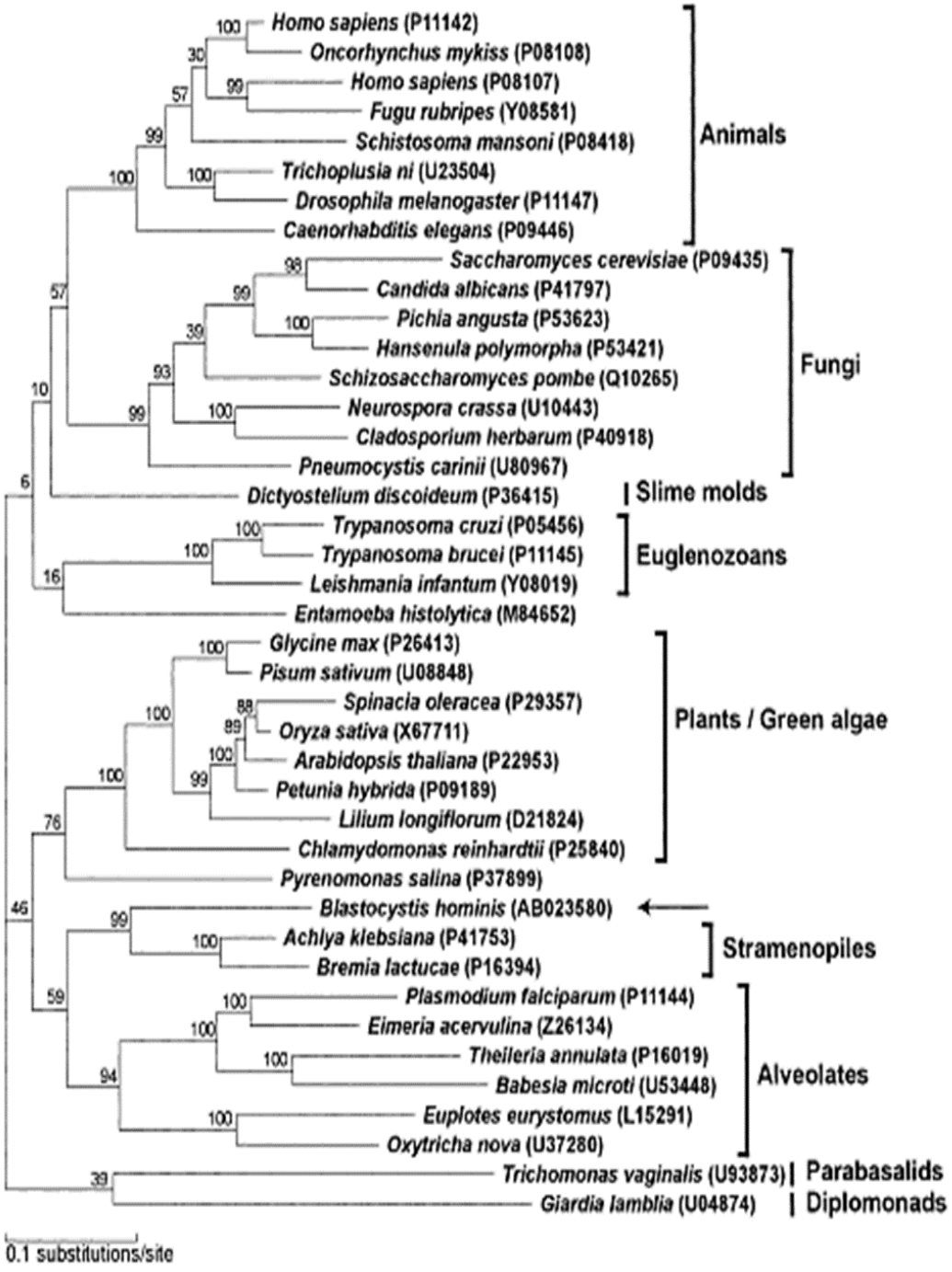


Fig. 1 Tomado de Arisue et al (9). Árbol eucarionte sin raíces de la proteína de choque térmico de tipo citosólico de 70 kDa (HSP70c) utilizando el método de máxima probabilidad (ML). Donde se observa a *Blastocystis* agrupado dentro de los Stramenopilos, cuyos organismos más cercanos son los Alveolatas.

Estudios posteriores de varios investigadores donde se analizaban secuencias del *SSUrDNA* de *Blastocystis* de humanos, identificaron una elevada variabilidad genética y además observaron que podían ocurrir coinfecciones con distintos subtipos de este microorganismo, pero que, a su vez, la mayoría de estos subtipos también se encontraban en otros mamíferos o aves. Esto indicaba que los nombres de las especies del binomio *Blastocystis*-hospedero; como *B. hominis* o *B. ratti*, eran insostenibles, ya que el mismo organismo estaba siendo llamado por múltiples nombres. Por ejemplo, un grupo de *B. hominis* demostró ser genéticamente indistinguible de *B. ratti*; ambos ahora conocidos como *Blastocistis* subtipo 4 (ST4). Para tratar de homologar la terminología, se hizo una propuesta que consistía en nombrar a *Blastocystis* acompañado del subtipo (ST) numerado. Esta clasificación ha simplificado la comunicación entre los investigadores que trabajan con *Blastocystis* de humanos (13). Actualmente se considera a *Blastocystis* como un género agrupado en por lo menos diecisiete subtipos (ST), basados en la secuencia del gen de la subunidad pequeña ribosomal, nueve de los cuales (ST1-ST9) se han encontrado en mayor medida en humanos y los restantes, en aves principalmente (14). Los *Blastocystis* de reptiles, anfibios e invertebrados aún conservan los nombres binomiales de Linneo. Esto se debe en gran medida a que, hasta la fecha, no se ha profundizado en investigaciones acerca de la variabilidad y el rango de hospederos que estos *Blastocystis* pudieran tener y no ha existido el mismo impulso que en los parásitos de humanos y aves, como para modificar la nomenclatura (14).

Originalmente identificado como una levadura, *Blastocystis* fue clasificado en el subphylum Sporozoa y posteriormente reclasificado en el subphylum Sarcodina con base en características morfológicas y fisiológicas. Sin embargo, la clasificación taxonómica actual, reafirmada por comparaciones filogenómicas, lo ubica como miembro del Reino SAR (Stramenopiles, Alveolata y Rhizaria), Phylum Stramenopiles, que es un grupo complejo y heterogéneo que integra a diversos organismos (unicelulares o multicelulares, heterotróficos o fotosintéticos) como oomicetes, mohos mucilaginosos, diatomeas y algas marrones. Algo que resulta importante resaltar, es que Blastocystis es el único Stramenopiles conocido que puede causar infecciones en humanos (Arisue et al., 2002; Silberman et al., 1996). de esta manera la clasificación actual de *Blastocystis* queda de la siguiente manera:

Reino: Sar

Phylium: Stramenopiles

Clase: Bigyra

Orden: Opalinata

Familia: Blastocystisdae

Género: *Blastocystis*

Especies: No aplicable

2.- Morfología

Blastocystis presenta una gran variabilidad morfológica, aunque se aceptan principalmente 4 estadios o formas principales: la vacuolar (o cuerpo central), la granular, la ameboide y la quística. La forma vacuolar, también referida como la forma de cuerpo o vacuola central es la forma predominante encontrada en las muestras de heces humanas, varía enormemente en su tamaño de 2 a 200 μm de diámetro, presenta una gran vacuola que ocupa hasta 90% del volumen celular, teniendo aparentemente una función de "almacén", en el citoplasma periférico se encuentran de uno a dos núcleos, organelos asociados a mitocondrias, aparato de Golgi y retículo endoplásmico (figura 2). (17,18). La forma granular comparte muchas similitudes con la vacuolar excepto que esta presenta numerosos gránulos en el citoplasma periférico e inclusive dentro de la vacuola central (Figura 2). La forma ameboide rara vez se encuentra en muestras fecales y en cultivo, aunque su descripción morfológica es motivo de conflicto y confusión en los reportes (19) y se ha sugerido que es una forma intermediaria entre la vacuolar y la quística, que ingiere bacterias para obtener los requerimientos suficientes para su enquistamiento, aunque esta hipótesis se encuentra aún en debate. También se ha propuesto que esta forma juega un papel importante en la patogenia de *Blastocystis* (20). La forma quística fue descrita y confirmada entre 1988 y 1999 y puede ser fácilmente confundida con detritus fecales; los quistes son esféricos u ovales y están protegidos por dos capas fibrosas, contienen de 1 a 4 núcleos, vacuolas y depósitos de lípidos y glucógeno (7,21,22). Estos quistes miden 15 μm de diámetro, pero presentan enormes diferencias en el tamaño cuando provienen de diferentes especies de hospederos animales. Se ha observado que pueden sobrevivir en agua durante 19 días a temperatura ambiente, pero son lábiles al calor y al frío extremo, contrariamente, las formas vacuolares y granular son

sensibles a los cambios de temperatura y exposición al aire, por lo que es muy probable que el estadio quístico sea la forma de transmisión de este parásito (23–25).

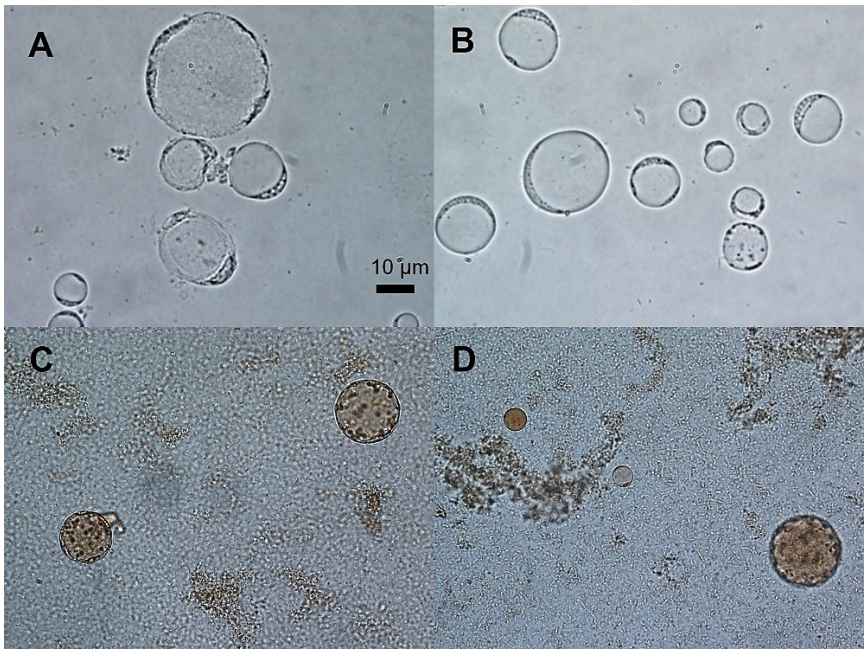


Figura 2. imágenes en microscopía de luz de *Blastocystis* en observadas con el objetivo a 60x. En A y B se observan en estadio vacuolar, con la vacuola que ocupa la mayor parte del citoplasma y algunos organelos en la periferia del citoplasma, también se ve la variabilidad de tamaño. En C y D Los *Blastocystis* están en estadio granular, observados con Lugol.

3.- Ciclo de vida

La infección por *Blastocystis*, al igual que en otras parasitosis se transmite de manera fecal oral. Estudios llevados a cabo con ratones BALB/c jóvenes (de menos de 8 semanas) inmunocompetentes y ratas demostraron que podían infectarse con *Blastocystis* inoculados con quistes por vía oral. Después de la infección con quistes, en los ratones se encontraron formas vacuolares y granulares en el ciego, pero sólo se observaron formas quísticas en el colon. Estos estudios confirmaron que el quiste fecal era la forma responsable de la transmisión externa y que el modo de transmisión era por la vía fecal-oral (26,27) . Una vez que se lleva a cabo la ingestión de los quistes, el parásito se desenquista en el intestino delgado y desarrolla la forma vacuolar, posteriormente se divide por fisión binaria y puede desarrollar las formas vacuolares/granular, aunque en cultivo se han observado otras formas: por ejemplo, formas ameboides, cuya función biológica y su eventual destino en el desarrollo del parasito requieren más investigación. El enquistamiento ocurre a través del paso del parásito por el intestino grueso y los quistes son excretados con las heces (figura 3). El ciclo

se completa cuando estos quistes son consumidos por humanos o animales a través de alimentos contaminados (28).

En cuanto a la localización de la infección por *Blastocystis*, diferentes estudios han mostrado que, en ratones y ratas infectados experimentalmente con quistes, los segmentos del tracto gastrointestinal afectados con mayor frecuencia son ciego y colon. Las formas morfológicas predominantes observadas son vacuolares y granulares. En el ciego las formas quísticas son las más observadas (26,27). Algunos estudios con roedores también han identificado formas ameboides de *Blastocystis* en el íleon, yeyuno y estómago (29). En otro estudio con cerdos inmunodeprimidos infectados experimentalmente se encontró a *Blastocystis* predominantemente en el intestino grueso, pero también se detectó en el intestino delgado, siendo las formas vacuolares/granulares las más observadas. Además se ha descrito que la localización de este protista en el tejido intestinal se halla principalmente en la luz o en el borde de la mucosa intestinal, junto con depósitos de mucina tanto en roedores infectados experimentalmente (29,30), como en cerdos infectados naturalmente (31,32).

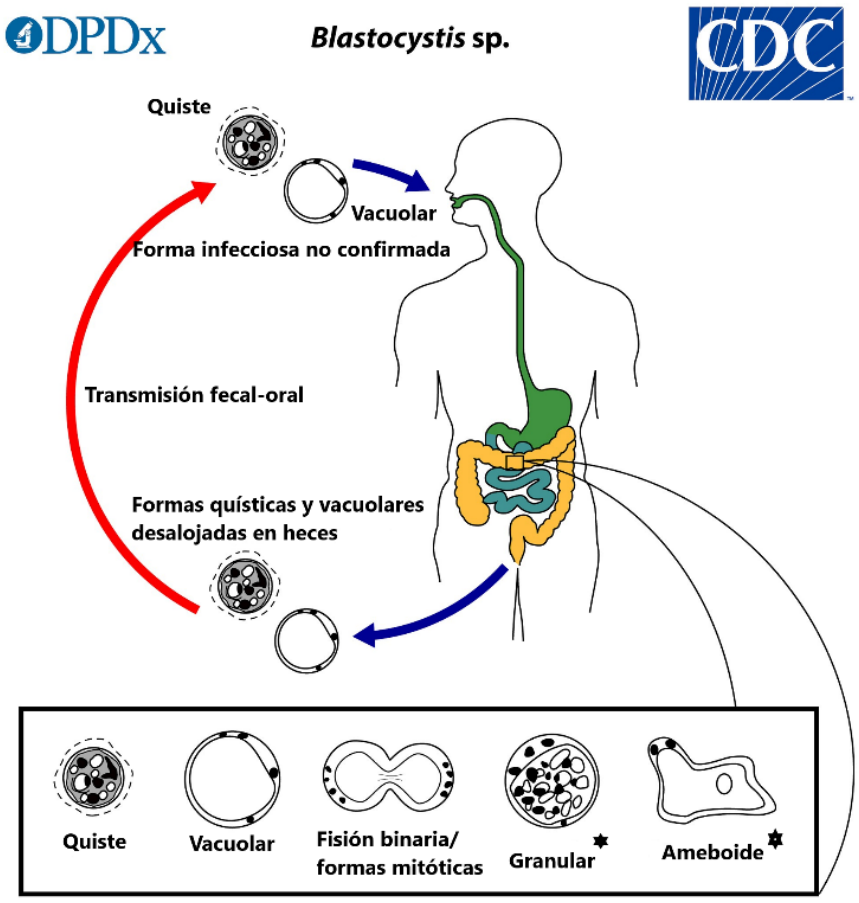


Figura 3. Ciclo de vida para *Blastocystis* (tomado y modificado) del DCD, haciendo énfasis en el hospedero humano. El quiste es la forma infectante liberada en las heces. Una vez que se ingieren los quistes, los parásitos se desenquistan en el intestino, desarrollando la forma vacuolar/granular la cual se reproduce por fisión binaria. se desconoce si la forma vacuolar puede dar lugar a formas ameboides o granulares en el intestino. Algunas formas vacuolares se enquistan y estos quistes son expulsados a través de la materia fecal en donde pierden una capa fibrilar que rodea a los quistes. Una vez en el ambiente, los quistes son transmitidos a humanos y a animales, completando de esta manera el ciclo vital del parásito. Las formas granular y ameboide (con asterisco) han sido observadas ocasionalmente en heces y cultivo de *Blastocystis*, pero su significado biológico no está entendido completamente.

4.- Epidemiología

Con respecto a su prevalencia en poblaciones humanas, *Blastocystis* se ha investigado en gran medida en el contexto de la enfermedad y algunos estudios lo han implicado en una serie de diferentes enfermedades intestinales y extraintestinales, incluyendo la enfermedad inflamatoria intestinal, el síndrome del intestino irritable (SII) y la urticaria (33–36). *Blastocystis* también se menciona como el agente causal de una enfermedad denominada blastocistosis, que se caracteriza por una serie de síntomas inespecíficos (entre otros, dolor abdominal, diarrea y distensión abdominal) que de hecho podrían atribuirse y estar asociados con cualquier otro tipo de microorganismos infecciosos (34). Los estudios transversales que prueban la hipótesis de que *Blastocystis* está vinculado al SII suponen en su mayoría que, si el organismo está asociado a la enfermedad, debería ser más común en pacientes con síntomas de SII. Los resultados de estos estudios han sido variados, ya que algunos han encontrado una mayor prevalencia de *Blastocystis* en los pacientes con SII y otros no encuentran ninguna diferencia o incluso una prevalencia menor. Hasta la fecha, las tasas de prevalencia notificadas son generalmente más altas en los países en vías de desarrollo y se han asociado con los niveles de saneamiento, la fuente de agua y el contacto con los animales (34,37,38). La prevalencia más alta reportada para *Blastocystis* es del 100% y proviene de un estudio reciente de 93 niños en la cuenca del río Senegal (39). En contraste con esto, las estimaciones actuales de la prevalencia de este microorganismo en los países desarrollados oscilan entre 0,5% y 30% dependiendo de los métodos utilizados y la población muestreada (40,41). En América la prevalencia de *Blastocystis* en algunos estudios es hasta del 44% y 58.9 %, siendo los ST1, ST2 y ST3 los subtipos más comúnmente encontrados (42,43). Por otro lado en estudios llevados a cabo en una cohorte aleatoria de 1106 muestras de heces del Flemish Gut Flora Project (FGFP) encontraron una prevalencia de portadores de *Blastocystis* del 30%, la presencia de *Blastocystis* no tuvo asociación con SII o malestares intestinales, además encontraron una prevalencia mayor del ST4 y los autores sugieren que la ambigüedad de *Blastocystis* con respecto a su papel en la salud humana, podría estar relacionada con el subtipo o incluso con la diversidad intra subtipo de *Blastocystis* (44).

Con respecto a México, se han publicado reportes que muestran una frecuencia variable dependiendo de la población estudiada, su condición de rural o urbana y la técnica de identificación utilizada. Un estudio realizado en una zona periurbana de la Ciudad de México, mediante estudios coproparasitoscópicos (CPS) mostró una frecuencia del 7% (45). Otro estudio en comerciantes de alimentos de un mercado establecido en la delegación

Xochimilco, también en la Ciudad de México y realizando estudios CPS reportó una frecuencia del 48% (46). De manera interesante, una revisión de más de 200,000 CPS realizados en entre 1990 y 2010 en el Hospital Infantil de México Federico Gómez mostró (HIMFG), mostró que la prevalencia de *Blastocystis* entre la década de 1990 y 1999 fue cercana al 2.5%, mientras que en la década de 2000 a 2010 este valor se incrementó a casi el 19%, los autores discutieron que esta diferencia muestra lo complejo que puede ser la identificación microscópica debida al alto polimorfismo de *Blastocystis* y que requiere de una constante capacitación técnica para la adecuada identificación de este microorganismo por estudios CPS (47). Publicaciones sobre la prevalencia de *Blastocystis* en poblaciones rurales han mostrado una frecuencia mayor del 60% (48,49). Estudios sobre epidemiología molecular en nuestro país en poblaciones urbanas y rurales, tanto en adultos como en niños mostraron que los principales STs que circulan son el ST1, ST2 y ST3, identificándose además en muy baja frecuencia (<5%) el ST7 (50,51).

Independientemente de que *Blastocystis* esté relacionado con la salud o la enfermedad gastrointestinal, es evidente que *Blastocystis* es mucho más común de lo que se había informado anteriormente. Debido al alto porcentaje de individuos positivos para este microorganismo y a su estabilidad temporal (cuando está presente), algunos investigadores han sugerido que *Blastocystis* podría ser parte de la microbiota intestinal del humano sano; además, los distintos subtipos y genotipos de *Blastocystis* presentes en los individuos, hace lógico suponer que diferentes ST pueden colonizar simultáneamente el intestino del hospedero sin causar sintomatología clínica (14,16,41); no obstante, hasta la fecha su papel patógeno continúa siendo un tema de debate científico.

5.- Patología

Aunque se ha especulado la posible contribución de este microorganismo en enfermedades funcionales del intestino humano, también es claro que la infección asintomática puede ser común. En otros parásitos patógenos como *Gardia lamblia* y *Entamoeba histolytica* se han observado muchas infecciones asintomáticas (52) y en el caso de *Entamoeba histolytica* se ha documentado que las infecciones sintomáticas son aproximadamente el 10% de los casos (53). Las características clínicas no específicas que se han asociado a *Blastocystis* son: nauseas, anorexia, dolor y distensión abdominal, flatulencia y diarrea crónica o aguda, siendo el dolor abdominal y la diarrea las más frecuentes, aunque en muchos casos estos síntomas

eventualmente se autolimitan. Sin embargo, la manifestación clínica más importante asociada a la presencia de este microorganismo es el Síndrome de Intestino Irritable (SII), (54–56)

Los hallazgos principales en mecanismos probables de patogenicidad en *Blastocystis*, se han llevado a cabo en estudios que involucran el co-cultivo de células epiteliales intestinales de rata o humano y con *Blastocystis* procedentes de estos mismos hospederos. Las células epiteliales intestinales han sido tratadas con parásitos completos, lisados parasitarios o productos de excreción/secreción del cultivo de parásitos (57). No obstante, una de las limitaciones para el estudio de *Blastocystis* es la falta de aislados axénicos de los diferentes subtipos, especialmente del ST3 que es uno de los más comunes en humanos. También es importante tener en cuenta que los cultivos axénicos utilizados en estos estudios son predominantemente de la forma vacuolar del parásito y pueden no necesariamente reflejar las formas morfológicas asociadas con el mecanismo de adhesión (31,32) y la patogénesis in vivo (58,59). Sin embargo, estos estudios han proporcionado información valiosa sobre los posibles mecanismos patógenos en la blastocystosis (figura 4).

En aislados de *Blastocystis* procedentes de roedores, se han identificado factores de virulencia como cisteín-proteasas, las cuales pueden estimular la sobreexpresión del gen de la interleucina 8 (IL-8) por medio de la activación del factor nuclear NF-κB (21). Además, en cultivo de líneas celulares epiteliales IEC-6, se ha observado que *Blastocystis* puede inducir apoptosis de manera independiente al contacto con las células, generar rearreglos en la distribución de F-actina, produciendo disminución de la resistencia transepitelial e incrementando la permeabilidad celular (22). Además, es probable que, como ocurre en otras infecciones por protozoos intestinales, la apoptosis en células del hospedero y la alteración en la función de la barrera epitelial, jueguen un papel relevante en la patogénesis de *Blastocystis*.

Puthia y colaboradores (21), han propuesto que este microorganismo puede inducir la producción de IL-8 por las células epiteliales del intestino, generando una gran afluencia de células inflamatorias en la mucosa intestinal, lo que a su vez conlleva a un daño en el tejido y a alteraciones gastrointestinales. Asimismo, se ha documentado que la invasión del epitelio intestinal por patógenos no es necesariamente por la inducción de un proceso inflamatorio, dado que *Blastocystis* no es un parásito invasivo, los productos secretados por el parásito (como las cisteín proteasas) (60), podrían iniciar el proceso inflamatorio por la activación de receptores en la superficie celular, así como ejercer una variedad de efectos dañinos en la

células del hospedero, resultando en efectos citopáticos, ruptura y alteración de la barrera epitelial enterocítica, así como la producción de citocinas proinflamatorias (60). Además de los efectos directos sobre el epitelio intestinal del huésped, *Blastocystis* también tiene propiedades inmunomoduladoras. Las cisteín proteasas presentes en lisados del parásito, así como otros componentes de excreción/secreción en el medio de cultivo del parásito, son capades de degradar la IgA de manera dependiente del pH y del tiempo, mostrando que la actividad de proteasa puede variar entre los subtipos, teniendo niveles de actividad más altos en ST7, respecto al ST4 (61,62).

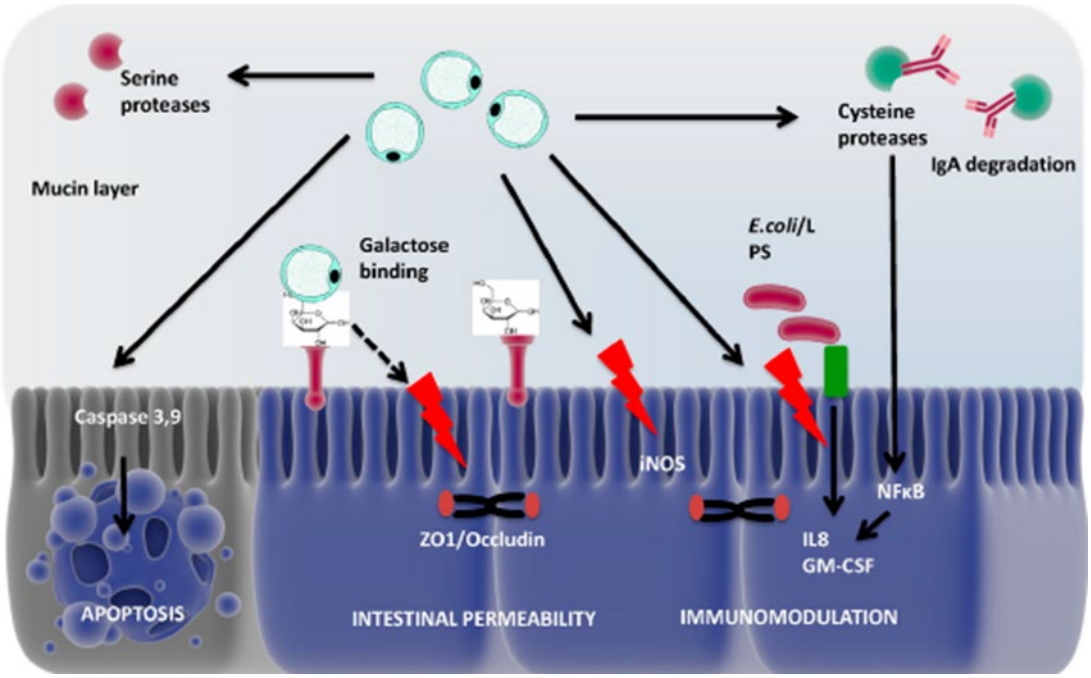


Figura 4. Esquema que sintetiza los hallazgos de estudios in vitro con *Blastocystis* y células del epitelio intestinal: El parásito puede liberar proteasas de cisteína y serína, que pueden activar al factor nuclear NF-κB y estimular la sobreexpresión del gen de la interleucina IL-8. También se ha observado que *Blastocystis* puede inducir apoptosis de manera independiente al contacto con las células, generar rearreglos en la distribución de F-actina, ocasionando disminución de la resistencia transepitelial e incrementando la permeabilidad celular (57)

Adicionalmente, se le ha atribuido a este parásito capacidad para inducir síntomas alérgicos y cutáneos, principalmente urticaria (36,63,64).

Cuando se examinaron secciones de tejido de intestinos de cerdo, infectados experimentalmente con *Blastocystis*. Fayer et al. (31) encontraron *Blastocystis* principalmente en el lumen, comúnmente asociado a restos de alimentos digeridos, y aunque a veces se encontraban cerca del epitelio o parecían adherirse él, no había células que penetraran al epitelio o a la lámina propia. Estas observaciones fueron confirmadas por Wang

et al (32), que no observaron ninguna patología evidente en secciones histológicas de biopsias de la mucosa intestinal de cerdos infectados con *Blastocystis*. En este último estudio, los parásitos se observaron cómo formas vacuolares/granulares encontradas dentro del material luminal o en la proximidad de las células epiteliales sin evidencia de adhesión o invasión.

Además, se ha argumentado que la capacidad de *Blastocystis* para inhibir la respuesta inmune del huésped puede favorecer su persistencia, adhesión y supervivencia dentro del tejido intestinal, facilitado por la liberación de moléculas en la interfaz huésped-parásito. Las cisteín proteasa asparaginil de 31-kDa llamada legumaina que se encuentra en la superficie celular del parásito, y la catepsina B, representan otros factores de virulencia. Se ha observado que el legumaina tiene un papel pro-supervivencia del parásito, ya que induce apoptosis de las células parasitarias cuando reaccionan con anticuerpos específicos (65). Estas dos proteasas fueron identificadas en sobrenadantes de cultivo de *Blastocystis* y pueden estar implicadas en el aumento de la permeabilidad intestinal, la degradación de la IgA, la inducción de la producción de IL-8, la degradación del moco, el daño tisular y la alteración de funciones gastrointestinales (66).

Por otro lado, en el 2011 fue publicado el genoma completo de *Blastocystis ST7*, aislado de un paciente de Singapur. Este genoma fue a su vez comparado con el genoma de 2 organismos filogenéticamente relacionados con *Blastocystis*; *Phytophthora sojae* (un parásito de plantas) y *Thalassiosira pseudonana* (un alga diatomea de vida libre). Posteriormente se realizó un análisis de predicción de proteínas que pudieran ser potencialmente secretadas por *Blastocystis*, dando un total de 307 proteínas, de las cuales 75 pudieran estar relacionadas con la patogenicidad, entre estas proteínas se encontraron diferentes moléculas de digestión: como proteasas y glicosidasas. Dentro de las glicosidasas se encontró que *Blastocystis* producía α-L-fucosidasa, una proteína de nuestro interés debido a que posee actividad mucolítica, capaz de liberar restos de fucosa del extremo no reductor de mucinas y otras glicoproteínas presentes en el moco intestinal (67). Se ha propuesto en otros microorganismos (principalmente en bacterias) que el moco intestinal sirve como fuente de nutrientes, lo que permite a las bacterias multiplicarse y colonizar la capa del moco intestinal creando un microambiente favorecedor (68,69). A pesar de los avances que se han realizado en la comprensión de las interacciones de *Blastocystis* con el epitelio intestinal y la inmunidad asociada a la mucosa intestinal, todavía hay muchas preguntas por responder.

6.- Alfa-L-fucosidasa

El epitelio intestinal de los humanos está cubierto por una densa capa de moco constituido por varias mucinas con estructuras complejas de carbohidratos en sus fracciones glicano, principalmente compuestas por estructuras tipo O-glicano, que desempeñan un papel crucial para el establecimiento y mantenimiento de la microbiota intestinal (70). Para sobrevivir, los microorganismos buscan fuentes de carbono durante la digestión secuencial de los carbohidratos ramificados, mediante la escisión de los residuos terminales de fucosilo de los O-glicanos (70). Algunos parásitos intestinales pueden expresar diferentes moléculas que interactúan con el epitelio intestinal del huésped, como las lectinas, enzimas glucósido hidrolasas y proteínas similares a la mucina, que pueden favorecer la unión y la invasión de las células huésped, así como la evasión de la respuesta inmune. Además, las enzimas parasitarias también pueden facilitar la infección degradando el moco del huésped o generando sitios de unión a las células huésped (71).

Dentro de las enzimas glucósido hidrolasas se encuentra la α-L-fucosidasa (ALFuc). Según Carbohydrate-Active EnZYmes Database" n.d. (CAZy), esta enzima cataliza la eliminación hidrolítica de los residuos de L-fucosa que se unen al extremo no reductor de las cadenas de glicanos, como las mucinas. De acuerdo con su secuencia de aminoácidos, las ALFucs se clasifican en dos familias de glucósido hidrolasas (GH), la GH29 y GH95. Las enzimas GH29 son una amplia familia de fucosidasas de retención con actividad sobre enlaces α(1,2)-, α(1,3)-, α(1,4)- y α(1,6)-L-fucosilo. La GH29 se ha dividido a su vez en subfamilias A y B, siendo la GH29A activa en una amplia gama de enlaces, mientras que la GH29B es específica para los enlaces α(1,3)- y α(1,4)-L-fucosilo. Mientras que las enzimas GH95 son una pequeña familia de α-fucosidasas invertidas con actividad sobre la α(1,2)-fucosil galactosa (72). La ALFuc se ha asociado con varios eventos celulares, como la respuesta inmune (73,74), transducción de señales (75), embriogénesis del desarrollo (76,77) y a procesos patológicos como cáncer (78,79) y fucosidosis (80,81). En la literatura hay bastante información acerca de varias ALFuc de bacterias intestinales, consideradas probióticos como: *Bifidobacterium* (82,83), Bacteroides (84,85) y *Lactobacillus* (86,87), por mencionar algunas. En estos trabajos también se ha descrito la relación entre las L-fucosidasas y la degradación de residuos fucosilados (88,89). Así mismo, en estos estudios se menciona que la capacidad de producir L-fucosidasas con especificidad de sustratos diferentes, tiene una función importante en la forma en que estos microorganismos pueden aprovechar los residuos de carbohidrato derivados de la acción de estás L-fucosidasas

Cuando se completó y analizó el genoma de *Blastocystis* ST7, la ALFuc fue identificada como una de las 307 proteínas predichas para ser potencialmente secretada por *Blastocystis*. En un estudio posterior se encontró que aproximadamente el 25% de los genes de *Blastocystis* se habían adquirido recientemente por transferencia lateral de genes (TLG) de donantes procariotas y eucariotas. Dentro de estos genes se identificaron 17 genes que participaban en el metabolismo de carbohidratos, incluyendo cinco genes implicados en la importación y el metabolismo de la L-fucosidasa, tales como L-fucosidasa, L-fucosa permeasa fucP y L-fuculose fosfato aldolasa, que son enzimas homólogas en *Bacteroides thetaiotaomicro*n, y L-fucosa deshidrogenasa y L-fuconolactase, y son similares a las que se encuentran en una vía metabólica en *Campylobacter jejuni*. Estos transportadores pueden ser relevantes para la colonización del intestino por el parásito, ya que están involucrados en la concentración y regulación de la fucosa (90).

III.- Planteamiento del problema.

Blastocystis es un parásito stramenopilo ubicuo, que puede parasitar mamíferos, reptiles y aves. Se considera que existen más de un billón de portadores en todo el mundo. La clasificación de STs de *Blastocystis* se realiza de acuerdo a la secuencia del gen de la subunidad pequeña ribosomal, hasta el momento se han descrito, por lo menos 25 STs. Los seres humanos pueden albergar a los STs 1–9 y ST12; más del 90% de los aislados humanos de *Blastocystis* corresponden a los STs 1–4 (16,91). No obstante, las características clínicas no específicas que se han asociado a la infección por *Blastocystis*, como: nauseas, anorexia, dolor, distensión abdominal, flatulencia, diarrea crónica o aguda y el SII, su participación como patógeno, sigue siendo controvertido (92). La importancia de *Blastocystis* en la salud pública sigue siendo desconocida, aunque en últimos estudios este microorganismo se ha encontrado más comúnmente en el tracto gastrointestinal de individuos sanos. Se ha observado que la diversidad y riqueza bacteriana intestinal son en su mayoría más altas en individuos positivos para *Blastocystis*. También se ha demostrado que *Blatocystis* ST7 disminuye los niveles de bacterias intestinales beneficiosas como *Bifidobacterium* y *Lactobacillus* (93,94).

Varios estudios en la literatura mencionan que la capacidad de *Bifidobacterium* y *Lactobacillus* para producir α-L-fucosidasas con especificidad de sustratos diferentes, tiene una función importante en la forma en que estos microorganismos pueden aprovechar los residuos de carbohidrato derivados de la acción de estás L-fucosidasas en los intestinos de

los humanos (82, 83, 86, 87). Aunado a esto, en *Blastocystis* se han descrito varios genes implicados en el metabolismo de los carbohidratos, como el gen de la ALFuc, L-fucosa permeasa fucP y L-fuculosa fosfato aldolasa. Todas estas enzimas pueden ser relevantes para la colonización del intestino por el parásito, ya que están involucrados en la concentración y regulación de la fucosa (90). Existe mucha información en la literatura relacionada con ALFuc en bacterias; sin embargo, hasta el momento, no se ha identificado ALFuc en medio de cultivo de parásitos, como producto de secreción de *Blastocystis*. Por estas razones es de nuestro interés, identificar la ALFuc en medio de cultivo de cepas axénicas de *Blastocystis* ST1 ATCC 50177 y ATCC 50610. Además de amplificar el gen completo de *alfuc* y traducir a proteínas para comparar las secuencias de ALFuc de los ST1, ST2 y ST3 de *Blastocystis.*

IV.- Hipótesis

Se han descrito varios genes implicados en el metabolismo de los carbohidratos de *Blastocystis* spp, entre ellos el gen *alfuc*. Asimismo, se ha predicho a la ALFuc, como una proteína producto de las secreciones del parásito. Por ello, se propone que sea factible identificar a la ALFuc en el medio de cultivo del parásito y al gen *alfuc* en aislados de *Blastocystis* ST1-3

V.- Objetivos

Objetivo general

Reconocer la ALFuc en medio de cultivo de cepas axénicas de *Blastocystis* ST1 ATCC 50177 y ATCC 50610 e Identificar el gen *alfuc* en aislados de *Blastocystis* ST1-3

Objetivos particulares

- Establecer el cultivo de *Blastocystis de* muestras fecales de diferentes portadores humanos
- Identificar la ALFuc en sobrenadantes de los cultivos de *Blastocystis*
- Secuenciar el gen completo *alfuc* y obtener la secuencia de proteínas
- Comparar las secuencias de proteínas de ALFuc e identificar dominios conservados
- Realizar modelado de predicción de estructura terciara de las ALFuc
- Realizar análisis filogenético con el gen *alfuc* de los distintos aislados obtenidos en el presente estudio

VI.- Metodología

1.- Cultivo de *Blastocystis* in vitro

Para realizar el cultivo de *Blastocystis*, se utilizaron las cepas axénicas de *Blastocystis* ST1 obtenidas de la *American Type Culture Collection* (ATCC), ATCC 50177 y ATCC 50610. Para establecer con éxito los cultivos in vitro de *Blastocystis*, se siguieron las recomendaciones del ATCC. Cuando las cepas axénicas estabilizaron su crecimiento en cultivo, se inocularon en 10 ml de Gibco Iscove's Modified Dulbecco's Medium (IMDM), suplementadas con 10% de suero de caballo inactivado (PAA, Pasching, Austria) y 1% de penicilina/estreptomicina (Gibco) en tubos cónicos de 15 ml. Los cultivos se cultivaron en un recipiente anaeróbico (Anaerocult Jar, MERCK, que contenía un sobre para anaerobiosis BD GasPak) y se incubaron a 37 ° C. El mantenimiento del cultivo se realizó por subcultivo cada 48-72 horas en 10 ml de IMDM fresco, suplementado con suero de caballo y penicilina/estreptomicina, como se mencionó anteriormente (95).

2.- Recolección de muestras de portadores sintomáticos de *Blastocystis* spp

El presente estudio se realizó de acuerdo con las recomendaciones de los Comités de Ética en Investigación e Investigación del Hospital General "Dr. Manuel Gea González" quienes aprobaron el protocolo (aprobación No. 12-77-2018). Se obtuvo el consentimiento informado por escrito de cada participante. Se recolectaron 18 muestras fecales frescas de portadores de *Blastocystis* que acudieron a consulta médica por trastornos gastrointestinales en el Hospital General "Dr. Manuel Gea González" estas muestras fueron examinadas mediante microscopía y se realizaron procedimientos de aislamiento del parásito. Se inocularon muestras de heces del tamaño de un guisante o muestras líquidas de 250 µl en un medio de Jones de 8 ml complementado con suero de caballo inactivado al 5% (PAA, Pasching, Austria), seis tubos de cultivo por muestra. Los tubos de cultivo se incubaron a 37°C durante 48 horas, y el éxito del aislamiento se confirmó mediante microscopía.

3.- Preparación de extractos de proteína *Blastocystis* y sobrenadantes de cultivo

El extracto de proteínas de *Blastocystis* (EpB) y las proteínas de excreción/secreción (Pe/s) de los sobrenadantes de cultivo, se obtuvieron a partir de los cultivos en IMDM (88) con las cepas axénicas ATCC de *Blastocystis* ST1 50177 y 50610 arriba mencionadas. Una vez que se obtuvo el mayor crecimiento de los parásitos (48-72 hrs), los cultivos se centrifugaron a 6000 xg durante 10 minutos para recuperar el sobrenadante de cultivo. Los parásitos se

lavaron tres veces con PBS 1X pH 7.2 y se centrifugaron a 6000 xg durante 10 minutos. Después, los parásitos fueron lisados por diez ciclos de congelación-descongelación a -70 ° C y 37 ° C, respectivamente, y sonicados (SONIC Ruptor 4000, OMNI International, USA) en hielo, utilizando diez pulsos durante un minuto, separados por un minuto de descanso (96). El EpB se clarificó por centrifugación a 16.000 xg a 4°C durante 10 minutos y se pasó a través de un filtro de 0,20 µm (Corning) (26). La concentración de proteínas se determinó mediante el ensayo de Bradford (Bio-rad), posteriormente los viales se almacenaron a -70 °C hasta que fue necesario. Los sobrenadantes de cultivo recolectados de la primera centrifugación de los parásitos (aproximadamente 15 ml) que contenían las Pe/s se centrifugaron a 1000 xg durante 10 min y se filtraron a través de membranas de 0,20 µm (Corning) y, a continuación, se aplicaron en membranas Amicon Ultra-15 100.000 NMWL (Millipore) a 4000 xg durante 40 min a 4 °C para eliminar proteínas con un peso molecular superior a 100 kDa (59). Posteriormente, estas Pe/s se concentraron membranas Amicon Ultra-4 10.000 NMWL a 4.000 xg durante 40 min a 4 °C, y las proteínas se resuspendieron en 2 mL de amortiguador Tris-HCl 50 mM pH 7,5 y se almacenaron a -70 °C, hasta su uso.

4.- Identificación de α-L-fucosidasa (ALFuc) de *Blastocystis*

Inicialmente, para identificar La ALFuc de *Blastocystis* en electroforesis, se utilizó la α-L-fucosidasa de 477 aminoácidos, anotada para *Blastocystis* ST7 disponible en la base de datos UniProt con el identificador D8M3D3. Con la secuencia de esta proteína, usando el software ExPASY-Compute pI/Mw, se predijo su masa molecular teórica de ~54kDa. Después, se realizó electroforesis de proteínas SDS-PAGE en geles prefabricados de gradiente de poliacrilamida del 4-20%. Para correr la electroforesis se utilizaron, 60ug de EpB y Pe/s y se diluyeron en amortiguador de carga 2x para proteínas que contenía Tris-HCl 150 mM, SDS al 3%, glicerol al 30%, azul de bromofenol al 0.1%. Después las muestras se colocaron en agua hirviendo durante 5 minutos y luego se corrieron en geles prefabricados de gradiente de poliacrilamida al 4-20% (geles Mini-PROTEAN TGX Bio-rad) a una temperatura constante de 100 voltios durante 1 hora. Posteriormente para visualizar las bandas de proteína, se utilizó la tinción azul brillante de Coomassie. Para identificar la banda de proteína de ~54kDa, correspondiente ALFuc de *Blastocystis* y calcular su movilidad relativa (Rf), se usaron pesos moleculares estándares de proteína (Kaleidoscope Bio-rad). Se midió la Rf de cada uno de los pesos moleculares, y de la proteína que migraba aproximadamente a la misma altura que

el peso molecular de 50kDa. Se utilizaron tres corridas electroforéticas para realizar el cálculo de Rf (97).

5.- *Western blot* para ALFuc

Para realizar la electroforesis, se utilizaron 20 µg de EpB y Pe/s, las proteínas se separaron en un gel prefabricado de gradiente de poliacrilamida al 4-20% (Mini-Protean TGX, Bio-Rad), usando las condiciones descritas anteriormente. Posteriormente, las proteínas de *Blastocystis* se transfirieron a membranas de PVDF (Immobilon-P, Millipore) a una constante de 100 voltios/h. Las bandas de proteínas fueron verificadas por tinción rojo de Ponceau.

Para encontrar las condiciones adecuadas en el *Western blot* se realizaron varios ensayos utilizando diferentes concentraciones de anticuerpos y diferentes soluciones de agente bloqueador de sitios inespecíficos. Finalmente las condiciones fueron las siguientes: los sitios de unión no específicos en la membrana se bloquearon incubando con *Blocking Buffer Reagent* (Roche, Alemania) en amortiguador Tris-Base 50 mM, pH 7.5, NaCl 0.15 M, Tween-20 0.03% (TBS-T) durante 1 hora a temperatura ambiente, seguido de tres lavados con TBS-T e incubación durante toda la noche a 4°C con anti-α-L-fucosidasa (FUCA2) (Santa Cruz Biotechnology Cat# sc-167929, RRID:AB_10840662), a una dilución de 1:2000 en TBS-T. Luego, la membrana se lavó tres veces con TBS-T y se incubó con anti-IgG de ratón biotinilada (Santa Cruz Biotechnology, CA) a una dilución de 1:4000 en TBS-T durante 2 horas a temperatura ambiente. Después de tres lavados, la membrana se incubó con estreptavidina-peroxidasa (Jackson Immunoreseach, Baltimore, PA) a una dilución de 1:10000 en TBS-T durante 1 hora a temperatura ambiente, seguida de tres lavados con TBS-T. Finalmente, la membrana se reveló con diaminobencidina (DAB) (Sigma-Aldrich, St. Louis, WI, USA).

6.- Extracción de DNA y amplificación de genes *alfuc*

Para la extracción de DNA se utilizó el crecimiento de un tubo de cultivo de 8 ml por cepa (CEPAS ATCC y cultivo xénico). Siete ml de sobrenadante de cultivo se retiraron cuidadosamente para cada tubo, y el mililitro restante se centrifugó a 10.000 xg durante 5 minutos, se utilizó QIAmp Fast DNA Stool Mini Kit (Qiagen, Venlo, Países Bajos), siguiendo las instrucciones del fabricante para extraer el DNA de *Blastocystis*. Las muestras de DNA se eluyeron en 50 µl de amortiguador de elución y se almacenaron a -70°C. En este

estudio se diseñaron cuatro nuevos pares de iniciadores (tabla 1) para amplificar una región de alrededor de 2.600 (pb), que incluía todo el gen *alfuc* para ST1-ST3, así como regiones parciales del gen de la subunidad 4 del factor de iniciación de la traducción eIF-3 (eif3S4) y del gen de la beta-estradiol-deshidrogenasa (hsd17B), también contenía, dos regiones intergénicas de flanqueo (figura 5). Para la PCR en volúmenes 25 µL, se usó amortiguador 1X para la PCR: $MgCl_2$ 2.4 mM, dNTP 0.5 mM, BSA 0,01 mg, Taq 1 U, 2-5 µL DNA polimerasa (Invitrogen) y agua libre de nucleasas. Para llevar a cabo la reacción de PCR se realizaron estandarizaciones con gradientes de temperaturas de alineamiento de los iniciadores, hasta encontrar la adecuada. Finalmente, las condiciones fueron: temperatura inicial de desnaturalización de 5 minutos, seguidos de 35 ciclos (desnaturalización a 95°C durante 30 segundos, alineamiento a 69°C durante 1 minuto, elongamiento a 72°C) y temperatura de extensión final de 72°C durante 7 minutos. Los productos de PCR se confirmaron mediante electroforesis en gel de agarosa al 1%. Para cada región del gen *alfuc*, se purificaron los productos de PCR utilizando el kit de limpieza de PCR AxyPrep (Axigen Biosciences, CA), siguiendo las instrucciones del fabricante. Posteriormente, las muestras se secuenciaron utilizando el método Sanger en el Laboratorio Nacional de Biodiversidad del Instituto de Biología de la Universidad Nacional Autónoma de México (NBL-BI-UNAM).

