Diversidad genética de *Blastocystis* y *Giardia intestinalis* en diferentes regiones de Colombia

Adriana Marcela Higuera Gélvez

"Tesis de doctorado presentada como requisito para optar al título de doctor en ciencias biomédicas y biológicas "

DOCTORADO EN CIENCIAS BIOMÉDICAS Y BIOLÓGICAS

UNIVERSIDAD DEL ROSARIO

BOGOTÁ, D.C.

2020

Diversidad genética de *Blastocystis* y *Giardia intestinalis* en diferentes regiones de Colombia

Estudiante

Adriana Marcela Higuera Gélvez

Microbióloga, Universidad de Los Andes

Magíster en Ciencias Biológicas, Universidad de Los Andes

Director

Juan David Ramírez González, PhD

Director Grupo de Investigaciones Microbiológicas-UR (GIMUR)

Facultad de Ciencias Naturales, Universidad del Rosario

DOCTORADO EN CIENCIAS BIOMÉDICAS Y BIOLÓGICAS

UNIVERSIDAD DEL ROSARIO

BOGOTÁ, D.C.

2020

AGRADECIMIENTOS

- En primera instancia, quisiera agradecer a Dios por darme la oportunidad de culminar esta etapa de mi vida, cumpliendo mi propósito de ser doctora en ciencias.

- A mis adorados padres y a mi hermanita, por ser mis principales motivadores desde siempre, por ser mi ejemplo de vida, por darme la fortaleza, palabras reconfortantes cuando ha sido necesario y por el apoyo para el cumplimiento de cada meta que me he propuesto.

- A mi esposo, quien me ha brindado su amor y apoyo incondicional en todo momento, por su colaboración y paciencia, que fueron indispensables para poder culminar esta etapa de mi vida.

- A mis dos lindas niñas, por ser el motor de mi existencia y la razón más importante de superación día a día.

- A Juan, porque más que ser mi director, fue un guía y un gran apoyo a lo largo de este camino, gracias por la confianza, por compartir ese enorme conocimiento y por permitirme ser parte de tan excelente grupo de investigación.

- A mis compañeros y amigos del grupo GIMUR, por su compañía, consejos y ayuda durante estos años. En especial, quiero darle las gracias a Marina, Lissa, Luz Helena, Giovanny y Carolina por ayudarme en los momentos en los que más los necesité, por escucharme y guiarme en cada instante. Gracias por aportarme tantas cosas valiosas profesional y personalmente.

- A todas los estudiantes de pregrado, maestría y pasantes que han sido parte de GIMUR en algún momento o que aún lo siguen siendo, por brindar esa energía renovadora que, con sus esfuerzos, dedicación y ganas de aprender, enriquece al grupo cada día.

- A mis jurados, los doctores James Richardson, Paula Hernández y Fabiana Lora, quienes han tenido una gran disposición para escuchar y aportar a mi trabajo de doctorado.

- Al comité doctoral del programa.

- A Minciencias, dentro del marco del Programa Nacional para promover la formación en investigación (convocatoria 757), quien fue el principal ente financiador de este doctorado. Y por financiar el proyecto de investigación “Desarrollo de una estrategia y plataforma de Multilocus Sequence Typing (MLST) para la caracterización molecular de siete parásitos intestinales” código 122271250521 y número de contrato 130–2016.

TABLA DE CONTENIDO

1. LISTA DE FIGURAS

2. LISTA DE TABLAS

3. LISTA DE PUBLICACIONES

3.1 Capítulo 1

1. Villamizar X, **Higuera A**, Herrera G, Vasquez-A LR, Buitron L, Muñoz LM, Gonzalez-C FE, Lopez MC, Giraldo JC, Ramírez JD. Molecular and descriptive epidemiology of intestinal protozoan parasites of children and their pets in Cauca, Colombia: a cross-sectional study. BMC Infectious Diseases. 2019 Dec;19(1):1-1. doi: 10.1186/s12879-019-3810-0.
2. **Higuera A**, Villamizar X, Herrera G, Giraldo JC, Vasquez-A LR, Urbano P, Villalobos O, Tovar C, Ramírez JD. Molecular detection and genotyping of intestinal protozoa from different biogeographical regions of Colombia. PeerJ. 2020 Mar 9;8:e8554. doi: 10.7717/peerj.8554.