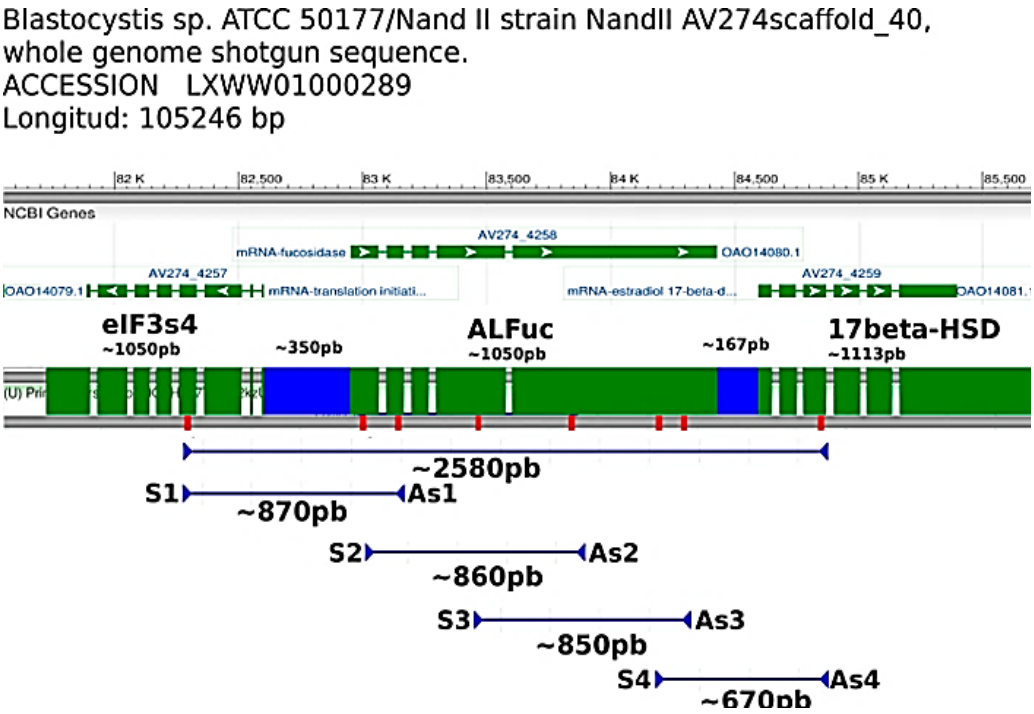


Figura 5. Esquema anotado en la página web del NCBI para el gen *alfuc* de *Blastocystis* ST1 ATCC 50177/Nand II cepa NandII AV274 scaffold_40, secuencia de escopeta del genoma completo, número de acceso: LXWW01000289.1, región: 82284-84859. En la imagen se muestra la región amplificada del gen *alfuc* completo para ST1-ST3 de 2580pb, que contiene regiones parciales de la subunidad 4 del factor de iniciación de la traducción eIF-3 (eif3S4) y la beta-17 estradiol deshidrogenasa (hsd17B), así como dos regiones intergénicas flanqueantes. Se observan: región 1 amplificada con el par de iniciadores ALFuc/S1 y ALFuc/As1, que abarca desde el eif3S4 hasta el extremo 5' del gen *alfuc*, región 2, amplificada con ALFuc/S2 y ALFuc/As2, que abarca la primera mitad del gen *alfuc*, región 3, amplificada con ALFUc/S3 y ALFUc/As3, abarca la segunda mitad del gen *alfuc* y la región 4, amplificada con ALFuc/S4 y ALFuc/As4, abarca desde el extremo 3' del gen *alfuc* hasta el extremo 5' del gen *hsd17B* del gen *hsd17B*.

Tabla 1. Iniciadores degenerados diseñados para amplificar el gen *alfuc* de *Blastocystis* e iniciadores para amplificar *SSUrRNA*

Nombre	**Secuencia de DNA (5'-3')**	**Longitud del fragmento**
ALFuc/S1 F	ACWCCCTCGTTTCCWCCMKYAG	~872 pb
AlFuc/As1 R	TTCAGYTTCACRGGGTTGAAGAT	
AlFuc/S2 F	GAGGAGCARGYYCARTGGCT	~860 pb
AlFuc/As2 R	GTGGTAGAACCAGCCSGGKC	
AlFuc/S3 F	TggAgACCggAgTACAACgA	~850 pb
AlFuc/As3 R	TCTTSACRCCRATggTggTKCCC	
AlFuc/S4 F	TCAACgTgRTSATgATgCAggA	~675 pb
ALFuc/As4 *R*	*TCCRTCVSTKgCWCCYgTSACC*	
Blast 505-532	*GGAGGTAGTGACAATAAATC*	~492pb
Blast 998-1017	*TGCTTTCGCACTTGTTCATC*	

7.- Subtipificación de *Blastocystis*

Para identificar el subtipo de *Blastocystis* de los diferentes aislados se purificó DNA de los parásitos en cultivo con el kit QIAmp Fast DNA Stool Mini Kit (Qiagen, Venlo, Países Bajos), siguiendo las instrucciones del fabricante. Después se utilizaron 200 ng de DNA de purificado para la reacción de PCR usando los iniciadores descritos por Santin et al (tabla 1) (98) y cuyo producto esperado de amplificación es de ~500pb. Este producto de amplificación se purificó con el kit para productos de PCR AxyPrep (Axigen Biosciences, CA), siguiendo las instrucciones del fabricante. Los productos de PCR se confirmaron mediante electroforesis en gel de agarosa al 1%. Posteriormente estos fragmentos purificados se enviaron a secuenciar utilizando el método Sanger en el Laboratorio Nacional de Biodiversidad del Instituto de Biología de la Universidad Nacional Autónoma de México (NBL-BI-UNAM).

8.- Estudios bioinformáticos

Las secuencias de aminoácidos predichas a partir de la traducción de los genes de *alfuc* se analizaron para predecir el péptido señal con SignalP-5 (http://www.cbs.dtu.dk/services/SignalP/versions.php) y determinar los dominios conservados de proteínas en la base de datos Pfam. También se calcularon las masas moleculares para las secuencias de ALFuc con el software ExPASY-Compute pI/Mw (ExPASy Bioinformatics Resource) (99).

9.- Predicción de estructura terciara de ALFuc

Para modelar la estructura 3D de *Blastocystis* ALFuc correspondiente a ST1, ST2 y ST3 obtenida por traducción del gen *alfuc* completo, se utilizó el servidor I-TASSER(100). Este servidor está en desarrollo activo con el objetivo de proporcionar las predicciones más precisas de la estructura y la función de las proteínas utilizando algoritmos de última generación. Tras el análisis, se seleccionaron los modelos con la mayor puntuación de confianza (C-score). Las estructuras tridimensionales se analizaron y visualizaron con PyMOL22(101).

10.- Análisis filogenético del gen *alfuc*

La calidad de los cromatogramas de las regiones de DNA secuenciadas se evaluó con phred/phrap/consed (*Laboratory of Phil Green of University of Washington*) (102–104). El

ensamblaje de las secuencias del gen *alfuc* se realizó, tomando como referencia las regiones genómicas que contienen el supuesto eif3S4-*alfuc*-hsd17B para ST1-ST3. Además, para genes de *SSUrRNA* validados, utilizamos dos conjuntos de datos de secuencia de *Blastocystis* ST1-ST3 y ST4-ST9 y para proteínas alfa-L-fucosidasa predichas (Archivo adicional 2: Tabla S1) (http://entamoeba.lshtm.ac.uk/blastorefseqs.htm). Para realizar el análisis se utilizó a *Proteromonas lacertae* como grupo externo [número de acceso *SSUrRNA*: U37108.1 y número de acceso *alfuc* previsto: NGBS01001212 región: c18708.. 20188]. Se calcularon alineaciones de secuencias múltiples en ClustalW versión 2.0 (105) y se recortaron con trimAl v1.2 (106) como se implementó en Phylemon 2.0 (107). Se utilizó ModelTest- NG v0.2.0 para seleccionar el modelo de sustitución más adecuado para las alineaciones de nucleótidos (*SSUrDNA*: HKY + G y gen *alfuc*: K2 + G) (108,109). Los árboles de máxima probabilidad se construyeron en la versión 10.1.8 con el software MEGA X (110) con 1000 réplicas de *bootstrap*. Los árboles bayesianos se calcularon con el Software Mr. Bayes versión 3.2.6 (111) durante cuatro millones de generaciones, la distribución de probabilidad posterior y la frecuencia de diagnóstico se muestrearon cada 1000 y 105 generaciones respectivamente, se implementó el modelo de sustitución durante el análisis por salto reversible, los árboles se resumieron con una muestra de post-burning del 50%. Los árboles filogenéticos fueron editados con el software Figtree. Las secuencias de los genes *alfuc* se anotaron en comparación con las regiones genómicas (Tabla 3) que codifican las secuencias de ALFuc de *Blastocystis* disponibles para ST1 [número de acceso: OAO14080.1], ST4 [número de acceso: XP_014526040.1] y ST7 [número de acceso: CBK22406.2].

VII.- RESULTADOS

1.- Identificación de ALFuc de *Blastocystis* en cultivos axénicos

Se analizó la secuencia de proteínas α-L-fucosidasa para *Blastocystis* ST7 disponible en la base de datos UniProt con el identificador D8M3D3, usando el software ExPASY-Compute pI/Mw, para predecir su masa molecular teórica, y se obtuvo como resultado una masa molecular de ~54.7 kDa (tabla 2). Este valor teórico se utilizó para identificar ALFuc de *Blastocystis* en EpB y Pe/s en las cepas axénicas (ATCC 50177 y 50610). La tinción del gel con azul brillante de Coomassie evidenció una fuerte banda de proteínas que migró alrededor de 50kDa tanto en EpB como en Pe/s (Figura 6). La masa molecular también se confirmó graficando el Rf de los pesos moleculares vs el log de la masa molecular de los pesos de

referencia (tabla 2). El valor obtenido de Rf = 0.52 para la banda de proteína que migraba aprox a 50kDa (ALFuc de *Blastocystis*), se interpoló en la recta, obteniéndose un valor de 1.70 y un resultado final del inverso del log de ~51.7 kDa (Figura 7). Posteriormente cuando se tuvo la secuencia completa traducida a proteína del gen completo de las dos cepas axénicas, se analizó por ExPASY-Compute pI/Mw, dando un resultado teórico de ~52.7 kDa (tabla 2). Este resultado es parecido al obtenido experimentalmente en electroforesis.

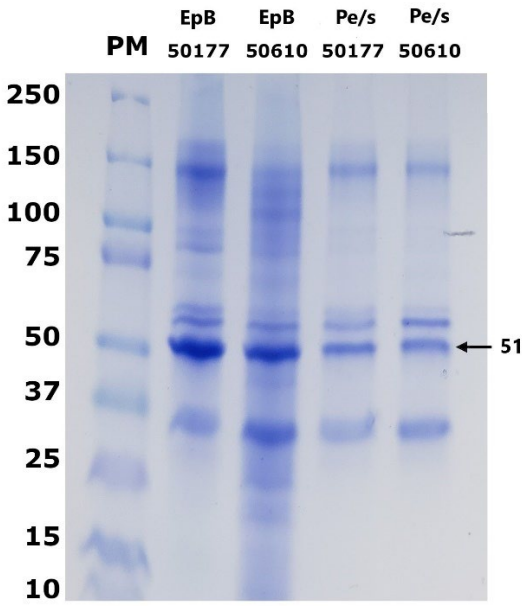


Figura 6. En la imagen se observa el gel SDS-PAGE, teñido con azul brillante de coomassie, de los EpB y Pe/s correspondientes a las cepas de referencia ATCC 50177 y 50610. La flecha indica la masa molecular esperada de 51kDa para la ALFuc de *Blastocystis*, calculada por el software ExPASY-Compute pI/Mw (ExPASy Bioinformatics Resource Portal)

Movilidad relativa Rf = Distancia recorrida por el frente de electroforesis / Distancia recorrida por la proteína

Tabla 2. Valores de Rf

RF	PM log
0.2	2.39
0.28	2.17
0.34	2
0.41	1.87
0.52	1.69
0.58	1.56
0.66	1.39
0.77	1.3
0.84	1.17
0.93	1

Gráfica de movilidad relativa (Rf)

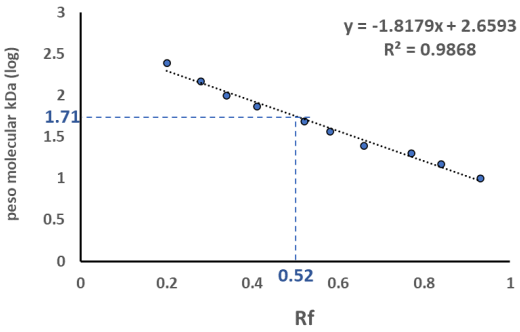


Figura 7. Gráfica de Movilidad relativa (Rf) vs log masa molecular aparente La línea punteada en azul indica la interpolación del valor obtenido de Rf para la ALFuc esperada de *Blastocystis* que fue de 0.52 y su correspondiente valor en Y= 1.71 y su inverso de log = 51.7 kDa, valor esperado para la ALFuc de *Blastocystis*.

2.- *Western blot* para ALFuc

Para identificar la proteína ALFuc en las fracciones EpB y Pe/s de cepas axénicas, utilizamos un anticuerpo monoclonal comercial de ratón, anti-FUCA2, que reconoce una región interna altamente conservada entre las α-L-fucosidasas secretadas. El anticuerpo mostró un claro reconocimiento de las bandas de proteínas que migraron a ~51.7 kDa en Pe/s de las dos cepas comerciales (ATCC-50177 y ATCC-50610); sin embargo, no reconoció las bandas de proteínas en Bpe (Figura 8). La detección de α-L-fucosidasa en el medio de cultivo sugiere que la proteína puede ser un producto de secreción de *Blastocystis*. El reconocimiento del anticuerpo por la ALFuc en EpB fue mucho menor, casi imperceptible, esto puede ser debido a que el anticuerpo se encuentra dirigido contra la α-L-fucosidasa secretada y en EpB es probable que la ALFuc se encuentre en gránulos y su conformación cambie y el anticuerpo tenga menos accesibilidad al epitope.

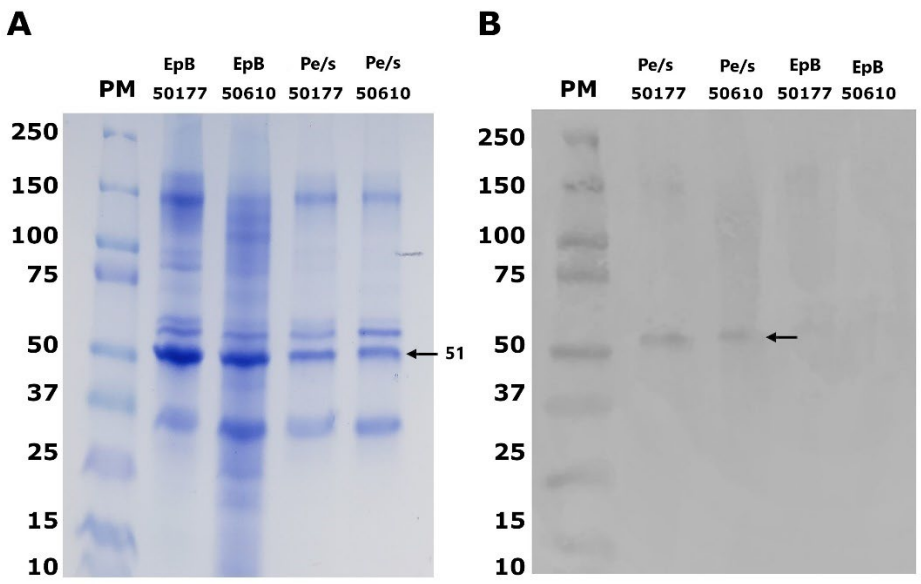


Figura 8. *Western blot* que muestra el reconocimiento de la ALFuc de *Blastocystis* por el anticuerpo comercial anti-α-L-fucosidasa secretora (FUCA2), en las proteínas de excresión/secresión (Pe/s) en las cepas de referencia ATCC 50177 y 50610 (flecha 51kDa). Se observa un muy ligero reconocimiento en el extracto de proteínas de *Blastocystis* en las 2 cepas.

3.- Estudios bioinformáticos

3.1- Comparación de secuencias del gene *alfuc* de *Blastocystis* ST1, ST2 y ST3.

Mediante la secuenciación de cuatro regiones superpuestas de aproximadamente 850 pb (regiones 1 a 3) y 650 pb (región 4) (figura 5), se obtuvo una región genómica de aproximadamente 2,500 pb para cada uno de los 18 aislados xénicos y las dos cepas axénicas de *Blastocystis* (Tabla 1, figura 9 y 10). Las secuencias obtenidas se compararon con regiones genómicas homólogas de ST1-ST4 y ST6-ST9, identificadas en los proyectos WGS (contigs de escopeta de genoma completo) para *Blastocystis* en la base de datos NCBI. Además, se utilizaron tres fucosidasas de *Blastocystis* (disponibles en la base de datos del NCBI) con diferentes longitudes. ST1, 450 aa [OAO14080.1]; ST4, 459 aa [KNB42597.1 y XP_014526040.1]; y ST7, 477 aa [CBK22406.2 y XP_012896454.1], para definir las regiones exónicas e intrónicas putativas del gen *alfuc*. La anotación del genoma de *Blastocystis* ST1 se utilizó para definir las regiones intergénicas y los genes parciales eif3S4 y hsd17B. Estas comparaciones reconocieron la existencia de dos ORFs putativos (marcos de lectura abiertos): un ORF1 basado en anotación del ST4, localizado en todos los ST (ST1-4 y ST6-ST9); y un ORF2 basado en la anotación del ST1, ubicado únicamente en ST1 y ST8. La variación principal entre los dos ORF fue que ORF2 comienza en el intrón uno y se puede traducir en una proteína más corta que carece de un péptido de señal (tabla 3).

Tabla 3. Secuencias predichas del gen *alfuc* a partir de los proyectos *Blastocystis* Whole Genome Shotgun (WGS)

Marcos de lectura abiertos (ORF) #	*Blastocystis* ST	Numero de acceso *	Posiciones basadas en la secuencia predicha del gen *alfuc* $										
			Exón 1	Intrón 1	exón 2	intrón 2	exón 3	intrón 3	exón 4	intrón 4	exón 5	intrón 5	exón 6
ORF 1	9	>JZRO01000168.1:6738-8241	1...36	37...68	69...173	174...203	204…270	271...300	301…373	374...402	-	-	-
	8	>JZRN01000723.1:c8640-10178	1...33	34...67	68...172	173...205	206...274	275...307	308...378	379...407	408...687	688...717	718…1625
	7	>CABX01000063.1:238291-239800	1...35	37...71	72…173	174...205	206...274	275...308	309…379	380…408	-	-	409…1564
	6	>JZRM01000146.1:c29656-31159	1...36	37...68	69...173	174...203	204…270	271...300	301…373	374...402	-	-	403...1570
	4	>JPUL02000258.1:4599-6131	1...36	37...67	68...169	170...201	202...270	271...301	302...372	373...401	402...678	679...7	709...1533
	2	>JZRJ01000159.1:c8839-103	1...34	35...70	71...174	175...208	209...275	276...306	307...379	380...410	411...690	691...71	722...1543
	1	>LXWW01000289.1:82890-84429	1...36	37...68	69...173	174...207	208...274	275...305	306...378	379...408	409...685	686...715	716...1540
	3	>JZRK01000455.1:2042-3575	1...36	37...64	65...169	170...200	201...267	268...298	299...371	372...400	401...680	681...709	710...1582
ORF 2	8	>JZRN01000723.1:c8640-10178	-	-	63…172	173...205	206...274	275...307	308...378	379...407	408...687	688...717	718…1625
	1	>LXWW01000289.1:82890-84429	-	-	62…173	174...207	208...274	275...305	306...378	379...408	409...685	686...715	716...1540

#.-Marcos de lectura abiertos basados en comparaciones de ALFuc de ST1, ST4 y ST7.
*.-Número de acceso del contig. Los dos puntos separan el rango en el que se ha comparado el *alfuc* predicho. La "c" antes del rango indica la orientación en el contig.
$.-Posición de las regiones exónicas e intrónicas en el gen *alfuc* predicho, basado en el número de acceso proporcionado en la columna tres.

Según el análisis funcional de Pfam, todas las secuencias de proteínas predichas pertenecen a la familia GH29 de glucósido hidrolasas, (IPR000933), y poseen dos dominios conservados: el dominio en el N-terminal característico de la familia α-L-fucosidasas del clan CL0058 y un dominio en el C-terminal de unión a galactosa (IPR008979), identificado como el dominio del factor de coagulación 5/8 (IPR000421) en algunas secuencias (Figura 11). Además, las proteínas comparten un motivo de secuencia débil con una protuberancia conservada en la lámina beta C-terminal. La función más probable de esta protuberancia es flexionar la lámina beta que contiene la protuberancia, permitiendo que la curvatura de la hoja que forme el sitio de unión al carbohidrato (112,113). Las proteínas ALFuc de *Blastocystis* predichas a partir de secuencias ORF1 pero no de orf2, presentaron el péptido señal en el N-terminal. El análisis del servidor web SignalP-5.0, identificó en secuencias ORF1 para *Blastocystis* ST1, una escisión peptídica (péptido señal) entre los aminoácidos 16 y 17 (VLA-RP) con una probabilidad de 0,6139. En ST2, una escisión en la posición de aminoácido 14-15 (ALA-KP), con una probabilidad de 0,79 y en ST3, la escisión en la posición de aminoácido 15-16 (VLS-KR), con la probabilidad de 0,79 (tabla 3). El análisis de secuencia no mostró predicción de regiones transmembranales.

Para comparar la identidad de las ALFuc de *Blastocystis*, se realizó alineamiento, usando ClustalW. Se incluyeron 10 secuencias obtenidas a partir de la amplificación y secuencias del gen ALFuc completo, además de las secuencias obtenidas de la misma manera de las cepas ST1 ATCC 50177 y 50610. También, se incluyeron una secuencia de ARNm ST1, una secuencia de Flemming ST2 y una secuencia ST3, estas últimas 3 secuencias tomadas de la base de datos NCBI. En el alineamiento pueden observarse los residuos conservados y similares, marcados con asteriscos y puntos, respectivamente. Así mismo se observan conservados los dominios: región del péptido señal, el dominio de α-L-fucosidasas de la familia GH29 y el dominio de unión a carbohidratos tipo C F5/8. Además, se muestra conservado el sitio catalítico característico de α-L-fucosidasas de la familia GH29, con el residuo nucleófilo asp (D) en rojo y el residuo general ácido/base Glu (E) en azul. Aunque las secuencias ST2_Joel40, ST3_Joel42 y ST3_joel 43 fueron secuenciadas en un 78%, 67% y 67%, respectivamente, se observa que contienen el dominio de las α-L-fucosidasas de la familia GH29 (figura 12). Para observar la conservación del sitio catalítico de α-L-fucosidasas de la familia GH29 se realizó alineamiento comparando las ALFuc de *Blastocystis* con α-L-fucosidasas procariotas conocidas de la familia GH29 tomadas de la base de datos del UnitPro: *Bacteroides thetaiotaomicron* [Q8A085], *Bifidobacterium bifidum*

[C5NS94] y *Streptomyces* sp [Q9Z4I9]. El alineamiento mostró la conservación del sitio catalítico característico de la familia GH29 con el residuo nucleófilo catalítico Asp (D) y el residuo ácido/base Glu (E) (Figura 13).

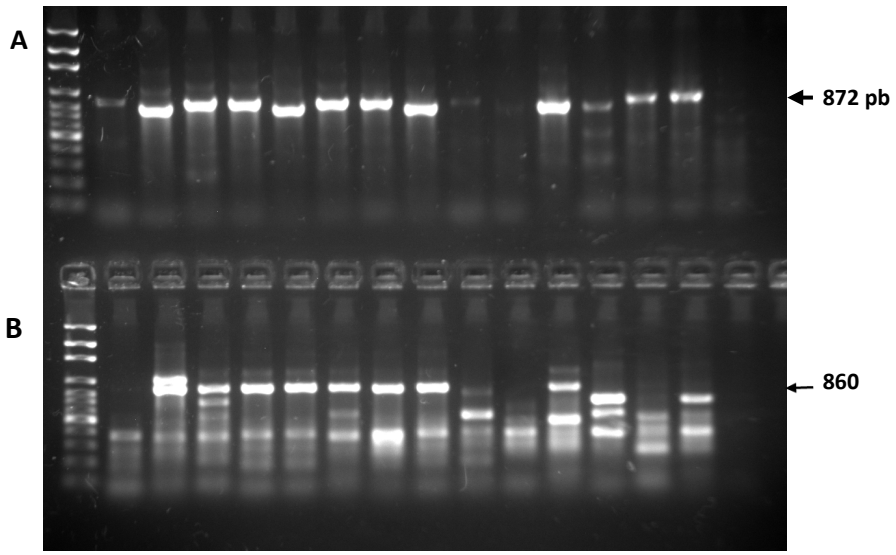


Figura 9. Gel de teñido con bromuro de etidio. En A se muestra el fragmento de la región 1, del gen de *alfuc*, obtenido con los iniciadores ALFuc/S1 y ALFuc/As1 diseñados en este estudio (imagen 6, tabla 1). La flecha indica el tamaño del amplicon esperado. En B se muestra el fragmento de la región 2, de *alfuc*, obtenido con los iniciadores ALFuc/S2 y ALFuc/As2 diseñados en este estudio (imagen 6, tabla 1). La flecha indica el tamaño del amplicon esperado

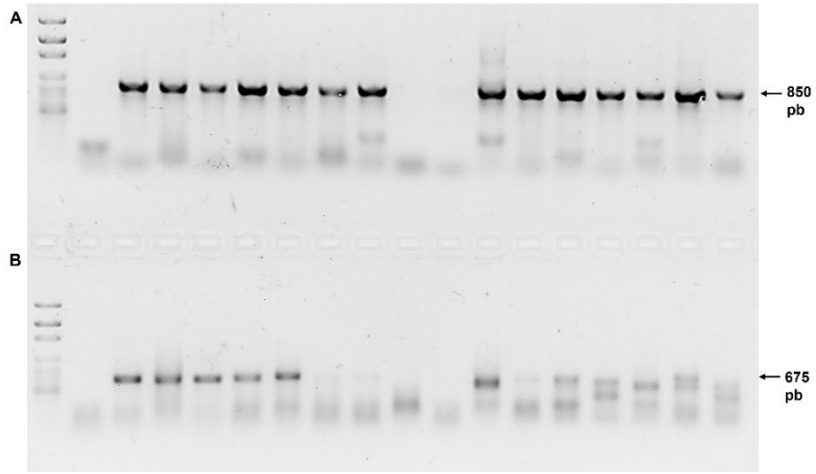


Figura 10. Gel de teñido con bromuro de etidio (imagen inversa, para visualizar mejor las bandas). En A se muestra el fragmento de la región 3, del gen de *alfuc*, obtenido con los iniciadores ALFuc/S3 y ALFuc/As3 diseñados en este estudio (imagen 6, tabla 1). La flecha indica el tamaño del amplicon esperado. En B se muestra el fragmento de la región 4, de *alfuc*, obtenido con los iniciadores ALFuc/S4 y ALFuc/As4 diseñados en este estudio (imagen 6, tabla 1). La flecha indica el tamaño del amplicon esperado.

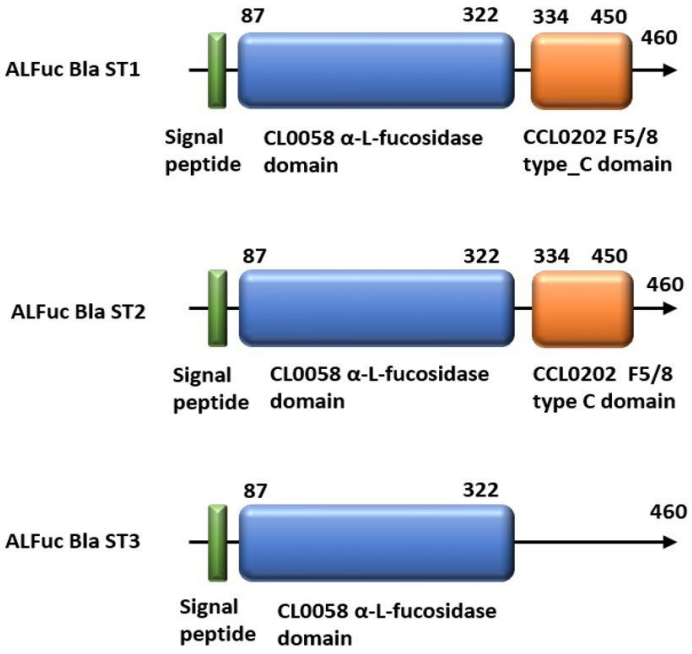


figura 11. Esquema que muestra la estructura de la ALFuc de *Blastocystis*, indicando el péptido señal, el dominio de conservado de α-L-fucosidasa y el dominio conservado tipo C F5/8. Los números superiores indican el número de aminoácido. Los ST 1 y 2 conservan los dos dominios, mientras en el ST3 el resultado del segundo dominio fue negativo.

Tabla 4. Proteínas ALFuc de *Blastocystis* y del péptido señal.

ALFuc	***Blastocystis* ST**	**Sitio de escisión del péptido (posición aa)**	**Probabilidad de escisión del péptido**	**Tamaño aa**	**Masa molecular teórica kDa**
ATCC 50177	1	16 and 17: VLA-RP	0.6139	460	52.7
ATCC 50610	1	16 and 17: VLA-RP.	0.6139	460	52.7
JOEL36	1	16 and 17: VLA-RP.	0.6139	460	52.7
CMX11	2	15 and 16: ALA-KP.	0.7958	460	52.7
Joel5B	2	15 and 16: ALA-KP.	0.7958	460	52.7
Joel19B	2	15 and 16: ALA-KP	0.7958	460	52.7
Joel29B	2	15 and 16: ALA-KP	0.6593	460	52.7
Flemming NCBI	2	15 and 16: ALA-KP	0.7782	459	52.6
CMX5	3	16 and 17: VLS-KR.	0.7936	461	52.7
ST3_NCBI	3	16 and 17: VLS-KR	0.7940	461	52.8
Joel17A	3	16 and 17: VLS-KR	0.7936	461	52.8
D8M3D3	7	ND	ND	477	54.7

```
ST1_mRNA_ALFuc     MISFAIFLLFAVLVLARPKLPIPTEEQVQWLERGISGFIHFGVNTFTGREWGLGNEDPNI 60
ST1_50177          MISFAIFLLFAVLVLARPKLPIPTEEQVQWLERGISGFIHFGVNTFTGREWGLGNEDPNI 60
ST1_50610          MISFAIFLLFAVLVLARPKLPIPTEEQVQWLERGISGFIHFGVNTFTGREWGLGNEDPNI 60
ST1_JOEL36         MISFAIFLLFAVLVLARPKLPIPTEEQVQWLERGISGFIHFGVNTFTGREWGLGNEDPNI 60
ST2_Joel40         MFSFVLFFL-IAIALAKPKLPLPTEEQAQWLERGISGFIHFGVNTFTGKEWGLGDEDPNI 59
ST2_Flemming_NCBI  MLSFVLFFL-IAIALAKPKLPLPTEEQAQWLERGISGFIHFGVNTFTGKEWGLGDEDPNI 59
ST2_Joel29B        MLSFVLFFL-IVIALAKPKLPLPTEEQAQWLERGISGFIHFGVNTFTGKEWGLGDEDPNI 59
ST2_CMX11          MFSFVLFFL-IAIALAKPKLPLPTEEQAQWLERGISGFIHFGVNTFTGKEWGLGDEDPNI 59
ST2_Joel19B        MFSFVLFFL-IAIALAKPKLPLPTEEQAQWLERGISGFIHFGVNTFTGKEWGLGDEDPNI 59
ST2_Joel5B         MFSFVLFFL-IAIALAKPKLPLPTEEQAQWLERGISGFIHFGVNTFTGKEWGLGDEDPNI 59
ST3_Joel42         MKNLLFILLFVAIVLSKRDLPYPTEEQAQWLERGVAGFIHFGVNTYTGKEWGDGTEDPKI 60
ST3_NCBI           MKNLLFILLFVAIVLSKRDLPYPTEEQAQWLERGVAGFIHFGVNTYTGKEWGDGTEDPKI 60
ST3_CMX5           MKNLLFILLFVAIVLSKRDLPYPTEEQAQWLERGVAGFIHFGVNTYTGNEWGDGTEDPKI 60
ST3_Joel17A        MKNLLFILLFVAIVLSKRDLPYPTEEQAQWLERGVAGFIHFGVNTYTGNEWGDGTEDPKI 60
ST3_Joel43         MKNLLFILLFVAIVLSKRDLPYPTEEQAQWLERGIAGFIHFGVNTYTGKEWGDGTEDPKI 60
                   * .: :::*  .:.*:: .** *****.******::*********:**.*** * ***:*

ST1_mRNA_ALFuc     FNPVKLNTTQWVEAFQSAGANEIILVTKHHDGFMLFPSEYSNHSVISSKWRDGKGDVVRD 120
ST1_50177          FNPVKLNTTQWVEAFQSAGANEIILVTKHHDGFMLFPSEYSNHSVISSKWRDGKGDVVRD 120
ST1_50610          FNPVKLNTTQWVEAFQSAGANEIILVTKHHDGFMLFPSEYSNHSVISSKWRDGKGDVVRD 120
ST1_JOEL36         FNPVKLNTTQWVEAFQSAGANEIILVTKHHDGFMLFPSEYSNHSVISSKWRDGKGDVVRD 120
ST2_Joel40         FNPVKLNTTQWVEAFQSAGANEIILVTKHHDGFMLFPSEYSNHSVISSKWRDGKGDVVRE 119
ST2_Flemming_NCBI  FNPVKLNTTQWVEAFQSAGANEIILVTKHHDGFMLFPSEYSNHSVISSKWRDGKGDVVRE 119
ST2_Joel29B        FNPVKLNTTQWVEAFQSAGANEIILVTKHHDGFMLFPSEYSNHSVISSKWRDGKGDVVRE 119
ST2_CMX11          FNPVKLNTTQWVEAFQSAGANEIILVTKHHDGFMLFPSEYSNHSVISSKWRDGKGDVVRE 119
ST2_Joel19B        FNPVKLNTTQWVEAFQSAGANEIILVTKHHDGFMLFPSEYSNHSVISSKWRDGKGDVVRE 119
ST2_Joel5B         FNPVKLNTTQWVEAFQSAGANEIILVTKHHDGFMLFPSEYSNHSVISSKWRDGKGDVVRE 119
ST3_Joel42         FNPVKLNTTQWIEAFRSIGAEEVILVAKHHDGFMLFPSAYSNHTVAYSTWRDGKGDVVRE 120
ST3_NCBI           FNPVKLNTTQWIEAFRSIGAEEVILVAKHHDGFMLFPSAYSNHTVAYSTWRDGKGDVVRE 120
ST3_CMX5           FNPVKLNTTQWIEAFRSIGAEEVILVAKHHDGFMLFPSAYSNHTVAYSTWRDGKGDVVRE 120
ST3_Joel17A        FNPVKLNTTQWIEAFRSIGAEEVILVAKHHDGFMLFPSAYSNHTVAYSTWRDGKGDVVRE 120
ST3_Joel43         FNPVKLNTTQWIEAFRSIGAEEVILVAKHHDGFMLFPSAYSNHTVAYSTWRDGKGDVVRE 120
                   ***********:***:* **:*:***:*********** ****:*  *.**********:

ST1_mRNA_ALFuc     FVDSCRLLGALPSFYLSPWDRFFYNMTWRPEYNDYYKKTLEVLTTRYGPIYELWWDGANA 180
ST1_50177          FVDSCRLLGALPSFYLSPWDRFFYNMTWRPEYNDYYKKTLEVLTTRYGPIYELWWDGANA 180
ST1_50610          FVDSCRLLGALPSFYLSPWDRFFYNMTWRPEYNDYYKKTLEVLTTRYGPIYELWWDGANA 180
ST1_JOEL36         FVDSCRLLGALPSFYLSPWDRFFYNMTWRPEYNDYYKKTLEVLTTRYGPIYELWWDGANA 180
ST2_Joel40         FVDSCHQLGALPSFYLSPWDRYFYNMTWRPEYNDYYKKTLDVLTSRYGPIYELWWDGANA 179
ST2_Flemming_NCBI  FVDSCHQLGALPSFYLSPWDRYFYNMTWRPEYNDYYKKTLDVLTSRYGPIYELWWDGANA 179
ST2_Joel29B        FVDSCHQLGALPSFYLSPWDRYFYNMTWRPEYNDYYKKTLDVLTSRYGPIYELWWDGANA 179
ST2_CMX11          FVDSCHQLGALPSFYLSPWDRYFYNMTWRPEYNDYYKKTLDVLTSRYGPIYELWWDGANA 179
ST2_Joel19B        FVDSCHQLGALPSFYLSPWDRYFYNMTWRPEYNDYYKKTLDVLTSRYGPIYELWWDGANA 179
ST2_Joel5B         FVDSCHQLGALPSFYLSPWDRYFYNMTWRPEYNDYYKKTLDVLTSRYGPIYELWWDGANA 179
ST3_Joel42         FVQSCRALGTKASLYLSPWDRFFYNMTWRPEYNEYYSHTLEELTTRYGPIYELWWDGANA 180
ST3_NCBI           FVQSCRALGTKASLYLSPWDRFFYNMTWRPEYNEYYSHTLEELTTRYGPIYELWWDGANT 180
ST3_CMX5           FVQSCRALGTKASLYLSPWDRFFYNMTWRPEYNEYYSHTLEELTTRYGPIYELWWDGANA 180
Joel17A_ST3        FVQSCRALGTKASLYLSPWDRFFYNMTWRPEYNEYYSHTLEELTTRYGPIYELWWDGANA 180
ST3_Joel43         FVQSCRALGTKASLYLSPWDRFFYNMTWRPEYNEYYSHTLEELTTRYGPIYELWWDGANA 180
                   **:**: **:  *:*******:***********:**.:**: **:**************:

ST1_mRNA_ALFuc     QPHMTHVYDWKGWYAILKKNQPQCLGGGCGGDN-DSFDCGPDTAWGKTESGLGREENWNF 239
ST1_50177          QPHMTHVYDWKGWYAILKKNQPQCLGGGCGGDN-DSFDCGPDTAWGKTESGLGREENWNF 239
ST1_50610          QPHMTHVYDWKGWYAILKKNQPQCLGGGCGGDN-DSFDCGPDTAWGKTESGLGREENWNF 239
ST1_JOEL36         QPHMTHVYDWKGWYAILKKNQPQCLGGGCGGDN-DSFDCGPDTAWGKTESGLGREENWNF 239
ST2_Joel40         KEHMTHVYDWKGWYEILKKNQPQCLGGGCGGDN-DSFDCGPDTAWGKTESGLGKEENWNF 238
ST2_Flemming_NCBI  KEHMTHVYDWKGWYEILKKNQPQCLGGGCGGDN-DSFDCGPDTAWGKTESGLGKEENWNF 238
ST2_Joel29B        KEHMTHVYDWKGWYAILKKNQPQCLGGGCGGDN-DSFDCGPDTAWGKTESGLGKEENWNF 238
ST2_CMX11          KEHMTHVYDWKGWYEILKKNQPQCLGGGCGGDN-DSFDCGPDTAWGKTESGLGKEENWNF 238
ST2_Joel19B        KEHMTHVYDWKGWYEILKKNQPQCLGGGCGGDN-DSFDCGPDTAWGKTESGLGKEENWNF 238
ST2_Joel5B         KEHMTHVYDWKGWYAILKKNQPQCLGGGCGGDN-DSFDCGPDTAWGKTESGLGKEENWNF 238
ST3_Joel42         QQHMTHVYDWQGWYKIIKRNQPQCLGGGCGGDEAGIFDCGPDTAWGQTESGLGKEENWNF 240
ST3_NCBI           QQHMTHVYDWQGWYKIIKRNQPQCLGGGCGGDEAGIFDCGPDTAWGQTESGLGKEENWNF 240
ST3_CMX5           QQHMTHVYDWQGWYKIIKRNQPQCLGGGCGGDEAGIFDCGPDTAWGQTESGLGKEENWNF 240
ST3_Joel17A        QQHMTHVYDWQGWYKIIKRNQPQCLGGGCGGDEAGIFDCGPDTAWGQTESGLGKEENWNF 240
ST3_Joel43         QQHMTHVYDWQGWYKIIKRNQPQCLGGGCGGDEAGIFDCGPDTAWGQTESGLGKEENWNF 240
                   : ********:*** *:*:**************: . **********:******:******

ST1_mRNA_ALFuc     HAPSVEFPGEELVFDPLFLDVSIRPGWFYHPNENPKSLEDLVHIYFRSVGLNYQLQLNVP 299
ST1_50177          HAPSVEFPGEELVFDPLFLDVSIRPGWFYHPNENPKSLEDLVHIYFRSVGLNYQLQLNVP 299
ST1_50610          HAPSVEFPGEELVFDPLFLDVSIRPGWFYHPNENPKSLEDLVHIYFRSVGLNYQLQLNVP 299
ST1_JOEL36         HAPSVEFPGEELVFDPLFLDVSIRPGWFYHPNEEPKSLEDLVHIYFRSVGLNYQLQLNVP 299
ST2_Joel40         HAPSVEFPGEDLVFDPLFLDVSIRPGWFYHANEEPKSLKELVHIYFRSVGLNYQLQLNVP 298
ST2_Flemming_NCBI  HAPSVEFPGEDLVFDPLFLDVSIRPGWFYHANEEPKSLKELVHIYFRSVGLNYQLQLNVP 298
ST2_Joel29B        HAPSVEFPGEDLVFDPLFLDVSIRPGWFYHPNEEPKSLKDLVHIYFRSVGLNYQLQLNVP 298
ST2_CMX11          HAPSVEFPGEDLVFDPLFLDVSIRPGWFYHANEEPKSLKELVHIYFRSVGLNYQLQLNVP 298
ST2_Joel19B        HAPSVEFPGEDLVFDPLFLDVSIRPGWFYHANEEPKSLKELVHIYFRSVGLNYQLQLNVP 298
ST2_Joel5B         HAPSVEFPGEDLVFDPLFLDVSIRPGWFYHPNEEPKSLKELVHIYFRSVGLNYQLQLNVP 298
ST3_Joel42         HTPSVEFPGKELVFSPLFLDVSIRPGWFYHANESPKTLKELVHIYFRSVGLNYQLQLNVP 300
ST3_NCBI           HTPSVEFPGKELVFSPLFLDVSIRPGWFYHANESPKTLKELVHIYFRSVGLNYQLQLNVP 300
ST3_CMX5           HTPSVEFPGKELVFSPLFLDVSIRPGWFYHANESPKTLKELVHIYFRSVGLNYQLQLNVP 300
ST3_Joel17A        HTPSVEFPGKELVFSPLFLDVSIRPGWFYHANESPKTLKELVHIYFRSVGLNYQLQLNVP 300
ST3_Joel43         HTPSVEFPGKELVFSPLFLDVSIRPGWFYHANESPKTLKELVHIYFRSVGLNYQLQLNVP 300
                   *:********::***.*************** **.**:*::********************
```

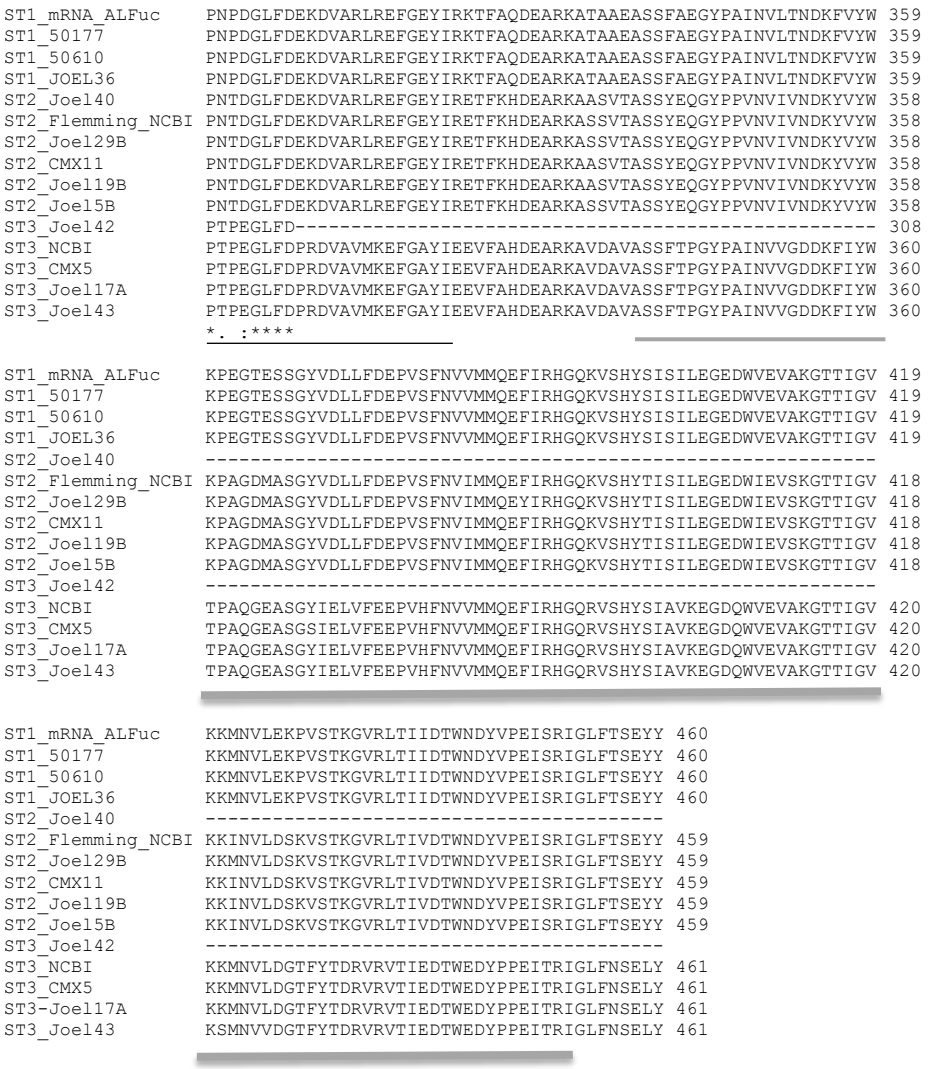


Figura 12. Alineamiento de secuencias ALFuc de *Blastocystis*, usando ClustalW. Se incluyeron 10 secuencias obtenidas a partir de la amplificación y secuencias del gen ALFuc completo, además de las secuencias obtenidas de la misma manera de las cepas ST1 ATCC 50177 y 50610. También, se incluyeron una secuencia de ARNm ST1, una secuencia de Flemming ST2 y una secuencia ST3, estas últimas 3 secuencias tomadas de la base de datos NCBI. Los residuos conservados y similares están marcados con asteriscos y puntos, respectivamente. La región del péptido señal está marcada con una barra negra, el dominio α-L-fucosidasa de la familia GH29 está marcado con una línea negra, el dominio de unión a carbohidratos tipo C F5/8 está indicado con una barra gris. El sitio catalítico con el residuo nucleófilo asp (D) en rojo y el residuo general ácido/base Glu (E) en azul. Aunque las secuencias ST2_Joel40, ST3_Joel42 y ST3_joel 43 fueron secuenciadas en un 78%, 67% y 67%, respectivamente, se observa que contienen el dominio de las α-L-fucosidasas de la familia GH29

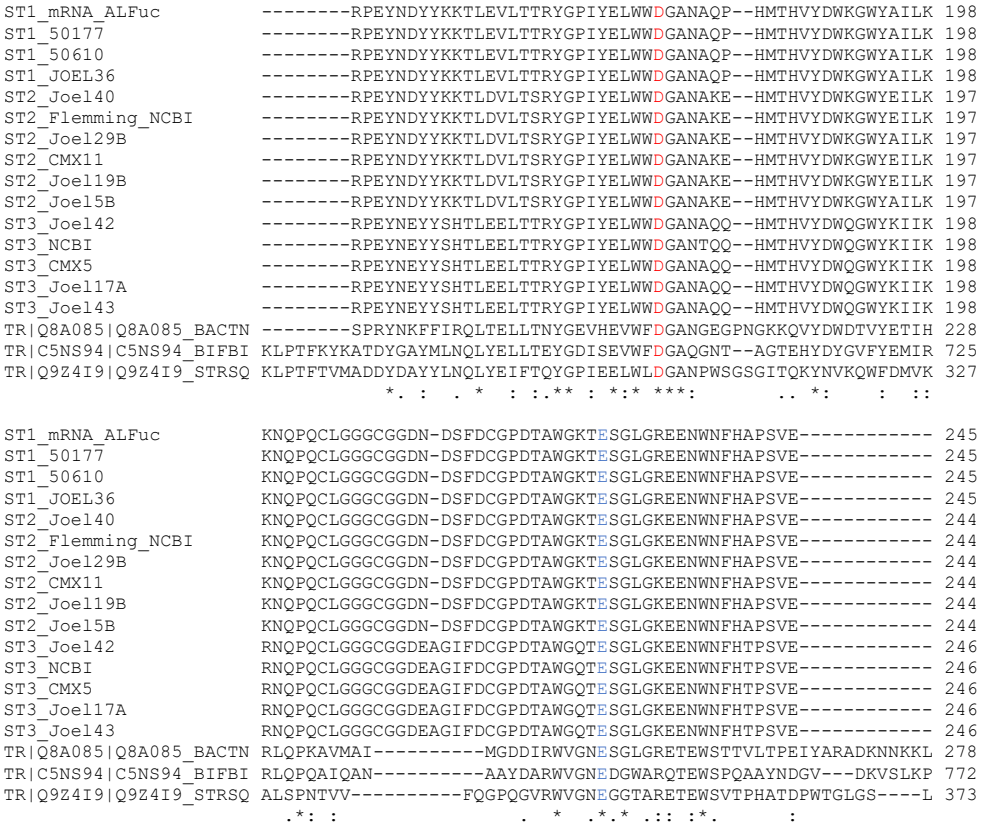


Figura 13. Alineamiento usando ClustalW. Sólo se muestra la región del sitio catalítico. Se observa la conseervación de los sitios catalíticos de la familia GH29 de las α-L-fucosidasas. Nuestras secuencias se alinearon con tres α-L-fucosidasas de la familia GH29. Estas tres secuencias se tomaron de la base de datos UnitPro: *Bacteroides thetaiotaomicron* Q8A085, *Bifidobacterium bifidum* C5NS94 y *Streptomyces* sp Q9Z4I9. El sitio catalítico se muestra con el residuo nucleófilo asp (D) en rojo y el residuo general ácido/base Glu (E) en azul

4.- Predicción de Estructura terciara de ALFuc

Las estructuras diseñadas fueron analizadas por el servidor I-TASSER. Este servidor genera algunas conformaciones estructurales y, a continuación, utiliza el programa SPICKER para agrupar todas las estructuras en función de la similitud de la estructura emparejada. Finalmente, para *Blastocystis* ALFuc ST1, ST2 y ST3 el servidor reportó cinco modelos principales, para cada uno de ellos, correspondientes a los cinco grupos más grandes. La seguridad de cada modelo se calculó mediante el C-score. Los valores del C-score muestran la precisión del modelo predicho que suele estar en el rango de -5 a 2. Además, cuanto mayor sea el valor del C-score, mayor será la calidad de la predicción. Las puntuaciones C de los modelos *Blastocystis* ALFuc ST1, ST2 y ST3 fueron de 1,22, 0,93 y 0,83, respectivamente. Por lo tanto, el ALFuc *Blastocystis* ST1 con la puntuación C de 1,22 mostró una mayor precisión entre los modelos predichos. La estructura de ALFuc ST1 está constituida por 7 arreglos α-hélice, de los cuales 5 forman parte del dominio GH29 α-L-fucosidasa y 6

estructuras β-plegada, 5 forman parte del dominio similar de unión a galactosa. La ALFuc ST2 está conformada de la misma manera que ST1 y ST3 posee 8 estructuras β-plegada, sin dominio similar de unión a galactosa. Además, los análisis de predicción de ligando arrojados por COFACTOR y COACH dieron como resultado los ligandos ZXD y DFU para las ALFuc ST1, ST2 y ST3, respectivamente, con un s-score de 0.89, 0.87 y 0.82 8 (tabla 5).

En la figura 14 se observan las estructuras terciarias del ALFuc de *Blastocystis* ST1, ST2 y ST3. En la imagen se puede ver el dominio de la α-L-fucosidasa, formado por cinco estructuras de α-hélice (color magenta); el dominio similar de unión a la galactosa, formado por cinco estructuras de β-cadena (color naranja). En ST3 el resultado de Pfam para este dominio fue negativo. Asimismo, se observa el sitio catalítico de la familia GH29, formado por el residuo nucleófilo D (esferas rojas) y el residuo ácido/base E (esferas azules)

Tabla 5. Características de estructuras terciarias de ALFuc de *Blastocystis* STs 1-3

ALFuc *Blastocystis*	**C-score (-5 a 2)**	**Estructuras α-hélice**		**Estructuras β-plegada**		**Ligando C-score (0-1)**
		Total	**Dominio GH-29 α-L-fucosidasa.**	**total**	**Dominio Similar de unión a galactosa**	
ST1	1.22	7	5	6	5	ZXD C-score= 0.89
ST2	0.93	7	5	6	5	ZXD C-score= 0.87
ST3	0.83	7	5	8	-	DFU C-score= 0.82

ZXD: (2S,3R,4S,5R)-2-(1-Metiletil) piperidina-3,4,5-Triol

DFU: (2S,3R,4S,5R)-2-Metlpiperidina-3,4,5-Triol

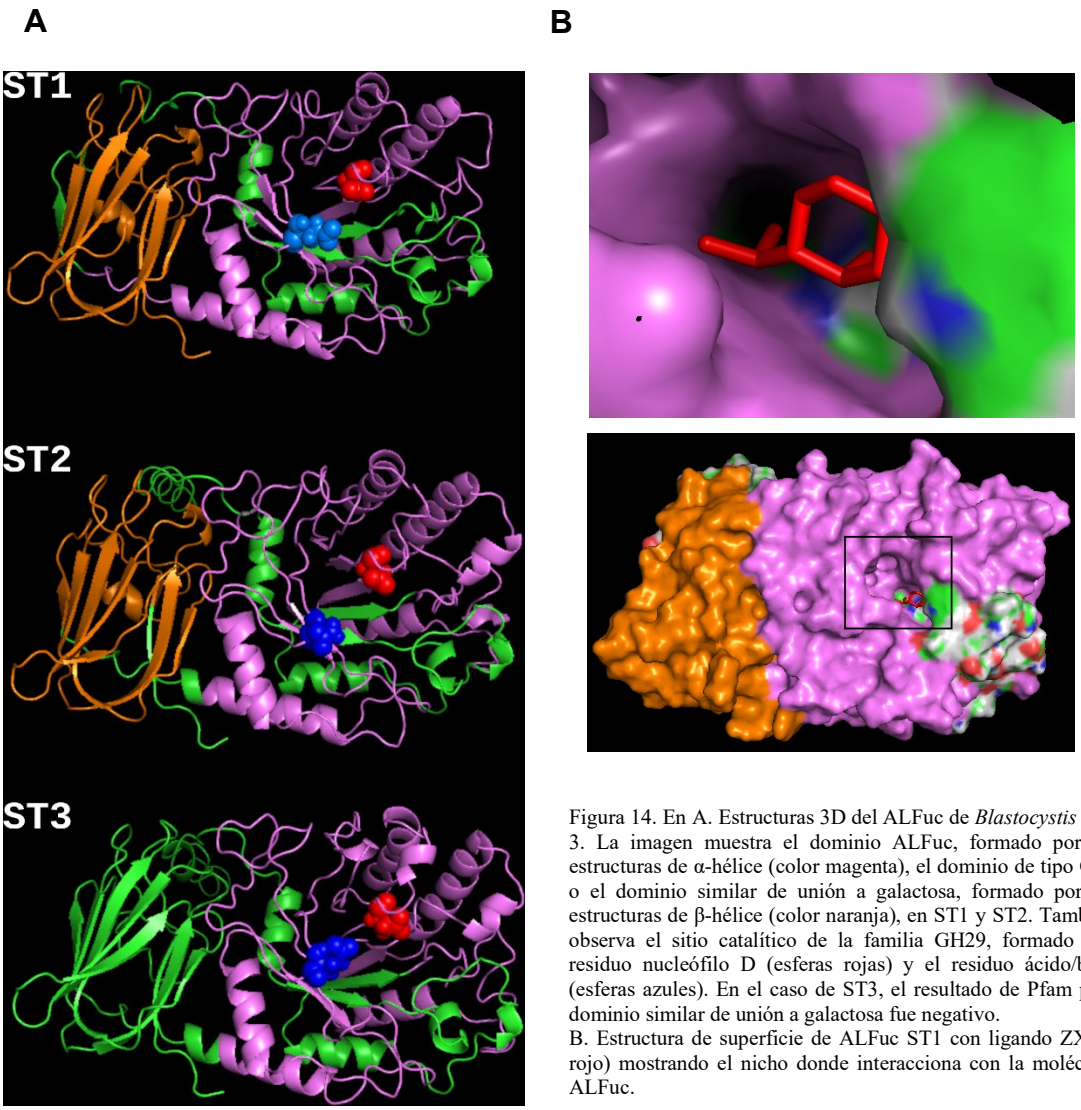


Figura 14. En A. Estructuras 3D del ALFuc de *Blastocystis* STs 1-3. La imagen muestra el dominio ALFuc, formado por cinco estructuras de α-hélice (color magenta), el dominio de tipo C F5/8 o el dominio similar de unión a galactosa, formado por cinco estructuras de β-hélice (color naranja), en ST1 y ST2. También se observa el sitio catalítico de la familia GH29, formado por el residuo nucleófilo D (esferas rojas) y el residuo ácido/base E (esferas azules). En el caso de ST3, el resultado de Pfam para el dominio similar de unión a galactosa fue negativo.
B. Estructura de superficie de ALFuc ST1 con ligando ZXD (en rojo) mostrando el nicho donde interacciona con la molécula de ALFuc.

5.- Análisis filogenético

Los ST identificados por PCR-secuenciación en los 18 aislamientos de *Blastocystis* fueron ST1 (3/18), ST2 (7/18) y ST3 (8/18) (figura 15). Estos ST fueron confirmados mediante la comparación de sus secuencias parciales de *SSUrDNA* con un conjunto de datos de secuencia de subtipos validados (Figura 16). Además, se exploraron las relaciones filogenéticas de *Blastocystis* para determinar si la ALFuc podía inferir la topología característica de *SSUrDNA* o la descrita con los genes *nad* de los genomas de orgánulos similares a las mitocondrias (MLO) de *Blastocystis*(114). Por lo tanto, se realizó análisis filogenético utilizando el gene *alfuc* completo de *Blastocystis* ST1-ST3 y ST4-ST9. Nuestros resultados confirmaron una identidad conservada en la topología del árbol entre *alfuc* y *SSUrDNA* en el análisis filogenético. La topología del árbol genético se asemeja a las descritas para

Blastocystis por los genes *SSUrDNA* (115) o MLO nad (114) y la proteína Miro (116). La topología consistía en ST1 y ST2 agrupados; ST3 agrupado o asociado a ST4 y ST8; y ST7 agrupados o asociados con ST6, y ST9 (Figura 17).

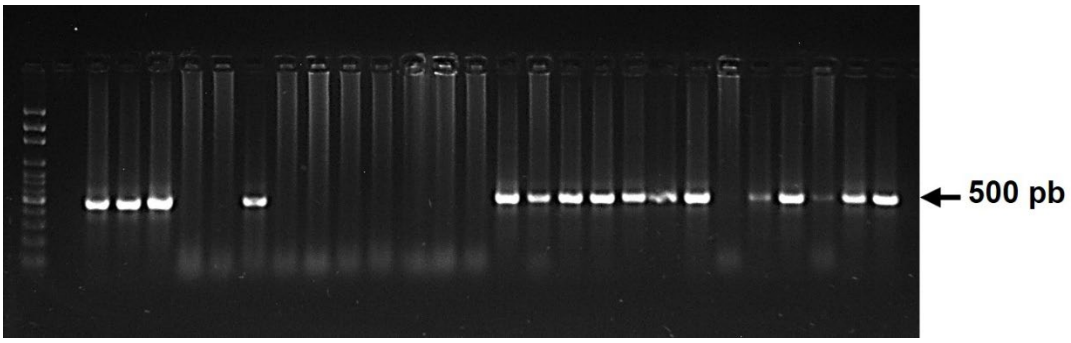


Figura 15. Gel de teñido con bromuro de etidio que muestra el fragmento específico de la *SSUrDNA* de *Blastocystis*. La flecha indica el tamaño del amplicon esperado, según los iniciadores reportados por Santin et al.

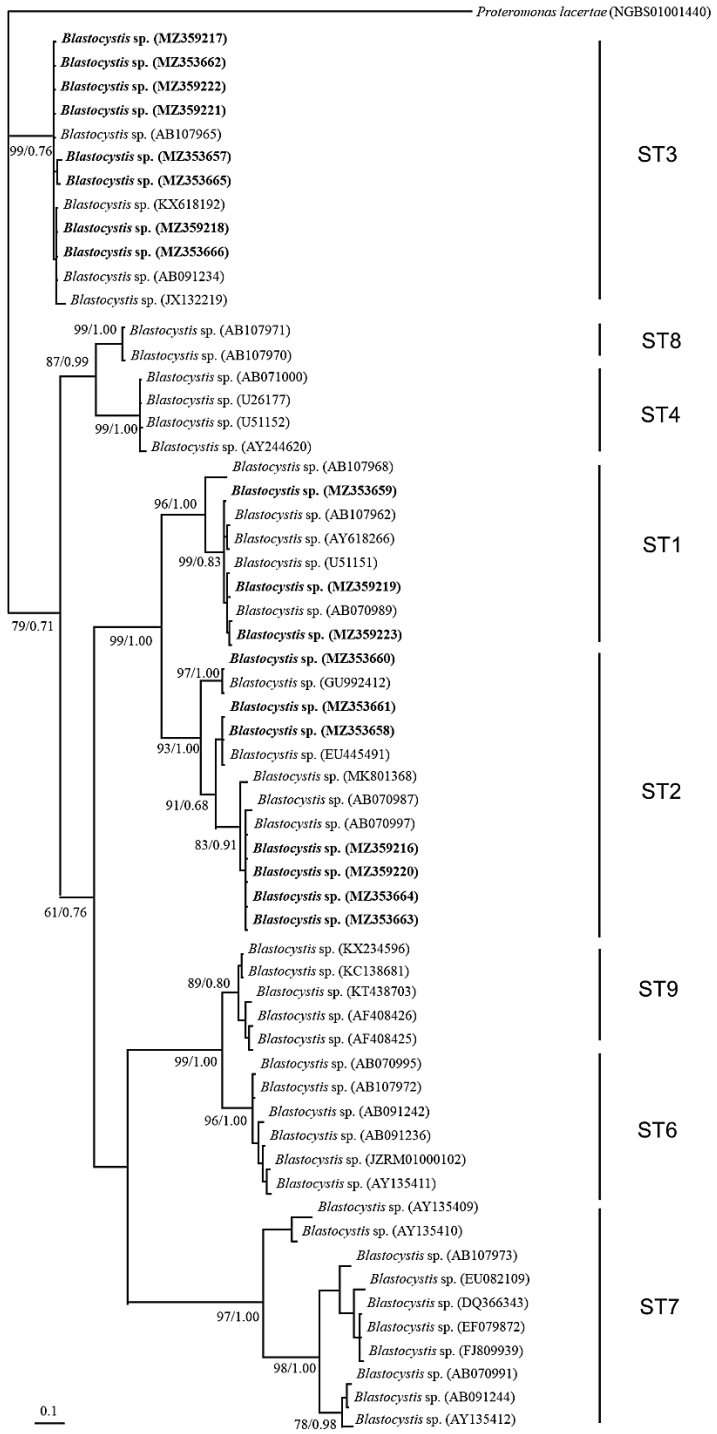


Figura 16. Inferencia filogenética de la secuencia parcial del gen *SSUrDNA* de *Blastocystis* STs 1-4 y STs 6-9. La región ortóloga de *Proteromonas lacertae* se utilizó como grupo externo. Los valores de los nodos indican las proporciones bootstrap y las probabilidades posteriores bayesianas en el siguiente orden: máxima verosimilitud/análisis bayesiano. Las secuencias obtenidas en la presente tesis se indican en negrita

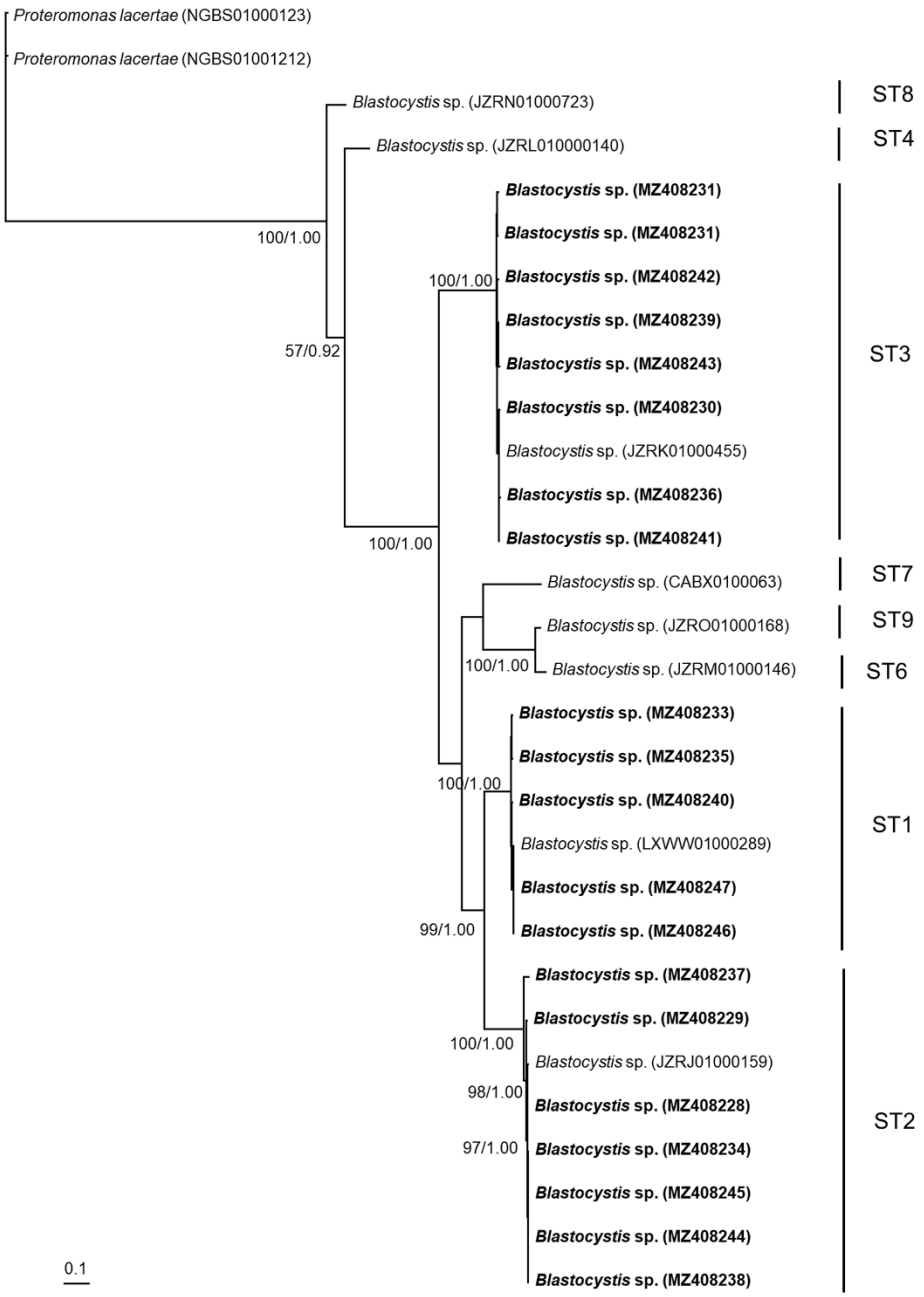


Figura 17. Inferencia filogenética del gene *alfuc* completo de *Blastocystis* STs 1-4 y STs 6-9. La región ortóloga de *Proteromonas lacertae* se utilizó como grupo externo. Los valores de los nodos indican las proporciones *bootstrap* y las probabilidades posteriores bayesianas en el siguiente orden: máxima verosimilitud/análisis bayesiano. Las secuencias obtenidas en la presente tesis se indican en negritas.

VIII.- Discusión

En la presente tesis, identificamos por primera vez la ALFuc de *Blastocystis* en medio de cultivo de parásitos, calculando su Rf y su masa molecular aproximada de 51.7 kDa en electroforesis. Este resultado concuerda con el análisis de predicción para la secuencia de ALFuc reportada para *Blastocystis* ST7 disponible en la base de datos UniProt con el identificador D8M3D3. Así mismo, esta misma proteína fue reconocida en *Western blot* por el anticuerpo comercial anti-FUCA2 dirigido contra la región interna conservada de la α-L-fucosidasa en su forma secretora; además, se muestran que la α-L-fucosidasa forma parte de los productos de secreción de *Blastocystis* en cultivo. Nuestros resultados son consistentes con informes en publicaciones anteriores que predicen la ALFuc como una proteína secretada (67, 90) y relacionada con el metabolismo de los carbohidratos (90).

El tracto gastrointestinal representa un nicho microecológico en el que la microbiota comensal y patógena interactúan de manera cooperativa-competitiva, delimitada por los nutrientes disponibles (117). Los glicanos de mucina derivados del huésped proporcionan una fuente de nutrientes para la microbiota comensal cerca de la capa de moco y regulan la infección por microbiota patógena (118). Algunas especies bacterianas comensales, como *Bacteroides thetaiotaomicron*, un glicófilo conocido, que secreta α-L-fucosidasas, pueden usar una amplia gama de glicanos dietéticos; por el contrario, las bacterias comensales o patógenas que carecen de enzimas con actividad glucósido hidrolasa, pueden verse favorecidas por el microambiente provocado por *B. thetaiotaomicron* para eliminar los subproductos de los glicanos dietéticos para establecer la colonización o invasión (119). Por ejemplo, la colonización de *E. coli* enterohemorrágica en el intestino es impulsada por moco rico en fucosa inducido por *B. thetaiotaomicron*. (120). Aunque ALFuc de *Blastocystis* y la de *B. thetaiotaomicron* comparten una estrecha homología, el metabolismo de la fucosa en *Blastocystis* es similar al patógeno *Campylobacter jejuni* (90). Investigaciones recientes basadas en estudios in vivo e in vitro apoyan ambas funciones biológicas; es decir, se ha informado que la colonización a largo plazo de *Blastocystis* ST3 induce una recuperación más rápida de la colitis en modelos animales mediante la estabilización del ecosistema intestinal (121). Por lo tanto, *Blastocystis* ST3 puede desempeñar un papel comensal en la microbiota intestinal. Por el contrario, *Blastocystis* ST7 induce la interrupción de la barrera epitelial por una actividad de proteasa de cisteína, lo que apoya la noción de que este ST actúa como un patógeno. La investigación futura sobre la función biológica de ALFuc podría

abordar si *Blastocystis* puede inducir una capa de moco rica en fucosa de manera comensal, similar a *B. thetaiotaomicron*, o requerir glicanos fucosilados para la adherencia e invasión como *C. jejuni* (84)

En esta tesis amplificamos el gen *alfuc* de *Blastocystis* e identificamos por primera vez tres ALFuc correspondientes a *Blastocystis* ST1, ST2 y ST3 que son los tres subtipos más frecuentes en humanos (14). El análisis de la secuencia de las tres ALFuc, realizado por SignalP-5 identificó el péptido señal en el extremo N-terminal con una probabilidad de escisión 0.6-0.8 %, lo que concuerda con el rango reportado para α-L-fucosidasas de otros protozoarios (122). En *Blastocystis* se había reportado el gen la α-L-fucosidasa, como uno de los genes transferidos de manera lateral y a su proteína como una predicción de secreción del parásito (90).

El mecanismo catalítico de la α-L-fucosidasa requiere la acción de dos residuos de aminoácidos, uno de los cuales desempeña el papel de ácido/base general y el otro actúa como nucleófilo (123). Para analizar la conservación del sitio catalítico de las α-L-fucosidasas de la familia GH29, realizamos alineamiento de nuestras secuencias de ALFuc de *Blastocystis* ST1, ST2 y ST3, con tres alfa-L-fucosidasas de la familia GH29 de la base de datos UniProtKB: *Bacteroides thetaiotaomicron* Q8A085, *Bifidobacterium bifidum* C5NS94, *Streptomyces s*p Q9Z4I9. Nuestros resultados mostraron que el sitio catalítico de las ALFuc de *Blastocystis* se conserva con el residuo nucleofílico Asp-176, Asp-175 y Asp-176, para ST1, ST2 y ST3 respectivamente; que corresponde a Asp-204 *Bacteroides thetaiotaomicron*, Asp-703 *Bifidobacterium bifidum* y Asp-303 *Streptomyces sp*. Asimismo, observamos la conservación del residuo ácido/base Glu-228, Glu-227 y Glu-229 para ST1, ST2 y ST3 respectivamente, y corresponde a Glu-249 *Bacteroides thetaiotaomicron*, Glu-746 *Bifidobacterium bifidum* y Glu-348 *Streptomyces sp*.

Además del dominio GH29 de ALFuc de *Blastocystis*, en ST1 y ST2 se identificó un dominio tipo C F5/8 del clan CL0202 (CAZy) o también llamado de discoidina. Esta gran superfamilia contiene dominios beta de unión a galactosa. Muchas de estas familias están involucradas en el reconocimiento de carbohidratos. Se ha informado que la mutante recombinante en el dominio tipo C F5/8, de *Bacillus circulans*, produce una *B-galactosidasa* con un dominio tipo C F5/8 con una menor actividad hidrolítica en comparación con la *B-galactosidasa* nativa, lo que sugiere que este motivo ayuda a las moléculas de galactosilactosa a orientarse adecuadamente dentro del sitio activo para que puedan hidrolizarse eficientemente para producir galactosa/glucosa inhibiendo la acumulación de galactooligosacáridos (GOS)

(124,125). Curiosamente, en nuestro estudio, el dominio F5/8 en AlFuc ST3 mostró una baja identidad para este dominio, posiblemente asociado a una menor actividad hidrolítica en comparación con ALFuc ST1 y ST2, aunque esto tendrá que ser evaluado en estudios posteriores con proteínas recombinantes. Como se ha mencionado, las ALFuc bacterianas son los homólogos más cercanos a las ALFuc de Blastocystis. En el microorganismo intestinal Akkermansia muciniphila, las proteínas con el dominio similar a la unión a la galactosa involucradas en la degradación de la mucina, desempeñan un papel importante en la salud intestinal del huésped, ya que regulan la actividad hidrolítica en regiones con tejidos heridos, mejorando la regeneración tisular y la cicatrización de heridas (127). En Ruminococcus gnavus, otro simbionte intestinal humano comúnmente asociado con enfermedades inflamatorias intestinales, la capacidad de crecer en mucina como única fuente de carbono depende de la cepa y está relacionada con grupos de genes, incluidas las proteínas ALFuc con la unión a la galactosa (128).

Blastocystis tiene una amplia diversidad genética, y hasta 17 variantes han sido descritos como distintas especies o incluso géneros (91). Además, algunos ST pueden colonizar múltiples huéspedes (126). Hasta el momento, se ha demostrado que todos los ST pueden infectar prácticamente a un amplio rango de huéspedes, desde insectos hasta humanos. Por lo tanto, no es sorprendente que *Blastocystis* se considere el microeucariota intestinal más extendido, con una falta casi completa de especificidad del huésped. Debido a su peso potencial en la salud humana, se han realizado esfuerzos significativos para asociar la diversidad genética de *Blastocystis* con manifestaciones clínicas gastrointestinales o firmas para la transmisión zoonótica.

Para la mayoría de los parásitos, se han desarrollado métodos de diagnóstico directos e indirectos. Los métodos directos incluyen aquellos basados en la morfología (microscopía) y la detección de DNA (típicamente PCR) o antígenos (IFA, antígeno ELISA, etc.), mientras que los métodos indirectos se basan principalmente en la detección de anticuerpos (129). A pesar de ser la principal herramienta de diagnóstico a nivel mundial, el uso de la microscopía para detectar *Blastocystis* tiene una utilidad limitada en los laboratorios de microbiología clínica y en la generación de datos con fines clínicos y epidemiológicos. Por estas razones y debido a la extensa diversidad genética críptica de *Blastocystis* (130), se han desarrollado una serie de estrategias para establecer los subtipos de *Blastocystis*. Actualmente, algunos autores sugieren un nivel de divergencia del 5% en la secuencia del gen que se esté analizando para definir un nuevo subtipo debido a que la variación dentro de los subtipos también puede

ser sustancial, hasta al menos el 3% (131). Por lo tanto, una sola secuencia "atípica" que parece ser distinta y potencialmente un nuevo subtipo podría eventualmente fusionarse en un subtipo adyacente a medida que haya más secuencias disponibles. Solo a medida que se acumulen más datos de subtipificación se probará la validez de este umbral arbitrario. La divergencia entre algunos subtipos existentes (por ejemplo, ST6 y ST9) es en realidad inferior al 5%. Sin embargo, cuando un muestreo es suficiente, existe la confianza de que se trate de linajes distintos en lugar de variantes del mismo subtipo. Se ha elegido una divergencia del 5% como un criterio bastante estricto y más datos pueden conducir a la revisión de nuevas definiciones de subtipos en el futuro. En humanos se han encontrado nueve subtipos distintos (ST1-ST9). Sin embargo, el 95% de las infecciones humanas muestreadas pertenecen a los ST1-ST4 (125) y solo el ST9 (hasta ahora), ha sido el único encontrado en humanos. Los cuatro ST más comunes en humanos también se han detectado en otros huéspedes. Con mayor frecuencia, estos huéspedes son otros primates, pero también se han encontrado en varios mamíferos con pezuñas, roedores e incluso aves (133). Por el contrario, los subtipos más raros en humanos (ST 5-8) se encuentran más comúnmente en otros huéspedes: ST5 en animales con pezuñas, ST 6 y 7 en aves y ST8 en primates no humanos. Se ha sugerido que estos subtipos más raros en humanos son de origen zoonótico y hay alguna evidencia que lo respalda. El ST8 se ha encontrado con frecuencia en cuidadores de zoológicos que trabajan con primates no humanos, y ST5 es frecuente en trabajadores de cerdos en Australia (134). Sin embargo, no hay razón para sospechar que las infecciones humanas que involucran a los ST comunes (STs 1-4) se originan en fuentes no humanas, excepto en casos raros (135). La exposición a animales infectados por *Blastocystis* por sí sola no es suficiente para provocar una infección. Por ejemplo, ST10 es muy común en el ganado, pero aún no se ha informado en humanos (136). Esto sugiere que otras variables además de la temperatura corporal están determinando la capacidad de *Blastocystis* para colonizar el intestino humano; la flora intestinal puede tener un impacto, por ejemplo. El grado de diversidad genética dentro de los subtipos es bastante variable. ST3 es probablemente el más diverso de los subtipos bien estudiados, que varía en aproximadamente un 3% en las secuencias *SSUrDNA*, mientras que ST4 muestra la menor variación, especialmente en humanos (136).

Para intentar mapear la epidemiología molecular de *Blastocystis* se han desarrollado varias herramientas. A principios de la década de 1990, un ensayo de PCR fue desarrollado y refinado para detectar subtipos utilizando iniciadores de sitio específico por secuencia (SES) (137). Esta estrategia de identificación implica el uso de siete reacciones de PCR, una

para cada uno de los subtipos 1-7, evitando la necesidad de secuenciación. Otra metodología descrita por Santin y col, consiste en un protocolo de PCR y secuenciación de un fragmento del gen *SSUrDNA* de ~500 pb amplificado por los iniciadores Blast 505-532 y Blast 998-1017. Esta estrategia de PCR-secuenciación es más sensible que SES y la región amplificada y secuenciada de *SSUrDNA* es conservada y suficientemente variable para permitir el análisis filogenético de *Blastocystis* (136).

Un método de código de barras de DNA (conocido en inglés como *barcoding*), desarrollado en 2006 por Scicluna et al. utiliza los iniciadores RD5 y BhRDr, que amplifican una región ~600 pb en el extremo 5 ′ del gen *SSUrRNA*, que posteriormente debe ser secuenciada para analizarse (130). El *barcoding* permite la detección de subtipos raros, más allá de los ST 1-7 y un mayor escrutinio de la diversidad genética y ha sido validado como un marcador de la diversidad genética general de *Blastocystis* (137). La comparación de los árboles filogenéticos obtenidos mediante *barcoding* con los obtenidos utilizando secuencias concatenadas obtenidas por MLST (loci que reflejan en el genoma del orgánulo similar a la mitocondria), han mostrado la conveniencia de utilizar la región *barcoding* como un marcador sustituto para la diversidad general del genoma en este organismo en particular. Los inconvenientes del código de barras en comparación con el método SES son que se requiere secuenciación y que las infecciones de subtipos mixtos pueden no ser siempre evidentes en los cromatogramas de secuencia y, aunque lo sean, pueden resultar difíciles de descifrar (139).

Los datos de *barcoding* de *Blastocystis* están disponibles una base de datos pública (http://pubmlst.org/blastocystis/), que incluye un repositorio de secuencias de *barcoding*, así como secuencias obtenidas por MLST. También tiene una instalación BLAST, donde se pueden cargar y analizar archivos fasta individuales o masivos para una rápida identificación del subtipo y el número de alelos, eliminando así la necesidad de análisis filogenéticos. Hasta la fecha, se han identificado 35 alelos de *SSUrDNA* de dentro de ST3, mientras que el número de alelos de *SSUrDNA* para ST4 y algunos otros subtipos sigue siendo mucho más limitado. Sin embargo, parte de la variación alélica incluida es el resultado de la secuenciación del ADN clonado; se ha reportado polimorfismo intragenómico *SSUrDNA* (130), y dicho polimorfismo probablemente pasará desapercibido cuando se estudien secuencias obtenidas directamente de productos de PCR.

Otro hallazgo destacable en esta tesis fue que la topología del árbol filogenético del gen alfuc, conserva la misma topología descrita por las subunidades del gen nad para los mismos STs (114), y la agrupación para los STs 1-4 y STs 6-9, con la topología clásica obtenida con los genes SSUrDNA (115). A diferencia de los genes arriba mencionados, el gen *alfuc* es uno de los 167 genes adquiridos por transferencia lateral (GATL) en el linaje *Blastocystis* desde su divergencia de otros stramenopiles y alveolata, estos genes desempeñaron un papel importante durante la adaptación intestinal, agregando nuevas funciones para competir y sobrevivir a condiciones adversas (90). El uso de *alfuc* como marcador para analizar aislados de *Blastocystis* puede revelar información sobre cómo este microorganismo tiene éxito o fracaso en la colonización de nichos tan diferentes como el tracto intestinal de grupos taxonómicos distantes (insectos vs. mamíferos) con diferentes requerimientos nutricionales (herbívoros vs. carnívoros) o incluso diferentes estados intestinales (sanos vs. enfermedad).

En este estudio, una región genómica que contiene el gen *alfuc* se amplificó utilizando cuatro pares de cebadores degenerados diseñados para obtener todo el gen para el análisis funcional comparativo en lugar de identificar un marcador filogenético. Por esta razón, se deben señalar las principales limitaciones y ventajas para evitar posibles sesgos o para ser abordadas por estudios futuros.

Los iniciadores fueron diseñados para amplificar las regiones de interés solo para ST1-ST3 debido a que la mayoría de las muestras analizadas en estudios mundiales son de origen humano, y más del 90% de las cepas humanas pertenecen a STs 1-4; por lo tanto, estos cebadores pueden ser ventajosos si el estudio involucra ST1-ST3 o ST filogenéticamente cercanos. Sin embargo, pueden surgir problemas cuando se analizan los ST más divergentes, especialmente los que colonizan anfibios, reptiles o insectos. A diferencia de estudios anteriores que utilizaron DNA extraído de muestras de heces para evaluar nuevos marcadores filogenéticos, en el presente estudio utilizamos DNA extraído de cultivos xénicos. El primer enfoque representa la limitación más desafiante para los estudios de campo que utilizan marcadores de copia única, ya que son más susceptibles al sesgo asociado con la baja carga de parásitos en la muestra. Por ejemplo, un informe reciente que analizó los DNA de muestras de heces humanas descartó 8 de los 12 nuevos marcadores evaluados para *Blastocystis* porque no fue posible obtener productos adecuados de amplificación por PCR (140). Aunque el aislamiento axénico aumenta el número de parásitos y, en consecuencia, el número de muestras positivas, este procedimiento favorece la selección de subtipos específicos (141);

un problema particular para los estudios que desean describir la variabilidad genética en una población particular. Para resolver este problema, nuevos estudios podrían implementarse con base en técnicas de secuenciación de próxima generación, utilizados para estudiar la diversidad genética de los subtipos de *Blastocystis* en humanos (44) y animales (142) y dirigirse a genes GATL como el *alfuc*; o implementar nuevas tecnologías disponibles para obtener secuencias de genes completas como ya se ha implementado para los genes *SSUrDNA* (143).

No hay duda de que los métodos basados en el ADN ahora nos permiten llevar a cabo estudios de investigación grandes y bien diseñados que dependen de la detección precisa y la caracterización molecular de *Blastocystis*. Dichos estudios son necesarios para producir datos que puedan arrojar luz sobre el papel de este organismo en la salud y la enfermedad humanas con miras a desarrollar potencialmente diagnósticos, biomarcadores y terapias, incluidos agentes antimicrobianos o probióticos, según corresponda.

IX.- Conclusiones

- La α-L-fucosidasa es un producto de secreción de *Blastocystis* y tiene una masa molecular en electroforesis de ~51.7 kDa
- Las ALFuc de *Blastocysis* ST1, ST2 y ST3 con ORF1, tienen conservado el péptido señal, con una probabilidad de escisión del péptido de 0.6-0.8 %. Tienen conservado el dominio IPR000933 de la familia GH29, de clan CL0058 de las α-L-fucosidasas. El dominio tipo C 5/8 (IPR000421) de unión a carbohidratos se encuentra conservado en ST1 y ST2. Ninguna de las proteínas tuvo predicción de segmentos transmembranales.
- Las ALfuc de *Blastocystis* ST1 y ST8 tienen variantes con ORF 2, una proteína de menor tamaño, sin péptido señal, pero conservan los dos dominios arriba mencionados.
- Las ALFuc de *Blastocystis* ST1-ST3 tienen conservado el sitio catalítico para la familia GH29, con el residuo nucleófilo asp (D) y el residuo general ácido/base Glu (E)
- El gen *alfuc* es un candidato prometedor como marcador filogenético, ya que muestra una topología idéntica al árbol filogenético del gen *SSUrDNA*.
- Finalmente, los genes adquiridos por GATL deben analizarse como marcadores filogenéticos potenciales porque han desempeñado un papel esencial en los eventos de adaptación en el intestino y pueden revelar brechas biológicas relacionadas con eventos de especiación del huésped o expansión del huésped.

X. Perspectivas

- Debido a que la cantidad de *Blastocystis* y su axenización es una limitante para estudiar proteínas, se propone la clonación del gen de *alfuc* en un vector de expresión y producir las proteínas recombinantes de cada ST

- Determinar la actividad de α-L-fucosidasa de las proteínas recombinantes, usando substratos conocidos fucosilados, para analizar si existen diferencias en actividad inter o intra subtipos de *Blastocystis*

- Identificar el tipo de α- enlace sobre cuál actúan cada una de ellas

- Realizar análisis de estructura para identificar sitios catalíticos.

- Extender el análisis filogenético del gen *alfuc* con otros subtipos de *Blastocystis*, para verificar si se mantiene la topología con el gen *SSUrDNA* y otro genes usados en variabilidad, para robustecer la propuesta de usar *alfuc* como marcador de variabilidad en *Blastocystis*

XI.- Referencias bibliográficas

1. Zhang FQ, Wang P, Feng X, Mi QM, Mei XF, Zhang ZC, et al. [Progress of researches on global prevalence of *Blastocystis* hominis human infections and its subtypes]. Zhongguo Xue Xi Chong Bing Fang Zhi Za Zhi. 2020 Jun 30;33(1):84–94.

2. Ascuña-Durand K, Salazar-Sánchez RS, Castillo-Neyra R, Ballón-Echegaray J. Relative frequency of *Blastocystis* subtypes 1, 2, and 3 in urban and periurban human populations of arequipa, peru. Trop Med Infect Dis. 2020 Nov 27;5(4).

3. Khaled S, Gantois N, Ly AT, Senghor S, Even G, Dautel E, et al. Prevalence and subtype distribution of *Blastocystis* sp. in senegalese school children. Microorganisms. 2020 Sep 12;8(9).

4. Tan KSW. New insights on classification, identification, and clinical relevance of *Blastocystis* spp. Clin Microbiol Rev. 2008 Oct;21(4):639–65.

5. Fedianina LV, Prodeus TV, Volkova IV. [*Blastocystis* hominis, Brumpt 1912 is a mysterious organism belonging to intestinal protozoa]. Med Parazitol (Mosk). 2011 Jun;(2):47–52.

6. Mehlhorn H. *Blastocystis* hominis, Brumpt 1912: are there different stages or species? Parasitol Res. 1988;74(4):393–5.

7. Tan KSW, Singh M, Yap EH. Recent advances in *Blastocystis* hominis research: hot spots in terra incognita. Int J Parasitol. 2002 Jun 15;32(7):789–804.

8. Silberman JD, Sogin ML, Leipe DD, Clark CG. Human parasite finds taxonomic home. Nature. 1996 Apr 4;380(6573):398.

9. Arisue N, Hashimoto T, Yoshikawa H, Nakamura Y, Nakamura G, Nakamura F, et al. Phylogenetic position of *Blastocystis* hominis and of stramenopiles inferred from multiple molecular sequence data. J Eukaryot Microbiol. 2002 Feb;49(1):42–53.

10. Hoevers JD, Snowden KF. Analysis of the ITS region and partial ssu and lsu rRNA genes of *Blastocystis* and Proteromonas lacertae. Parasitology. 2005 Aug;131(Pt 2):187–96.

11. Snowden K, Logan K, Blozinski C, Hoevers J, Holman P. Restriction-fragment-length polymorphism analysis of small-subunit rRNA genes of *Blastocystis* isolates from animal hosts. Parasitol Res. 2000 Jan;86(1):62–6.

12. Ho LC, Armiugam A, Jeyaseelan K, Yap EH, Singh M. *Blastocystis* elongation factor-1alpha: genomic organization, taxonomy and phylogenetic relationships. Parasitology. 2000 Aug;121 (Pt 2):135–44.

13. Stensvold CR, Suresh GK, Tan KSW, Thompson RCA, Traub RJ, Viscogliosi E, et al. Terminology for *Blastocystis* subtypes--a consensus. Trends Parasitol. 2007 Mar;23(3):93–6.

14. Stensvold CR, Clark CG. Current status of *Blastocystis*: A personal view. Parasitol Int. 2016 Dec;65(6 Pt B):763–71.

15. Adl SM, Bass D, Lane CE, Lukeš J, Schoch CL, Smirnov A, et al. Revisions to the classification, nomenclature, and diversity of eukaryotes. J Eukaryot Microbiol. 2019 Jan;66(1):4–119.

16. Stensvold CR, Tan KSW, Clark CG. *Blastocystis*. Trends Parasitol. 2020 Mar;36(3):315–6.

17. Parija SC, Jeremiah S. *Blastocystis*: Taxonomy, biology and virulence. Trop Parasitol. 2013 Jan;3(1):17–25.

18. He N, Zhang YQ, Hong ML, Cong M. [Morphological and ultrastructural observation of *Blastocystis* hominis]. Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi. 2001;19(3):169–72.

19. Zhang X, Qiao J, Dong X, Li Y, Li X, Li C. [Study on morphology of *Blastocystis* hominis in culture and from diarrhea patients]. Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi. 2003;21(2):116–8.

20. Tan TC, Suresh KG. Predominance of amoeboid forms of *Blastocystis* hominis in isolates from symptomatic patients. Parasitol Res. 2006 Feb;98(3):189–93.

21. Puthia MK, Lu J, Tan KSW. *Blastocystis* ratti contains cysteine proteases that mediate interleukin-8 response from human intestinal epithelial cells in an NF-kappaB-dependent manner. Eukaryotic Cell. 2008 Mar;7(3):435–43.

22. Puthia MK, Sio SWS, Lu J, Tan KSW. *Blastocystis* ratti induces contact-independent apoptosis, F-actin rearrangement, and barrier function disruption in IEC-6 cells. Infect Immun. 2006 Jul;74(7):4114–23.

23. Yoshikawa H, Wu Z, Nagano I, Takahashi Y. Molecular comparative studies among *Blastocystis* isolates obtained from humans and animals. J Parasitol. 2003 Jun;89(3):585–94.

24. Chen XQ, Singh M, Howe J, Ho LC, Tan SW, Yap EH. In vitro encystation and excystation of *Blastocystis* ratti. Parasitology. 1999 Feb;118 (Pt 2):151–60.

25. Singh M, Suresh K, Ho LC, Ng GC, Yap EH. Elucidation of the life cycle of the intestinal protozoan *Blastocystis* hominis. Parasitol Res. 1995;81(5):446–50.

26. Yoshikawa H, Yoshida K, Nakajima A, Yamanari K, Iwatani S, Kimata I. Fecal-oral transmission of the cyst form of *Blastocystis* hominis in rats. Parasitol Res. 2004 Dec;94(6):391–6.

27. Suresh K, Ng GC, Ramachandran NP, Ho LC, Yap EH, Singh M. In vitro encystment and experimental infections of *Blastocystis* hominis. Parasitol Res. 1993;79(6):456–60.

28. Suresh K, Ng GC, Ho LC, Yap EH, Singh M. Differentiation of the various stages of *Blastocystis* hominis by acridine orange staining. Int J Parasitol. 1994 Jul;24(4):605–6.

29. Yao F, Qiao J, Zhao Y, Zhang X, Yang J, Li X. [Experimental infection of mice with *Blastocystis* hominis]. Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi. 2005 Dec 30;23(6):444–8.

30. Moe KT, Singh M, Howe J, Ho LC, Tan SW, Chen XQ, et al. Experimental *Blastocystis* hominis infection in laboratory mice. Parasitol Res. 1997;83(4):319–25.

31. Fayer R, Elsasser T, Gould R, Solano G, Urban J, Santin M. *Blastocystis* tropism in the pig intestine. Parasitol Res. 2014 Apr;113(4):1465–72.

32. Wang W, Bielefeldt-Ohmann H, Traub RJ, Cuttell L, Owen H. Location and pathogenic potential of *Blastocystis* in the porcine intestine. PLoS ONE. 2014 Aug 5;9(8):e103962.

33. Boorom KF. Is this recently characterized gastrointestinal pathogen responsible for rising rates of inflammatory bowel disease (IBD) and IBD associated autism in Europe and the United States in the 1990s? Med Hypotheses. 2007 Mar 26;69(3):652–9.

34. Tan KSW, Mirza H, Teo JDW, Wu B, Macary PA. Current Views on the Clinical Relevance of *Blastocystis* spp. Curr Infect Dis Rep. 2010 Jan;12(1):28–35.

35. Poirier P, Wawrzyniak I, Vivarès CP, Delbac F, El Alaoui H. New insights into *Blastocystis* spp.: a potential link with irritable bowel syndrome. PLoS Pathog. 2012 Mar 15;8(3):e1002545.

36. Casero RD, Mongi F, Sánchez A, Ramírez JD. *Blastocystis* and urticaria: Examination of subtypes and morphotypes in an unusual clinical manifestation. Acta Trop. 2015 Aug;148:156–61.

37. Lee LI, Chye TT, Karmacharya BM, Govind SK. *Blastocystis* sp.: waterborne zoonotic organism, a possibility? Parasit Vectors. 2012 Jun 28;5:130.

38. Wawrzyniak I, Poirier P, Viscogliosi E, Dionigia M, Texier C, Delbac F, et al. *Blastocystis*, an unrecognized parasite: an overview of pathogenesis and diagnosis. Ther Adv Infect Dis. 2013 Oct;1(5):167–78.

39. El Safadi D, Gaayeb L, Meloni D, Cian A, Poirier P, Wawrzyniak I, et al. Children of Senegal River Basin show the highest prevalence of *Blastocystis* sp. ever observed worldwide. BMC Infect Dis. 2014 Mar 25;14:164.

40. Bart A, Wentink-Bonnema EMS, Gilis H, Verhaar N, Wassenaar CJA, van Vugt M, et al. Diagnosis and subtype analysis of *Blastocystis* sp. in 442 patients in a hospital setting in the Netherlands. BMC Infect Dis. 2013 Aug 23;13:389.

41. Scanlan PD, Stensvold CR. *Blastocystis*: getting to grips with our guileful guest. Trends Parasitol. 2013 Nov;29(11):523–9.

42. Aruni Chura J, Macchioni F, Furzi F, Balboa V, Mercado É, Gómez J, et al. Cross-sectional study on intestinal parasite infections in different ecological zones of the Department of La Paz, Bolivia. One Health. 2021 Dec;13:100271.

43. Falcone AC, Zonta ML, Unzaga JM, Navone GT. Parasitic risk factors in migrant horticultural families from Bolivia settled in the rural area of La Plata, Buenos Aires, Argentina. One Health. 2020 Dec 20;11:100179.

44. Tito RY, Chaffron S, Caenepeel C, Lima-Mendez G, Wang J, Vieira-Silva S, et al. Population-level analysis of *Blastocystis* subtype prevalence and variation in the human gut microbiota. Gut. 2019 Jul;68(7):1180–9.

45. Diaz E, Mondragon J, Ramirez E, Bernal R. Epidemiology and control of intestinal parasites with nitazoxanide in children in Mexico. Am J Trop Med Hyg. 2003;68(4):384-385.

46. Cruz Licea V, Plancarte Crespo A, Morán Alvarez C, Valencia Rojas S, Rodríguez Sásnchez G, Vega Franco L. Blastocystis hominis among food vendors in Xochimilco markets. Rev Latinoam Microbiol. 2003;45(1-2):12-15.

47. Partida-Rodríguez O, Serrano-Vázquez A, Nieves-Ramírez ME, Moran P, Rojas L, Portillo T, González E, Hernández E, Finlay BB, Ximenez C.Human Intestinal Microbiota: Interaction Between Parasites and the Host Immune Response. Arch Med Res. 2017;48(8):690-700. doi: 10.1016/j.arcmed.2017.11.015.

48. Rojas-Velázquez L, Morán P, Serrano-Vázquez A, et al. Genetic Diversity and Distribution of Blastocystis Subtype 3 in Human Populations, with Special Reference to a Rural Population in Central Mexico. Biomed Res Int. 2018;2018:3916263. Published 2018 Mar 18. doi:10.1155/2018/3916263

49. Rojas-Velázquez L, Maloney JG, Molokin A, et al. Use of next-generation amplicon sequencing to study Blastocystis genetic diversity in a rural human population from Mexico. Parasit Vectors. 2019;12(1):566. Published 2019 Nov 27. doi:10.1186/s13071-019-3814-z

50. Villegas-Gómez I, Martínez-Hernández F, Urrea-Quezada A, et al. Comparison of the genetic variability of Blastocystis subtypes between human carriers from two contrasting climatic regions of México. Infect Genet Evol. 2016;44:334-340.

51. Guangorena-Gómez JO, Lozano-Ochoa II, Rivera-Medina IL, Méndez-Hernández A, Espinosa-Fematt JA, Muñoz-Yáñez C. Relationship Among Blastocystis, the Firmicutes/Bacteroidetes Ratio and Chronic Stress in Mexican University Students. Curr Microbiol. 2022;79(3):72. Published 2022 Jan 24.

52. Tellevik MG, Moyo SJ, Blomberg B, Hjøllo T, Maselle SY, Langeland N, et al. Prevalence of Cryptosporidium parvum/hominis, Entamoeba histolytica and Giardia lamblia among Young Children with and without Diarrhea in Dar es Salaam, Tanzania. PLoS Negl Trop Dis. 2015 Oct 9;9(10):e0004125.

53. Pritt BS, Clark CG. Amebiasis. Mayo Clin Proc. 2008 Oct;83(10):1154–9; quiz 1159.

54. Cifre S, Gozalbo M, Ortiz V, Soriano JM, Merino JF, Trelis M. *Blastocystis* subtypes and their association with Irritable Bowel Syndrome. Med Hypotheses. 2018 Jul;116:4–9.

55. Khademvatan S, Masjedizadeh R, Rahim F, Mahbodfar H, Salehi R, Yousefi-Razin E, et al. *Blastocystis* and irritable bowel syndrome: Frequency and subtypes from Iranian patients. Parasitol Int. 2017 Apr;66(2):142–5.

56. Vasquez-Rios G, Machicado JD, Terashima A, Marcos LA. Irritable bowel syndrome and intestinal parasites: a view from South America. Rev Gastroenterol Peru. 2016 Jun;36(2):153–8.

57. Ajjampur SSR, Tan KSW. Pathogenic mechanisms in *Blastocystis* spp. - Interpreting results from in vitro and in vivo studies. Parasitol Int. 2016 Dec;65(6 Pt B):772–9.

58. Moe KT, Singh M, Gopalakrishnakone P, Ho LC, Tan SW, Chen XQ, et al. Cytopathic effect of *Blastocystis* hominis after intramuscular inoculation into laboratory mice. Parasitol Res. 1998 Jun;84(6):450–4.

59. Iguchi A, Yoshikawa H, Yamada M, Kimata I, Arizono N. Expression of interferon gamma and proinflammatory cytokines in the cecal mucosa of rats experimentally infected with *Blastocystis* sp. strain RN94-9. Parasitol Res. 2009 Jul;105(1):135–40.

60. Sio SWS, Puthia MK, Lee ASY, Lu J, Tan KSW. Protease activity of *Blastocystis* hominis. Parasitol Res. 2006 Jul;99(2):126–30.

61. Puthia MK, Vaithilingam A, Lu J, Tan KSW. Degradation of human secretory immunoglobulin A by *Blastocystis*. Parasitol Res. 2005 Nov;97(5):386–9.
62. Mirza H, Tan KSW. *Blastocystis* exhibits inter- and intra-subtype variation in cysteine protease activity. Parasitol Res. 2009 Jan;104(2):355–61.
63. Shirvani G, Fasihi-Harandi M, Raiesi O, Bazargan N, Zahedi MJ, Sharifi I, et al. Prevalence and Molecular Subtyping of *Blastocystis* from Patients with Irritable Bowel Syndrome, Inflammatory Bowel Disease and Chronic Urticaria in Iran. Acta Parasitol. 2020 Mar;65(1):90–6.
64. Bahrami F, Babaei E, Badirzadeh A, Riabi TR, Abdoli A. *Blastocystis*, urticaria, and skin disorders: review of the current evidences. Eur J Clin Microbiol Infect Dis. 2020 Jun;39(6):1027–42.
65. Wu B, Yin J, Texier C, Roussel M, Tan KS-W. *Blastocystis* legumain is localized on the cell surface, and specific inhibition of its activity implicates a pro-survival role for the enzyme. J Biol Chem. 2010 Jan 15;285(3):1790–8.
66. Wawrzyniak I, Texier C, Poirier P, Viscogliosi E, Tan KSW, Delbac F, et al. Characterization of two cysteine proteases secreted by *Blastocystis* ST7, a human intestinal parasite. Parasitol Int. 2012 Sep;61(3):437–42.
67. Denoeud F, Roussel M, Noel B, Wawrzyniak I, Da Silva C, Diogon M, et al. Genome sequence of the stramenopile *Blastocystis*, a human anaerobic parasite. Genome Biol. 2011 Mar 25;12(3):R29.
68. Koropatkin NM, Cameron EA, Martens EC. How glycan metabolism shapes the human gut microbiota. Nat Rev Microbiol. 2012 Apr 11;10(5):323–35.
69. Martens EC, Chiang HC, Gordon JI. Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. Cell Host Microbe. 2008 Nov 13;4(5):447–57.
70. Tailford LE, Crost EH, Kavanaugh D, Juge N. Mucin glycan foraging in the human gut microbiome. Front Genet. 2015 Mar 19;6:81.
71. Martínez-Ocaña J, Maravilla P, Olivo-Díaz A. Interaction between human mucins and parasite glycoproteins: the role of lectins and glycosidases in colonization by intestinal protozoa. Rev Inst Med Trop Sao Paulo. 2020 Sep 4;62:e64.
72. Sakurama H, Tsutsumi E, Ashida H, Katayama T, Yamamoto K, Kumagai H. Differences in the substrate specificities and active-site structures of two α-L-fucosidases (glycoside

hydrolase family 29) from Bacteroides thetaiotaomicron. Biosci Biotechnol Biochem. 2012 May 7;76(5):1022–4.

73. Benešová E, Lipovová P, Krejzová J, Kovaľová T, Buchtová P, Spiwok V, et al. Alpha-L-fucosidase isoenzyme iso2 from Paenibacillus thiaminolyticus. BMC Biotechnol. 2015 May 27;15:36.

74. Liang E-Y, Li G-H, Wang W-G, Qiu X-M, Ke P-F, He M, et al. Clinical relevance of serum α-l-fucosidase activity in the SARS-CoV-2 infection. Clin Chim Acta. 2021 Aug;519:26–31.

75. Occhiodoro T, Hopwood JJ, Morris CP, Anson DS. Correction of alpha-L-fucosidase deficiency in fucosidosis fibroblasts by retroviral vector-mediated gene transfer. Hum Gene Ther. 1992 Aug;3(4):365–9.

76. Romero-Aguirregomezcorta J, Matás C, Coy P. α-L-fucosidase enhances capacitation-associated events in porcine spermatozoa. Vet J. 2015 Jan;203(1):109–14.

77. Aguilera AC, Boschin V, Robina I, Elías-Rodríguez P, Sosa MA. Epididymal α-l-fucosidase and its possible role in remodelling the surface of bull spermatozoa. Theriogenology. 2017 Dec;104:134–41.

78. Ishida S, Kayamori K, Sakamoto K, Yukimori A, Kugimoto T, Harada H, et al. Alpha-L-fucosidase-1 is a diagnostic marker that distinguishes mucoepidermoid carcinoma from squamous cell carcinoma. Pathol Int. 2019 Feb 6;69(2):76–85.

79. Bonin S, Parascandolo A, Aversa C, Barbazza R, Tsuchida N, Castellone MD, et al. Reduced expression of α-L-Fucosidase-1 (FUCA-1) predicts recurrence and shorter cancer specific survival in luminal B LN+ breast cancer patients. Oncotarget. 2018 Mar 16;9(20):15228–38.

80. Zhang X, Zhao S, Liu H, Wang X, Wang X, Du N, et al. Identification of a novel homozygous loss-of-function mutation in FUCA1 gene causing severe fucosidosis: A case report. J Int Med Res. 2021 Apr;49(4):3000605211005975.

81. Chkioua L, Amri Y, Saheli C, Fenni F, Boudabous H, Ben Turkia H, et al. Fucosidosis in Tunisian patients: mutational analysis and homology-based modeling of FUCA1 enzyme. BMC Med Genomics. 2021 Aug 23;14(1):208.

82. Curiel JA, Peirotén Á, Landete JM, Ruiz de la Bastida A, Langa S, Arqués JL. Architecture Insight of Bifidobacterial α-L-Fucosidases. Int J Mol Sci. 2021 Aug 6;22(16).

83. Ashida H, Fujimoto T, Kurihara S, Nakamura M, Komeno M, Huang Y, et al. 1,6-α-L-Fucosidases from Bifidobacterium longum subsp. infantis ATCC 15697 Involved in the

Degradation of Core-fucosylated N-Glycan. Journal of Applied Glycoscience. 2020 Feb 20;67(1):23–9.

84. Luijkx YMCA, Bleumink NMC, Jiang J, Overkleeft HS, Wösten MMSM, Strijbis K, et al. Bacteroides fragilis fucosidases facilitate growth and invasion of *Campylobacter jejuni* in the presence of mucins. Cell Microbiol. 2020 Dec;22(12):e13252.

85. Liu P, Zhang H, Wang Y, Chen X, Jin L, Xu L, et al. Screening and characterization of an α-L-fucosidase from Bacteroides fragilis NCTC9343 for synthesis of fucosyl-N-acetylglucosamine disaccharides. Appl Microbiol Biotechnol. 2020 Sep;104(18):7827–40.

86. Rodríguez-Díaz J, Monedero V, Yebra MJ. Utilization of natural fucosylated oligosaccharides by three novel alpha-L-fucosidases from a probiotic Lactobacillus casei strain. Appl Environ Microbiol. 2011 Jan;77(2):703–5.

87. Rodríguez-Díaz J, Rubio-del-Campo A, Yebra MJ. Lactobacillus casei ferments the N-Acetylglucosamine moiety of fucosyl-α-1,3-N-acetylglucosamine and excretes L-fucose. Appl Environ Microbiol. 2012 Jul;78(13):4613–9.

88. Prabhu SK, Li C, Zong G, Zhang R, Wang L-X. Comparative studies on the substrate specificity and defucosylation activity of three α-l-fucosidases using synthetic fucosylated glycopeptides and glycoproteins as substrates. Bioorg Med Chem. 2021 Jul 15;42:116243.

89. Zeuner B, Muschiol J, Holck J, Lezyk M, Gedde MR, Jers C, et al. Substrate specificity and transfucosylation activity of GH29 α-l-fucosidases for enzymatic production of human milk oligosaccharides. N Biotechnol. 2018 Mar 25;41:34–45.

90. Eme L, Gentekaki E, Curtis B, Archibald JM, Roger AJ. Lateral gene transfer in the adaptation of the anaerobic parasite *Blastocystis* to the gut. Curr Biol. 2017 Mar 20;27(6):807–20.

91. Stensvold CR, Clark CG. Pre-empting Pandora's Box: *Blastocystis* Subtypes Revisited. Trends Parasitol. 2020 Mar;36(3):229–32.

92. Krogsgaard LR, Andersen LO 'Brien, Johannesen TB, Engsbro AL, Stensvold CR, Nielsen HV, et al. Characteristics of the bacterial microbiome in association with common intestinal parasites in irritable bowel syndrome. Clin Transl Gastroenterol. 2018 Jun 19;9(6):161.

93. Yason JA, Liang YR, Png CW, Zhang Y, Tan KSW. Interactions between a pathogenic *Blastocystis* subtype and gut microbiota: in vitro and in vivo studies. Microbiome. 2019 Mar 11;7(1):30.

94. Beghini F, Pasolli E, Truong TD, Putignani L, Cacciò SM, Segata N. Large-scale comparative metagenomics of *Blastocystis*, a common member of the human gut microbiome. ISME J. 2017 Dec;11(12):2848–63.
95. Ho LC, Singh M, Suresh G, Ng GC, Yap EH. Axenic culture of *Blastocystis* hominis in Iscove's modified Dulbecco's medium. Parasitol Res. 1993;79(7):614–6.
96. Abou Gamra MM, Elwakil HS, El Deeb HK, Khalifa KE, Abd Elhafiz HE. The potential use of 29 kDa protein as a marker of pathogenicity and diagnosis of symptomatic infections with *Blastocystis* hominis. Parasitol Res. 2011 May;108(5):1139–46.
97. . Matsumoto H, Haniu H, Komori N. Determination of Protein Molecular Weights on SDS-PAGE. Methods Mol Biol. 2019;1855:101-105.
98. Santín M, Gómez-Muñoz MT, Solano-Aguilar G, Fayer R. Development of a new PCR protocol to detect and subtype *Blastocystis* spp. from humans and animals. Parasitol Res. 2011 Jul;109(1):205–12.
99. Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, et al. Protein identification and analysis tools in the ExPASy server. Methods Mol Biol. 1999;112:531–52.
100. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. Nat Methods. 2015;12(1):7-8.
101. Schrödinger LLC. The {PyMOL} Molecular Graphics System. Version 1.80 LLC, New York, NY. 2015.
102. Gordon D, Abajian C, Green P. Consed: a graphical tool for sequence finishing. Genome Res. 1998 Mar;8(3):195–202.
103. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 1998 Mar;8(3):186–94.
104. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. 1998 Mar;8(3):175–85.
105. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007 Nov 1;23(21):2947–8.
106. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009 Aug 1;25(15):1972–3.

107. Sánchez R, Serra F, Tárraga J, Medina I, Carbonell J, Pulido L, et al. Phylemon 2.0: a suite of web-tools for molecular evolution, phylogenetics, phylogenomics and hypotheses testing. Nucleic Acids Res. 2011 Jul;39(Web Server issue):W470-4.
108. Flouri T, Izquierdo-Carrasco F, Darriba D, Aberer AJ, Nguyen LT, Minh BQ, et al. The phylogenetic likelihood library. Syst Biol. 2015 Mar;64(2):356–62.
109. Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T. ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. Mol Biol Evol. 2020 Jan 1;37(1):291–4.
110. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018 Jun 1;35(6):1547–9.
111. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012 May;61(3):539–42.
112. Murzin AG, Bateman A. Distant homology recognition using structural classification of proteins. Proteins. 1997;Suppl 1:105–12.
113. Dunbrack RL. Sequence comparison and protein structure prediction. Curr Opin Struct Biol. 2006 Jun;16(3):374–84.
114. Jacob AS, Andersen LO, Bitar PP, Richards VP, Shah S, Stanhope MJ, et al. *Blastocystis* mitochondrial genomes appear to show multiple independent gains and losses of start and stop codons. Genome Biol Evol. 2016 Dec 14;8(11):3340–50.
115. Alfellani MA, Taner-Mulla D, Jacob AS, Imeede CA, Yoshikawa H, Stensvold CR, et al. Genetic diversity of *Blastocystis* in livestock and zoo animals. Protist. 2013 Jul;164(4):497–509.
116. Gentekaki E, Curtis BA, Stairs CW, Klimeš V, Eliáš M, Salas-Leiva DE, et al. Extreme genome diversity in the hyper-prevalent parasitic eukaryote *Blastocystis*. PLoS Biol. 2017 Sep 11;15(9):e2003769.
117. Pacheco AR, Moel M, Segrè D. Costless metabolic secretions as drivers of interspecies interactions in microbial ecosystems. Nat Commun. 2019 Jan 9;10(1):103.
118. Guo Y, Kitamoto S, Kamada N. Microbial adaptation to the healthy and inflamed gut environments. Gut Microbes. 2020 Nov 9;12(1):1857505.
119. Sakurama H, Tsutsumi E, Ashida H, Katayama T, Yamamoto K, Kumagai H. Differences in the substrate specificities and active-site structures of two α-L-fucosidases

(glycoside hydrolase family 29) from Bacteroides thetaiotaomicron. Biosci Biotechnol Biochem. 2012;76(5):1022-1024.

120. Pacheco AR, Curtis MM, Ritchie JM, Munera D, Waldor MK, Moreira CG, et al. Fucose sensing regulates bacterial intestinal colonization. Nature. 2012 Dec 6;492(7427):113–7.

121. Billy V, Lhotská Z, Jirků M, Kadlecová O, Frgelecová L, Parfrey LW, et al. *Blastocystis* Colonization Alters the Gut Microbiome and, in Some Cases, Promotes Faster Recovery From Induced Colitis. Front Microbiol. 2021 Apr 7;12:641483.

122. Intra J, Perotti M-E, Pavesi G, Horner D. Comparative and phylogenetic analysis of alpha-L-fucosidase genes. Gene. 2007 May 1;392(1–2):34–46.

123. Tarling CA, He S, Sulzenbacher G, Bignon C, Bourne Y, Henrissat B, et al. Identification of the catalytic nucleophile of the family 29 alpha-L-fucosidase from Thermotoga maritima through trapping of a covalent glycosyl-enzyme intermediate and mutagenesis. J Biol Chem. 2003 Nov 28;278(48):47394–9.

124. Song J, Imanaka H, Imamura K, Minoda M, Yamaguchi S, Nakanishi K. The discoidin domain of Bacillus circulans β-galactosidase plays an essential role in repressing galactooligosaccharide production. Biosci Biotechnol Biochem. 2013 Jan 7;77(1):73–9.

125. Yin H, Pijning T, Meng X, Dijkhuizen L, van Leeuwen SS. Biochemical Characterization of the Functional Roles of Residues in the Active Site of the β-Galactosidase from Bacillus circulans ATCC 31382. Biochemistry. 2017 Jun 20;56(24):3109–18.

126. Jiménez PA, Jaimes JE, Ramírez JD. A summary of *Blastocystis* subtypes in North and South America. Parasit Vectors. 2019 Jul 29;12(1):376.

127. Kosciow K, Deppenmeier U. Characterization of a phospholipid-regulated β-galactosidase from *Akkermansia muciniphila* involved in mucin degradation. Microbiologyopen. 2019 Aug;8(8):e00796.

128. Crost EH, Tailford LE, Le Gall G, Fons M, Henrissat B, Juge N. Utilisation of mucin glycans by the human gut symbiont *Ruminococcus gnavus* is strain-dependent. PLoS ONE. 2013 Oct 25;8(10):e76341.

129. Verweij JJ, Stensvold CR, Molecular testing for clinical diagnosis and epidemiological investigations of intestinal parasitic infections. Clin Microbiol Rev. 2014;27: 371–418

130. Stensvold CR, Alfellani MA, Clark CG. Levels of genetic diversity vary dramatically between *Blastocystis* subtypes, Infect. Genet Evol. 2011;12: 263–273.

131. Clark CG, Giezen MDG, Alfellani MA, Stensvold CR. Recent developments in *Blastocystis* research. Adv Parasitol. 2013; 82: 1–32.

132. Alfellani MA, Stensvold CR, Vidal-Lapiedra A, Onuoha ESU, Fagbenro-Beyioku AF, Clark CG. Variable geographic distribution of *Blastocystis* subtypes and its potential implications. Acta Trop. 2013; 126 :11–18.

133. Alfellani MA, Jacob AS, Ortíz Perea N, Krecek RC, Taner-Mulla D, Verweij JJ, Levecke B, Tannich E, Clark CG, Stensvold CR. Diversity and distribution of *Blastocystis* sp. subtypes in non-human primates. Parasitology. 2013; 140: 966–971.

134. Wang W, Owen H, Traub RJ, Cuttell L, Inpankaew T, Bielefeldt-Ohmann H. Molecular epidemiology of *Blastocystis* in pigs and their in-contact humans in Southeast Queensland, Australia, and Cambodia. Vet Parasitol. 2014; 203 : 264–269.

135. Yoshikawa H, Wu Z, Kimata I, Iseki M, Ali IKM, Hossain MB, Zaman V, Haque R, Takahashi Y, Polymerase chain reaction-based genotype classification among human *Blastocystis* hominis populations isolated from different countries. Parasitol Res. 2004; 92: 22–29.

136. Santín M, Gómez-Muñoz M. T, Solano-Aguilar G, Fayer R. (). Development of a new PCR protocol to detect and subtype *Blastocystis* spp. from humans and animals. Parasitology research. 2011; 109(1): 205–212.

137. Scicluna SM, Tawari B, Clark CG. DNA barcoding of *Blastocystis*. Protist. 2006; 157:77–85.

138. Stensvold CR. Comparison of sequencing (barcode region) and sequence-tagged-site PCR for *Blastocystis* subtyping. J Clin Microbiol. 2013; 51: 190–194.

139. Meloni D, Poirier P, Mantini C, Noël C, Gantois N, Wawrzyniak I, Delbac F, Chabé M, Delhaes L, Dei-Cas E, Fiori PL, El Alaoui H, Viscogliosi E. Mixed human intra- and inter-subtype infections with the parasite *Blastocystis* sp. Parasitol Int 2012; 61: 719–722.

140. Higuera A, Muñoz M, López MC, Reyes P, Urbano P, Villalobos O, et al. Succinate dehydrogenase gene as a marker for studying *Blastocystis* genetic diversity. Heliyon. 2020 Nov 2;6(11):e05387.

141. Clark CG, Stensvold CR. *Blastocystis*: isolation, xenic cultivation, and cryopreservation. Curr Protoc Microbiol. 2016 Nov 18;43:20A.1.1-20A.1.8.

142. Maloney JG, da Cunha MJR, Molokin A, Cury MC, Santin M. Next-generation sequencing reveals wide genetic diversity of *Blastocystis* subtypes in chickens including potentially zoonotic subtypes. Parasitol Res. 2021 Jun;120(6):2219–31.

143. Maloney JG, Molokin A, Santin M. Use of Oxford Nanopore MinION to generate full-length sequences of the *Blastocystis* small subunit (SSU) rRNA gene. Parasit Vectors. 2020 Nov 25;13(1):595.

Anexo 1. Artículo requisito para la obtención del grado

REVISTA DO INSTITUTO DE MEDICINA TROPICAL DE SÃO PAULO

JOURNAL OF THE SÃO PAULO INSTITUTE OF TROPICAL MEDICINE

ORIGINAL ARTICLE

http://doi.org/10.1590/S1678-9946202062064

Interaction between human mucins and parasite glycoproteins: the role of lectins and glycosidases in colonization by intestinal protozoa

Joel Martínez-Ocaña[1], **Pablo Maravilla**[2], **Angélica Olivo-Díaz**[3]

ABSTRACT

Intestinal mucins are the first line of defense against microorganisms. Although knowledge about the mechanisms involved in the establishment of intestinal protozoa is limited, there is evidence that these parasites produce lectin-like molecules and glycosidases, that exert both, constitutive and secretory functions, promoting the establishment of these microorganisms. In the present review, we analyse the main interactions between mucins of the host intestine and the four main protozoan parasites in humans and their implications in intestinal colonization. There are lectin-like molecules that contain complex oligosaccharide structures and N-acetylglucosamine (GlcNAc), mannose and sialic acid as main components, which are excreted/secreted by *Giardia intestinalis*, and recognized by the host using mannose-binding lectins (MBL). *Entamoeba histolytica* and *Cryptosporidium* spp. express the lectin galactose/N-acetyl-D-galactosamine, which facilitates their adhesion to cells. In *Cryptosporidium,* the glycoproteins gp30, gp40/15 and gp900 and the glycoprotein lectin CpClec are involved in protozoan adhesion to intestinal cells, forming an adhesion-attack complex. *G. intestinalis* and *E. histolytica* can also produce glycosidases such as β-N-acetyl-D-glucosaminidase, α-d-glucosidase, β-d-galactosidase, β-l-fucosidase, α-N-acetyl-d-galactosaminidase and β-mannosidase. In *Blastocystis*, α-D-mannose, α-D-glucose, GlcNAc, α-D-fucose, chitin and sialic acid that have been identified on their surface. Fucosidases, hexosaminidases and polygalacturonases, which may be involved in the mucin degradation process, have also been described in the *Blastocystis* secretoma. Similarly, symbiotic coexistence with the intestinal microbiota promotes the survival of parasites facilitating cell invasion and nutrients obtention. Furthermore, it is necessary to identify and characterize more glycosidases, which have been only partially described by *in silico* analyses of the parasite genome.

KEYWORDS: Mucins. Lectins. Glycosidases. Protozoa. Intestinal parasites.

INTRODUCTION

The normal human gastrointestinal tract is inhabited by a vast and diverse community of microorganisms, such as bacteria, yeasts and parasites, and it is thought that a well-balanced microbiota is important for normal digestion and maintenance of the intestinal ecosystem. The mucosal epithelial tissue, after exposure to the external environment, has developed multiple functions and defense mechanisms in response to their susceptibility to microorganisms. Among the functions of mucosal epithelium are the absorption of nutrients and the role of cofactors in the intestine; gas exchange in the respiratory system; transparency of the cornea and lubrication; and detection of chemical elements. On the other hand, mucous membranes also have close relationships with the innate and the adaptive

[1]Hospital General "Dr. Manuel Gea González", Departamento de Ecología de Agentes Patógenos, Ciudad de México, Mexico

[2]Hospital General "Dr. Manuel Gea González", Subdirección de Investigación, Ciudad de México, Mexico

[3]Hospital General "Dr. Manuel Gea González", Departamento de Biología Molecular e Histocompatibilidad, Hospital General "Dr. Manuel Gea González", Ciudad de México, Mexico

Correspondence to: Angélica Olivo-Díaz
Hospital General "Dr. Manuel Gea González", Departamento de Biología Molecular e Histocompatibilidad, Calzada de Tlalpan 4800, Col. Sección XVI, Ciudad de México, 14080, Mexico
Tel: 5255 4000 3000, Ext. 6105

E-mail: aolivod@yahoo.com

Received: 15 May 2020

Accepted: 17 August 2020

immune systems, allowing a rapid response to threats to the mucous surface[1]. Therefore, mucosal epithelia throughout the body can be considered dynamic and vital entities forboth, the normal function of the body and its interaction with internal and external environments. Considering that the gastrointestinal tract is one of most exposed epithelia to environmental aggressors, it is important to study its interactions with different microorganisms. Thus, some intestinal parasites that cause diseases in humans can express different molecules that interact with the intestinal epithelium of the host, such as lectins, glycosidases and mucin-like proteins, which can favour the binding with and the invasion of host cells as well as the evasion of the immune response. In addition, parasitic enzymes can also facilitate infection by degrading the host mucus or by generating binding sites to host cells[2].

The present review focuses on the interaction between some molecules, such as lectin, lectin-like, mucin-like and glycosidase molecules which are related to the attachment and colonization by *Giardia intestinalis*, *Entamoeba histolytica*, *Cryptosporidium* spp. and *Blastocystis* spp. as the main parasitic protozoa that affect the intestinal mucosa of the human host, highlighting the possible role that these molecules play in the pathogenicity of these microorganisms.

Mucins: the first line of defense of the intestinal epithelium

The gastrointestinal mucosa is formed by a layer of columnar epithelial cells covered by a mucus coating. The mucus forms two layers, a sterile inner thin layer and a thicker non-sterile easy-to-remove layer. The viscous properties of mucus are generated by the secreted mucins, composed of large and complex glycoproteins bound to *O*-carbohydrates, representing approximately 70% of their composition and assembled as homo-oligomers[3]. In the gastrointestinal region, there are five secreted mucins that oligomerize (gel-forming) (MUC2, MUC5AC, MUC5B, MUC6 and MUC19) and a secreted mucin that does not oligomerize (non-gel-forming) (MUC7) (Table 1)[3,4]. Homo-oligomers are formed through intermolecular disulfide bonds between the cysteine-rich domains found at the amino and carboxyl ends of mucins. In the endoplasmic reticulum, N-terminal glycosylation and C-terminal dimerization occur, and then, in the Golgi apparatus, *O*-glycosylation and N-terminal multimerization of the dimers continue, followed by the packing of the mucins into granules for subsequent secretion. The glycosyltransferases within the Golgi of the goblet cells determine the identity of the oligosaccharides that are secreted. The expression of these mucins can be modulated by the innate and the adaptive immunity and differs between sites in the gastrointestinal tract. In the stomach, the mucus is composed of MUC5AC and MUC6 and is produced by calyx cells. In the small intestine, the mucus layer is thinner (mainly composed of MUC2) and consists of an inner layer of ~15-30 µm and an outer layer of ~100-400 µm, being thicker in the ileum. Mucus in the small intestine is produced by calyx cells and Paneth cells. In the large intestine, mucus (mainly MUC2) is predominantly produced by cells in the calyx and consists of an inner layer of ~100 µm and a thick outer layer of ~100 µm[5].

Table 1 - Distribution of mucins in the gastrointestinal tract.

Gastrointestinal Tract	Mucins	References
Stomach	MUC1, MUC5AC, MUC6	1,3-6
Small intestine	MUC1, MUC2, MUC3, MUC4, MUC6, MUC12, MUC13, MUC17	1,3-6
Large intestine	MUC1, MUC2, MUC3, MUC4, MUC12, MUC13, MUC17	1,3-6

A mucin protein (apomucin) is composed of a tandem repeat arrangement with elevated proportions of proline, threonine and serine (PTS domain). Carbohydrates are bound to the PTS domains, constituting approximately 50% of the molecular weight and with sizes and numbers that vary according to the type of mucin. PTS domains are encoded by regions known as variable number of tandem repeats (VNTR), which are polymorphic and variable in different individuals and between distinct species[6]. Furthermore, the genes encoding glysosyl transferases are genetically variable, increasing intra- and inter-species diversity. This diversity is an important contributor to the specificity of commensal and pathogenic microorganisms, since the external composition of the mucins can have multiple ligands with which microorganisms can bind and can be used as energy sources[3].

Lectins, mucin-like molecules and glycosidases in intestinal parasites

The definition of lectin is that it is a protein that has at least one non-catalytic domain that binds with a carbohydrate in a specific and reversible way[7]. In parasitic protozoa, lectins play important roles in different host-parasite interactions, as they facilitate the attachment to

various glycoproteins, glycolipids and proteoglycans found on the surface of the host's gastrointestinal tract, which is the site of contact for invasion and colonization (Table 2)[8].

Mucin-like glycoproteins have also been identified in several parasites, and these molecules are composed of amino acids and carbohydrates, which resemble mammalian mucins, in that they are highly glycosylated and rich in Thr, Ser and Pro. Proteophosphoglycans (PPGs) are the structural equivalents of mammalian mucins and have been detected in *Entamoeba histolytica*[9], as well as in *Cryptosporidium parvum*[10,11].

In addition, parasites that live in the mucosal epithelium require mechanisms to penetrate into the host's mucins, which can damage the mucous lining of the intestinal tract, as observed in *Giardia, Entamoeba*, and *Trichomonas* infections; these mechanisms may include the production of enzymes that degrade the protective epithelial mucus, the generation of ligands for the binding with and the invasion of the host's tissues and the release of nutrients that the parasite can consume[2,12].

Giardia intestinalis

Giardia intestinalis colonizes the duodenum, jejunum, and ileum and is the etiological agent of human giardiasis, one of the most common gastrointestinal diseases worldwide. The clinical manifestations of giardiasis vary from no symptoms to acute or chronic diarrhoea with abdominal pain, flatulence, weight loss, intestinal lesions and malabsorption syndrome, which can last for several months. *Giardia intestinalis* exhibits two stages, the cyst (resistant form) and the trophozoite (the parasitic form). *Giardia* cysts have a hard wall and are impervious to small molecules, which protect them from the environment, but they can break through quickly when the parasite excysts and colonizes the duodenum[13]. The *Giardia* cyst wall contains β-1,3-linked N-acetylgalactosamine (GalNAc), a homopolymer carbohydrate and at least three cyst wall proteins (CWPs) composed of Leu-rich repeats and a C-terminal conserved Cys-rich region. The Leu-rich repeat domain of CWP1 is a lectin that binds with the curled fibrils of the GalNAc homopolymer[14].

There is evidence of two main mechanisms that allow *Giardia* trophozoites to adhere to intestinal cells: a) a cytoskeletal mechanism involving ventral flagella or ventral discs, and b) a lectin-mediated mechanism[15]. Additionally, *Giardia* may produce α-N-acetyl-D-glucosaminidase, β-N-acetyl-D-glucosaminidase, α-D-galactosidase, β-l-fucosidase, α-mannosidase and β-xylosidase, enzymes that can degrade the intestinal mucins of the host, facilitating the invasion by the parasite[16].

Alternatively, it has been described that MBL and ficolins (innate pattern recognition receptors [PRRs]) of the host are capable of recognizing GlcNAc on the surface of the parasite, activating the lectin-complement pathway, as demonstrated by the inhibition of this route with different concentrations of mannose *in vitro*[17].

Furthermore, in the excretion/secretion products of *G. intestinalis*, two 58 and 63 kDa glycoproteins were identified by the periodic acid-Schiff staining. However, when these products were treated with sodium metaperiodate or N-glycosidase F, a deglycosylation

Table 2 - Lectin or glycan-like structures and glycosidases secreted by the main protozoan parasites.

Parasite	Lectin/glycan-like structures	Glycosidases	References
Giardia lamblia	Polymer b-1,3-linked GalNAc specific in cyst, Secretion N-glycan structures, including GlcNAC, mannose and sialic acid as main components	β-N-acetyl-d-glucosaminidase, α-d-glucosidase, β-d-galactosidase, β-l-fucosidase, α-N-acetyl-d-galactosaminidas, β-mannosidase	12,14,16-18
Entamoeba histolytica	Gal/GalNAc-specific surface	β-N-acetyl-d-glucosaminidase, α-d-glucosidase, β-d-galactosidase, β-l-fucosidase, α-N-acetyl-d-galactosaminidas, β-mannosidase, α 1,6-glucosidase	12,16,19-24
Cryptosporidium parvum	Gal/GalNAc-specific surface (p30), gp15, gp40, gp900, CpClec	INA	37-46
Blastocystis spp	α-D-mannose, α-D-glucose, GlcNAc, α-L-fucose, chitin, and sialic acid specific surface	L-Fucosidase, Hexosaminidase, Polygalacturonase,	53-58

INA = information not available

enzyme, serum from the mice immunized with excretion/secretion products detected three proteins of 36, 58 and 63 kDa, indicating the possibility of a glycosidic epitope recognition by antibodies[18].

Entamoeba histolytica

Entamoeba histolytica is a mitochondria-devoided protist and the parasite responsible for human amebiasis; in the colon it may cause dysentery, but it can also spread through the blood circulation to other organs, mainly the liver, causing abscesses. The parasite has two life cycle forms: motile trophozoites and cyst forms. Trophozoites are the proliferative form crucial for the development of the disease. The cysts are the infective form, produced in the intestine and eliminated in the feces and subsequently ingested through contaminated food or water.

E. histolytica mainly expresses a Gal/GalNAc-specific surface lectin on its surface, which is associated with its virulence, thus facilitating adherence to the host's mucins and colonic cells. Certain virulent clones of *E. histolytica* can internalize and degrade the mucins of the host's colon and secrete them, facilitating the passage in to the mucosal epithelium through the protective mucus. Avirulent clones have a much lower rate of internalization and mucosal discharge[12,19,20]. Adhesion of trophozoites to colonic mucin is almost completely inhibited by the addition of Gal or GalNAc *in vitro*, although some attachment may still occur at high concentrations of these molecules, indicating the participation of other components in this interaction[20,21]. To lyse host epithelial cells, the parasite-host interaction through this lectin is necessary, as indicated by the addition of Gal or GalNAc inhibiting the general virulence of the parasite *in vitro*.

The Gal/GalNAc lectin is composed of a 260 kDa heterodimer of heavy (*hgl*) (170 kDa) and light (*lgl*) (35/31 kDa) subunits linked by disulfide bonds noncovalently associated with an intermediate subunit (*igl*) of 150 kDa. The *lgl* and *igl* subunits are anchored to the membrane by acyl-glycosyl-phosphatidylinositol (GPI). The *hgl* subunit has a carbohydrate recognition domain (CRD), within a cysteine-rich region (amino acids 356 to 1143), and has a short cytoplasmic tail, which is implicated in intracellular signalling and is homologous to the β2 and β7 integrins.

In *E. histolytica,* the *hgl* subunit gene family consists of five genes with 89-95% of identity at the amino acid level. The *lgl* subunit is composed of a gene family of 6-7 members encoding isoforms, sharing 79-85% identity, with different post-translational modifications. The *igl* subunit has two genes encoding proteins with 84% amino acid identity, but do not share similarity with the 260-kDa protein[20,22].

The *lgl* subunit does not appear to be involved in carbohydrate recognition, as mAb and polyclonal antibodies directed against the light subunit isoforms have no significant effect on adherence or cytotoxicity. When the *lgl* protein is purified, it is always bound with the *hgl* protein, however, when the *lgl* protein lacks the GPI anchor, it does not bind with *hgl*, or with carbohydrates (Gal), indicating that *lgl*, when not part of the heterodimeric lectin, lacks this capacity[23]. Additionally, when amebas were transfected with a truncated *lgl* gene, which encodes a protein produced in excess, there was a significant decrease in their ability to adhere and kill mammalian cells, to form rosettes with human red blood cells (which is mediated by the Gal/GalNAc *E. histolytica* lectin incubated with red cells) and to phagocyte erythrocytes, and they were unable to cap anti-Gal lectin complexes on their surface (the process known as capping is the rapid redistribution of surface antigen-anti-body complexes in trophozoites), which indicates that the light subunit has the role of grouping these complexes, an activity that is essential for virulence[24].

On the other hand, *igl* lacks a carbohydrate recognition motif; however, it has galactose-binding activity and contains several CXXC and CXC domains, which indicates protein-protein interactions; monoclonal antibodies against *igl* significantly inhibit ameba adherence, phagocytosis and toxicity[25]. In addition, 96% of hamsters immunized with the C-terminal fragment of *igl* were protected against liver abscess formation, and the antibodies generated inhibited up to 80% the adhesion of trophozoites to mammalian cells *in vitro*, supporting the use of this fragment as a vaccine to prevent amebiasis[26].

Furthermore, the CRD of the Gal/GalNAc lectin has been shown to function as a pathogen-associated-molecular-pattern-like molecule (PAMP), as it binds with Toll-like receptors TLR-2 and with TLR-4 on human colon cells, and it can induce the expression of elements of the innate immune response, subsequently leading to the production of pro-inflammatory cytokines. This is exacerbated when there is previous exposure to pathogenic enterobacteria, leading to the development of an invasive disease[27]. Another interesting activity described for Gal/GalNAc lectin is that the *E. histolytica* Igl1 subunit has both, hemolytic and cytotoxic activities and they depend directly on the level of expression of the *Lgl1* gene[28].

The 220-kDa lectin (L220) has specificity for GlcNAc, but hemagglutination may be inhibited by hyaluronic acid, chitotriose, chitin, GalNAc and Gal. L220 was immunolocalized on the surface of the trophozoites,

indicating their participation in adhesion to the target cells[20].

The 112-kDa protein is also involved in adhesion and is located within the plasma membrane and into trophozoite vacuoles. The final protein is encoded by two nearby genes that give rise to a 49 and a 75 kDa polypeptide. The 49-kDa polypeptide is a cysteine protease (EhCP112), which can digest azocasein, gelatine, type I collagen, fibronectin and haemoglobin, while the 75-kDa polypeptide is an adhesin (EhADH); both polypeptides are located on the surface of trophozoite as they are membrane proteins. The activity of the cysteine protease allows mucin degradation and parasite access to the epithelium of the host colon, which causes increased mucin secretion to prevent invasion of the parasite. EhCP112 also has a cathepsin (Cys-His-Asn) catalytic domain and an RGD (Arg-Gly-Asp) domain, suggesting that this molecule can bind with integrins[20,27,29].

Several glycosidases secreted by or found in the lysates of *E. histolytica* trophozoites have been described, including β-N-acetyl-D-glucosaminidase, α-D-glucosidase, β-D-galactosidase, xylosidase, glucuronidase, N-acetylneuraminate lyase, neuraminidase, hyaluronidase, β-L-fucosidase, α-N-acetyl-galactosaminidase, sialidase, β-amylase and α-mannosidase[16,30-32]. Of these, sialidase, α-glucosidase, β-amylase and β-N-acetyl-D-glucosaminidase seem to be involved in ameba pathogenesis, being the most active β-N-acetyl-D-glucosaminidase, which seems to have a central role in the degradation of mucin[31]. Glycosidases act through the degradation of mucin carbohydrates, thus exposing the protein skeleton to parasite proteases[31,33]. Cysteine proteases secreted by *E. histolytica* are mainly involved in the final disruption of colonic mucin, allowing ameba adherence and invasion[34].

Cryptosporidium spp.

Cryptosporidium is an obligate intracellular protozoan parasite, a member of the Apicomplexa phylum, which was initially recognized as an opportunistic pathogen in immunocompromised patients, but can also be found in immunocompetent hosts, usually causing asymptomatic or self-limited infections[35]. The species that mainly infect humans are *C. hominis* and *C. parvum*[36]. *Cryptosporidium* life cycle is complex, starting with the ingestion of food or water contaminated with oocysts, followed by an excystation process in the small intestine, going through various stages of asexual and sexual reproduction in the intestinal epithelial cells, to finally produce oocysst that are going to be released into the intestine and excreted into the environment[37].

The genomes of *Cryptosporidium* spp. indicate that two kinds of proteins are involved in adhesion: mucin-like glycoproteins and thrombospondin-related adhesive proteins (TRAPs)[38].

Glycoprotein 900 (GP900) was the first cloned and characterized mucin-like protein of *C. parvum*, and it is located at the micronemes of the sporozoites, where it is released to mediate invasion. This protein shares homology with human MUC2 and MUC5 mucins and has several domains: two mucin-like, polythreonine, trinucleotide repeat domains; two cysteine-rich domains; one large degenerate 8-mer amino acid repeat domain; one transmembrane domain; and a signal peptide[10]. The glycoprotein gp900 is involved in parasite attachment and invasion of host cells.

The surface gp60 protein from sporozoites and merozoites is also called gp15/40/60, is a mucin-like molecule and is processed into mature gp40 and gp15 glycopeptides, both having a core of *O*-linked α-GalNAc, and gp15 having also a CPI anchor. This protein presents great variability and has been used to classify *Cryptosporidium* isolates into subtypes[36].

P23 is a 23-kDa sporozoite and merozoite surface glycoprotein belonging to a family of proteins of 23-27 kDa, each of which has a core of *O*-linked α-GalNAc. P23 is antigenically conserved in isolates from different parts of the world and is deposited on trails during the initial stages of infection. This protein is involved in attachment and invasion[37].

P12 is a glycoprotein detected by phage display and has an N-terminal signal peptide, a transmembrane region, an N-glycosylation site, a casein kinase II phosphorylation site and two N-myristoylation sites; it is located in the apical region of sporozoites and on the surface of oocysts[39].

A *Cryptosporidium parvum* protein that contains the C-type lectin domain (CTLD) was described and named CpClec[40]. It is a type 1 transmembrane protein that contains, in addition to a CTLD, a mucin-like domain predicted to be *O*-glycosylated and a Tyr-based sorting motif in the cytoplasmic tail. The expression of the CpClec protein is regulated during development, being located in the apical region and in the dense granules of sporozoites and merozoites and in the feeder organelle in intracellular stages, suggesting possible roles in the union, invasion and/or intracellular development in the host cell. In addition, CpClec mediates the infection of *C. parvum* through Ca^{2+}-dependent binding with sulfated proteoglycans on host intestinal epithelial cells[41].

P30 is a lectin in *C. parvum* and *C. hominis* that specifically binds with the Gal/GalNAc molecules of the host cell. This protein is found in the apical region of sporozoites, is involved in cell binding and invasion and is associated with gp900 and gp40 for this purpose[42].

The Cpa135 protein is expressed and secreted by oocysts and sporozoites and, after the invasion of the sporozoites into the host cell, it overlays the parasitophore vacuole. It is a modular protein that has a signal peptide at the N-terminus, a domain similar to ricin B between residues 170 and 290, an LCCL domain between residues 738 and 828, and a region rich in cysteines at the C-terminus[43].

The circumsporozoite-like glycoprotein (CSL) is a conserved molecule with a molecular weight of 1,300 kDa that is found in the apical complex of both sporozoites and merozoites and is an important surface ligand that mediates the binding with and facilitates the invasion of intestinal epithelial cells. CSL contains a sporozoite and merozoite ligand that is involved in parasite attachment to intestinal epithelial cells during the infection process in a dose-dependent, saturable, auto-shift manner and, once bound, significantly diminishes the permissiveness of the cell to infection by *C. parvum* sporozoites. The host receptor recognized by CSL was characterized and detected on the microvillar surface of the Caco-2 human intestinal epithelial cell line. It was identified as an 85-kDa surface protein that is capable of significantly inhibiting the binding and invasion of sporozoites that are incubated with it, prior to being inoculated into Caco-2 cells. Monoclonal antibodies directed to CSL were shown to completely neutralize sporozoite infectivity *in vitro* and limit an infection in an *in vivo* mouse model[44].

The thrombospondin-related adhesive proteins (TRAPs) contain thrombospondin type 1 (TSP1), the genes of which have been conserved among apicomplexans and contain one or more TSP-1 domains, which are implicated in zoite gliding motility and invasion of host cells[45]. TRAP-C1 (*Cryptosporidium* 1 thrombospondin-related adhesive protein) is a 76-kDa protein located in the apical pole of sporozoites that is structurally related to the micronemal proteins MIC2, TRAP and Etp100, sharing structural homology with members of the thrombospondin adhesive protein family in other apicomplexans[46]. TRAPs and other structurally related proteins constitute a family of functionally homologous proteins involved in parasite motility and cell penetration by binding with sulfated glycosaminoglycans (GAGs) on the host cell surface[47]. It is believed that gliding motility and cell invasion, which depend on the microfilaments of the parasite, result from a capping activity of the parasite surface ligands directly associated with TRAP-C1 molecules[48].

Blastocystis spp.

Blastocystis is the most common unicellular parasite that infects the intestinal tract of humans, as well as many animals; it is prevalent worldwide and is estimated to affect between 1 and 2 billion people, although its clinical importance remains controversial[49].

Multiple morphological forms of *Blastocystis* can be found in different individuals, such as vacuolar, multivacuolar, avacuolar, granular, ameboid and cystic forms. However, only four stages have been formally accepted: vacuolar, granular, trophozoite and cyst, and *Blastocystis* also have great genetic diversity[50].

At first, *Blastocystis* was named *Blastocystis hominis* because it was detected mainly in humans; however, it was later determined that there were other hosts for this parasite, and the name was changed to *Blastocystis* spp. Subsequently, with the use of small-subunit ribosomal RNA gene (SSU-rDNA) analysis, 17 subtypes (STs) were identified, and their presence worldwide has been widely documented. Of the 17 STs, nine (STs 1-9) can infect humans, with STs 1, 2, 3 and 4 being those found in more than 90% of human isolates[51]. Regarding clinical presentations, *Blastocystis* carriers infected with ST1 or ST3 had referred intestinal disorders, while ST2 has been associated with asymptomatic infections[52].

In late 20th century, during a survey focused on determining the surface composition of *Blastocystis* using lectins, alpha-D-mannose, alpha-D-glucose, N-acetyl-alpha-D-glucosamine, alpha-L-fucose, chitin and sialic acid were detected[53].

It has been argued that the ability of *Blastocystis* to inhibit the host immune response may favor its persistence, adhesion and survival within the intestinal tissue, which can be facilitated by the release of molecules in the host-parasite interface. Cysteine proteases play central roles in the virulence of *Blastocystis*; for example, the 31-kDa asparaginil cysteine protease called legumain is found on the parasite's cell surface, and cathepsin B, which is 31 kDa, represents another virulence factor. It has been observed that legumain plays a pro-survival role, as it induces apoptosis of parasite cells when they react with specific antibodies[54]. These two proteases were identified in parasite culture supernatants and may be involved in increasing intestinal permeability, degradation of human secretory immunoglobulin A, induction of IL-8 production, mucus degradation, tissue damage and gastrointestinal disturbances[55].

Little is known about the *Blastocystis* enzymes involved in the degradation of the host's intestinal barriers; however, the parasite genome has been sequenced, and 75 probable proteins have been detected, including proteases, glycosidases, lectins, glycosyl transferases and protease inhibitors, which may facilitate *Blastocystis* binding to the host. Based on these sequencing data, it was proposed that eight glycosyltransferases and one carbohydrate-binding

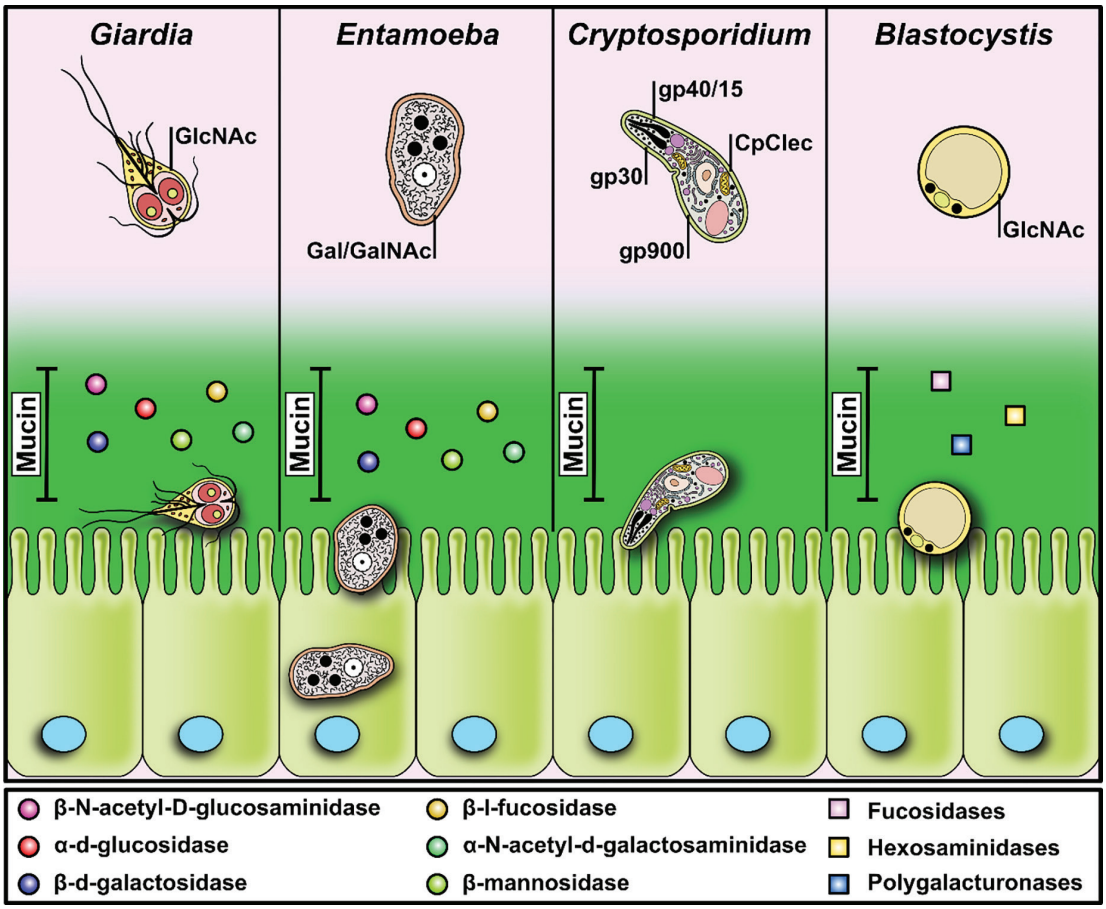


Figure 1 - Schematic representation of the interactions between the main protozoa intestinal pathogens and the human intestinal mucosa. In *Giardia*, lectin-like structures have been identified on the surface of the parasite, including GlcNAc, mannose and sialic acid as main components that can also be excreted/secreted by the parasite. *E. histolytica* and *Cryptosporidium spp.* express specific Gal/GalNAc (named gp30 in cryptosporidium) lectins that facilitate cell adhesion. In addition, *Cryptosporidium* gp40/15, gp900 and a CpClec lectin-like glycoprotein have been l described and are involved in adhesion to intestinal cells, forming an adhesion-attack complex. *G. lamblia* and *Entamoeba hystolytica* can also produce glycosidases such as: β-N-acetyl-D-glucosaminidase, α-d-glucosidase, β-d-galactosidase, β-l-fucosidase, α-N-acetyl-d-galactosaminidase and β-mannosidase. In *Blastocystis,* α-D-mannose, α-D-glucose, GlcNAc, α-D-Fucose, chitin and sialic acid have been identified on the surface of the parasite. Fucosidases, hexosaminidases and polygalacturonases, which may be involved in the mucin degradation process, have also been described in the predicted *Blastocystis* secretoma.

protein were secreted. The sequence of several hydrolases has also been determined, and these hydrolases can be used by the parasite for the degradation of host tissues. Among these hydrolases, fucosidase, hexosaminidase and polygalacturonase can degrade host glycoproteins, while proteases can act in several of the processes involved in the invasion and virulence of the parasite[56].

The function of glycoconjugates found on the surfaces of parasitic protozoa, involves more than serving as a simple additional barrier to the cell membrane; they play important roles in the adhesion and invasion of the parasite and its evasion of the host immune response. The variation in the composition of the surface glycoconjugates of different *Blastocyst* STs has been determined by electron microscopy, and the modifications observed in the images may be associated with differences in ST pathogenicity[57].

When phylogeny-based methods and strict validation criteria are used, it was found that approximately 25% of *Blastocystis* genes were recently acquired by lateral gene transfer (LGT) from prokaryotic and eukaryotic donors and that this process is ongoing. They reported 167 genes belonging to 74 gene families that corresponded to recent LGT events. The acquired genes come mainly from the bacteria that inhabit the intestine and are central to the adaptation of the colonization in the intestinal tract. Seventeen genes participating in carbohydrate metabolism were identified, including five genes involved in the import and metabolism of L-fucose, such as alpha L-fucosidase (FUCA), L-fucose permease fucP and L-fuculose phosphate aldolase, which are homologous enzymes in *Bacteroides thetaiotaomicron*, and L-fucose dehydrogenase and L-fuconolactase, and are similar to those found in a metabolic

pathway in *Campylobacter jejuni*. On the other hand, at least seven genes with the ability to participate in the evasion of the host's immune response inducing pathogenicity have been described. In particular, a gene encoding a beta-1,3-galactosyltransferase (β1,3GalT) appears to have been acquired from animals and subsequently duplicated several times in the *Blastocystis* genome. The authors hypothesize that these genes code for proteins that are potentially involved in "molecular mimicry" and serve as camouflage on the cell surface, thereby avoiding recognition by the host[58]. After LGT, extensive duplication of genes is very common and occurs in *Blastocystis* and in other organisms. Fucose transporters represent one of the families that have undergone a large increase, with seven to nine copies in all STs. These transporters may be relevant for the colonization of the intestine by the parasite, as they are involved in the concentration and regulation of fucose. Despite the advances that have been made in the understanding of the interactions of *Blastocystis* with the intestinal epithelium and the immunity associated with the intestinal mucosa, there are still many questions to be answered.

CONCLUSION

The present review focuses on the interactions of the host mucous membrane with lectins expressed on the surface and excreted/secreted and the glycosidases secreted by parasites and their roles during adhesion and colonization of human intestinal mucin cells. However, information on these molecules, particularly glycosidases in parasites, is scarce. In contrast, a number of references about bacterial glycosidases have been published. The studies described herein indicate that the host/parasite interface is dynamic and there is a constant exchange of molecules that can prevent or facilitate host colonization. In addition, the symbiotic coexistence with the intestinal microbiota promotes an adequate microenvironment for the survival of parasites, either by the direct action of the parasites excreted/secreted lectins and glycosidases into the intestinal microenvironment or by the use of hydrolysed carbohydrates as energy sources for their own metabolic pathways. Furthermore, it is necessary to characterize and evaluate more glycosidases, which thus far have only cursorily described the genome in several parasites.

AUTHOR'S CONTRIBUTIONS

All authors prepared the bibliographic survey, manuscript preparation and approved the final version of the manuscript.

CONFLICT OF INTERESTS

The authors declare that they have no conflict of interests.

FUNDING

This work was supported by Consejo Nacional de Ciencia y Tecnología (CONACYT) grant Nº 168619. Joel Martínez Ocaña is a doctoral student from Programa de Doctorado em Ciencias Biológicas, Universidad Nacional Autonoma de Mexico
(UNAM).

REFERENCES

1. Linden SK, Sutton P, Karlsson NG, Korolik V, McGuckin MA. Mucins in the mucosal barrier to infection. Mucosal Immunol. 2008;1:183-97.
2. Theodoropoulos G, Hicks SJ, Corfield AP, Miller BG, Carrington SD. The role of mucins in host-parasite interactions: Part II - helminth parasites. Trends Parasitol. 2001;17:130-5.
3. Johansson ME, Sjövall H, Hansson GC. The gastrointestinal mucus system in health and disease. Nat Rev Gastroenterol Hepatol. 2013;10:352-61.
4. Kim YS, Ho SB. Intestinal goblet cells and mucins in health and disease: recent insights and progress. Curr Gastroenterol Rep. 2010;12:319-30.
5. McGuckin MA, Lindén SK, Sutton P, Florin TH. Mucin dynamics and enteric pathogens. Nat Rev Microbiol. 2011;9:265-78.
6. Dhanisha SS, Guruvayoorappan C, Drishya S, Abeesh P. Mucins: structural diversity, biosynthesis, its role in pathogenesis and as possible therapeutic targets. Crit Rev Oncol Hematol. 2018;122:98-122.
7. Singh RS, Bhari R, Kaur HP. Characteristics of yeast lectins and their role in cell-cell interactions. Biotechnol Adv. 2011;29:726-31.
8. Singh RS, Walia AK, Kanwar JR. Protozoa lectins and their role in host-pathogen interactions. Biotechnol Adv. 2016;34:1018-29.
9. Moody-Haupt S, Patterson JH, Mirelman D, McConville MJ. The major surface antigens of Entamoeba histolytica trophozoites are GPI-anchored proteophosphoglycans. J Mol Biol. 2000;297:409-20.
10. Barnes DA, Bonnin A, Huang JX, Gousset L, Wu J, Gut J, et al. A novel multi-domain mucin-like glycoprotein of Cryptosporidium parvum mediates invasion. Mol Biochem Parasitol. 1998;96:93-110.
11. Cevallos AM, Zhang X, Waldor MK, Jaison S, Zhou X, Tzipori S, et al. Molecular cloning and expression of a gene encoding

Cryptosporidium parvum glycoproteins gp40 and gp15. Infect Immun. 2000;68:4108-16.

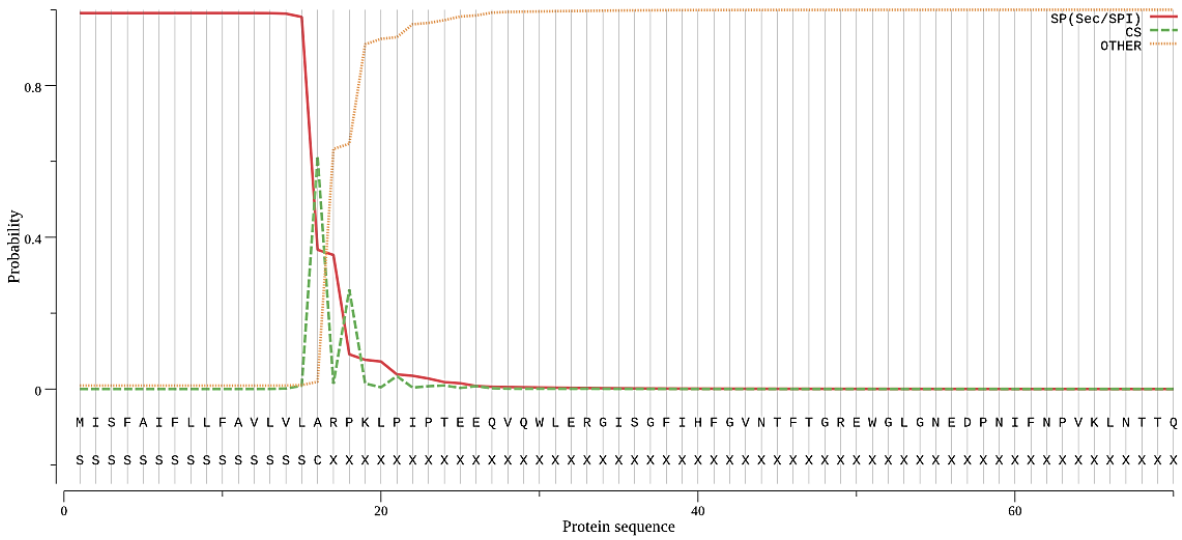
12. Hicks SJ, Theodoropoulos G, Carrington SD, Corfield AP. The role of mucins in host-parasite interactions: Part I - Protozoan parasites. Parasitol Today. 2000;16:476-81.
13. Samuelson J, Robbins P. A simple fibril and lectin model for cyst walls of Entamoeba and perhaps Giardia. Trends Parasitol. 2011;27:17-22.
14. Chatterjee A, Carpentieri A, Ratner DM, Bullitt E, Costello CE, Robbins PW, et al. Giardia cyst wall protein 1 is a lectin that binds to curled fibrils of the GalNAc homopolymer. PLoS Pathog. 2010;19:e1001059.
15. Sousa MC, Gonçalves CA, Bairos VA, Poiares-Da-Silva J. Adherence of Giardia lamblia trophozoites to Int-407 human intestinal cells. Clin Diagn Lab Immunol. 2001;8:258-65.
16. Connaris S, Greenwell P. Glycosidases in mucin-dwelling protozoans. Glycoconj J. 1997;14:879-82.
17. Li E, Tako EA, Singer SM. Complement activation by Giardia duodenalis parasites through the lectin pathway contributes to mast cell responses and parasite control. Infect Immun. 2016;84:1092-9.
18. Jiménez JC, Morelle W, Michalsky JC, Dei-Cas E. Excreted/secreted glycoproteins of G. intestinalis play an essential role in the antibody response. Parasitol Res. 2007;100:715-20.
19. Chadee K, Johnson ML, Orozco E, Petri WA Jr, Ravdin JI. Binding and internalization of rat colonic mucins by the galactose/N-acetyl-D-galactosamine adherence lectin of Entamoeba histolytica. J Infect Dis. 1988;158:398-406.
20. Aguirre-García M, Gutiérrez-Kobeh L, López Vancell R. Entamoeba histolytica: adhesins and lectins in the trophozoite surface. Molecules. 2015;20:2802-15.
21. Chadee K, Petri WA Jr, Innes DJ, Ravdin JI. Rat and human colonic mucins bind to and inhibit the adherence lectin of Entamoeba histolytica. J Clin Invest. 1987;80:1245-54.
22. Coudrier E, Amblard F, Zimmer C, Roux P, Olivo-Marin JC, Rigothier MC, et al. Myosin II and the Gal-GalNAc lectin play a crucial role in tissue invasion by Entamoeba histolytica. Cell Microbiol. 2005;7:19-27.
23. Petri WA Jr, Haque R, Mann BJ. The bittersweet interface of parasite and host: lectin-carbohydrate interactions during human invasion by the parasite Entamoeba histolytica. Annu Rev Microbiol. 2002;56:39-64.
24. Katz U, Ankri S, Stolarsky T, Nuchamowitz Y, Mirelman D. Entamoeba histolytica expressing a dominant negative N-truncated light subunit of its gal-lectin are less virulent. Mol Biol Cell. 2002;13:4256-65.
25. Cheng XJ, Hughes MA, Huston CD, Loftus B, Gilchrist CA, Lockhart LA, et al. Intermediate subunit of the Gal/GalNAc lectin of Entamoeba histolytica is a member of a gene family containing multiple CXXC sequence motifs. Infect Immun. 2001;69:5892-8.
26. Min X, Feng M, Guan Y, Man S, Fu Y, Cheng X, et al. Evaluation of the C-terminal fragment of Entamoeba histolytica Gal/GalNAc lectin intermediate subunit as a vaccine candidate against amebic liver abscess. PLoS Negl Trop Dis. 2016;10:e0004419.
27. Galván-Moroyoqui JM, Del Carmen Domínguez-Robles M, Meza I. Pathogenic bacteria prime the induction of Toll-like receptor signalling in human colonic cells by the Gal/GalNAc lectin Carbohydrate Recognition Domain of Entamoeba histolytica. Int J Parasitol. 2011;41:1101-12.
28. Kato K, Makiuchi T, Cheng X, Tachibana H. Comparison of hemolytic activity of the intermediate subunit of Entamoeba histolytica and Entamoeba dispar lectins. PLoS One. 2017;12:e0181864.
29. Ocádiz R, Orozco E, Carrillo E, Quintas LI, Ortega-López J, García-Pérez RM, et al. EhCP112 is an Entamoeba histolytica secreted cysteine protease that may be involved in the parasite-virulence. Cell Microbiol. 2005;7:221-32.
30. Betanzos A, Bañuelos C, Orozco E. Host invasion by pathogenic amoebae: epithelial disruption by parasite proteins. Genes (Basel). 2019;10:618.
31. Nok AJ, Rivera W. Characterization of sialidase from Entamoaeba hystolitica and possible pathogenic role in amebiasis. Parasitol Res. 2003;89:302-7.
32. Espinosa-Cantellano M, Martínez-Palomo A. Pathogenesis of intestinal amebiasis: from molecules to disease. Clin Microbiol Rev. 2000;13:318-31.
33. Thibeaux R, Weber C, Hon CC, Dillies MA, Avé P, Coppée JY, et al. Identification of the virulence landscape essential for Entamoeba histolytica invasion of the human colon. PLoS Pathog. 2013;9:e1003824.
34. Que X, Reed SL. Cysteine proteinases and the pathogenesis of amebiasis. Clin Microbiol Rev. 2000;13:196-206.
35. Gibson AR, Striepen B. Cryptosporidium. Curr Biol. 2018;28:R193-4.
36. Peng MM, Xiao L, Freeman AR, Arrowood MJ, Escalante AA, Weltman AC, et al. Genetic polymorphism among Cryptosporidium parvum isolates: evidence of two distinct human transmission cycles. Emerg Infect Dis. 1997;3:567-73.
37. Bouzid M, Hunter PR, Chalmers RM, Tylera KM. Cryptosporidium pathogenicity and virulence. Clin Microbiol Rev. 2013;26:115-34.
38. Wanyiri J, Ward H. Molecular basis of Cryptosporidium-host cell interactions: recent advances and future prospects. Future Microbiol. 2006;1:201-8.
39. Yao L, Yin J, Zhang X, Liu Q, Li J, Chen L, et al. Cryptosporidium parvum: identification of a new surface adhesion protein on sporozoite and oocyst by screening of a phage-display cDNA library. Exp Parasitol. 2007;115:333-8.
40. Bhalchandra S, Ludington J, Coppens I, Ward HD. Identification and characterization of Cryptosporidium parvum Clec, a novel

C-type lectin domain-containing mucin-like glycoprotein. Infect Immun. 2013;81:3356-65.

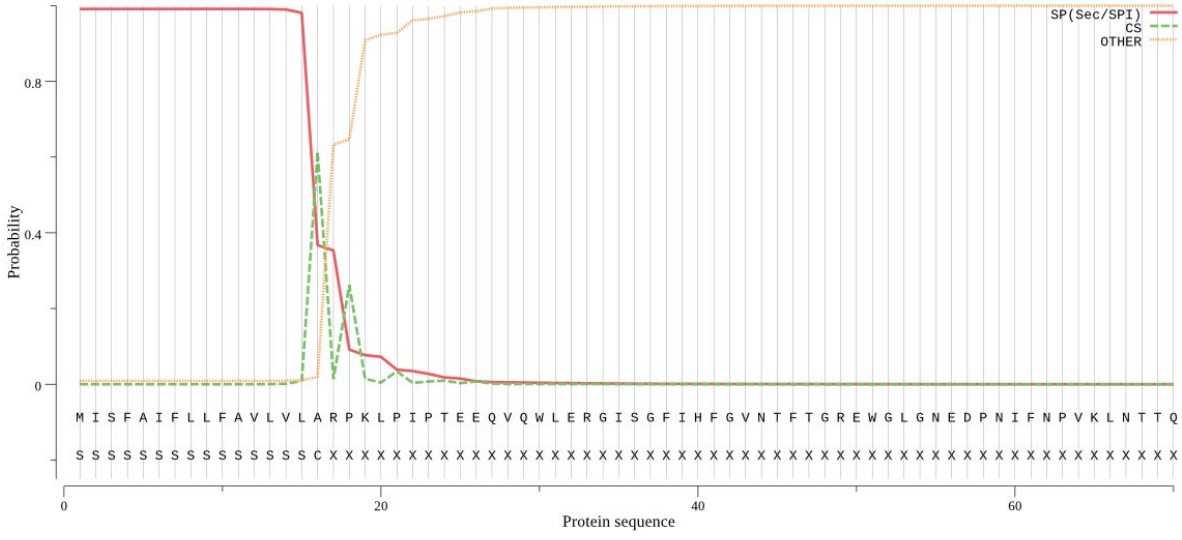
41. Ludington JG, Ward HD. The Cryptosporidium parvum C-type lectin CpClec mediates infection of intestinal epithelial cells via interactions with sulfated proteoglycans. Infect Immun. 2016;84:1593-1602.
42. Bhat N, Joe A, PereiraPerrin M, Ward HD. Cryptosporidium p30, a galactose/N-acetylgalactosamine-specific lectin, mediates infection in vitro. J Biol Chem. 2007;282:34877-87.
43. Tosini F, Agnoli A, Mele R, Gomez Morales MA, Pozio E. A new modular protein of Cryptosporidium parvum, with ricin B and LCCL domains, expressed in the sporozoite invasive stage. Mol Biochem Parasitol. 2004;134:137-47.
44. Langer RC, Schaefer DA, Riggs MW. Characterization of an intestinal epithelial cell receptor recognized by the Cryptosporidium parvum sporozoite ligand CSL. Infect Immun. 2001;69:1661-70.
45. Deng M, Rutherford MS, Abrahamsen MS. Host intestinal epithelial response to Cryptosporidium parvum. Adv Drug Deliv Rev. 2004;56:869-84.
46. Boulter-Bitzer JI, Lee H, Trevors JT. Molecular targets for detection and immunotherapy in Cryptosporidium parvum. Biotechnol Adv. 2007;25:13-44.
47. Naitza S, Spano F, Robson KJ, Crisanti A. The thrombospondin-related protein family of Apicomplexan parasites: the gears of the cell invasion machinery. Parasitol Today. 1998;14:479-84.
48. Kappe S, Bruderer T, Gantt S, Fujioka H, Nussenzweig V, Ménard R. Conservation of a gliding motility and cell invasion machinery in Apicomplexan parasites. J Cell Biol. 1999;147:937-44.
49. Roberts T, Stark D, Harkness J, Ellis J. Update on the pathogenic potential and treatment options for Blastocystis sp. Gut Pathog. 2014;6:17.
50. Tan KS. New insights on classification, identification, and clinical relevance of Blastocystis spp. Clin Microbiol Rev. 2008;21:639-65.
51. Stensvold CR, Clark CG. Current status of Blastocystis: a personal view. Parasitol Int. 2016;65:763-71.
52. Boorom KF, Smith H, Nimri L, Viscogliosi E, Spanakos G, Parkar U, et al. Oh my aching gut: irritable bowel syndrome, Blastocystis, and asymptomatic infection. Parasit Vectors. 2008;1:40.
53. Lanuza MD, Carbajal JA, Borrás R. Identification of surface coat carbohydrates in Blastocystis hominis by lectin probes. Int J Parasitol. 1996;26:527-32.
54. Wu B, Yin J, Texier C, Roussel M, Tan KS. Blastocystis legumain is localized on the cell surface, and specific inhibition of its activity implicates a pro-survival role for the enzyme. J Biol Chem. 2010;285:1790-8.
55. Wawrzyniak I, Texier C, Poirier P, Viscogliosi E, Tan KS, Delbac F, El Alaoui H. Characterization of two cysteine proteases secreted by Blastocystis ST7, a human intestinal parasite. Parasitol Int. 2012;61:437-42.
56. Denoeud F, Roussel M, Noel B, Wawrzyniak I, Da Silva C, Diogon M, et al. Genome sequence of the stramenopile Blastocystis, a human anaerobic parasite. Genome Biol. 2011;12:R29.
57. Yason JA, Tan KS. Membrane surface features of Blastocystis subtypes. Genes (Basel). 2018;9:417.
58. Eme L, Gentekaki E, Curtis B, Archibald JM, Roger AJ. Lateral gene transfer in the adaptation of the anaerobic parasite Blastocystis to the gut. Curr Biol. 2017;27:807-20.

Anexo 2. Gráficas de predicción del péptido señal

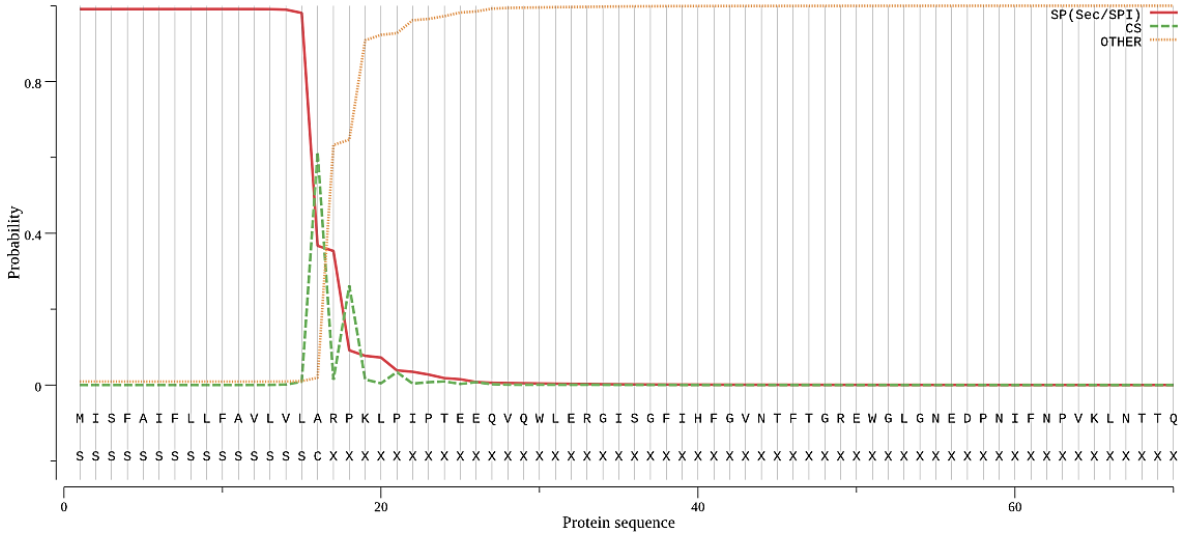
SignalP-5.0 prediction (Eukarya): 50177



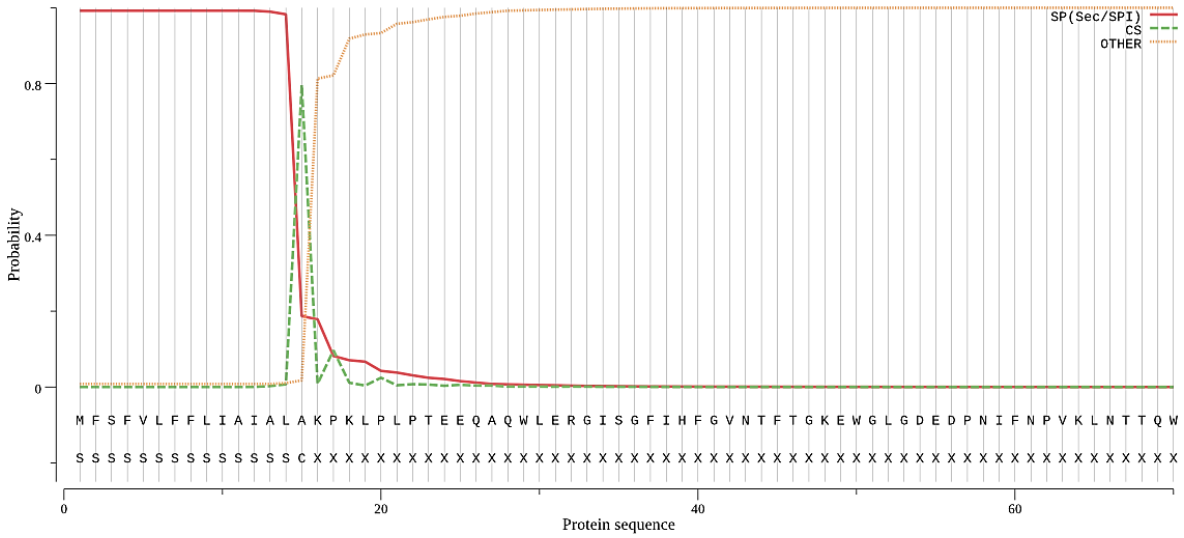
SignalP-5.0 prediction (Eukarya): 50610



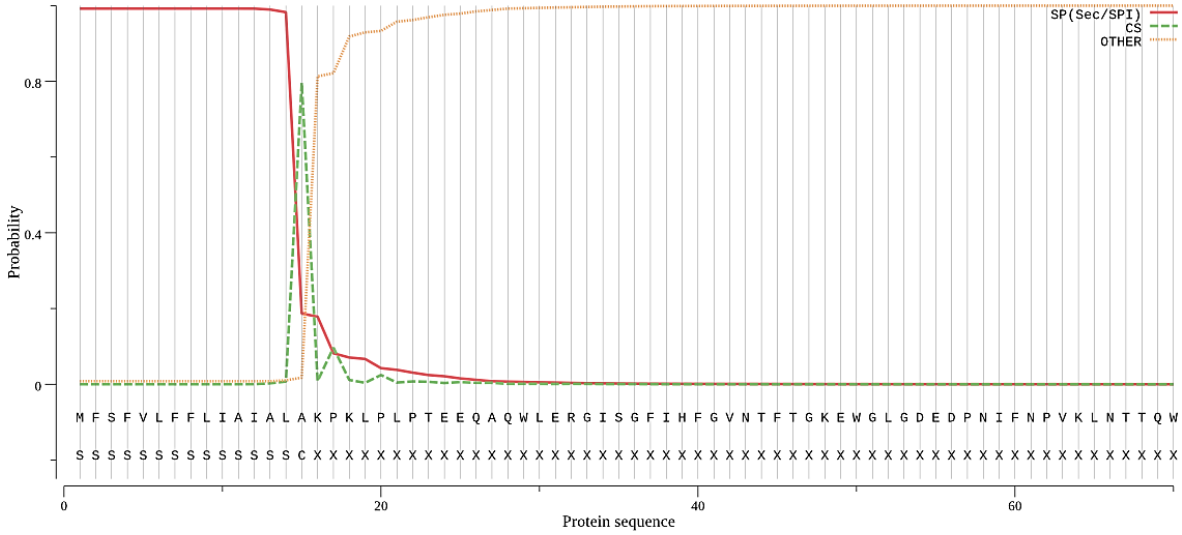
SignalP-5.0 prediction (Eukarya): JOEL36_ST1



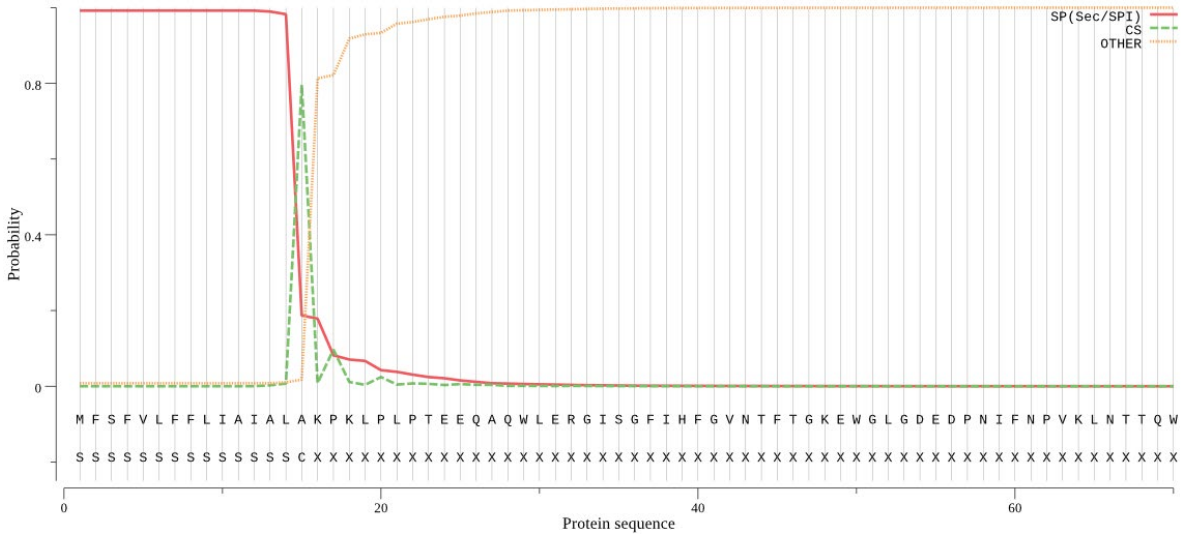
SignalP-5.0 prediction (Eukarya): CMX11_ST2



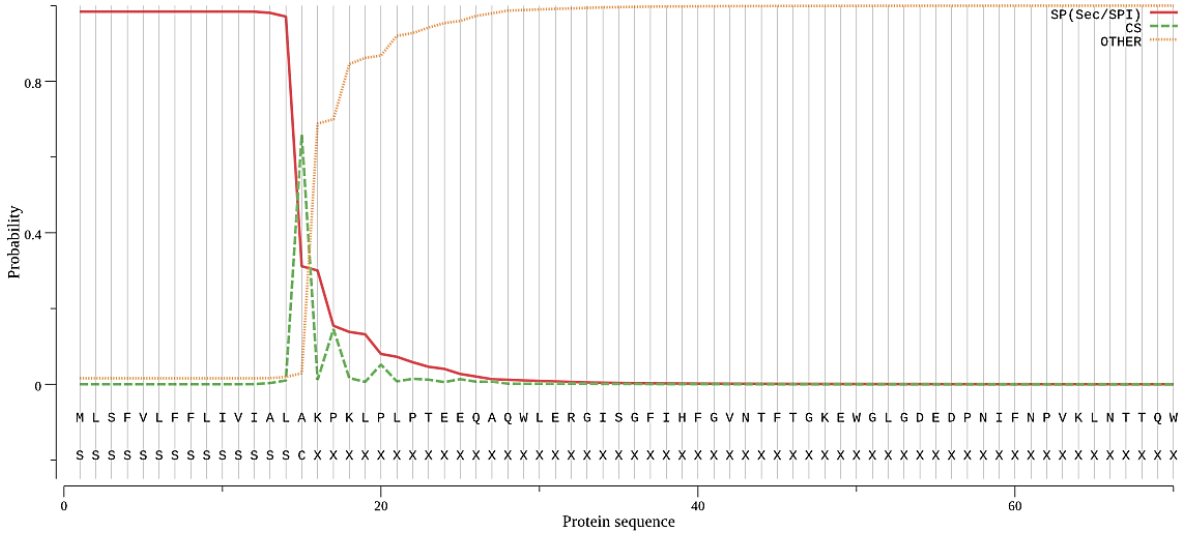
SignalP-5.0 prediction (Eukarya): Joel5B_ST2



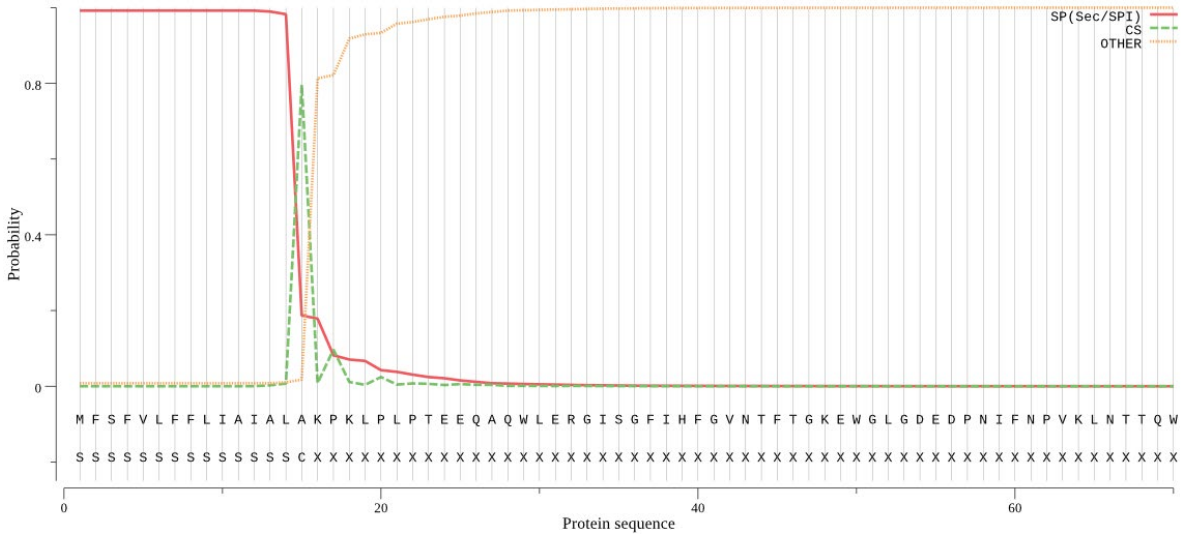
SignalP-5.0 prediction (Eukarya): Joel19B_ST2



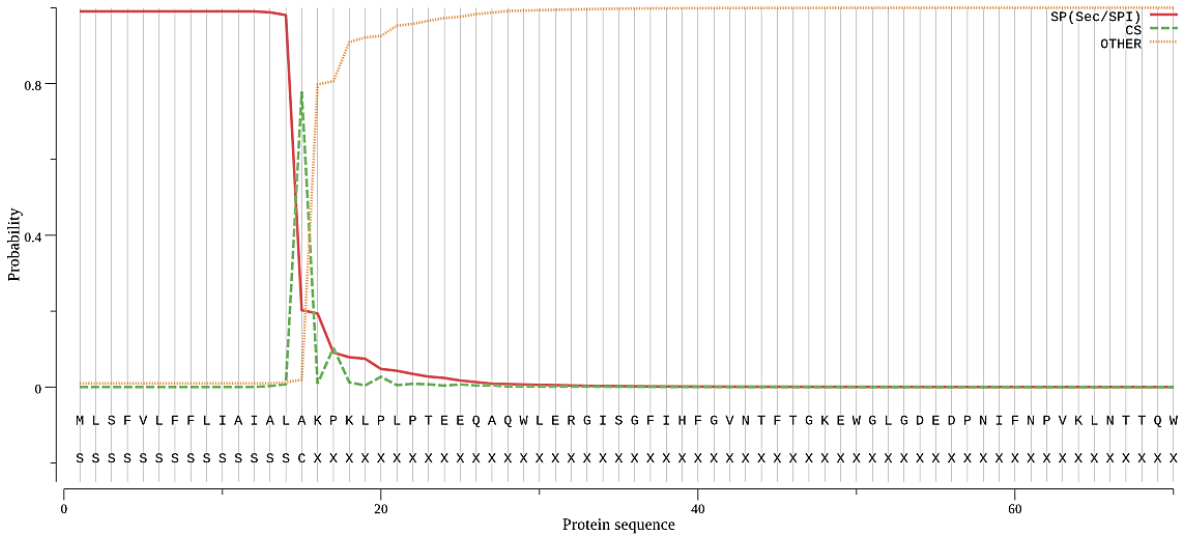
SignalP-5.0 prediction (Eukarya): Joel29B_ST2



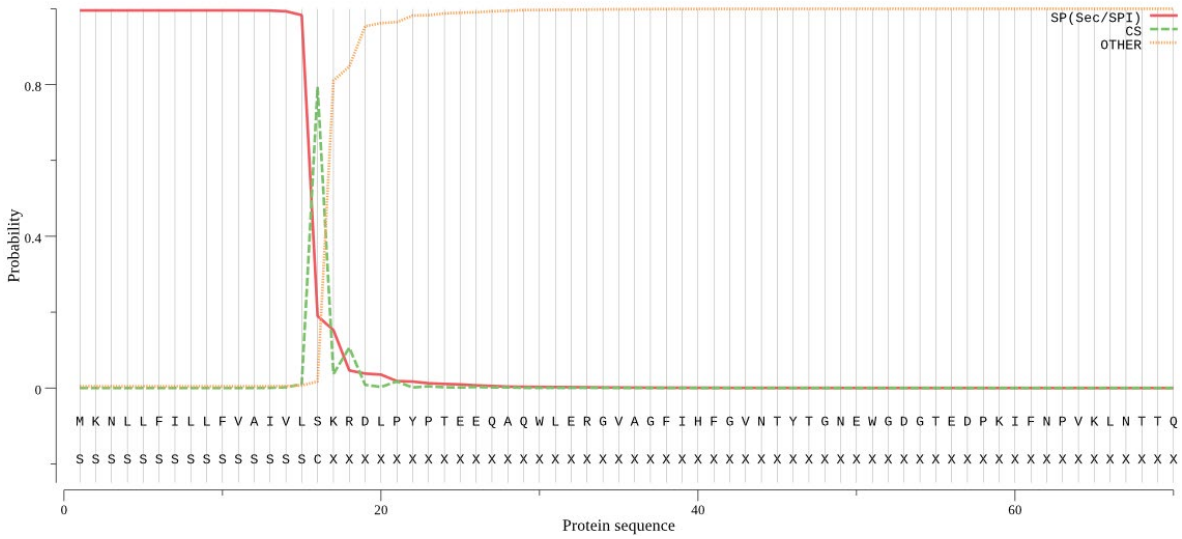
SignalP-5.0 prediction (Eukarya): Sequence



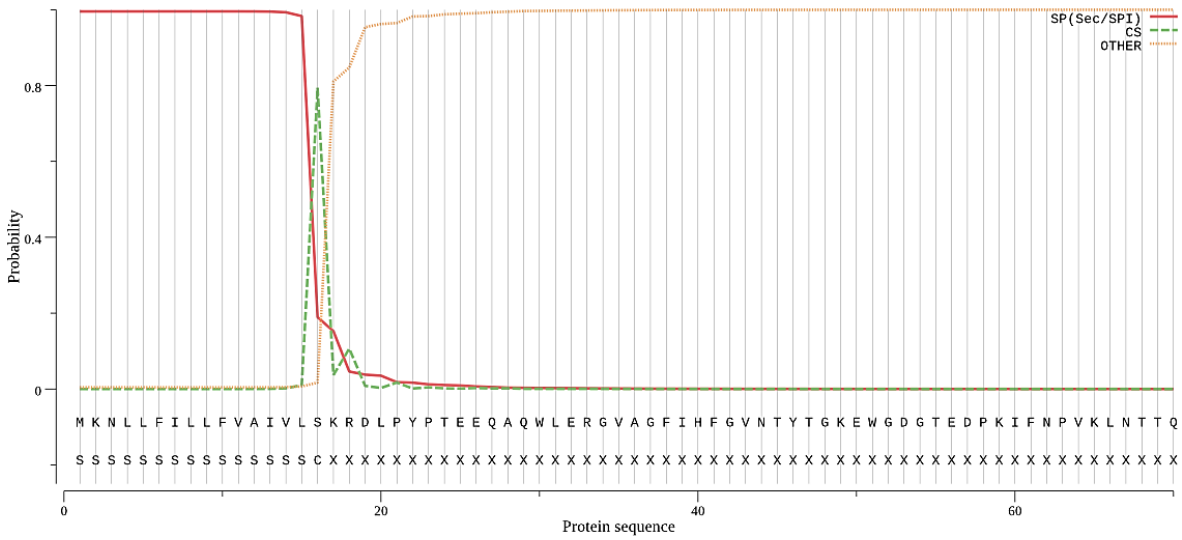
SignalP-5.0 prediction (Eukarya): Flemming_ST2_NCBI



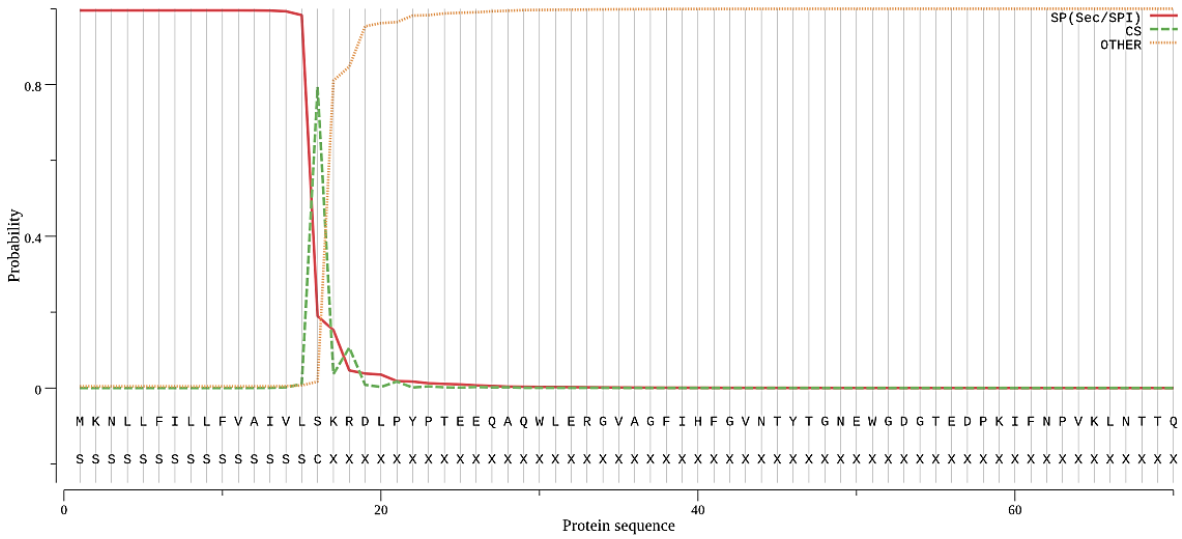
SignalP-5.0 prediction (Eukarya): CMX5_ST3



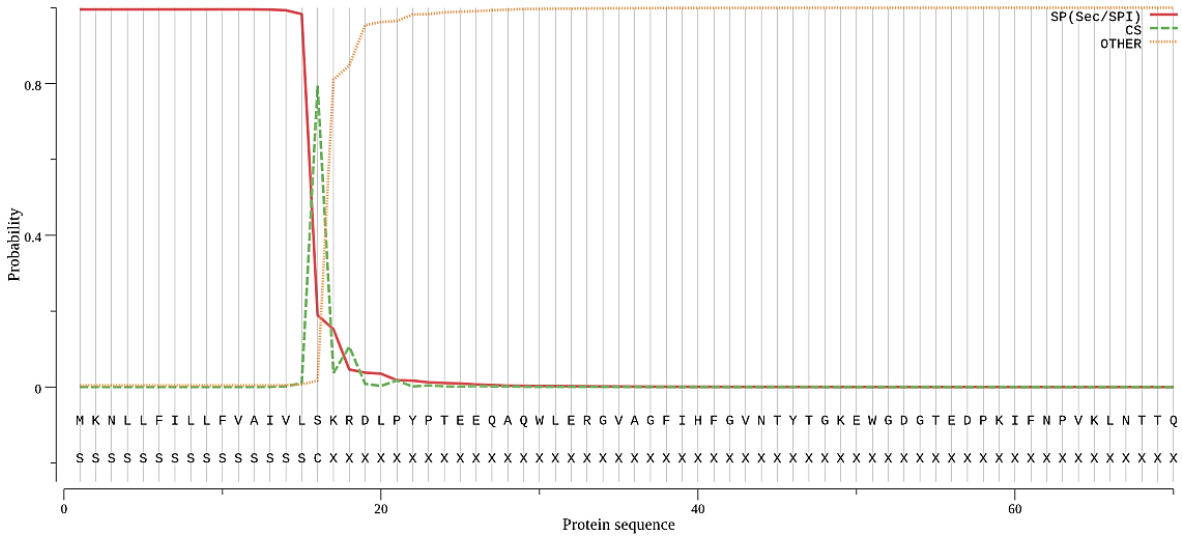
SignalP-5.0 prediction (Eukarya): ST3_NCBI



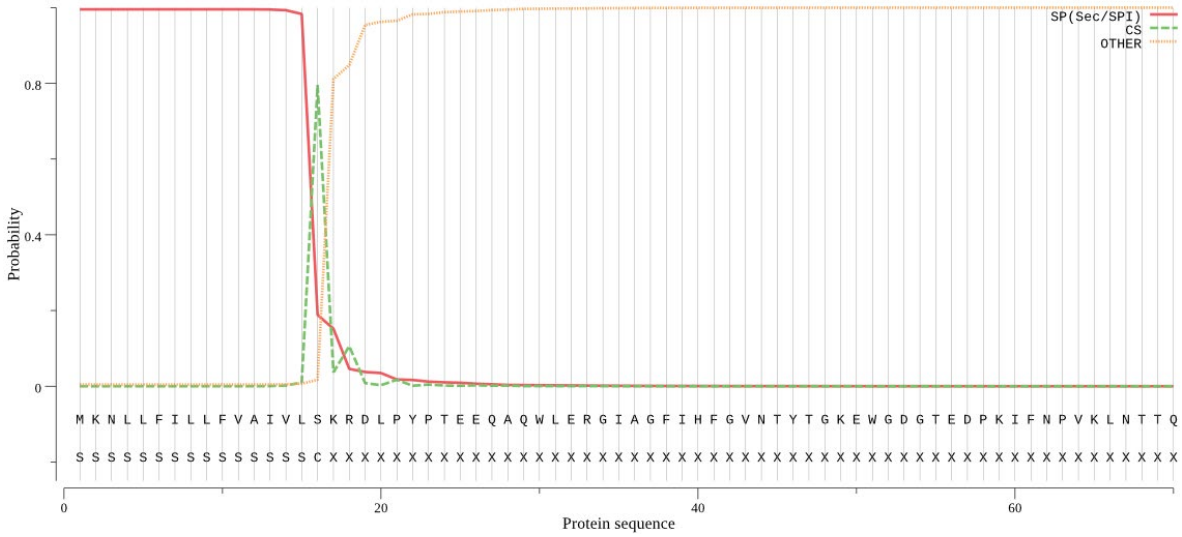
SignalP-5.0 prediction (Eukarya): Joel17A_ST3



SignalP-5.0 prediction (Eukarya): Joel42_ST3



SignalP-5.0 prediction (Eukarya): Joel43_ST3

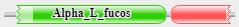


Anexo 3. Predicción de dominios conservados en ALFuc de Blastocystis

Sequence search results

Show the detailed description of this results page.

We found **2** Pfam-A matches to your search sequence (**all** significant)



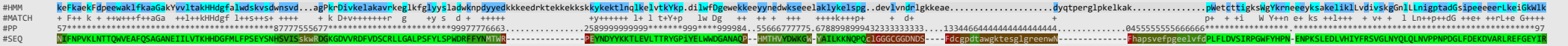
Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
Alpha_L_fucos	Alpha-L-fucosidase	Domain	CL0058	27	324	59	322	**87**	**346**	350	96.8	1.8e-27	n/a	Hide
F5_F8_type_C	F5/8 type C domain	Domain	CL0202	334	452	336	445	**2**	**118**	127	30.6	3.1e-07	n/a	Show



Pfam is part of the ELIXIR infrastructure

Pfam is an Elixir service **Read more**

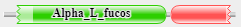
Comments or questions on the site? Send a mail to **pfam-help@ebi.ac.uk**.

European Molecular Biology Laboratory

Sequence search results

Show the detailed description of this results page.

We found **2** Pfam-A matches to your search sequence (**all** significant)



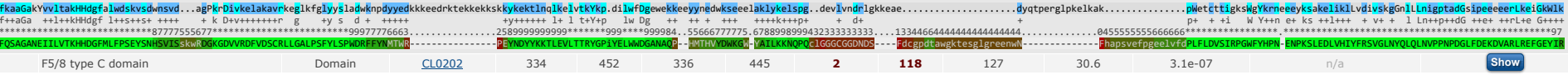
Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
Alpha_L_fucos	Alpha-L-fucosidase	Domain	CL0058	27	324	59	322	**87**	**346**	350	96.7	1.8e-27	n/a	Hide
F5_F8_type_C	F5/8 type C domain	Domain	CL0202	334	452	336	445	**2**	**118**	127	30.6	3.1e-07	n/a	Show



Pfam is part of the ELIXIR infrastructure

Pfam is an Elixir service **Read more**

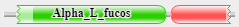
Comments or questions on the site? Send a mail to **pfam-help@ebi.ac.uk**.

European Molecular Biology Laboratory

Sequence search results

Show the detailed description of this results page.

We found **2** Pfam-A matches to your search sequence (**all** significant)



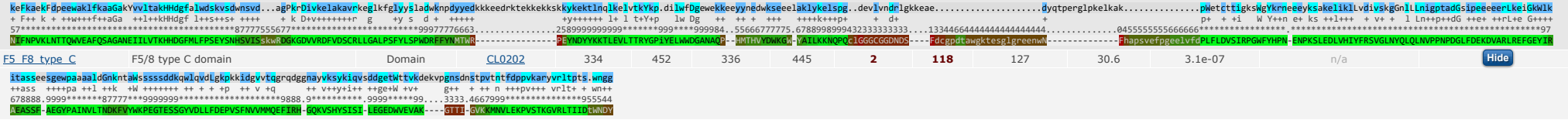
Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope Start	Envelope End	Alignment Start	Alignment End	HMM From	HMM To	HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
Alpha_L_fucos	Alpha-L-fucosidase	Domain	CL0058	27	324	59	322	**87**	**346**	350	96.8	1.8e-27	n/a	Hide
F5_F8_type_C	F5/8 type C domain	Domain	CL0202	334	452	336	445	**2**	**118**	127	30.6	3.1e-07	n/a	Hide



Sequence search results

Show the detailed description of this results page.

We found **2** Pfam-A matches to your search sequence (**1** significant and **1** insignificant)

Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
Alpha_L_fucos	Alpha-L-fucosidase	Domain	CL0058	22	325	59	323	**87**	**346**	350	101.1	8.3e-29	n/a	Hide

```
#HMM    keFkaekFdpeewaklfkaaGakVvvltakHHdgfalwdskvsd....wnsvdagPkrDivkelakavrkeglkfglyyslad.wknpdyyedkkkeedrktekkekkskkykektlnqlkelvtkYkp.dilwfDgewekkeeyynedwkseeelaklykelspgdevlvndrlgkkeae.....................dyqtperglpke.lkakpWetcttigksWgYkrneeeyksakeliklLvdivskgGnllLnigptadGsipeeeeerLkeiGkwlk
#MATCH  k F++ k + ++w+++f++ Ga  v+l+akHHdgf l++s++s+    +++ + + k D+v+e++++r+ g k  ly s  d +  +++++                +y+e++ + l+el t+Y+p   lw Dg   +  +++ +  +++   +k++k+++p+  +l    +g +ea                        +++tp  + p++ l  +p+  + +i   W Y+ n e+ k+ kel+++  + v+ +  l Ln++pt++G ++++ ++++ke G+++
#PP     67*****************************************655555554.55*****************************977776663.............259999***************999****98888..455666777774.8899999998..44333.3333333344456775555556666666666666666666634445*****************.9*****************************************986
#SEQ    KIFNPVKLNTTQWIEAFRSIGAEEVILVAKHHDGFMLFPSAYSNitvaYSTWR-DGKGDVVREFVQSCRALGTKASLYLSPWDRFFYVMTWR---------------PEYNEYYSHTLEELTTRYGPiYELWWDGANAQ--DHMTHVYDWQGK-YKIIKRNQPQ---cLGGG-cGGDEAGifdcgpdtawgqtesglgkeenwNFHTPSVEFPGKeLVFSPLFLDVSIRPGWFYHAN-ESPKTLKELVHIYFRSVGLNYQLQLNVPPTPEGLFDPRDVAVMKEFGAYIE
```

Insignificant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
F5_F8_type_C	F5/8 type C domain	Domain	CL0202	336	453	339	443	**4**	**115**	127	23.1	6.2e-05	n/a	Hide

```
#HMM    asseesgewpaaaldGnkntaWsssssddkqwlqvdLgkpkkidgvvtqgrqdggnayvksykiqvsddgetWttvkdekvpgnsdnstpvtntfdppvkaryvrltpts.w
#MATCH  ass  ++++pa +++  +k  +W+++++++++  +++ +++p  ++ v +q      +++v++y+i v++ g++W +v+     g++  + ++ n +d    +  vr+t+++ w
#PP     6777.88898888888655555***9999999*********************9888.9*********55.569****99.....3333.46679999999999*******99944
#SEQ    ASSF-TPGYPAINVVGDDKFIYWTPAQGEASGSIELVFEEPVHFNVVMMQEFIRH-GQRVSHYSIAVKG-GDQWVEVAR----STTI-SVKGMNVLDGTFYTDRVRVTIEDtK
```

Sequence search results

Show the detailed description of this results page.

We found **2** Pfam-A matches to your search sequence (**all** significant)

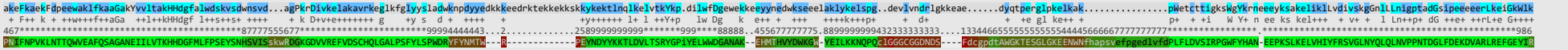
Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
Alpha_L_fucos	Alpha-L-fucosidase	Domain	CL0058	21	323	57	321	**86**	**346**	350	97.5	1.0e-27	n/a	Hide

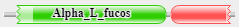


F5_F8_type_C	F5/8 type C domain	Domain	CL0202	334	451	334	444	1	**118**	127	30.6	3.0e-07	n/a	Hide

Sequence search results

Show the detailed description of this results page.

We found **2** Pfam-A matches to your search sequence (**all** significant)



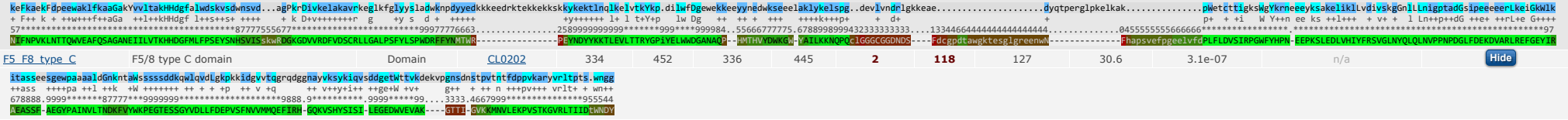
Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
Alpha_L_fucos	Alpha-L-fucosidase	Domain	CL0058	27	324	59	322	**87**	**346**	350	96.6	2.0e-27	n/a	Hide
F5_F8_type_C	F5/8 type C domain	Domain	CL0202	334	452	336	445	**2**	**118**	127	30.6	3.1e-07	n/a	Hide



Sequence search results

Show the detailed description of this results page.

We found **1** Pfam-A match to your search sequence (**all** significant)

Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
Alpha_L_fucos	Alpha-L-fucosidase	Domain	CL0058	21	323	57	321	**86**	**346**	350	98.4	5.4e-28	n/a	Hide

#HMM akeFkackFdpeewaklfkaaGakYvvltakHHdgfalwdskvsdwnsvd...agPkrDivkelakavrkeglkfglyyslad.wknpdyyedkkkeedrktekkekkskkykektlnqlkelvtkYkp.dilwfDgewekkeeyynedwkseeelaklykelspg..devlvndrlgkkeae......dyqtpergl.pkelkak..............pWetcttigksk.gYkrneeeyksakeliklLvdivskgGnllLnigptadGsipeeeeerLkeiGkkwlk
#MATCH + F++ k + ++w+++f++aGa ++l++kHHdgf l++s+++s+ ++++ + k D+v+e+++++++ g +y s d + ++++ + +yv++++++ l+ l ++Y+p lw Dg k e++ + +++ ++++k+++p+ + d+ + +e gl ke++ + p+ + +i W Y+ n ee ks kel+++ + v+ + l Ln++p+ dG +ee+ ++rL+e G++++
#PP 467***8777755567 7**********************99994444443...2.....2589999999999******999****88888..455677777775.889999999943233333333....13344665555555555554444566666777777777*****************.**986
#SEQ PNIFNPVKLNTTQWVEAFQSAGANEIILVTKHHDGFMLFPSEYSNHSVISSkw.DGKGDVVREFVDSCHQLGALPSFYLSPWDRYFYNMTW----R----------PEYNDYYKKTLDVLTSRYGP1YELWWDGANAK--EHMTHVVYDWKG.-YEILKKNQPQ2lGGGCGGDNDS----PdcgpdtAWGKTESGLGKEENWNfhapsvefpged1vfdPLFLDVSIRPGWFYHAN-EEPKSLKELVHIYFRSVGLNYQLQLNVPPNTDGLFDEKDVARLREFGEYIQ

Sequence search results

Show the detailed description of this results page.

We found **1** Pfam-A match to your search sequence (**all** significant)

Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
Alpha_L_fucos	Alpha-L-fucosidase	Domain	CL0058	22	308	58	307	**86**	**330**	350	91.3	8.3e-26	n/a	Hide

Sequence search results

Show the detailed description of this results page.

We found **2** Pfam-A matches to your search sequence (**1** significant and **1** insignificant)

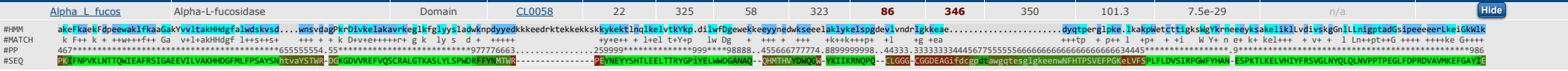
Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

Show or hide all alignments.

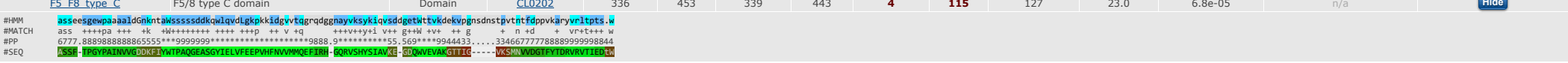
Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
Alpha_L_fucos	Alpha-L-fucosidase	Domain	CL0058	22	325	58	323	**86**	**346**	350	101.3	7.5e-29	n/a	Hide



Insignificant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
F5_F8_type_C	F5/8 type C domain	Domain	CL0202	336	453	339	443	**4**	**115**	127	23.0	6.8e-05	n/a	Hide



Sequence search results

Show the detailed description of this results page.

We found **2** Pfam-A matches to your search sequence (**all** significant)

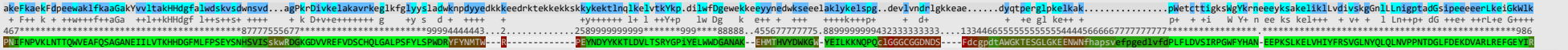
Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
Alpha_L_fucos	Alpha-L-fucosidase	Domain	CL0058	21	323	57	321	**86**	**346**	350	97.5	1.0e-27	n/a	Hide

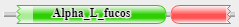


F5_F8_type_C	F5/8 type C domain	Domain	CL0202	334	451	334	444	1	**118**	127	30.6	3.0e-07	n/a	Hide

Sequence search results

Show the detailed description of this results page.

We found **2** Pfam-A matches to your search sequence (**all** significant)



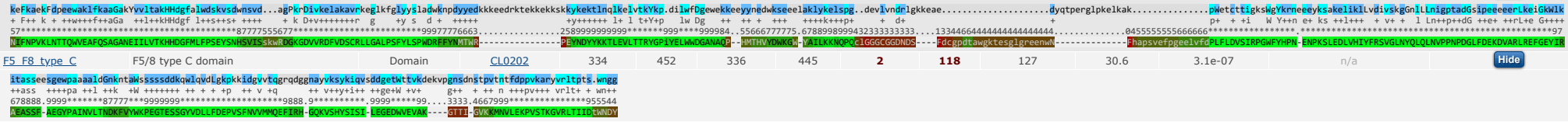
Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope Start	Envelope End	Alignment Start	Alignment End	HMM From	HMM To	HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
Alpha_L_fucos	Alpha-L-fucosidase	Domain	CL0058	27	324	59	322	**87**	**346**	350	96.8	1.8e-27	n/a	Hide
F5_F8_type_C	F5/8 type C domain	Domain	CL0202	334	452	336	445	**2**	**118**	127	30.6	3.1e-07	n/a	Hide



Sequence search results

Show the detailed description of this results page.

We found **2** Pfam-A matches to your search sequence (**1** significant and **1** insignificant)

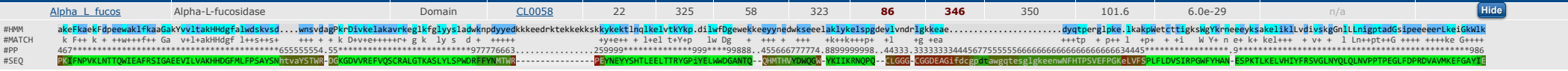
Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

Show or hide all alignments.

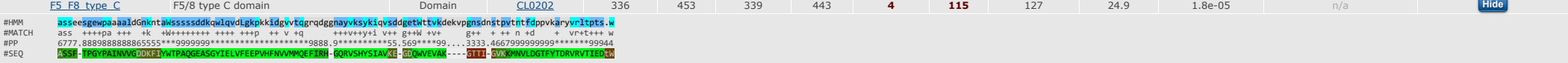
Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
Alpha_L_fucos	Alpha-L-fucosidase	Domain	CL0058	22	325	58	323	**86**	**346**	350	101.6	6.0e-29	n/a	Hide



Insignificant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
F5_F8_type_C	F5/8 type C domain	Domain	CL0202	336	453	339	443	**4**	**115**	127	24.9	1.8e-05	n/a	Hide



Sequence search results

Show the detailed description of this results page.

We found **2** Pfam-A matches to your search sequence (**all** significant)

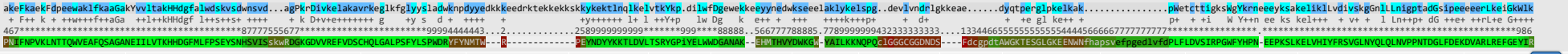
Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
Alpha_L_fucos	Alpha-L-fucosidase	Domain	CL0058	21	323	57	321	**86**	**346**	350	98.0	7.6e-28	n/a	Hide



F5_F8_type_C	F5/8 type C domain	Domain	CL0202	334	451	334	444	1	**118**	127	30.6	3.0e-07	n/a	Hide

Pfam is part of the ELIXIR infrastructure

Pfam is an Elixir service **Read more**

Comments or questions on the site? Send a mail to **pfam-help@ebi.ac.uk**.

European Molecular Biology Laboratory

Sequence search results

Show the detailed description of this results page.

We found **2** Pfam-A matches to your search sequence (**1** significant and **1** insignificant)

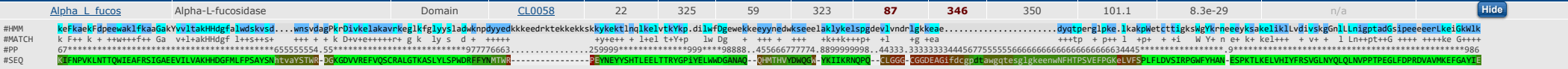
Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

Show or hide all alignments.

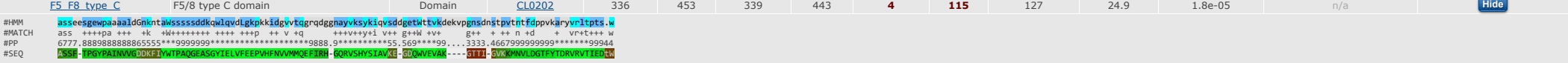
Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
Alpha_L_fucos	Alpha-L-fucosidase	Domain	CL0058	22	325	59	323	**87**	**346**	350	101.1	8.3e-29	n/a	Hide



Insignificant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
F5_F8_type_C	F5/8 type C domain	Domain	CL0202	336	453	339	443	**4**	**115**	127	24.9	1.8e-05	n/a	Hide



Pfam is part of the ELIXIR infrastructure

Pfam is an Elixir service Read more

Comments or questions on the site? Send a mail to **pfam-help@ebi.ac.uk**.

European Molecular Biology Laboratory

Sequence search results

Show the detailed description of this results page.

We found **2** Pfam-A matches to your search sequence (**all** significant)

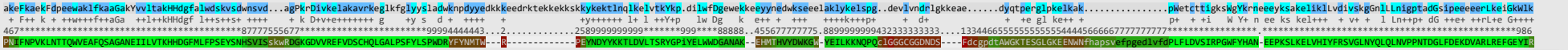
Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
Alpha_L_fucos	Alpha-L-fucosidase	Domain	CL0058	21	323	57	321	**86**	**346**	350	97.5	1.0e-27	n/a	Hide



F5_F8_type_C	F5/8 type C domain	Domain	CL0202	334	451	334	444	1	**118**	127	30.6	3.0e-07	n/a	Hide

Sequence search results

Show the detailed description of this results page.

We found **2** Pfam-A matches to your search sequence (**all** significant)

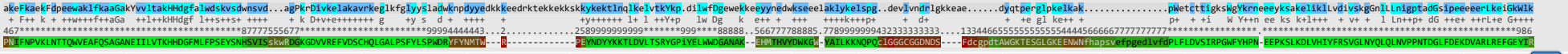
Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

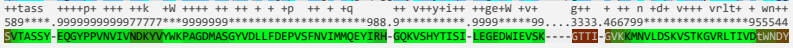
Significant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
Alpha_L_fucos	Alpha-L-fucosidase	Domain	CL0058	21	323	57	321	**86**	**346**	350	97.4	1.1e-27	n/a	Hide



F5_F8_type_C	F5/8 type C domain	Domain	CL0202	334	451	334	444	1	**118**	127	31.3	1.8e-07	n/a	Hide



Pfam is part of the ELIXIR infrastructure

Pfam is an Elixir service **Read more**

Comments or questions on the site? Send a mail to **pfam-help@ebi.ac.uk**.

European Molecular Biology Laboratory

Anexo 4. Cálculo de masa molecular

Compute pI/Mw

Compute pI/Mw

Theoretical pI/Mw (average) for the user-entered sequence:

```
        10         20         30         40         50         60
MISFAIFLLF AVLVLARPKL PIPTEEQVQW LERGISGFIH FGVNTFTGRE WGLGNEDPNI

        70         80         90        100        110        120
FNPVKLNTTQ WVEAFQSAGA NEIILVTKHH DGFMLFPSEY SNHSVISSKW RDGKGDVVRD

       130        140        150        160        170        180
FVDSCRLLGA LPSFYLSPWD RFFYNMTWRP EYNDYYKKTL EVLTTRYGPI YELWWDGANA

       190        200        210        220        230        240
QPHMTHVYDW KGWYAILKKN QPQCLGGGCG GDNDSFDCGP DTAWGKTESG LGREENWNFH

       250        260        270        280        290        300
APSVEFPGEE LVFDPLFLDV SIRPGWFYHP NENPKSLEDL VHIYFRSVGL NYQLQLNVPP

       310        320        330        340        350        360
NPDGLFDEKD VARLREFGEY IRKTFAQDEA RKATAAEASS FAEGYPAINV LTNDKFVYWK

       370        380        390        400        410        420
PEGTESSGYV DLLFDEPVSF NVVMMQEFIR HGQKVSHYSI SILEGEDWVE VAKGTTIGVK

       430        440        450        460
KMNVLEKPVS TKGVRLTIID TWNDYVPEIS RIGLFTSEYY
```

Theoretical pI/Mw: 5.02 / 52717.42

Compute pI/Mw

Compute pI/Mw

Theoretical pI/Mw (average) for the user-entered sequence:

```
        10         20         30         40         50         60
MISFAIFLLF AVLVLARPKL PIPTEEQVQW LERGISGFIH FGVNTFTGRE WGLGNEDPNI

        70         80         90        100        110        120
FNPVKLNTTQ WVEAFQSAGA NEIILVTKHH DGFMLFPSEY SNHSVISSKW RDGKGDVVRD

       130        140        150        160        170        180
FVDSCRLLGA LPSFYLSPWD RFFYNMTWRP EYNDYYKKTL EVLTTRYGPI YELWWDGANA

       190        200        210        220        230        240
QPHMTHVYDW KGWYAILKKN QPQCLGGGCG GDNDSFDCGP DTAWGKTESG LGREENWNFH

       250        260        270        280        290        300
APSVEFPGEE LVFDPLFLDV SIRPGWFYHP NENPKSLEDL VHIYFRSVGL NYQLQLNVPP

       310        320        330        340        350        360
NPDGLFDEKD VARLREFGEY IRKTFAQDEA RKATAAEASS FAEGYPAINV LTNDKFVYWK

       370        380        390        400        410        420
PEGTESSGYV DLLFDEPVSF NVVMMQEFIR HGQKVSHYSI SILEGEDWVE VAKGTTIGVK

       430        440        450        460
KMNVLEKPVS TKGVRLTIID TWNDYVPEIS RIGLFTSEYY
```

Theoretical pI/Mw: 5.02 / 52717.42

Compute pI/Mw

Compute pI/Mw

Theoretical pI/Mw (average) for the user-entered sequence:

```
        10         20         30         40         50         60
MKNLLFILLF VAIVLSKRDL PYPTEEQAQW LERGVAGFIH FGVNTYTGKE WGDGTEDPKI

        70         80         90        100        110        120
FNPVKLNTTQ WIEAFRSIGA EEVILVAKHH DGFMLFPSAY SNHTVAYSTW RDGKGDVVRE

       130        140        150        160        170        180
FVQSCRALGT KASLYLSPWD RFFYNMTWRP EYNEYYSHTL EELTTRYGPI YELWWDGANT

       190        200        210        220        230        240
QQHMTHVYDW QGWYKIIKRN QPQCLGGGCG GDEAGIFDCG PDTAWGQTES GLGKEENWNF

       250        260        270        280        290        300
HTPSVEFPGK ELVFSPLFLD VSIRPGWFYH ANESPKTLKE LVHIYFRSVG LNYQLQLNVP

       310        320        330        340        350        360
PTPEGLFDPR DVAVMKEFGA YIEEVFAHDE ARKAVDAVAS SFTPGYPAIN VVGDDKFIYW

       370        380        390        400        410        420
TPAQGEASGY IELVFEEPVH FNVVMMQEFI RHGQRVSHYS IAVKEGDQWV EVAKGTTIGV

       430        440        450        460
KKMNVLDGTF YTDRVRVTIE DTWEDYPPEI TRIGLFNSEL Y
```

Theoretical pI/Mw: 5.07 / 52888.55

Compute pI/Mw

Compute pI/Mw

Theoretical pI/Mw (average) for the user-entered sequence:

```
        10         20         30         40         50         60
MISFAIFLLF AVLVLARPKL PIPTEEQVQW LERGISGFIH FGVNTFTGRE WGLGNEDPNI

        70         80         90        100        110        120
FNPVKLNTTQ WVEAFQSAGA NEIILVTKHH DGFMLFPSEY SNHSVISSKW RDGKGDVVRD

       130        140        150        160        170        180
FVDSCRLLGA LPSFYLSPWD RFFYNMTWRP EYNDYYKKTL EVLTTRYGPI YELWWDGANA

       190        200        210        220        230        240
QPHMTHVYDW KGWYAILKKN QPQCLGGGCG GDNDSFDCGP DTAWGKTESG LGREENWNFH

       250        260        270        280        290        300
APSVEFPGEE LVFDPLFLDV SIRPGWFYHP NENPKSLEDL VHIYFRSVGL NYQLQLNVPP

       310        320        330        340        350        360
NPDGLFDEKD VARLREFGEY IRKTFAQDEA RKATAAEASS FAEGYPAINV LTNDKFVYWK

       370        380        390        400        410        420
PEGTESSGYV DLLFDEPVSF NVVMMQEFIR HGQKVSHYSI SILEGEDWVE VAKGTTIGVK

       430        440        450        460
KMNVLEKPVS TKGVRLTIID TWNDYVPEIS RIGLFTSEYY
```

Theoretical pI/Mw: 5.02 / 52717.42

Compute pI/Mw

Compute pI/Mw

Theoretical pI/Mw (average) for the user-entered sequence:

```
        10         20         30         40         50         60
MKNLLFILLF VAIVLSKRDL PYPTEEQAQW LERGVAGFIH FGVNTYTGNE WGDGTEDPKI

        70         80         90        100        110        120
FNPVKLNTTQ WIEAFRSIGA EEVILVAKHH DGFMLFPSAY SNHTVAYSTW RDGKGDVVRE

       130        140        150        160        170        180
FVQSCRALGT KASLYLSPWD RFFYNMTWRP EYNEYYSHTL EELTTRYGPI YELWWDGANA

       190        200        210        220        230        240
QQHMTHVYDW QGWYKIIKRN QPQCLGGGCG GDEAGIFDCG PDTAWGQTES GLGKEENWNF

       250        260        270        280        290        300
HTPSVEFPGK ELVFSPLFLD VSIRPGWFYH ANESPKTLKE LVHIYFRSVG LNYQLQLNVP

       310        320        330        340        350        360
PTPEGLFDPR DVAVMKEFGA YIEEVFAHDE ARKAVDAVAS SFTPGYPAIN VVGDDKFIYW

       370        380        390        400        410        420
TPAQGEASGS IELVFEEPVH FNVVMMQEFI RHGQRVSHYS IAVKEGDQWV EVAKGTTIGV

       430        440        450        460
KKMNVLDGTF YTDRVRVTIE DTWEDYPPEI TRIGLFNSEL Y
```

Theoretical pI/Mw: 5.02 / 52768.36

Compute pI/Mw

Compute pI/Mw

Theoretical pI/Mw (average) for the user-entered sequence:

```
        10         20         30         40         50         60
MFSFVLFFLI AIALAKPKLP LPTEEQAQWL ERGISGFIHF GVNTFTGKEW GLGDEDPNIF

        70         80         90        100        110        120
NPVKLNTTQW VEAFQSAGAN EIILVTKHHD GFMLFPSEYS NHSVISSKWR DGKGDVVREF

       130        140        150        160        170        180
VDSCHQLGAL PSFYLSPWDR YFYNMTWRPE YNDYYKKTLD VLTSRYGPIY ELWWDGANAK

       190        200        210        220        230        240
EHMTHVYDWK GWYEILKKNQ PQCLGGGCGG DNDSFDCGPD TAWGKTESGL GKEENWNFHA

       250        260        270        280        290        300
PSVEFPGEDL VFDPLFLDVS IRPGWFYHAN EEPKSLKELV HIYFRSVGLN YQLQLNVPPN

       310        320        330        340        350        360
TDGLFDEKDV ARLREFGEYI RETFKHDEAR KAASVTASSY EQGYPPVNVI VNDKYVYWKP

       370        380        390        400        410        420
AGDMASGYVD LLFDEPVSFN VIMMQEFIRH GQKVSHYTIS ILEGEDWIEV SKGTTIGVKK

       430        440        450
INVLDSKVST KGVRLTIVDT WNDYVPEISR IGLFTSEYY
```

Theoretical pI/Mw: 5.06 / 52717.33

Compute pI/Mw

Compute pI/Mw

Theoretical pI/Mw (average) for the user-entered sequence:

```
        10         20         30         40         50         60
MLSFVLFFLI AIALAKPKLP LPTEEQAQWL ERGISGFIHF GVNTFTGKEW GLGDEDPNIF

        70         80         90        100        110        120
NPVKLNTTQW VEAFQSAGAN EIILVTKHHD GFMLFPSEYS NHSVISSKWR DGKGDVVREF

       130        140        150        160        170        180
VDSCHQLGAL PSFYLSPWDR YFYNMTWRPE YNDYYKKTLD VLTSRYGPIY ELWWDGANAK

       190        200        210        220        230        240
EHMTHVYDWK GWYEILKKNQ PQCLGGGCGG DNDSFDCGPD TAWGKTESGL GKEENWNFHA

       250        260        270        280        290        300
PSVEFPGEDL VFDPLFLDVS IRPGWFYHAN EEPKSLKELV HIYFRSVGLN YQLQLNVPPN

       310        320        330        340        350        360
TDGLFDEKDV ARLREFGEYI RETFKHDEAR KASSVTASSY EQGYPPVNVI VNDKYVYWKP

       370        380        390        400        410        420
AGDMASGYVD LLFDEPVSFN VIMMQEFIRH GQKVSHYTIS ILEGEDWIEV SKGTTIGVKK

       430        440        450
INVLDSKVST KGVRLTIVDT WNDYVPEISR IGLFTSEYY
```

Theoretical pI/Mw: 5.06 / 52699.31

Compute pI/Mw

Compute pI/Mw

Theoretical pI/Mw (average) for the user-entered sequence:

```
        10         20         30         40         50         60
MFSFVLFFLI AIALAKPKLP LPTEEQAQWL ERGISGFIHF GVNTFTGKEW GLGDEDPNIF

        70         80         90        100        110        120
NPVKLNTTQW VEAFQSAGAN EIILVTKHHD GFMLFPSEYS NHSVISSKWR DGKGDVVREF

       130        140        150        160        170        180
VDSCHQLGAL PSFYLSPWDR YFYNMTWRPE YNDYYKKTLD VLTSRYGPIY ELWWDGANAK

       190        200        210        220        230        240
EHMTHVYDWK GWYAILKKNQ PQCLGGGCGG DNDSFDCGPD TAWGKTESGL GKEENWNFHA

       250        260        270        280        290        300
PSVEFPGEDL VFDPLFLDVS IRPGWFYHPN EEPKSLKELV HIYFRSVGLN YQLQLNVPPN

       310        320        330        340        350        360
TDGLFDEKDV ARLREFGEYI RETFKHDEAR KASSVTASSY EQGYPPVNVI VNDKYVYWKP

       370        380        390        400        410        420
AGDMASGYVD LLFDEPVSFN VIMMQEFIRH GQKVSHYTIS ILEGEDWIEV SKGTTIGVKK

       430        440        450
INVLDSKVST KGVRLTIVDT WNDYVPEISR IGLFTSEYY
```

Theoretical pI/Mw: 5.10 / 52701.33

Compute pI/Mw

Compute pI/Mw

Theoretical pI/Mw (average) for the user-entered sequence:

```
        10         20         30         40         50         60
MKNLLFILLF VAIVLSKRDL PYPTEEQAQW LERGVAGFIH FGVNTYTGNE WGDGTEDPKI

        70         80         90        100        110        120
FNPVKLNTTQ WIEAFRSIGA EEVILVAKHH DGFMLFPSAY SNHTVAYSTW RDGKGDVVRE

       130        140        150        160        170        180
FVQSCRALGT KASLYLSPWD RFFYNMTWRP EYNEYYSHTL EELTTRYGPI YELWWDGANA

       190        200        210        220        230        240
QQHMTHVYDW QGWYKIIKRN QPQCLGGGCG GDEAGIFDCG PDTAWGQTES GLGKEENWNF

       250        260        270        280        290        300
HTPSVEFPGK ELVFSPLFLD VSIRPGWFYH ANESPKTLKE LVHIYFRSVG LNYQLQLNVP

       310        320        330        340        350        360
PTPEGLFDPR DVAVMKEFGA YIEEVFAHDE ARKAVDAVAS SFTPGYPAIN VVGDDKFIYW

       370        380        390        400        410        420
TPAQGEASGY IELVFEEPVH FNVVMMQEFI RHGQRVSHYS IAVKEGDQWV EVAKGTTIGV

       430        440        450        460
KKMNVLDGTF YTDRVRVTIE DTWEDYPPEI TRIGLFNSEL Y
```

Theoretical pI/Mw: 5.02 / 52844.46

Compute pI/Mw

Compute pI/Mw

Theoretical pI/Mw (average) for the user-entered sequence:

```
        10         20         30         40         50         60
MFSFVLFFLI AIALAKPKLP LPTEEQAQWL ERGISGFIHF GVNTFTGKEW GLGDEDPNIF

        70         80         90        100        110        120
NPVKLNTTQW VEAFQSAGAN EIILVTKHHD GFMLFPSEYS NHSVISSKWR DGKGDVVREF

       130        140        150        160        170        180
VDSCHQLGAL PSFYLSPWDR YFYNMTWRPE YNDYYKKTLD VLTSRYGPIY ELWWDGANAK

       190        200        210        220        230        240
EHMTHVYDWK GWYEILKKNQ PQCLGGGCGG DNDSFDCGPD TAWGKTESGL GKEENWNFHA

       250        260        270        280        290        300
PSVEFPGEDL VFDPLFLDVS IRPGWFYHAN EEPKSLKELV HIYFRSVGLN YQLQLNVPPN

       310        320        330        340        350        360
TDGLFDEKDV ARLREFGEYI RETFKHDEAR KAASVTASSY EQGYPPVNVI VNDKYVYWKP

       370        380        390        400        410        420
AGDMASGYVD LLFDEPVSFN VIMMQEFIRH GQKVSHYTIS ILEGEDWIEV SKGTTIGVKK

       430        440        450
INVLDSKVST KGVRLTIVDT WNDYVPEISR IGLFTSEYY
```

Theoretical pI/Mw: 5.06 / 52717.33

Compute pI/Mw

Compute pI/Mw

Theoretical pI/Mw (average) for the user-entered sequence:

```
        10         20         30         40         50         60
MLSFVLFFLI VIALAKPKLP LPTEEQAQWL ERGISGFIHF GVNTFTGKEW GLGDEDPNIF

        70         80         90        100        110        120
NPVKLNTTQW VEAFQSAGAN EIILVTKHHD GFMLFPSEYS NHSVISSKWR DGKGDVVREF

       130        140        150        160        170        180
VDSCHQLGAL PSFYLSPWDR YFYNMTWRPE YNDYYKKTLD VLTSRYGPIY ELWWDGANAK

       190        200        210        220        230        240
EHMTHVYDWK GWYAILKKNQ PQCLGGGCGG DNDSFDCGPD TAWGKTESGL GKEENWNFHA

       250        260        270        280        290        300
PSVEFPGEDL VFDPLFLDVS IRPGWFYHPN EEPKSLKDLV HIYFRSVGLN YQLQLNVPPN

       310        320        330        340        350        360
TDGLFDEKDV ARLREFGEYI RETFKHDEAR KASSVTASSY EQGYPPVNVI VNDKYVYWKP

       370        380        390        400        410        420
AGDMASGYVD LLFDEPVSFN VIMMQEYIRH GQKVSHYTIS ILEGEDWIEV SKGTTIGVKK

       430        440        450
MNVLDSKVST KGVRLTIVDT WNDYVPEISR IGLFTSEYY
```

Theoretical pI/Mw: 5.09 / 52715.37

Compute pI/Mw

Compute pI/Mw

Theoretical pI/Mw (average) for the user-entered sequence:

```
        10         20         30         40         50         60
MISFAIFLLF AVLVLARPKL PIPTEEQVQW LERGISGFIH FGVNTFTGRE WGLGNEDPNI

        70         80         90        100        110        120
FNPVKLNTTQ WVEAFQSAGA NEIILVTKHH DGFMLFPSEY SNHSVISSKW RDGKGDVVRD

       130        140        150        160        170        180
FVDSCRLLGA LPSFYLSPWD RFFYNMTWRP EYNDYYKKTL EVLTTRYGPI YELWWDGANA

       190        200        210        220        230        240
QPHMTHVYDW KGWYAILKKN QPQCLGGGCG GDNDSFDCGP DTAWGKTESG LGREENWNFH

       250        260        270        280        290        300
APSVEFPGEE LVFDPLFLDV SIRPGWFYHP NEEPKSLEDL VHIYFRSVGL NYQLQLNVPP

       310        320        330        340        350        360
NPDGLFDEKD VARLREFGEY IRKTFAQDEA RKATAAEASS FAEGYPAINV LTNDKFVYWK

       370        380        390        400        410        420
PEGTESSGYV DLLFDEPVSF NVVMMQEFIR HGQKVSHYSI SILEGEDWVE VAKGTTIGVK

       430        440        450        460
KMNVLEKPVS TKGVRLTIID TWNDYVPEIS RIGLFTSEYY
```

Theoretical pI/Mw: 4.98 / 52732.43

Anexo 5. Artículos publicados durante la elaboración de esta tesis

PLOS ONE

RESEARCH ARTICLE

Blastocystis Isolates from Patients with Irritable Bowel Syndrome and from Asymptomatic Carriers Exhibit Similar Parasitological Loads, but Significantly Different Generation Times and Genetic Variability across Multiple Subtypes

Gie-Bele Vargas-Sanchez[1☯], Mirza Romero-Valdovinos[2☯], Celedonio Ramirez-Guerrero[3], Ines Vargas-Hernandez[4], Maria Elena Ramirez-Miranda[2], Joel Martinez-Ocaña[2], Alicia Valadez[5], Cecilia Ximenez[5], Eduardo Lopez-Escamilla[2], Maria Elena Hernandez-Campos[1], Guiehdani Villalobos[6], Fernando Martinez-Hernandez[2], Pablo Maravilla[2]*

1 Escuela Superior de Medicina, Instituto Politecnico Nacional, Plan de San Luis y Diaz Mirón s/n, Casco de Santo Tomas, 11340, Mexico City, DF, Mexico, **2** Hospital General "Dr. Manuel Gea Gonzalez", Calzada de Tlalpan 4800, Seccion XVI, 14080, Mexico City, DF, Mexico, **3** Hospital Infantil de Mexico Federico Gomez, Dr. Marquez 162, Doctores, 06720, Mexico City, DF, Mexico, **4** Centro Especializado de Atención Primaria "Melchor-Ocampo, Bicentenario", Secretaria de Salud del Estado Mexico, Estado de Mexico, 54880, Mexico, **5** Departamento de Medicina Experimental, Facultad de Medicina, Universidad Nacional Autonoma de Mexico, 04510, Mexico City, DF, Mexico, **6** Departamento de Ecologia Evolutiva, Instituto de Ecologia, Universidad Nacional Autonoma de Mexico, Avenida Universidad 3000, 04510, Mexico City, DF, Mexico

☯ These authors contributed equally to this work.
* maravillap@yahoo.com

OPEN ACCESS

Citation: Vargas-Sanchez G-B, Romero-Valdovinos M, Ramirez-Guerrero C, Vargas-Hernandez I, Ramirez-Miranda ME, Martinez-Ocaña J, et al. (2015) *Blastocystis* Isolates from Patients with Irritable Bowel Syndrome and from Asymptomatic Carriers Exhibit Similar Parasitological Loads, but Significantly Different Generation Times and Genetic Variability across Multiple Subtypes. PLoS ONE 10 (4): e0124006. doi:10.1371/journal.pone.0124006

Academic Editor: Alejandro Escobar-Gutiérrez, Instituto de Diagnostico y Referencia Epidemiologicos, MEXICO

Received: November 24, 2014

Accepted: March 9, 2015

Published: April 29, 2015

Data Availability Statement: All data are fully available without restriction in http://dx.doi.org/10.6084/m9.figshare.1344544.

Funding: This work was sponsored by Consejo Nacional de Ciencia y Tecnologia (CONACYT) grant numbers 182089 and 168619.

Abstract

Blastocystis spp is a common intestinal parasite of humans and animals that has been associated to the etiology of irritable bowel syndrome (IBS); however, some studies have not found this association. Furthermore, many biological features of *Blastocystis* are little known. The objective of present study was to assess the generation times of *Blastocystis* cultures, from IBS patients and from asymptomatic carriers. A total of 100 isolates were obtained from 50 IBS patients and from 50 asymptomatic carriers. Up to 50 mg of feces from each participant were cultured in Barret's and in Pavlova's media during 48 h. Initial and final parasitological load were measured by microscopy and by quantitative PCR. Amplicons were purified, sequenced and submitted to GenBank; sequences were analysed for genetic diversity and a Bayesian inference allowed identifying genetic subtypes (ST). Generation times for *Blastocystis* isolates in both media, based on microscopic measures and molecular assays, were calculated. The clinical symptoms of IBS patients and distribution of *Blastocystis* ST 1, 2 and 3 in both groups was comparable to previous reports. Interestingly, the group of cases showed scarce mean nucleotide diversity (π) as compared to the control group (0.011±0.016 and 0.118±0.177, respectively), whilst high gene flow and small genetic differentiation indexes between different ST were found. Besides, Tajima's D test

Competing Interests: The authors have declared that no competing interests exist.

showed negative values for ST1-ST3. No statistical differences regarding parasitological load between cases and controls in both media, as searched by microscopy and by qPCR, were detected except that parasites grew faster in Barret's than in Pavlova's medium. Interestingly, slow growth of isolates recovered from cases in comparison to those of controls was observed ($p<0.05$). We propose that generation times of *Blastocystis* might be easily affected by intestinal environmental changes due to IBS probably because virulent strains with slow growth may be selected, reducing their genetic variability.

Introduction

Irritable bowel syndrome (IBS) is defined as a functional bowel disorder in which abdominal pain or discomfort is associated with disordered defecation or with a change in bowel habits in the absence of an organic cause [1–2]. The pathophysiology of IBS remains elusive and no mechanism is unique to, or characteristic of, IBS. There are probably several interconnected factors which occur to varying degrees in patients that account for the clinical symptoms of IBS, these include altered gut reactivity (colonic and/or small bowel motility) in response to luminal or psychological stimuli, hypersensitive viscera or gut, enhanced visceral perception and pain [3–4].

Blastocystis spp., an intestinal parasite, is one of the most common parasites worldwide in humans and although its ability to cause disease has been questioned, some reports have demonstrated that this microorganism is associated to the development of IBS [5–9]. Up to 17 subtypes (ST) have been described in this parasite, based on its small subunit rDNA analysis (SSUrDNA), humans are colonized mainly by ST1–ST4; their prevalence seems to vary from country to country and among communities within the same country. Some analyses have shown a lack of association between ST and IBS [10–12]. In addition, a relevant phenotypic variation in size, growth and clumping was documented among ST3 isolates from patients with gastrointestinal alterations and in asymptomatic carriers [13]. Many culture media, as well as physical and chemical parameters have been tested for *Blastocystis* growth; some of them generated improvement of its diagnosis [14–18], but no links between *in vitro* generation times indifferent *Blastocystis* ST, and their genetic variation, have been reported. Therefore, the objective of present study was to assess generation times and genetic polymorphism of *Blastocystis* cultures from IBS patients and from asymptomatic carriers.

Materials and Methods

Samples

Fecal samples of 50 IBS patients (cases) and 50 unrelated healthy volunteers (controls) with *Blastocystis* as a unique infection, diagnosed by flotation-concentration Ferreira's technique were studied. For the control' group, all participants were recruited when attending the Blood Bank of the Hospital General "Dr. Manuel Gea Gonzalez" for a voluntary and altruistic blood donation; after a clinical check-up they were invited to participate in the present study by providing fecal samples for a coprological diagnosis. Cases were recruited during their medical consultations. Informed consent was obtained from each participant before their recruitment. IBS patients were diagnosed according to the Rome III criteria [2], which refer recurrent abdominal pain or uncomfortable sensation at least 3 days per month in the last 3 months associated with 2 or more of the following: i) improvement with defecation, ii) onset associated with

a change in frequency of stool release, iii) onset associated with a change in form or appearance of stools; in order to discard the presence of Rotavirus and pathogenic enterobacteria, all samples were tested with a commercial Rota Test (CerTest, Spain) and VITEK-2 system (BioMérieux, France), respectively.

Culture

Using templates of the Kato-Katz technique [19], 50 mg of feces of each participant were cultured in 2mL of Barret's medium [20–21] and of Pavlova's medium [16, 22]; both complemented with 10% horse serum, as well as Penicillin G (1,000IU/mL) and Streptomycin (100μg/mL). Sterile 2.5mL screw cap tubes were used for incubation at 37°C; all assays were performed in duplicate. Two aliquots of ~30 and 500μL were recovered at the beginning and at 48h. The first aliquot was used to measure the concentration of *Blastocystis* cells by Neubauer chamber, whilst the second was used to obtain DNA. At the end of the experiment, to facilitate counting parasites, cultures were fixed in 10% formalin.

DNA extraction and quantitative PCR (qPCR) assays

Total DNA from each stool sample cultured was obtained using a conventional extraction protocol with proteinase K, phenol/chloroform solution and precipitated with isopropanol [23]; DNA extracts were stored at -20°C until molecular analysis. Absolute quantification was performed by qPCR in a Light Cycler480-II system (Roche, Germany), using as positive control a sample with a known number of parasites (8,920 *Blastocystis* organisms/mg feces) and a DNA concentration of 160ng/μL, determined by UV-spectrophotometry. A set of primers that amplify a region of the *Blastocystis* small subunit of ribosomal DNA (SSUrDNA) gene, previously reported by Poirier et al. [24] were used and the amplification was performed with SYBR green kit (Fermentas, CA), 3.5 mM $MgCl_2$, 0.2μM of each primer and 200ng of DNA extracted. The amplification conditions were: an initial denaturation cycle of 95°C for 10min, 45 amplification cycles 95°C 20s, 65°C 20s and 72°C 20s. To establish the coefficient of correlation of the qPCR assays, 10 fold dilutions of DNA control were used in duplicated experiments. For data analysis, the melting curves at 95°C 5s and of 65°C to 97°C for 1min was performed; afterwards a concentration curve was obtained and all DNA samples were amplified and compared with this control.

Blastocystis subtype (ST) determination and genetic variation analysis

High Resolution Melting (HRM) analysis was conducted in order to support the ST identification, in some DNA samples that were amplified in 96-well plates, and HRM curves acquisition was performed, similarly to qPCR, under the same amplification conditions and system, but using a Light Cycler HRM master kit (Roche, Germany) to assign a melting temperature. DNA samples with ST identified by sequencing (ST1, ST2, ST3 and ST7) were included as references for wild-type curves. According to Wittwer et al. [25] and Gonzalez-Bosquet et al. [26], normalization and background subtraction were first performed by fitting an exponential to the background surrounding the HRM transitions of interest; the normalized HRM curves were temperature-overlaid, to eliminate slight temperature errors between wells or runs; after, different plots of these normalized and temperature-overlaid curves were obtained by deducting the fluorescence difference of each curve from the average wild-type curve at all temperature points. Thus, HRM profile with a plot interpreted by the software to be different from the averaged wild-type curve, were considered to be suspicious of harboring a nucleotide change, mutation or variant that, in our particular case, corresponded to each ST. Direct sequencing was performed to identify the ST in DNA samples. Direct qPCR product sequencing was

performed according to Poirier et al [24], after purification with a QIAquick PCR Purification Kit (Qiagen, Germany). Sequencing was performed by a commercial supplier using the same primers as those used for qPCR (BL18SPPF1 and BL18SR2PP) that amplify a DNA fragment of 320 to 342bp, depending on the subtype.

All sequences were subjected to a BLAST search in the GenBank database; multiple alignments were performed using the CLUSTAL W [27] and Muscle [28] programmes with manual adjustment in MEGA 5.05 software [29]. Hasegawa Kishino Yano model with gamma distribution and invariable sites was the best fit model of nucleotide substitution for SSUrDNA, determined using the Akaike Information Criterion in Modeltest version 3.7 software [30]. The phylogenetic reconstruction using Bayesian inference was performed with the Mr. Bayes 3.1.2 program [31–33]. The analysis was done for 10 million generations with sampling trees every 100 generations. Trees with scores lower than those at the stationary phase (burn-in) were discarded, and the trees that reached the stationary phase were collected and used to build majority consensus trees. Other sequences were obtained from GenBank and used as subtype references. A genetic diversity analysis within and between populations was performed using DnaSPv4 [34] and included nucleotide diversity (π), haplotype polymorphism (θ), gene flow (*Nm*), genetic differentiation index (F_{ST}) and Tajima's D test. These indexes refer to the following: π, the average proportion of nucleotide differences between all possible pairs of sequences in the sample; θ, the proportion of nucleotide sites that are expected to be polymorphic in any suitable sample from this region of the genome. Both indexes are used to assess polymorphisms at the DNA level and to monitor diversity within or between ecological populations and examine the genetic variation in related species or their evolutionary relationships. F_{ST} is a typical genetic statistic used to measure differentiation between or among populations. The common used values for genetic differentiation are as follows: 0 to 0.05, small; 0.05 to 0.15, moderate; 0.15 to 0.25, great; values above 0.25 indicate huge genetic differentiation. The gene flow or migration index (*Nm*) refers to the movement of organisms among subpopulations; those strongly differentiated have an $Nm<<1$, whereas an $Nm>4$ behaves as a single panmictic unit [35].

Generation time

The generation times of *Blastocystis* isolates in the two media, based on microscopic measures and molecular assays, were calculated according to Zhang et al. [17], with the following equation: $T_g = (T_2\text{-}T_1)/(\log_2(n_2/n_1))$, where T_g denotes the generation time, n_1 represents the number of cultured parasitic organisms at the initial time (T_1), and n_2 represents the number of parasitic cells at subsequent time (T_2). Thus, $(T_2\text{-}T_1) = 48$ hours of *in vitro* culture.

Due to absolute quantification by qPCR, it was necessary to consider i) the size of the *Blastocystis* genome ~18.8Mbp [36]; ii) 1pg of DNA ~978Mbp [37] and iii) the concentration of DNA control was 160ng/μL; thus, by cross multiplications, the number of copies of the genetic marker amplified were estimated.

Statistical analysis

Descriptive statistics are expressed as mean and standard deviation (SD). Analysis by Student's t test for unrelated and related samples and Mantel–Haenszel test were applied; odds ratio (OR) and 95% confidence intervals (95% CI) were also obtained. Data analysis was performed with SSPS software Version 15.0 (SPSS Institute, Chicago, IL).

Ethics statement

All procedures are in accordance with the provisions of the Regulations of the General Health Law in the Field of Health Research in Mexico: Title II, Chapter II, from research communities;

the Ethics and Research Committees of the General Hospital "Dr. Manuel Gea Gonzalez" approved the protocol with the reference number 12-87-2012, and written informed consent was obtained from each participant before recruitment.

Results

Table 1 summarizes clinical symptoms and *Blastocystis* ST for cases and controls. All IBS patients presented abdominal pain and bloating, while flatulence and alternating diarrhea or constipation were the least frequent. All amplified DNA fragments were purified and sequenced; some DNA from the different ST (ST1, 2, 3 and 7) were amplified and melting analysis exhibited distinct temperature peaks, depending on ST; however, some of them were too close to distinguish among ST, therefore the sequence analysis leaded to identify the ST; 49 sequences were obtained for cases and 48 for controls (GenBank accession numbers KP055659-KP055754; however, sequence of the sample "control 45" was <200bp and its accession number was not assigned). ST1 was the most frequent in the cases' group (41%) followed by ST3 (33%); in contrast, for controls' group, the order of ST was inverted, because ST3 was the most frequent (54%) followed by ST1 (21%); distribution of ST2 and ST7 was similar between both groups. Interestingly, the group of cases showed a scarce mean nucleotide diversity and haplotype polymorphism as compared to control group ($p<0.001$), being ST2 the most diverse (Table 2). In contrast, ST1 was more variable in the cases' group. Regarding F_{ST} and *Nm* indexes, a small genetic differentiation and a high gene flow between different ST was found. Result of Tajima's D test showed negative values for ST1-ST3. The phylogenetic tree (Fig 1) supported the grouping of our sequences in the ST1, 2, 3 and 7 clades.

During the counting of parasites in the cultured samples, the predominant stage was the vacuolar form in both media; after 48h of culture in both groups, few amoeboid and granular forms were observed, with a light increase of the granular form in Pavlova's medium. No differences were found regarding the parasitological load between cases and controls in both media by microscopy and by qPCR, perhaps due to a high variance within data (Fig 2); however, some statistical differences between media and among groups were found, such as parasites showed faster growth in Barret's medium than in Pavlova's medium ($p<0.001$, t-test related

Table 1. Characteristics of the carriers and of *Blastocystis* ST identified in IBS patients and in the control group.

	IBS group	Control group
Age, years	46±11.5	38±11.3
Male/female, n	11/39	10/40
Clinical symptoms	% (n)	% (n)
Abdominal pain	100 (50)	0 (50)
Bloating	100 (50)	0 (50)
Flatulence	45 (24)	0 (50)
Diarrhea-Constipation	40 (20)	0 (50)
***Blastocystis* ST**	% (n = 49)	% (n = 48)
1	41 (20)	21 (10)
2	24 (12)	23 (11)
3	33 (16)	54 (26)[a]
7	2 (1)	2 (1)

[a] $p = 0.043$,
OR(95%IC) = 0.43(0.17–1.06).

doi:10.1371/journal.pone.0124006.t001

Table 2. Population genetic indexes between different *Blastocystis* ST sequences.

Blastocystis ST	π		Θ		F_{ST}	*Nm*	Tajima's D (*p* value)
	IBS	Control	IBS	Control			
1	0.031	0.019	0.905	0.891	0.011	7.18	-2.256 ($p<0.01$)
2	0.002	0.324	0.682	0.972	0.039	6.48	-2.144 ($p<0.05$)
3	0.001	0.013	0.342	0.628	0.021	11.22	-2.562 ($p<0.01$)
Mean±SD	0.011±0.016	0.118±0.177	0.643±0.283	0.830±0.179			
(*p*; 95%CI[a])	$p<0.001$; 95%CI = -0.16–0.06		$p<0.001$; 95%CI = -0.28–0.09				

[a]95%CI, 95% Confidence interval.

doi:10.1371/journal.pone.0124006.t002

samples). Interestingly, slow growth for isolates recovered from cases in comparison with controls was observed ($p<0.05$ by microscopy and $p = 0.001$ by qPCR; Fig 3).

Discussion

Clinical symptoms in IBS patients infected with *Blastocystis* were concordant to those reported in Mexican populations [8, 38–40] and in other countries [6–7, 41–43]. Also, distribution of *Blastocystis* ST 1, 2 and 3 was comparable to other reports [8, 10, 38–40], showing no association with symptoms or the development of IBS, in contrast with those findings reported by Ramírez et al. [44], who found that all patients infected with *Blastocystis* ST2 presented diarrhea, while than asymptomatic carriers exhibited only ST1.

Recently, a study focused to assess the genetic variation and differentiation of *Blastocystis* subtypes recovered from symptomatic children from Michoacan, Mexico showed that by analysing partial sequences of the SSUrDNA the mean nucleotide diversity was π = 0.0179±0.0112 [40]; interestingly in the present work, a similar mean π = 0.011±0.016 for cases was calculated; in contrast, our control group showed a mean π almost ten times higher than the cases' group (π = 0.118±0.177), indicating that parasites from IBS patients exhibited less genetic variable than controls. Besides, F_{ST} and *Nm* indexes showed a small genetic differentiation and a high gene flow between cases and control groups for ST 1, 2 and 3, suggesting a large flow of parasites among carriers, regardless their ST. Besides, the negative values of Tajima's D test advise a recent expansion process or an effect of purifying selection [35] that it might occur, particularly in those isolates from symptomatic *Blastocystis* carriers.

No statistical differences were observed in the parasitological load between cases and controls; besides, the HRM analyses did not allowed distinguishing unambiguously among ST, probably due to the relationship between the sensitivity of the technique and the high genetic polymorphism among ST [25]. The most significant findings in the present study were the dissimilarities in the T_g values from IBS patients and control groups, since a slow growth for isolates recovered from cases in comparison with controls was observed in the two media used; although Barret's medium allowed a faster growth, measured by two different techniques, independently of their ST. Differences between T_g values obtained by microscopy and by qPCR, can be explained because both techniques measure dissimilar features and by the high sensitivity of the molecular assay, since with qPCR it has been possible to detect as few as 10^2 *Blastocystis* parasites/g of stools as referred by Poirier et al. [24]. Tan et al. [45] studied some phenotypic characteristics of ten asymptomatic and ten symptomatic human-derived *Blastocystis* ST3 isolates, and found that asymptomatic isolates grew more rapidly (mean T_g = 9.3h) than symptomatic isolates (mean T_g = 15.3h) using Jones' medium supplemented with 10% horse serum. Recently, Ragavan et al. [13] studied 12 isolates of ST3 from four IBS patients, four

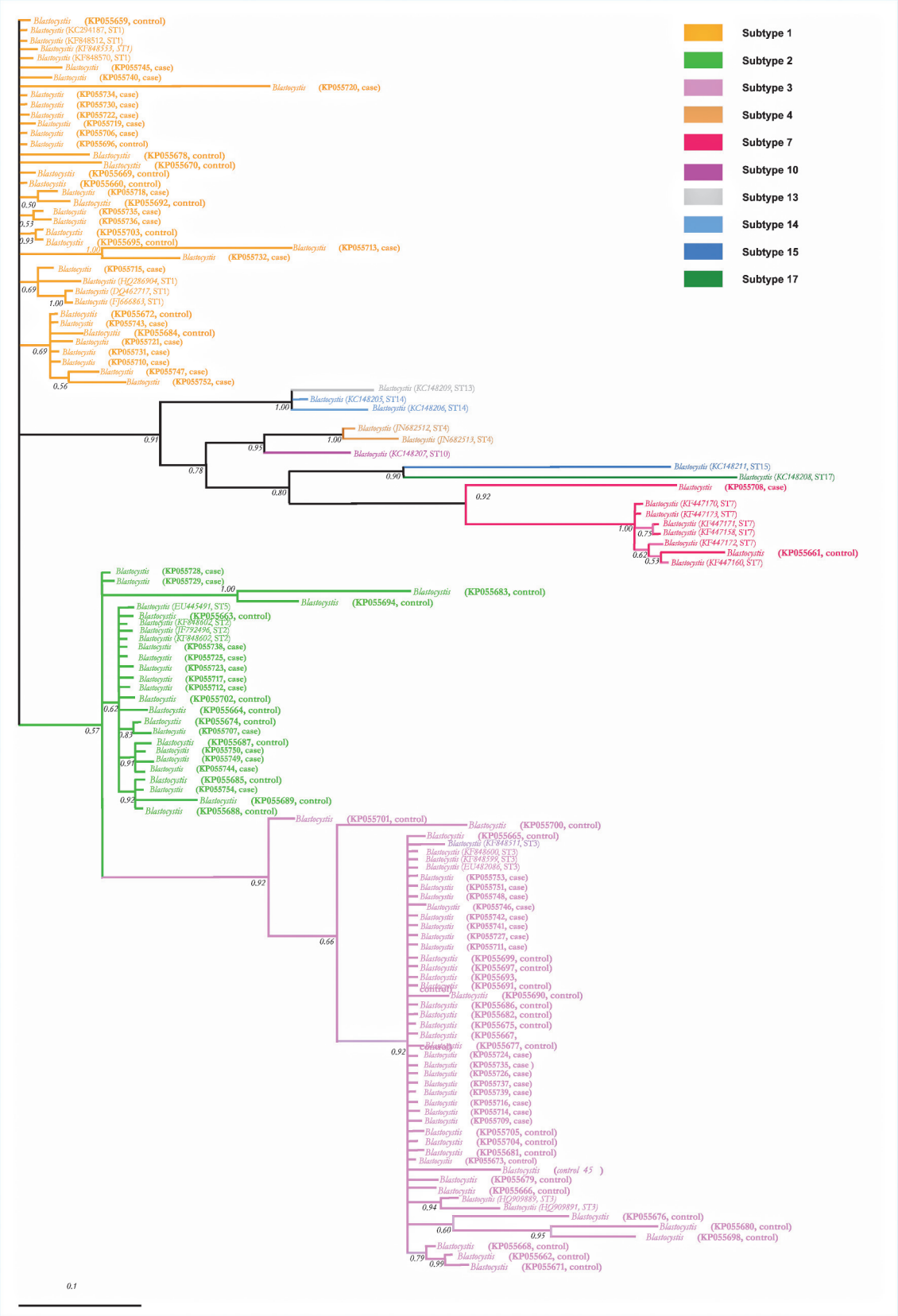

Subtype 1
Subtype 2
Subtype 3
Subtype 4
Subtype 7
Subtype 10
Subtype 13
Subtype 14
Subtype 15
Subtype 17
0.1

Fig 1. Phylogenetic inference of *Blastocystis* spp. Bayesian phylogenetic tree using a fragment of SSUrDNA sequences; the values of the nodes indicate posterior probabilities values using 10 million generations. GenBank accession numbers are included, as well as if correspond to a case or a control; each ST clade is shown in different branch colors.

doi:10.1371/journal.pone.0124006.g001

gastrointestinal symptomatic patients and four asymptomatic carriers, finding that the asymptomatic isolates grew faster than symptomatic and IBS isolates which were similar with those reported by Tan et al. [45]; these data are in accordance with those of the present study. The media of Jones, Barret and Pavlova are simple sera-saline media; but the last one, phosphate salts and yeast extract are added to increase osmolarity [16, 20, 46]. In addition, Zhang et al. [17] evaluated the growth of *Blastocystis* in three commercially available liquid media (RPMI 1640, 199 Medium and Dulbecco's modified Eagle's medium) under a different pH, inoculum sizes and amounts of calf serum, finding that all culture factors modified the T_g in each case. Thus, we consider that T_g in *Blastocystis* is easily affected by intestinal environment changes (such as osmolarity, variations in the microbiota composition [47, 48], host diet, etc.) eventually leading to loss of infection, explaining the spontaneous remission observed by Sanchez-Aguillon et al. [49] during a cohort study.

Nutrient availability, an effective immune response and the host genetics can influence the response against pathogen challenges [50–52]; for example, it has been suggested that certain

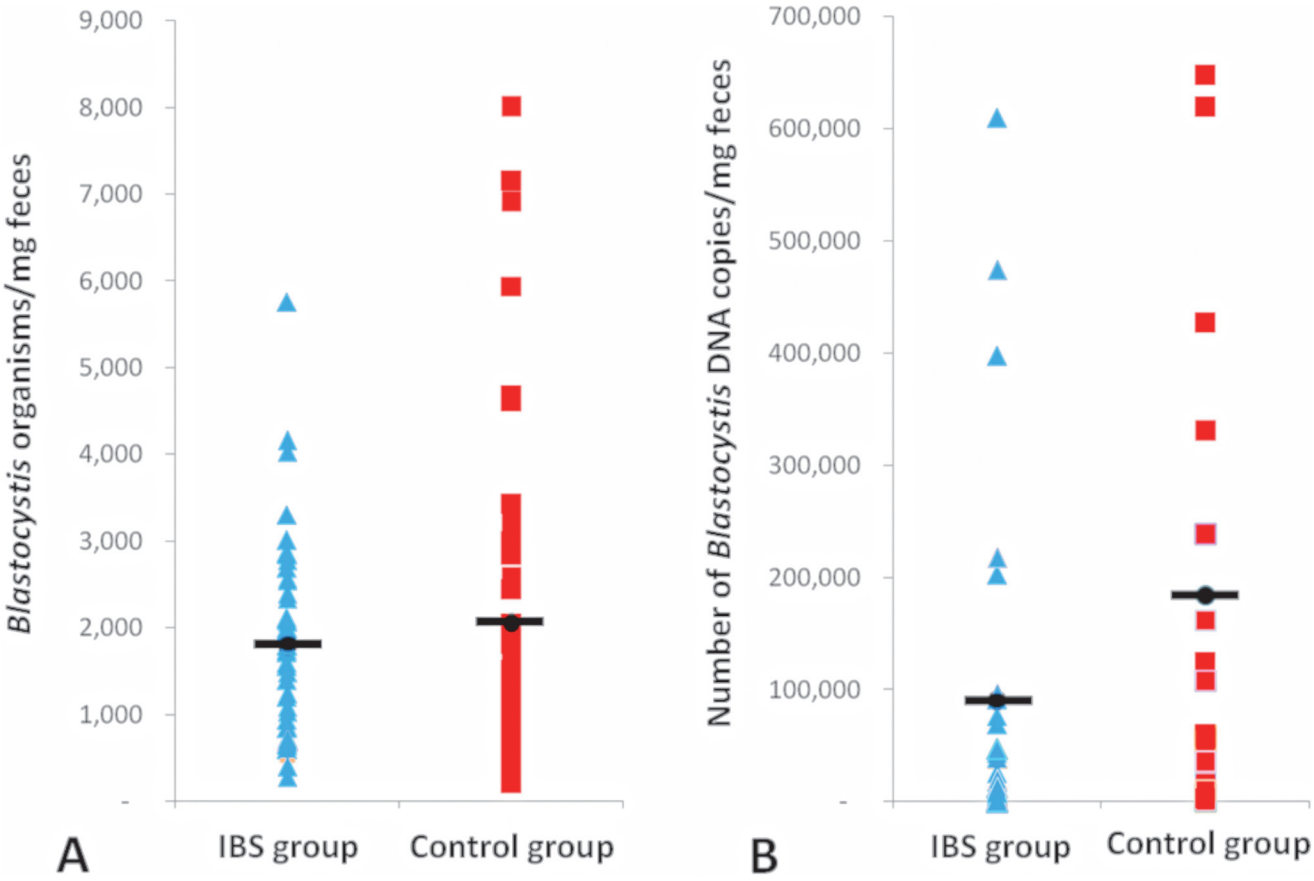


Fig 2. Parasitological loads estimated at the beginning of the study between cases and controls. Plots drew with all sample data according to the number of *Blastocystis* organisms/mg feces by microscopy (A) and number of *Blastocystis* DNA copies/mg feces by qPCR (B). The black bars mean the average in each group.

doi:10.1371/journal.pone.0124006.g002

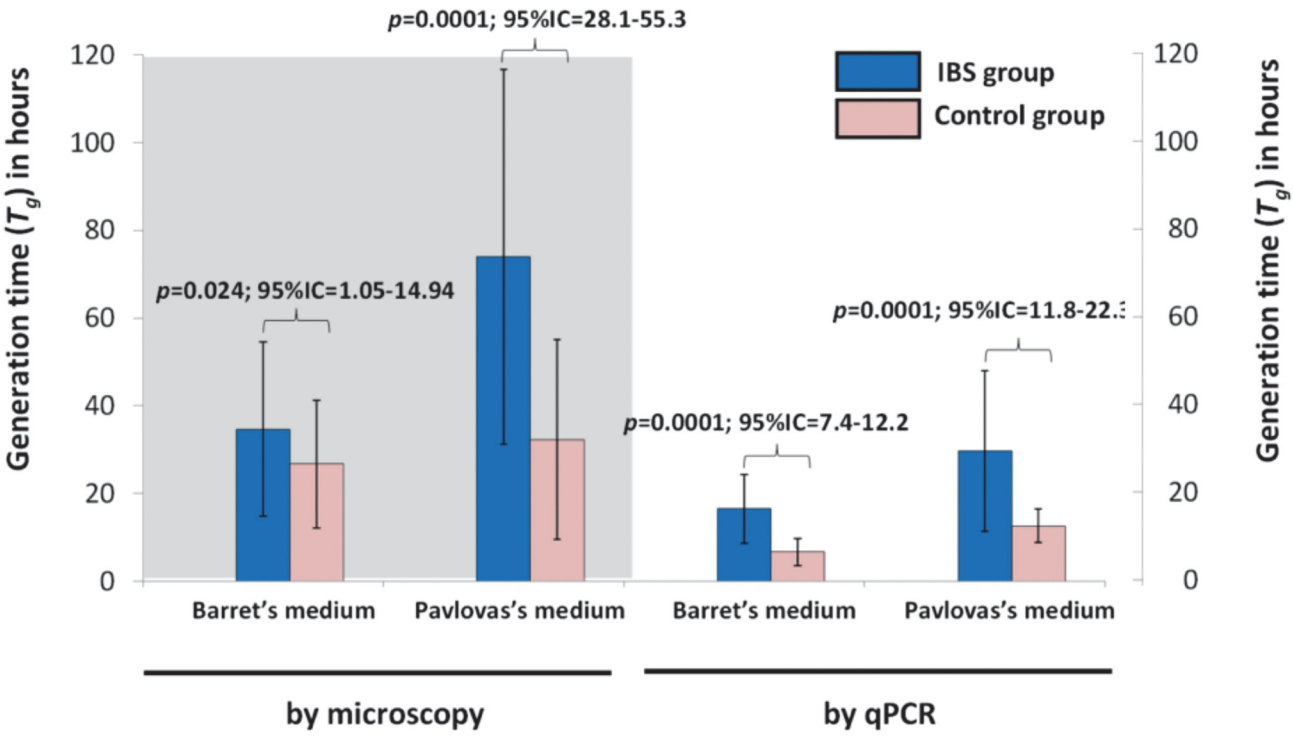


Fig 3. Generation time (T_g) values of *Blastocystis* isolates between case and control groups. Average and standard deviation of T_g in Barret's and Pavlova media, based on microscopic measures and qPCR assays.

doi:10.1371/journal.pone.0124006.g003

single nucleotide polymorphisms of interleukin-8 and -10 could change individual susceptibility increasing the relative risk in the development of IBS in *Blastocystis* carriers [53], Therefore differences for T_g observed in early cultures, free of the host immune response pressure and with sufficient nutrients, allow to assume differences in the cellular protein metabolism, not yet studied between isolates recovered from symptomatic and asymptomatic carriers. Besides, some cellular and physiological features, not codified by changes in the DNA sequence, such as DNA methylation, histone modifications and RNA-associated silencing, have been associated to virulence and evasion of the immune response in some protozoan parasites, such as *Entamoeba histolytica* and *Giardia lamblia* [54–56]. Thus, many aspects of the host-pathogen relationship in *Blastocystis* infections require to be studied, in order to clarify the pathogenic role of this microorganism.

Although it has been argued that genetic polymorphism inside each ST, could be related with the pathogenicity of *Blastocystis* [11, 57], little has been discussed about whether genetic differences in this microorganism may be associated with its growth, and if this might give some clue about its pathogenicity. In addition, it has been argued that some parasites exploit their hosts in a prudent way, taking the resources that they need without causing noticeable damage. Prudent exploitation yields sustainable benefits to the parasite as long as the host remains healthy. Other parasites attack their host more quickly and vigorously. Rapid exploitation may allow the parasites to achieve higher reproductive rates, but damage to the host reduces parasites' opportunity for sustainable yield. Following this economic line of thought, each parasite faces a trade-off when increasing the rate at which host resources are used; however, simple economic considerations will certainly not explain all aspects of parasitism and the severity of disease (virulence) [58]. The virulence trade-off hypothesis assumes that higher within-host replication rates lead to higher damage [59], however new opposite examples are

emerging [60], i.e. Sistrom et al [61] analyzed the whole-genome sequences of *Trypanosoma brucei* complex from diverse hosts and regions, finding that a strong purifying selection associated to cytoskeleton structure and regulatory genes related with virulence; therefore, it is likely that carriers of *Blastocystis* virulent strains grow slowly (which would explain the chronic symptoms associated with this parasite) and by a purifying selection process, these strains could be directed towards specialization, showing reduced genetic variability. This idea is supported by our Tajima's D test data and by a recent article, that it was focused in a comparative analysis of the population structure of eukaryotic pathogens using Multilocus Sequence Typing, finding that *Blastocystis* dataset corresponded to a population of recent origin having undergone a radiation process [62].

Author Contributions

Conceived and designed the experiments: MR-V MEH-C FM-H PM. Performed the experiments: G-BV-S MR-V CR-G IV-H MER-M JM-O AV EL-E. Analyzed the data: G-BV-S MR-V CX GV FM-H PM. Contributed reagents/materials/analysis tools: IV-H MER-M JM-O GV FM-H. Wrote the paper: G-BV-S MR-V CR-G IV-H MER-M JM-O AV CX EL-E MEH-C GV FM-H PM. Obtained permission and approval of health authorities: G-BV-S PM.

References

1. Brandt LJ, Bjorkman D, Fennerty MB, Locke GR, Olden K, Peterson W, et al. Systematic review on the management of irritable bowel syndrome in North America. Am J Gastroenterol. 2002; 97: S7–26. PMID: 12425586
2. Longstreth GF, Thompson WG, Chey WD, Houghton LA, Mearin F, Spiller RC. Functional bowel disorders. Gastroenterol. 2006; 130: 1480–1491. PMID: 16678561
3. Thompson WG, Heaton KW, Smyth GT, Smyth C. Irritable bowel syndrome in general practice: prevalence, characteristics, and referral. Gut. 2000; 46: 78–82. PMID: 10601059
4. Mertz HR. Irritable bowel syndrome. N Engl J Med. 2003; 349: 2136–2146. PMID: 14645642
5. Giacometti A, Cirioni O, Antonicelli L, D´Amato G, Silvestri C, Del Prete MS, et al. Prevalence of intestinal parasites among individuals with allergic diseases. J Parasitol. 2003; 89: 490–492. PMID: 12880246
6. Yakoob J, Jafri W, Jafri N, Khan R, Islam M, Beg MA, et al. Irritable bowel syndrome: in search of an etiology: role of *Blastocystis hominis*. Am J Trop Med Hyg. 2004; 70: 383–385. PMID: 15100450
7. Yakoob J, Jafri W, Beg MA, Abbas Z, Naz S, Islam M, et al. Irritable bowel syndrome: is it associated with genotypes of *Blastocystis hominis*. Parasitol Res. 2010; 106: 1033–1038. doi: 10.1007/s00436-010-1761-x PMID: 20177906
8. Jimenez-Gonzalez DE, Martinez-Flores WA, Reyes-Gordillo J, Ramirez-Miranda ME, Arroyo-Escalante S, Romero-Valdovinos M, et al. *Blastocystis* infection is associated with irritable bowel syndrome in a Mexican patient population. Parasitol Res. 2012; 110: 1269–1275. doi: 10.1007/s00436-011-2626-7 PMID: 21870243
9. Poirier P, Wawrzyniak I, Vivarès CP, Delbac F, El Alaoui H. New insights into *Blastocystis* spp.: a potential link with irritable bowel syndrome. PLoS Pathog. 2012; 8: e1002545. Available: http://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1002545. doi: 10.1371/journal.ppat.1002545 PMID: 22438803
10. Alfellani MA, Stensvold CR, Vidal-Lapiedra A, Onuoha ES, Fagbenro-Beyioku AF, Clark CG. Variable geographic distribution of *Blastocystis* subtypes and its potential implications. Acta Trop. 2013; 126: 11–18. doi: 10.1016/j.actatropica.2012.12.011 PMID: 23290980
11. Scanlan PD, Stensvold CR. *Blastocystis*: getting to grips with our guileful guest. Trends Parasitol. 2013; 29: 523–529. doi: 10.1016/j.pt.2013.08.006 PMID: 24080063
12. Krogsgaard LR, Engsbro AL, Stensvold CR, Nielsen HV, Bytzer P. The prevalence of intestinal parasites is not greater among individuals with irritable bowel syndrome: a population-based case-control study. Clin Gastroenterol Hepatol. 2015; 13: 507–513.e2. doi: 10.1016/j.cgh.2014.07.065 PMID: 25229421

13. Ragavan ND, Govind SK, Chye TT, Mahadeva S. Phenotypic variation in *Blastocystis* sp. ST3. Parasit Vectors. 2014; 7: 404. Available: http://www.parasitesandvectors.com/content/7/1/404. doi: 10.1186/1756-3305-7-404 PMID: 25174569

14. Leelayoova S, Taamasri P, Rangsin R, Naaglor T, Thathaisong U, Mungthin M. In-vitro cultivation: a sensitive method for detecting *Blastocystis hominis*. Ann Trop Med Parasitol. 2022; 96: 803–807.

15. Termmathurapoj S, Leelayoova S, Aimpun P, Thathaisong U, Nimmanon T, Taamasri P, et al. The usefulness of short-term in vitro cultivation for the detection and molecular study of *Blastocystis hominis* in stool specimens. Parasitol Res. 2004; 93: 445–447. PMID: 15243800

16. Irikov OA, Antokhin AI, Romanov YA. Study of the dynamics of *Blastocystis hominis* reproduction in vitro. Bull Exp Biol Med. 2009; 148:99–102. PMID: 19902108

17. Zhang X, Qiao J, Wu X, Da R, Zhao L, Wei Z. In vitro culture of *Blastocystis hominis* in three liquid media and its usefulness in the diagnosis of blastocystosis. Int J Infect Dis. 2012; 16: e23–28. Available: http://www.ijidonline.com/article/S1201-9712(11)00191-3/abstract. doi: 10.1016/j.ijid.2011.09.012 PMID: 22047715

18. Santos HJ, Rivera WL. Comparison of direct fecal smear microscopy, culture, and polymerase chain reaction for the detection of *Blastocystis* sp. in human stool samples. Asian Pac J Trop Med. 2013; 6: 780–784. doi: 10.1016/S1995-7645(13)60138-8 PMID: 23870466

19. Katz N, Chaves A, Pellegrino J. A simple device for quantitative stool thick-smear technique in *Schistosomiasis mansoni*. Rev Inst Med Trop Sao Paulo. 1972; 14: 397–400. PMID: 4675644

20. Barret HP. A method for the cultivation of *Blastocystis*. Ann Trop Med Parasitol. 1921; 15: 113–116.

21. Clark CG, Diamond LS. Methods for cultivation of luminal parasitic protists of clinical importance. Clin Microbiol Rev. 2002; 15: 329–341. PMID: 12097242

22. Zerpa R, Huicho L, Náquira C, Espinoza I. A simplified culture method for *Blastocystis hominis*. Rev Mex Patol Clin. 2000; 47: 17–19.

23. Sambrook J, Fitsch EF, Maniatis T. Molecular cloning: a laboratory manual. 3rd ed. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 2001.

24. Poirier P, Wawrzyniak I, Albert A, El Alaoui H, Delbac F, Livrelli V, et al. Development and evaluation of a real-time PCR assay for detection and quantification of *Blastocystis* parasites in human stool samples: prospective study of patients with hematological malignancies. J Clin Microbiol. 2011; 49: 975–983. doi: 10.1128/JCM.01392-10 PMID: 21177897

25. Wittwer CT, Reed GH, Gundry CN, Vandersteen JG, Pryor RJ. High-Resolution Genotyping by Amplicon Melting Analysis Using LCGreen. Clin Chem. 2003; 49: 853–860. PMID: 12765979

26. Gonzalez-Bosquet J, Calcei J, Wei JS, Garcia-Closas M, Sherman ME, Hewitt S, et al. Detection of Somatic Mutations by High-Resolution DNA Melting (HRM). Analysis in Multiple Cancers. PLoS ONE; 2011 6: e14522. Available: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0014522. doi: 10.1371/journal.pone.0014522 PMID: 21264207

27. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: Improving the sensitivity and progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994; 22: 4673–4680. PMID: 7984417

28. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004; 32: 1792–1797. PMID: 15034147

29. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011; 28: 2731–2739. doi: 10.1093/molbev/msr121 PMID: 21546353

30. Posada D, Crandall KA. Modeltest: testing the model of DNA substitution. Bioinformatics. 1998; 14: 817–818. PMID: 9918953

31. Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP. Bayesian inference of phylogeny and its impact on evolutionary biology. Science. 2001; 14: 2310–2314.

32. Holder M, Lewis PO. Phylogeny estimation: traditional and Bayesian approaches. Nat Rev Genet. 2003; 4: 275–284. PMID: 12671658

33. Ronquis F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 2003; 19: 1572–1574. PMID: 12912839

34. Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics. 2003; 19: 2496–2497. PMID: 14668244

35. Hartl DL, Clark AG. Principles of Population Genetics. Massachusetts: Sinauer Associates Inc. Publishers; 1997

36. Denoeud F, Roussel M, Noel B, Wawrzyniak I, Da Silva C, Diogon M, et al. Genome sequence of the stramenopile *Blastocystis*, a human anaerobic parasite. Genome Biol. 2011; 12:R29. Available: http://genomebiology.com/content/12/3/R29 doi: 10.1186/gb-2011-12-3-r29 PMID: 21439036

37. Dolezel J, Bartos J, Voglmayr H, Greilhuber J. Nuclear DNA content and genome size of trout and human. Cytometry A. 2003; 51: 127–128. PMID: 12541287

38. Ramirez-Miranda ME, Hernandez-Castellanos R, Lopez-Escamilla E, Moncada D, Rodriguez-Magallan A, Pagaza-Melero C, et al. Parasites in Mexican patients with irritable bowel syndrome: a case-control study. Parasit Vectors. 2010; 3: 96. Available: http://www.parasitesandvectors.com/content/3/1/96. doi: 10.1186/1756-3305-3-96 PMID: 20942938

39. Ramírez-Miranda ME, Jiménez-González DE, Rodríguez-Campa ME, González-Angulo A, Hernández-Castellanos R, Arroyo-Escalante S, et al. Irritable bowel syndrome: frequency and phylogenetic relationship of *Blastocystis* sp. from Mexican patients. Rev Gastroenterol Mex. 2011; 76: 309–315. PMID: 22188955

40. Villalobos G, Orozco-Mosqueda G, Lopez-Perez M, Lopez-Escamilla E, Córdoba-Aguilar A, Rangel-Gamboa L, et al. Suitability of internal transcribed spacers (ITS) as markers for the population genetic structure of *Blastocystis* spp. Parasit Vectors. 2014; 7: 461. Available: http://www.parasitesandvectors.com/content/7/1/461. doi: 10.1186/s13071-014-0461-2 PMID: 25274498

41. Qadri SM, al-Okaili GA, al-Dayel F. Clinical significance of *Blastocystis hominis*. J Clin Microbiol. 1989; 27: 2407–2409. PMID: 2808664

42. Boorom KF, Smith H, Nimri L, Viscogliosi E, Spanakos G, Parkar U, et al. Oh my aching gut: irritable bowel syndrome, *Blastocystis*, and asymptomatic infection. Parasit Vectors. 2008; 1: 40. Available: http://www.parasitesandvectors.com/content/1/1/40. doi: 10.1186/1756-3305-1-40 PMID: 18937874

43. Stensvold CR, Lewis HC, Hammerum AM, Porsbo LJ, Nielsen SS, Olsen KE, et al. *Blastocystis*: unravelling potential risk factors and clinical significance of a common but neglected parasite. Epidemiol Infect. 2009; 137: 1655–1663. doi: 10.1017/S0950268809002672 PMID: 19393117

44. Ramirez JD, Sanchez LV, Bautista DC, Corredor AF, Florez AC, Stensvold CR. *Blastocystis* subtypes detected in humans and animals from Colombia. Infect Genet Evol. 2013; 22: 223–228. doi: 10.1016/j.meegid.2013.07.020 PMID: 23886615

45. Tan TC, Suresh KG, Smith HV. Phenotypic and genotypic characterisation of *Blastocystis hominis* isolates implicates subtype 3 as a subtype with pathogenic potential. Parasitol Res. 2008; 104: 85–93. doi: 10.1007/s00436-008-1163-5 PMID: 18795333

46. Pavlova EA. Culture methods for *Entamoeba histolytica*. Medzkaya Parasit. 1938; 7: 224–227.

47. Hyland NP, Quigley EM, Brint E. Microbiota-host interactions in irritable bowel syndrome: epithelial barrier, immune regulation and brain-gut interactions. World J Gastroenterol. 2014; 20: 8859–8866. doi: 10.3748/wjg.v20.i27.8859 PMID: 25083059

48. Dupont HL. Review article: evidence for the role of gut microbiota in irritable bowel syndrome and its potential influence on therapeutic targets. Aliment Pharmacol Ther. 2014; 39: 1033–1042. doi: 10.1111/apt.12728 PMID: 24665829

49. Sanchez-Aguillon F, Lopez-Escamilla E, Velez-Perez F, Martinez-Flores WA, Rodriguez-Zulueta P, Martinez-Ocaña J, et al. Parasitic infections in a Mexican HIV/AIDS cohort. J Infect Dev Ctries. 2013; 7: 763–766. doi: 10.3855/jidc.3512 PMID: 24129632

50. Lochmiller RL, Deerenberg C. Trade-offs in evolutionary immunology: just what is the cost of immunity?. Oikos 2000; 88: 87–98.

51. Doeschl-Wilson AB, Brindle W, Emmans G, Kyriazakis I. Unravelling the relationship between animal growth and immune response during micro-parasitic infections. PloS One, 2009; 4: e7508. Available: http://dx.plos.org/10.1371/journal.pone.0007508. doi: 10.1371/journal.pone.0007508 PMID: 19838306

52. Bedi B, McNair NN, Förster I, Mead JR. IL-18 cytokine levels modulate innate immune responses and cryptosporidiosis in mice. J Eukaryot Microbiol. 2015; 62: 44–50. doi: 10.1111/jeu.12164 PMID: 25155632

53. Olivo-Diaz A, Romero-Valdovinos M, Gudiño-Ramirez A, Reyes-Gordillo J, Jimenez-Gonzalez DE, Ramirez-Miranda ME, et al. Findings related to IL-8 and IL-10 gene polymorphisms in a Mexican patient population with irritable bowel syndrome infected with *Blastocystis*. Parasitol Res. 2012; 111: 487–491. doi: 10.1007/s00436-012-2830-0 PMID: 22287022

54. Tovy A, Ankri S. Epigenetics in the unicellular parasite *Entamoeba histolytica*. Future Microbiol. 2010; 5: 1875–1884. doi: 10.2217/fmb.10.140 PMID: 21198420

55. Gomez C, Ramirez ME, Calixto-Galvez M, Medel O, Rodriguez MA. Regulation of gene expression in protozoa parasites. Biomed Biotechnol. 2010; 2010: 726045. Available: http://www.hindawi.com/journals/bmri/2010/726045/. doi: 10.1155/2010/726045 PMID: 20204171

56. Adam RD, Nigam A, Seshadri V, Martens CA, Farneth GA, Morrison HG, et al. The *Giardia lamblia* vsp gene repertoire: characteristics, genomic organization, and evolution. BMC Genomics. 2010; 11: 424. Available: http://www.biomedcentral.com/1471-2164/11/424. doi: 10.1186/1471-2164-11-424 PMID: 20618957

57. Stensvold CR, Alfellani M, Clark CG. Levels of genetic diversity vary dramatically between *Blastocystis* subtypes. Infect Genet Evol. 2012; 12: 263–273. doi: 10.1016/j.meegid.2011.11.002 PMID: 22116021

58. Frank SA. Models of parasite virulence. Q Rev Biol. 1996; 71: 37–78. PMID: 8919665

59. Read AF. The evolution of virulence. Trends Microbiol. 1994; 2: 73–76. PMID: 8156274

60. Little TJ, Chadwick W, Watt K. Parasite variation and the evolution of virulence in a Daphnia-microparasite system. Parasitology. 2008; 135: 303–308. PMID: 18005474

61. Sistrom M, Evans B, Bjornson R, Gibson W, Balmer O, Mäser P, et al. Comparative genomics reveals multiple genetic backgrounds of human pathogenicity in the *Trypanosoma brucei* Complex. Genome Biol Evol. 2014; 6: 2811–2819. doi: 10.1093/gbe/evu222 PMID: 25287146

62. Tomasini N, Lauthier JJ, Ayala FJ, Tibayrenc M, Diosque P. How often do they have sex? A comparative analysis of the population structure of seven eukaryotic microbial pathogens. PLoS One. 2014; 9: e103131. Available: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0103131. doi: 10.1371/journal.pone.0103131 PMID: 25054834

REVISTA DO INSTITUTO DE MEDICINA TROPICAL DE SÃO PAULO

JOURNAL OF THE SÃO PAULO INSTITUTE OF TROPICAL MEDICINE

ORIGINAL ARTICLE

http://doi.org/10.1590/S1678-9946202264040

Identification of α-L-fucosidase (ALFuc) of *Blastocystis* sp. subtypes ST1, ST2 and ST3

Joel Martínez-Ocaña[1], Williams Arony Martínez-Flores[1], Angélica Olivo-Díaz[1], Mirza Romero-Valdovinos[1], Fernando Martínez-Hernández[1], Guillermo Aguilar-Osorio[2], Ana Flisser[3], Pablo Maravilla[1]

ABSTRACT

Blastocystis sp. is a common intestinal microorganism. The α-L-fucosidase (ALFuc) is an enzyme long associated with the colonization of the gut microbiota. However, this enzyme has not been experimentally identified in *Blastocystis* cultures. The objective of the present study was to identify ALFuc in supernatants of axenic cultures of *Blastocystis* subtype (ST)1 ATCC-50177 and ATCC-50610 and to compare predicted ALFuc proteins of *alfuc* genes in sequenced STs1–3 isolates in human *Blastocystis* carriers. Excretion/secretion (Es/p) and cell lysate proteins were obtained by processing *Blastocystis* ATCC cultures and submitting them to SDS–PAGE and immunoblotting. In addition, 18 fecal samples from symptomatic *Blastocystis* human carriers were analyzed by sequencing of amplification products for subtyping. A complete identification of the *alfuc* gene and phylogenetic analysis were performed. Immunoblotting showed that the amplified band corresponding to ALFuc (~51 kDa) was recognized only in the ES/p. Furthermore, prediction analysis of ALFuc 3D structures revealed that the domain α-L-fucosidase and the GH29 family's catalytic sites were conserved; interestingly, the galactose-binding domain was recognized only in ST1 and ST2. The phylogenetic inferences of ALFuc showed that STs1–3 were clearly identifiable and grouped into specific clusters. Our results show, for the first time through experimental data that ALFuc is a secretion product of *Blastocystis* sp., which could have a relevant role during intestinal colonization; however, further studies are required to clarify this condition. Furthermore, the *alfuc* gene is a promising candidate for a phylogenetic marker, as it shows a conserved classification with the *SSU-rDNA* gene.

KEYWORDS: α-L-fucosidase. *Blastocystis* sp. Glycoside hydrolase. Phylogenetic marker. Subtypes.

INTRODUCTION

Blastocystis is an anaerobic stramenopile that colonizes the intestinal tract of several taxa[1]. Previous reports focused on the speciation of isolates was based on host species. However, a subtype (ST) classification system based on *SSUr-DNA* genes was established whereas molecular typing revealed a disparity in host-based classification[2]. In addition, guidelines were proposed to correctly identify new *Blastocystis* STs to avoid confusion in the literature; at present, 25 subtypes meet the currently recommended criteria for unique subtype designations (STs1–17, ST21, STs23–26, and STs27–29), while STs18–20 and ST22 have been considered insufficient[2,3]. Although their pathogenicity in humans remains controversial, more than one billion *Blastocystis* carriers could exist worldwide[4]. STs 1–3 have been reported as the most

[1]Hospital General "Dr. Manuel Gea Gonzalez", Departamento de Ecología de Agentes Patógenos, Ciudad de México, México

[2]Universidad Autonoma de México, Instituto de Química, Ciudad de México, México

[3]Universidad Autonoma de México, Facultad de Medicina, Departamento de Microbiologia y Parasitologia, Ciudad de México, México

Correspondence to: Williams Arony Martínez-Flores
Hospital General "Dr. Manuel Gea Gonzalez", Departamento de Ecología de Agentes Patógenos, Calzada de Tlalpan, 4800, Colonia, Sección XVI, 14080, Ciudad de México, México

E-mail: Williams.arony@salud.gob.mx

Pablo Maravilla
Hospital General "Dr. Manuel Gea Gonzalez", Departamento de Ecología de Agentes Patógenos, Calzada de Tlalpan, 4800, Colonia, Sección XVI, 14080, Ciudad de México, México

E-mail: maravillap@yahoo.com

Received: 17 January 2022

Accepted: 25 April 2022

prevalent in human populations, however, other STs have also been found in humans and animals (ST4–10, ST12, ST14, and ST16). On the other hand, some STs can infect birds and mammals[4], however, isolates from amphibian and reptiles appear to be restricted to these groups[5,6].

Intestinal microorganisms can express/secrete different molecules that interact with the host intestinal mucosa, among them, the α-L-fucosidase (ALFuc), a glycoside hydrolase (GH)[7]. According to the "Carbohydrate-Active EnZYmes Database" n.d. (CAZy), ALFuc is an enzyme that catalyzes the hydrolytic removal of L-fucose residues that bind to the non-reducing end of glycan chains, such as mucins[8]. According to their amino acid sequence, α-L-fucosidases are classified into two families: GH29 and GH95. GH29 enzymes are a broad family of retention fucosidases active on α*(1,2)-*, α*(1,3)-*, α*(1,4)-*, and α*(1,6)-L-fucosyl* ligands. GH29 has been further divided into subfamilies A and B, with GH29A being active on a wide range of ligands, whereas GH29B is specific for α*(1,3)-* and α*(1,4)-L-fucosyl* ligands. GH95 enzymes are a small family of inverted fucosidases active on α*(1,2)-fucosyl galactose*[9].

A significant link has been established between human α*-L-fucosidase* (termed FUCA2) and *Helycobacter pylori* adhesion, growth and pathogenicity[10]. Furthermore, it has been documented that growth and invasion of *Campylobacter jejuni fuc*+, 129, 108, and NCTC 11168 strains are increased in the presence of active L-fucosidases released by *Bacteriodes fragilis*, demonstrating that *C. jejuni* is dependent on external fucosidases for further growth and invasion[11]. It has also been shown that the ability of bifidobacteria to metabolize fucosylated compounds (found in breast milk) via fucosidases is an essential mechanism for shaping the intestinal microbiome in humans during the first months of life[12].

There is plenty of information on ALFuc in bacteria. However, ALFuc has not been experimentally identified in *Blastocystis* cultures. Therefore, the objective of the present study was to identify ALFuc in supernatants from commercial axenic cultures of *Blastocystis* ST1, and to compare predicted ALFuc proteins of *alfuc* genes after sequencing STs1–3 isolates from human *Blastocystis* carriers.

MATERIALS AND METHODS

Ethics statement

The current study was approved by the Research and Ethics Committee of the "Dr. Manuel Gea Gonzalez" General Hospital, with reference number 12-77-2018. Written consent was obtained from all participants.

Blastocystis in vitro cultures

The axenic commercial *Blastocystis* ST1 cultures ATCC-50177 and ATCC-50610 were obtained from The American Type Culture Collection (ATCC) and used in this study to perform a protein analysis. For the propagation of axenic strains of *Blastocystis* ST1 ATCC 50177 and 50610, the data sheets recommend the use of *Blastocystis* egg biphasic medium ATCC 1671 supplemented with 10% of horse serum. The constitution of the medium is as follows: 130 mM NaCl, 1 mM $CaCl_2$, 2 mM KCl, 0.4 mM $MgCl_2$, 10 mM Na_2HPO_4, 4 mM $NaHCO_3$, and 2 mM KH_2PO_4. For its preparation, egg yolks were emulsified and sieved; 4 mL were deposited per tube, then eggs were solidified at 60 °C by placing the tubes in an inclined position. Subsequently, 4 mL of Stone's modified Locke's solution were added. The tubes were then sterilized. At the time of use, 10% of horse serum was added and 1 mL of the suspension of each axenic strain of *Blastocystis* was deposited. Cultures were kept in an anaerobic jar with BD GasPak (BD, Franklin Lakes, NJ, USA) for 48 h at 37 °C. After that, they were inoculated into 10 mL of Iscove's modified Dulbecco's medium (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) supplemented with 10% of inactivated horse serum (PAA Laboratories GmbH, Pasching, Austria) and 1% of penicillin/streptomycin (Gibco, Thermo Fisher Scientific, Waltham, MA, USA). The cultures were grown in an anaerobic flask (Merk Millipore, Burlington, MA, USA), containing a BD GasPak (BD, Franklin Lakes, NJ, USA) and incubated at 37 °C for 48 h[13]. Afterward, microscopic observations were performed only on some cultures to confirm cellular growth.

Eighteen fecal samples from adult *Blastocystis* carriers who attended medical consultation at the "Dr. Manuel Gea Gonzalez" General Hospital due to gastrointestinal disorders were screened by microscopy and xenic isolation procedures were carried out within 6 h of deposition[13]. Pea-sized pieces of stool or 250 µL liquid samples were inoculated into 8 mL of Jone's medium supplemented with 5% of inactivated horse serum[13] (PAA Laboratories GmbH, Pasching, Austria) using six culture tubes per sample. The culture tubes were incubated at 37 °C for 48 h (during exponential growth), and the success of the isolation was confirmed using microscopy.

Preparation of *Blastocystis* protein extracts and culture supernatants

For the excretion/secretion protein (Es/p) analysis, ATCC-50177 and ATCC-50610 cultures were centrifuged at 6,000 x g for 10 min, and supernatants (~15 mL)

and cellular pellets (~850 *Blastocystis* cells/mL using a Neubauer chamber) were separately recovered. Then, the supernatants were centrifuged at 16,000 x g at 4 °C for 10 min and passed through a 0.20 µm filter (Corning, Merk Millipore, Burlington, MA, USA) before being placed in an Amicon Ultra-15 100,000 NMWL (Merk Millipore, Burlington, MA, USA) and centrifuged at 4,000 x g for 40 min at 4 °C to remove proteins with a molecular weight greater than 100 kDa. Finally, ES/p was concentrated for 40 min at 4 °C using an Amicon Ultra-4 10,000 NMWL at 4,000 x g, and proteins were resuspended in 2 mL of 50 mM Tris-HCl pH 7.5 and stored at -70 °C until use[14].

For the analysis of *Blastocystis* cell lysate (Bcl), the previously obtained cellular pellets were washed three times with phosphate-buffered saline (PBS) 1X pH 7.2 and centrifuged at 6,000 x g for 10 min. The parasites were lysed by ten freeze-thaw cycles at -70 °C and 37 °C, sonicated on ice using ten one-minute pulses, separated by one minute of relaxation, divided into aliquots, and stored at -70 °C until required[13]. For both ES/p and Bcl, the total protein concentration was determined by a Bradford assay (Bio-Rad Laboratories, Hercules, CA, USA)[15].

Identification of *Blastocystis* ALFuc.

The annotated α-L-fucosidase protein for *Blastocystis* ST7 [UniProtKB-D8M3D3] available in the UniProt database[16] was used to predict its theoretical molecular mass with ExPASY-Compute pI/Mw[17] for further comparisons. After standardizing and testing different concentrations of Es/p and Bcl, the optimal technical conditions were established, and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed to identify the putative ALFuc; to this end 60 µg of Bcl and Es/p were loaded in 2x loading buffer containing 3% of SDS, 150 mM Tris-HCl, pH 6.8, 30% glycerol and 0.1% bromophenol blue. Samples were placed in boiling water for 5 min and then on precast 4-20% Mini-PROTEAN® TGX™ Precast Protein Gels (Bio-Rad Laboratories, Hercules, CA, USA) at 100 volts/h. Coomassie brilliant blue staining was used to identify the putative ALFuc protein band and calculate its Rf using Precision Plus Protein Standards (Bio-Rad Laboratories, Hercules, CA, USA). Three electrophoretic runs were used to perform the relative mobility (Rf) calculation[18].

Immunoblotting for ALFuc

To perform electrophoresis and then transfer the proteins, 20 mg of Es/p and Bcl each were analyzed using the same conditions described for SDS–PAGE. Then the proteins were transferred to polyvinylidene difluoride (PVDF) membranes (Merk Millipore, Burlington, MA, USA) at 100 volts/h. The presence of protein bands was verified by using Ponceau red staining. To standardize the immunoblot conditions, several assays were performed, testing different concentrations of antibodies as well as blocking agents, until the optimal conditions were found. Nonspecific binding sites were blocked in the membrane by incubating it with 1X Blocking Reagent (Roche, Merk Millipore, Burlington, MA, USA) in 50 mM Tris-Base pH 7.5, 0.15 M NaCl, 0.03% Tween-20 (TBS-T) for 1h at room temperature, followed by three TBS-T washes and overnight incubation with anti-α-L-fucosidase (Santa Cruz Biotechnology, Dallas, TX, USA) at a 1:2,000 dilution in blocking buffer with 0.03% TBS-T at 4 °C. The membrane was then washed three times with TBS-T and incubated with biotinylated goat anti-mouse anti-IgG (Santa Cruz Biotechnology, Dallas, TX, USA) at 1:4,000 dilutions for 2 h at room temperature. After three washes, the membrane was incubated with streptavidin-peroxidase (Jackson Immunoresearch, West Grove, PA, USA) at a dilution of 1:10,000 for 1 h at room temperature, followed by three washes with TBS-T. Finally, the membrane was treated with diaminobenzidine (DAB) (Sigma-Aldrich, Merk Millipore, Burlington, MA, USA)[19].

DNA extraction and PCR for the *alfuc* gene

One milliliter of each of the 18 fecal samples, ATCC-50177 and ATCC-50610 cultures, was recovered for DNA extractions. A QIAmp Fast DNA Stool Mini Kit (Qiagen, Hilden, Germany) was used following the manufacturer's instructions to extract the DNA from the *Blastocystis* pellets. The DNA samples were eluted in 50 µL of elution buffer and stored at -70 °C. The extracted DNA was used for polymerase chain reactions (PCR) to obtain a partial sequence of the *SSUrDNA* gene[20] to identify the *Blastocystis* STs in xenic cultures, according to current guidelines[13]. Four new primer pairs were designed in this study to amplify four overlapping regions and to assemble a sequence of approximately 2,500 base pairs (bp) (Table 1), which included the whole *alfuc* gene for STs1–3 and partial regions of the eIF-3 translation initiation factor subunit 4 (*eif3S4*) and the estradiol 17-beta-dehydrogenase (*hsd17B*), as well as two flanking intergenic regions (accession Nº LXWW01000289.1, region: 82284-84859)[21]. In addition to the two strains ATCC 50177 and 50610, a PCR was performed with eighteen isolates obtained from patients to amplify the complete sequence of the *alfuc* gene, using 25 µL volumes: 1X PCR buffer, 2.4 mM $MgCl_2$, 0.5 mM dNTPs, 0.01 mg BSA, 1 U Taq DNA Polymerase

Table 1 - Primers used to amplify partial regions of the *SSUr-RNA* gene and the *alfuc* gene of *Blastocystis*.

Locus	Primer name	Position[$]	Size[¥] (bp)	Direction	5'-3' sequence
SSU-rRNA[&]	Blast 505-532	637-1411	~492	Forward	GGAGGTAGTGACAATAAATC
	Blast 998-1017			Reverse	TGCTTTCGCACTTGTTCATC
α-L-fucosidase[£]	ALFUC/S1	1-928	~872	Forward	AC**W**CCCTCGTTTCC**W**CC**MKY**AG
	ALFUC/As1			Reverse	TTCAG**Y**TTCAC**R**GGGTTGAAGAT
	ALFUC/S2	767-1639	~860	Forward	GAGGAGCA**R**G**YY**CA**R**TGGCT
	ALFUC/As2			Reverse	GTGGTAGAACCAGCC**S**GG**K**C
	ALFUC/S3	1240-2093	~850	Forward	TGGAGACCGGAGTACAACGA
	ALFUC/As3			Reverse	TCTT**S**AC**R**CC**R**ATGGTGGT**K**CCC
	ALFUC/S4	1971-2678	~675	Forward	TCAACGTG**RTS**ATGATGCAGGA
	ALFUC/As4			Reverse	TCC**R**TC**VSTK**GC**W**CC**Y**GT**S**ACC

& = Primer set as described by Santin *et al.*[20] to detect and subtype *Blastocystis* of human and animal origin; £ = Primer set designed in this study to amplify four overlapping regions from partial eif3S4, complete *alfuc* and partial hsd17B of *Blastocystis* STs1–3; $ = Positions of the sequenced regions in the alignments of the complete *SSU-rRNA* gene of *Blastocystis* STs1–17 and non-human/other mammal/bird sources; positions of the four overlapping regions according to sequence alignments of ST1 genomic regions [accession number: LXWW01000289, region: 82284..84859], ST2 [accession number: JZRJ01000159, region: c8456..10980] and ST3 [accession number: JZRK01000455, region: 1512..3972]; ¥ = Approximate expected amplification sequence size on different loci. Bold nucleotides in the primer sequences indicate degenerated nucleotides to anneal targeted positions of STs1–3.

(Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA), 2-5 µL of DNA and nuclease-free water. The PCR products were confirmed by 1% agarose gel electrophoresis. For each *alfuc* region, 200 µL of PCR products were purified using the AxyPrep PCR Clean-up kit (Axigen Biosciences, Union City, CA, USA) following the manufacturer's instructions. The sequences were obtained using the Sanger method at the National Biodiversity Laboratory, Institute of Biology of the Universidad Nacional Autonoma de Mexico (NBL-BI-UNAM).

Three-dimensional structure prediction

The I-TASSER[22] server was used to model the 3D structure of *Blastocystis* ALFuc corresponding to ST1, ST2, and ST3 obtained by translation of the complete *alfuc* gene. This server is under active development intending to provide the most accurate predictions of protein structure and function using state-of-the-art algorithms. After analysis, the models with the highest confidence scores (C-score) were selected. The three-dimensional structures were then analyzed and visualized using PyMOL 2.5.2[23].

Phylogenetic analysis

The quality of chromatograms of the sequenced regions was evaluated with the programs phred, phrap and consed, versions 0.11220.c, 1.090518 and 29, respectively (Seattle, WA, USA). The assembly of the *alfuc* gene sequences was guided using the genomic regions containing putative *eif3S4-alfuc-hsd17B* for STs1–3 as a reference. In addition, we used two sequence datasets from *Blastocystis* STs1–3 and STs4–9 for the validated SSUr-RNA and the predicted alpha-L-fucosidase genes and proteins (Supplementary Table S1). *Proteromonas lacertae* [SSU-rRNA accession Nº U37108.1[21] and predicted *alfuc* accession Nº NGBS01001212[21] region: c18708.20188] was used as an outgroup. Multiple sequence alignments were computed in ClustalW version 2.0 and trimmed with trimAl version 1.2 as implemented in Phylemon version 2.0. ModelTest-NG version 0.2.0 was used to select the best-fit substitution model for nucleotide alignments (SSUr-DNA: HKY+G and *alfuc* gene: K2+G). According to the common methods of phylogenetic inference, maximum likelihood trees were constructed in MEGA X software version 10.1.8 with 1,000 bootstrap replicates. Bayesian trees were computed in Mr. Bayes Software version 3.2.6 for four million generations; the posterior probability distribution and diagnostic frequency were sampled every 1,000 and 105 generations, respectively; and a substitution model was implemented during the analysis by reversible jump. The trees were summarized with a post-burning sample of 50%. Figtree version 1.4.4 was used to edit the phylogenetic trees. The *alfuc* gene sequences were annotated by comparison with the genomic regions (Supplementary Table S1) that code for *Blastocystis* ALFuc sequences, available for ST1 [accession Nº OAO14080.1][24], ST4 [accession Nº XP_014526040.1][24] and ST7 [accession Nº CBK22406.2][24]. The predicted amino acid sequences were further analyzed to theoretically calculate both, the molecular mass and the

isoelectric point with ExPASY-Compute pI/Mw software[17] to predict the signal peptides with SignalP-5 and to identify the conserved domains in the Pfam database[25].

RESULTS

Identification of ALFuc from *Blastocystis* axenic cultures

To identify the ALFuc of *Blastocystis* in SDS-PAGE, the theoretical molecular mass of 51.5 kDa was calculated from the sequence annotated for α-L-fucosidase of *Blastocystis* ST7 in GenBank, using ExPASY-Compute pI/Mw. To obtain this theoretical result, a strong protein band that migrated ~50 kDa was identified in the electrophoresis by staining with Coomassie blue, both in Bcl and in Es/p in the two strains 50177 and 50610 analyzed (Figure 1A). The molecular mass of this protein was calculated by plotting a graph of Rf vs log apparent molecular mass and interpolating the value of Rf = 0.52, obtaining the calculated molecular mass of 51 kDa (Supplementary Figure S1A). This result was consistent with the theoretical molecular mass initially calculated by ExPASY-Compute pI/Mw. In addition, the immunoblot for ALFuc showed a clear recognition of the commercial antibody[26] against this same protein band of ~51 kDa in Es/p of both 50177 and 50610 strains, and to a much lesser extent in their Bcl (Figure 1B).

Comparison of *alfuc* gene sequences in ST1, ST2 and ST3.

To evaluate the variability of the *alfuc* gene in the most common *Blastocystis* STs in humans, subtyping and phylogenetic inferences were performed. The most common STs identified within the 18 isolates were ST1 (3/18), ST2 (7/18), and ST3 (8/18) were confirmed by comparing their partial *SSUrDNA* sequences with a sequence dataset of validated subtypes (Figure 2). The sequenced genomic region containing the *alfuc* gene was obtained for each of the 18 xenic isolates and the two axenic strains. The sequence was then annotated as described above and then compared to the homologous regions in STs1–4 and STs6–9 to define putative exonic and intronic regions in the *alfuc* gene, trim the intergenic regions and detect *eif3S4* and *hsd17B* partial genes. These comparisons recognized the existence of two putative ORFs (open reading frames): ORF1 based on ST4 annotation, identified in all STs (STs1–4 and STs6–9), and ORF2 based on ST1 annotation, identified only in ST1 and ST8. The primary variation between the two ORFs was that ORF2 starts at intron one and can be translated into a shorter protein without a signal peptide (Supplementary Table S1). These proteins were analyzed using Pfam[25] to identify the presence of conserved domains in the proteins obtained in this study through amplification of the complete *alfuc* gene and its subsequent translation into amino acids.

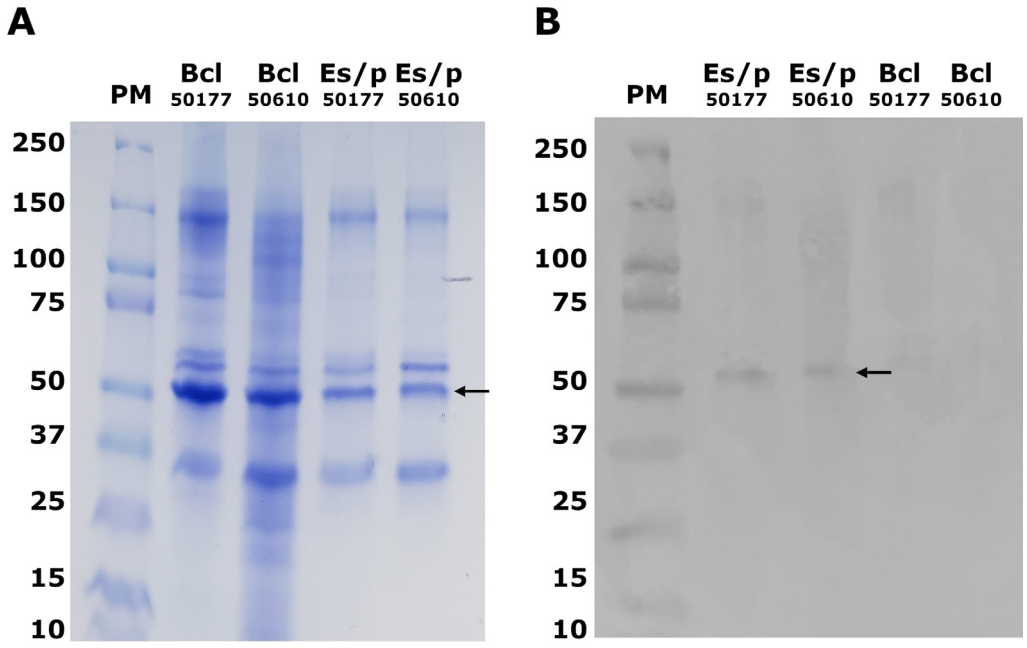


Figure 1 - SDS-PAGE and Immunoblot of ALFuc *Blastocystis*: A) SDS-PAGE, stained with Coomassie blue of cells lysate (Bcl) and *Blastocystis* excretion/secretion proteins (Es/p), from axenic strains ATCC 50177 and ATCC 50610 are shown. The arrow indicates the molecular mass calculated by the electrophoretic mobility (Rf) of 51 kDa for *Blastocystis* ALFuc; B) the recognition of *Blastocystis* ALFuc by a commercial anti-α-L-fucosidase antibody is observed. The arrow indicates the protein recognized by the antibody with an expected molecular weight of 51 kDa which is consistent with the calculated molecular mass on SDS-PAGE for *Blastocystis* ALFuc.

According to this analysis, we identified the characteristic domain (IPR000933) of the GH29 family of glycoside hydrolases, the conserved domain of the α-L-fucosidase family of the CL0058 clan, and a similar galactose-binding domain (IPR008979), identified as the C-terminal domain (IPR000421) of coagulation factors 5/8 (Supplementary Figure S1B). In addition, in these same 13 proteins with ORF1, the signal peptide was identified at the N-terminal end. In proteins with ORF2, no signal peptide was present. The SignalP-5.023[27] analysis identified the signal peptide through peptide excision only in sequences with ORF1. It was found between amino acids 16 and 17 (VLA-RP) with a probability of 0.6139 in ST1 and 0.79 in ST2, in amino acid positions 14–15 (ALA-KP) with a 0.79 probability for ST3, and in positions 15–16 (VLS-KR). Analysis for transmembrane regions showed negative results for all proteins.

To strengthen the result obtained by the Pfam analysis confirming that our proteins belong to the GH29 family of α-L-fucosidases, we performed a ClustalW alignment on the proteins with three α-L-fucosidases characterized as belonging to the GH29 family: *Bacteroides thetaiotaomicron* (*B. theta.*) [UniProtKB-Q8A085][16], *Bifidobacterium bifidum* [UniProtKB-C5NS94][16] and *Streptomyces sp*. [UniProtKB-Q9Z4I9][16]. Our alignment analysis showed a high conservation of the characteristic catalytic site of the α-L-fucosidases of the GH-29 family, consisting of the characteristic catalytic nucleophilic residue Asp (D) and the acid/base residue Glu (E) (Figure 3).

Prediction of three-dimensional structures

The designed structures were analyzed by the I-TASSER server. This server generates some structural conformations, then used the SPICKER program to group all the structures based on the similarity of their paired structures. Finally, for *Blastocystis* ALFuc ST1, ST2 and ST3 the server reported five main models each of them, corresponding to the five largest clusters. The confidence of each model was calculated using the C-score. The C-score values showed the accuracy of the predicted model, which is usually in the range of -5 to 2. In addition, the higher the value of the C-score, the better the quality of the prediction. The C scores of the *Blastocystis* ALFuc models ST1, ST2, and ST3 were 1.22, 0.93, and 0.83, respectively. Therefore, ALFuc *Blastocystis* ST1 with a C score of 1.22 showed greater accuracy among the predicted models. Figure 4 shows the tertiary structures of the ALFuc of *Blastocystis* ST1, ST2, and ST3. In the image, one can see the domain of the α-L-fucosidase, consisting of five α-helix structures (magenta color); the similar domain of binding to galactose, formed by five structures of β-strand (orange color). In ST3, the Pfam result for this domain was negative. Likewise, the catalytic site of the GH29 family was observed, formed by the nucleophilic residue D (red spheres) and the acid/base residue E (blue spheres).

Phylogenetic analysis

The phylogenetic relationships of *Blastocystis* were explored to determine whether α-L-fucosidase could infer the characteristic of the *SSU-rDNA* topology or the one described for *nad* genes from the *Blastocystis* mitochondrion-related organelle (MRO) genomes. Therefore, we undertook phylogenetic analyses using two partial regions and the complete *alfuc* genes from *Blastocystis* STs1–3 and STs4–9. Partial region two spans most of the α-L-fucosidase domain, and region three contains the galactose-binding-like domain. Our results confirmed a conserved identity in the tree topology among *alfuc* and SSU-rDNA in the phylogenetic analysis. The gene tree topology (Figure 3) is similar to that described for *Blastocystis* by SSU-rDNA[28], MRO *nad* genes[29], and Miro protein[30]. The topology consisted of ST1 and ST2 clustered together; ST3 clustered or associated with ST4 and ST8, and ST7 clustered or associated with ST6 and ST9 (Figure 5). Interestingly, in the phylogenetic analysis of the two partial regions of the *alfuc* gene, ST3 and ST7 shuffled their associations from their common clusters, while the STs1/2, STs4/8 and STs 6/9 clusters remained consistent (data not shown).

DISCUSSION

In the present study, ALFuc was identified in Es/p by predicting its molecular mass in electrophoresis, further using a commercial antibody directed against the secreted α-L-fucosidase, suggesting that this enzyme could be part of the secretion products of *Blastocystis*. This finding is consistent with previous reports predicting ALFuc as a secreted protein related to carbohydrate metabolism in *Blastocystis*[31], but this claim has not been supported by experimental evidence so far.

Likewise, through the amplification of the *alfuc* gene, and its subsequent translation into proteins, the amino acid sequences of the ALFucs ST1-3 were obtained. These sequences were analyzed using Pfam to identify the domains. The result of this analysis showed the identification of the characteristic domain of α-L-fucosidases, GH29 enzymes, and the similar galactose-binding domain (only in ST1 and ST2). To support these findings, the catalytic sites of

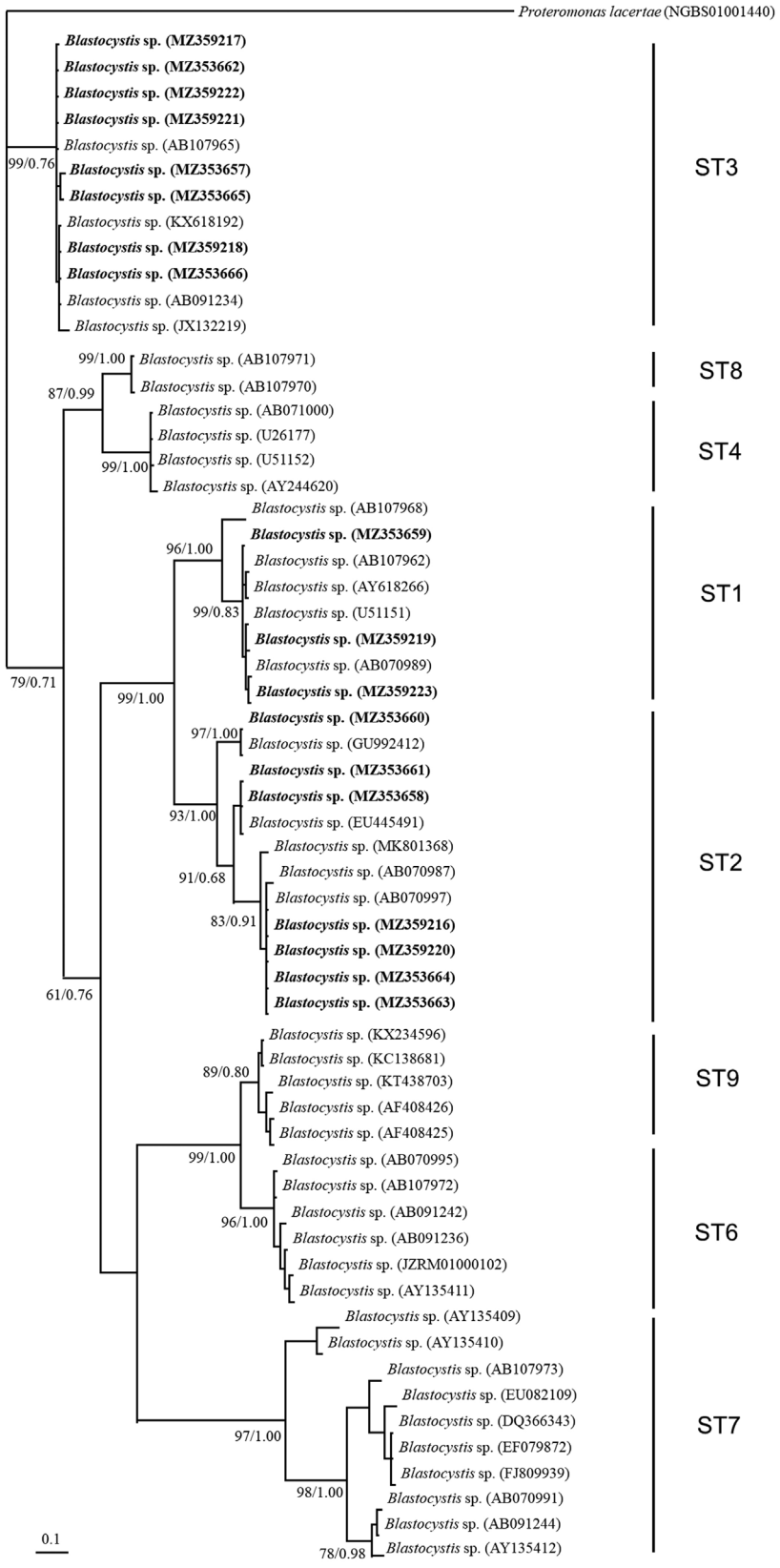


Figure 2 - Phylogenetic inference of *SSUrDNA* gene partial sequence of *Blastocystis* STs1-4 and STs 6-9. The orthologous region of *Proteromonas lacertae* was used as the outgroup. The values of the nodes indicate the bootstrap proportions and Bayesian posterior probabilities in the following order: maximum likelihood/Bayesian analysis. The sequences obtained in the present study are indicated in bold.

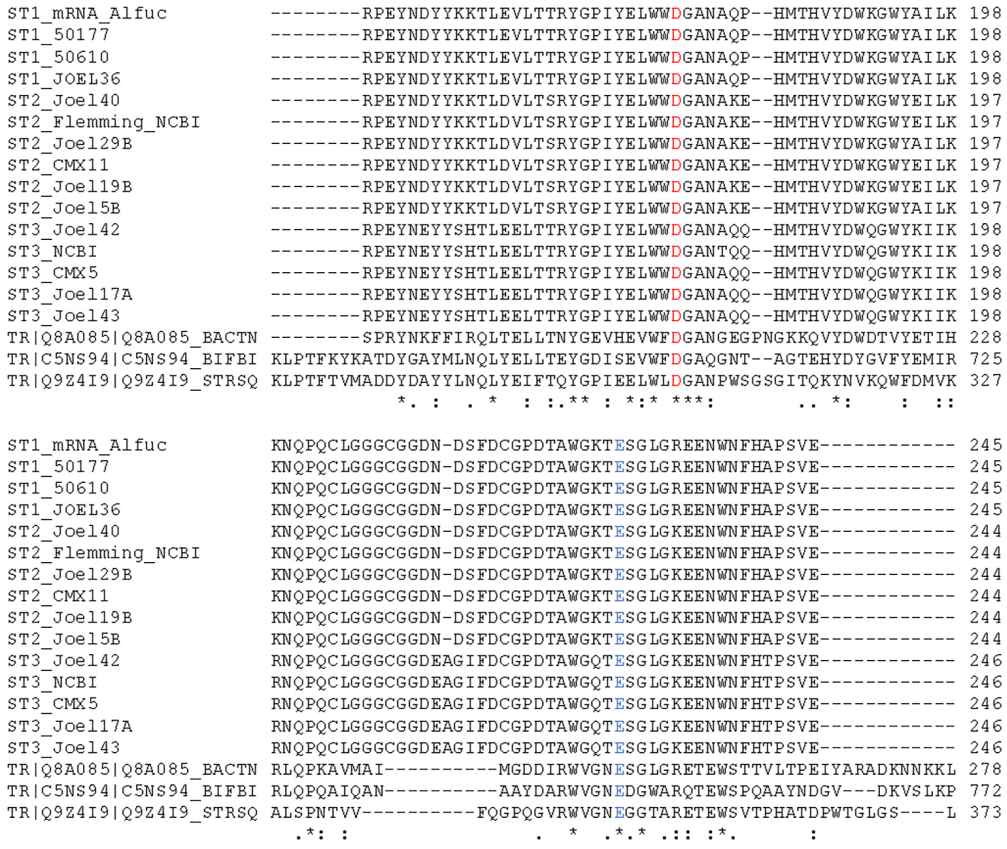


Figure 3 - Sequence alignment by ClustalW. Only the region corresponding to the catalytic sites of the α-L-fucosidases of the GH29 family is shown. To compare the region of catalytic sites, our sequences were aligned with three α-L-fucosidases of the GH29 family from the UniProt database: *Bacteroides thetaiotaomicron* Q8A085, *Bifidobacterium bifidum* C5NS94, and *Streptomyces sp* Q9Z4I9. The catalytic site is shown with the nucleophilic asp residue (D) in red and the general acid/base Glu residue (E) in blue.

our ALFucs were compared with α-L-fucosidases of the GH29 family of homologous bacteria deposited in the databases. In the present study, we identified the conserved nucleophilic residue D and the acid/base residue E in our sequences. Our results are consistent with the catalytic mechanism described in bacteria for α-L-fucosidase GH29, which requires two amino acid residues, one of which plays the overall acid/base role, while the other acts as a nucleophile[32].

In our study, we could only identify a similar galactose-binding domain in ST1 and ST2. In *Akkermansia muciniphila,* proteins with a similar galactose-binding domain involved in mucin degradation play a significant role in host intestinal health, as they downregulate hydrolytic activity in regions with injured tissues, improving tissue regeneration and wound healing[33]. In *Ruminococcus gnavus*, another human gut symbiont commonly associated with inflammatory bowel disease, the ability to grow with mucin as the sole carbon source depends on the strain and is related to gene groups, including ALFuc proteins with galactose-like binding domain[34]. In addition, it has been observed that the absence of beta domains of analogous galactose-like binding with a topology like a gelatin roll is related to lower hydrolytic activity of β-galactosidase in *Bacillus circulans*[35]. It has been suggested that this domain helps galactosyl-lactose molecules to be correctly targeted within the active site to efficiently hydrolyze thus producing galactose/glucose and inhibiting the accumulation of galacto-oligosaccharides (GOS)[35,36]. In the present study, the ALFuc of *Blastocystis* ST3 showed a negative result for a similar domain of galactose binding, possibly related to affinity/activity for the carbohydrate as mentioned in the examples above. But these data will have to be analyzed with the characterized protein.

Blastocystis ST have been described "as divergent as species or even genera"[4] capable of colonizing a wide variety of hosts[1]. Therefore, it is not surprising that *Blastocystis* is considered the most widespread intestinal micro-eukaryote, with an almost complete lack of host specificity. Due to its potential impact on human health, significant efforts have been made to associate the genetic diversity of *Blastocystis* with gastrointestinal clinical manifestations of zoonotic transmission by analyzing

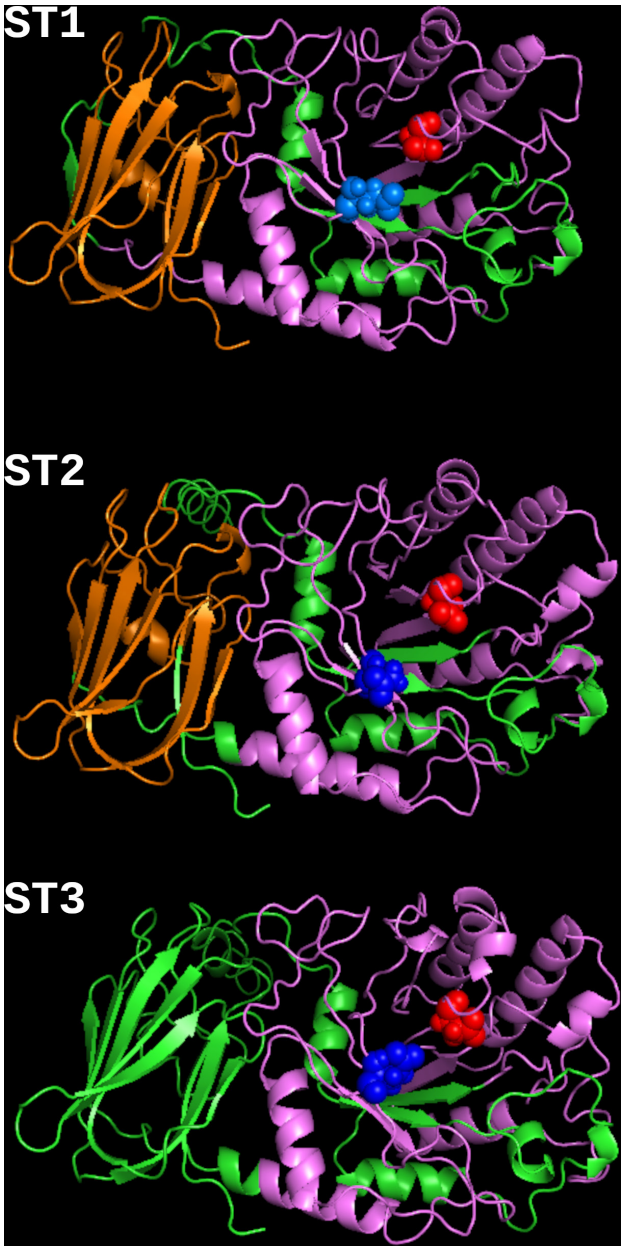


Figure 4 - 3D structures of the ALFuc of *Blastocystis* STs1-3. The image shows the ALFuc domain, formed by five α-helix structures (magenta color), the C-type domain F5/8 or similar galactose-binding domain, formed by five β-strand structures (orange color), in ST1 and ST2. In addition, the catalytic site of the GH29 family, consisting of the nucleophilic residue D (red spheres) and the acid/base residue E (blue spheres), is observed. For ST3, the Pfam result for the similar galactose-binding domain was negative.

different molecular markers between geographically distant populations or isolated from symptomatic and asymptomatic patients[37-39]. Although these studies have demonstrated surprising findings, it was not possible to identify a clear association with pathogenic signatures. A remarkable finding in our study is the conserved topology in the phylogenetic tree of the *alfuc* gene, with the topology described by the subunits of the *nad* gene for the same STs, and the grouping for STs1–4 and STs6–9, with the classical clustering obtained with the *SSUr-RNA* genes[28].

In this study, a genomic region containing the *alfuc* gene was amplified using four pairs of degenerate primers designed to obtain the entire gene for comparative functional analysis rather than to identify a phylogenetic marker. The main limitations and advantages of this strategy should be pointed out to avoid possible biases or to be addressed by future studies.

The primers were designed to amplify regions of interest only for ST1–3. As most of the samples analyzed in global studies are of human origin and "more than 90% of human strains belong to ST1–4"[4], these primers may be advantageous if the study entails ST1–3 or a phylogenetically close ST. However, problems can arise when analyzing the most divergent STs, especially those that colonize amphibians, reptiles, or insects. Unlike previous studies that used DNA extracted from stool samples to evaluate new phylogenetic markers, we used DNA extracted from xenic cultures. The first approach represents the most challenging limitation for field studies using single-copy markers, as they are more susceptible to biases associated with a low parasite load in the sample, e.g., a recent report that analyzed DNA from human stool samples ruled out eight of the 12 new markers evaluated for *Blastocystis* because it was not possible to obtain quality PCR amplification products[2]. Although xenic isolation increases the number of parasites and, consequently, the number of positive samples, this procedure favors the selection of specific subtypes[13], a particular problem for studies designed to describe genetic variability in a particular population. To solve this problem, new studies could implement the next-generation sequencing approaches used to study the genetic diversity of *Blastocystis* subtypes in humans and animals[40] and target new genes such as *alfuc* or implement new available technologies to obtain complete genetic sequences[4], as has already been implemented for SSU-rRNA genes.

CONCLUSION

Our experimental results showed, for the first time that ALFuc is a secretion product of *Blastocystis* sp., which could have a relevant role during intestinal colonization; however, further studies are required to clarify this condition. Furthermore, the *alfuc* gene is a promising candidate to act a phylogenetic marker according to the resulting subtypes, and needs to be investigated to clarify whether the same behavior occurs with other subtypes.

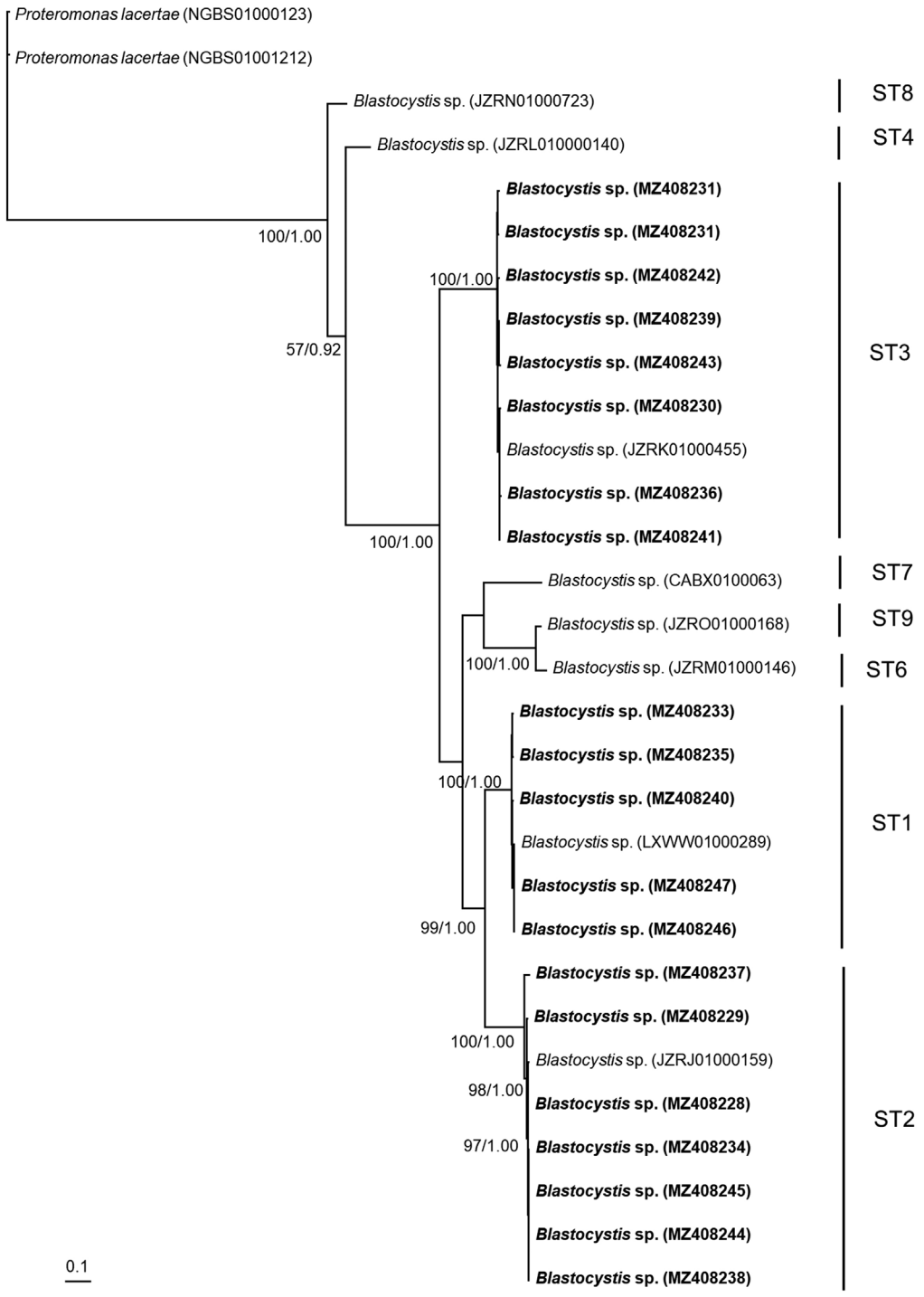


Figure 5 - Phylogenetic inference of the complete *alfuc* genes from *Blastocystis* STs1–4 and STs 6–9. The orthologous region of *Proteromonas lacertae* was used as the outgroup. The values of the nodes indicate the bootstrap proportions and Bayesian posterior probabilities in the following order: maximum likelihood/Bayesian analysis. The sequences obtained in the present study are indicated in bold.

ACKNOWLEDGMENTS

We thank Laura Margarita Márquez-Valdelamar and Nelly López from the National Biodiversity Laboratory at the Institute of Biology of the *Universidad Nacional Autónoma de México* (NBL-BI-UNAM) for their support in obtaining the DNA sequences. We also thank Tobías Portillo Bobadilla, member of *Red de Apoyo a la Investigación* (RAI), *Universidad Nacional Autónoma de México* e *Instituto Nacional de Ciencias Médicas y Nutrición*

Salvador Zubirán, Mexico City, for his support in the analysis of the predicted ALFuc sequences. Os nomes dos locais estão em itálico e deveriam ter sido deixados em espanhol

AUTHORS' CONTRIBUTIONS

JMO, WAMF, and PM conceptualized the study. JMO, WAMF, and MRV performed the experimental work and collected the data. JMO, WAMF, and FMH analyzed the data. JMO and WAMF performed the bioinformatics analyses. JMO and WAMF drafted the manuscript, which was reviewed by PM, AOD, AF, and GAO. PM supervised the project, experimental design, data collection and analysis, and manuscript preparation. All authors contributed to the article and approved the submitted version.

CONFLICT OF INTERESTS

The authors declare that they have no competing interests.

FUNDING

This work was partially supported by *El Consejo Nacional de Ciencia y Tecnología* (CONACyT) grant Nº 168619. JMO is a PhD student of the *Programa en Ciencias Biológicas*, *Universidad Nacional Autónoma de México* (UNAM).

REFERENCES

1. Tan KS. New insights on classification, identification, and clinical relevance of Blastocystis spp. Clin Microbiol Rev. 2008;21:639-65.
2. Stensvold CR, Clark CG. Pre-empting Pandora's box: blastocystis subtypes revisited. Trends Parasitol. 2020;36:229-32.
3. Maloney JG, Santin M. Mind the gap: new full-length sequences of blastocystis subtypes generated via Oxford Nanopore Minion Sequencing allow for comparisons between full-length and partial sequences of the small subunit of the ribosomal RNA gene. Microorganisms. 2021;9:997.
4. Stensvold CR, Tan KS, Clark CG. Blastocystis. Trends Parasitol. 2020;36:315-6.
5. Maloney JG, Molokin A, Cunha MJ, Cury MC, Santin M. Blastocystis subtype distribution in domestic and captive wild bird species from Brazil using next generation amplicon sequencing. Parasite Epidemiol Control. 2020;9:e00138.
6. Yoshikawa H, Koyama Y, Tsuchiya E, Takami K. Blastocystis phylogeny among various isolates from humans to insects. Parasitol Int. 2016;65:750-9.
7. Martínez-Ocaña J, Maravilla P, Olivo-Díaz A. Interaction between human mucins and parasite glycoproteins: the role of lectins and glycosidases in colonization by intestinal protozoa. Rev Inst Med Trop Sao Paulo. 2020;62:e64.
8. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 2014;42:D490-5.
9. Sakurama H, Tsutsumi E, Ashida H, Katayama T, Yamamoto K, Kumagai H. Differences in the substrate specificities and active-site structures of two α-L-fucosidases (glycoside hydrolase family 29) from Bacteroides thetaiotaomicron. Biosci Biotechnol Biochem. 2012;76:1022-4.
10. Liu TW, Ho CW, Huang HH, Chang SM, Popat SD, Wang YT, et al. Role for alpha-L-fucosidase in the control of Helicobacter pylori-infected gastric cancer cells. Proc Natl Acad Sci U S A. 2009;106:14581-6.
11. Luijkx YM, Bleumink NM, Jiang J, Overkleeft HS, Wösten MM, Strijbis K, et al. Bacteroides fragilis fucosidases facilitate growth and invasion of Campylobacter jejuni in the presence of mucins. Cell Microbiol. 2020;22:e13252.
12. Curiel JA, Peirotén A, Landete JM, Ruiz de la Bastida A, Langa S, Arqués JL. Architecture insight of Bifidobacterial α-L-Fucosidases. Int J Mol Sci. 2021;22:8462.
13. Clark CG, Stensvold CR. Blastocystis: isolation, xenic cultivation, and cryopreservation. Curr Protoc Microbiol. 2016;43:20A.1.1-8.
14. Wawrzyniak I, Texier C, Poirier P, Viscogliosi E, Tan KS, Delbac F, et al. Characterization of two cysteine proteases secreted by Blastocystis ST7, a human intestinal parasite. Parasitol Int. 2012;61:437-42.
15. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. 1976;72:248-54.
16. UniProtKB: D8M3D3 (D8M3D3_BLAHO). [cited 2022 Apr 25]. Available from: https://www.uniprot.org/uniprot/D8M3D3
17. Swiss Institute of Bioinformatics. Expasy: Compute pI/Mw tool. [cited 2022 Apr 25]. Available from: https://web.expasy.org/compute_pi/
18. Matsumoto H, Haniu H, Komori N. Determination of protein molecular weights on SDS-PAGE. Methods Mol Biol. 2019;1855:101-5.
19. Tsang VC, Peralta JM, Simons AR. Enzyme-linked immunoelectrotransfer blot techniques (EITB) for studying the specificities of antigens and antibodies separated by gel electrophoresis. Methods Enzymol. 1983;92:377-91.
20. Santín M, Gómez-Muñoz MT, Solano-Aguilar G, Fayer R. Development of a new PCR protocol to detect and subtype Blastocystis spp. from humans and animals. Parasitol Res. 2011;109:205-12.
21. National Center for Biotechnology Information. Blastocystis sp. ATCC 50177/Nand II strain NandII AV274scaffold_40, whole

genome shotgun sequence. [cited 2022 Apr 25]. Available from: https://www.ncbi.nlm.nih.gov/nuccore/LXWW01000289.1

22. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. Nat Methods. 2015;12:7-8.
23. Schrödinger. PyMOL. [cited 2022 Apr 25]. Available from: https://pymol.org/2/
24. National Center for Biotechnology Information. Protein. [cited 2022 Apr 25]. Available from: https://www.ncbi.nlm.nih.gov/protein/
25. European Molecular Biology Laboratory. Pfam. [cited 2022 Apr 25]. Available from: http://pfam.xfam.org/
26. Ezawa I, Sawai Y, Kawase T, Okabe A, Tsutsumi S, Ichikawa H, et al. Novel p53 target gene FUCA1 encodes a fucosidase and regulates growth and survival of cancer cells. Cancer Sci. 2016;107:734-45.
27. Technical University of Denmark. DTU health tech: SignalP - 5.0. [cited 2022 Apr 25]. Available from: https://services.healthtech.dtu.dk/service.php?SignalP-5.0
28. Alfellani MA, Taner-Mulla D, Jacob AS, Imeede CA, Yoshikawa H, Stensvold CR, et al. Genetic diversity of Blastocystis in livestock and zoo animals. Protist. 2013;164:497-509.
29. Jacob AS, Andersen LO, Bitar PP, Richards VP, Shah S, Stanhope MJ, et al. Blastocystis mitochondrial genomes appear to show multiple independent gains and losses of start and stop codons. Genome Biol Evol. 2016;8:3340-50.
30. Gentekaki E, Curtis BA, Stairs CW, Klimeš V, Eliáš M, Salas-Leiva DE, et al. Extreme genome diversity in the hyper-prevalent parasitic eukaryote Blastocystis. PLoS Biol. 2017;15:e2003769.
31. Eme L, Gentekaki E, Curtis B, Archibald JM, Roger AJ. Lateral gene transfer in the adaptation of the anaerobic parasite Blastocystis to the gut. Curr Biol. 2017;27:807-20.
32. Tarling CA, He S, Sulzenbacher G, Bignon C, Bourne Y, Henrissat B, et al. Identification of the catalytic nucleophile of the family 29 alpha-L-fucosidase from Thermotoga maritima through trapping of a covalent glycosyl-enzyme intermediate and mutagenesis. J Biol Chem. 2003;278:47394-9.
33. Kosciow K, Deppenmeier U. Characterization of a phospholipid-regulated β-galactosidase from Akkermansia muciniphila involved in mucin degradation. Microbiologyopen. 2019;8:e00796.
34. Crost EH, Tailford LE, Le Gall G, Fons M, Henrissat B, Juge N. Utilisation of mucin glycans by the human gut symbiont Ruminococcus gnavus is strain-dependent. PLoS One. 2013;8:e76341.
35. Song J, Imanaka H, Imamura K, Minoda M, Yamaguchi S, Nakanishi K. The discoidin domain of Bacillus circulans β-galactosidase plays an essential role in repressing galactooligosaccharide production. Biosci Biotechnol Biochem. 2013;77:73-9.
36. Yin H, Pijning T, Meng X, Dijkhuizen L, van Leeuwen SS. Biochemical characterization of the functional roles of residues in the active site of the β-Galactosidase from Bacillus circulans ATCC 31382. Biochemistry. 2017;56:3109-18.
37. Villalobos G, Orozco-Mosqueda GE, Lopez-Perez M, Lopez-Escamilla E, Córdoba-Aguilar A, Rangel-Gamboa L, et al. Suitability of internal transcribed spacers (ITS) as markers for the population genetic structure of Blastocystis spp. Parasit Vectors. 2014;7:461.
38. Alarcon-Valdes P, Villalobos G, Martinez-Flores WA, Lopez-Escamilla E, Gonzalez-Arenas NR, Romero-Valdovinos M, et al. Can the pyruvate: ferredoxin oxidoreductase (PFOR) gene be used as an additional marker to discriminate among Blastocystis strains or subtypes? Parasit Vectors. 2018;11:564.
39. Higuera A, Muñoz M, López MC, Reyes P, Urbano P, Villalobos O, et al. Succinate dehydrogenase gene as a marker for studying Blastocystis genetic diversity. Heliyon. 2020;6:e05387.
40. Maloney JG, Cunha MJ, Molokin A, Cury MC, Santin M. Next-generation sequencing reveals wide genetic diversity of Blastocystis subtypes in chickens including potentially zoonotic subtypes. Parasitol Res. 2021;120:2219-31.

Supplementary Material available from:
https://doi.org/10.48331/scielodata.LVLQEX