3.2 Capítulo 2

1. **Higuera A**, Muñoz M, López MC, Reyes P, Urbano P, Villalobos O, Ramírez JD. Development of a multilocus sequence typing scheme for giardia intestinalis. Genes. 2020 Jul;11(7):764. doi: 10.3390/genes11070764.

3.3 Capítulo 3

Artículo sometido

1. **Higuera A**, Muñoz M, López MC, Reyes P, Urbano P, Villalobos O, Ramírez JD. Succinate dehydrogenase gene as a marker for studying *Blastocystis* genetic diversity. Sometido a la revista Helyion.

4. LISTA DE ABREVIATURAS

Abreviatura	Término
gdh	Glutamato deshidrogenasa
tpi	Triosa fosfato isomerasa
bg	Beta giardina
SNPs	Polimorfismos de nucleótido simple
ssuARNr	Subunidad pequeña del ARN ribosomal
ADN	Ácido desoxirribonucleico
ARN	Ácido ribonucleico
ST	Subtipo
ITS	Espaciador interno transcrito
ICBF	Instituto colombiano de bienestar familiar
ASML	Análisis de secuencias multilocus
RFLPs	Fragmentos de restricción de longitude polimórfica
RAPDs	Amplificación aleatoria de ADN polimórfico
MLG	Análisis de genotipificación multilocus
NGS	Secuenciación de nueva generación
MLST	Tipificación de secuencias multilocus
SDHA	Succinato deshidrogenasa subunidad A
PFOR	Piruvato ferredoxin oxidoreductasa

5. RESUMEN

En el mundo, *Blastocystis* ha sido reportado como el microorganismo eucariota más común en el intestino de humanos y animales, con prevalencias incluso hasta del 100%, junto con *Giardia intestinalis*, quien ha sido considerado el principal agente causal de cuadros diarreicos en humanos, afectando aproximadamente 200 millones de individuos a nivel mundial.

En general, el diagnóstico de ambos microorganismos se basa frecuentemente en la observación de formas típicas en heces y solo a nivel de investigación se busca su tipificación por medio del uso de técnicas moleculares. Además, cabe resaltar que, pocos genes han sido utilizados para evaluar sus características a nivel genético y que la mayoría de estudios moleculares se han enfocado únicamente a la tipificación y pocos han incluído la evaluación de la diversidad genética y estructura de poblaciones intra e inter grupos, dejando de lado el estudio de su biología, taxonomía, distribución, potencial zoonótico de transmisión e incluso obtener evidencia que permita esclarecer si su estructura poblacional es de tipo sexual o clonal.

En Colombia la situación no es muy diferente a lo reportado en otros países, principalmente aquellos en vía de desarrollo, donde las condiciones comportamentales, socioeconómicas y ambientales favorecen la transmisión de algunas enfermedades intestinales, y, zonas con alta endemicidad propenden por el aumento en la variación genética de estos microorganismos, que, posiblemente, junto con el ambiente de fuerte competencia a nivel intestinal, hacen más factible la ampliación hacia nuevos hospederos y por ende mantienen la transmisión zoonótica, tanto con animales domésticos como silvestres, cumpliendo un papel fundamental en el mantenimiento de los ciclos epidemiológicos y haciendo cada vez más difícil el control y prevención de estas infecciones. Por esta razón, éste estudio buscó, no solamente detectar molecularmente estos microorganismos en diferentes regiones del país, sino también conocer sus genotipos circulantes, realizando un primer acercamiento a la epidemiología molecular de estos parásitos intestinales en las zonas evaluadas, y posteriormente, evaluar la diversidad existente a nivel intra taxa tanto de *G. intestinalis* como de *Blastocystis,* por medio del análisis de nuevos blancos genéticos que nos permitieron demostrar una gran variación genética, particularmente relacionada con posibles eventos de intercambio genético en *G. intestinalis* y variación intra Subtipo (ST) en el caso de *Blastocystis*.

Los resultados producidos por esta tesis permitieron: a) describir una primera aproximación de la frecuencia de cada uno de los microorganismos parasitarios evaluados, como *Blastocystis, G. intestinalis, Cryptosporidium* y el complejo *Entamoeba histolytica/dispar/moshkovskii* en cinco regiones biogeográficas de Colombia, sugiriendo que la región del Caribe tiene una mayor frecuencia para cada uno de estos. Además, gracias a los ensamblajes de *Giardia*, los STs de *Blastocystis* y especies de *Cryptosporidium* halladas,

mostrar la posible transmisión zoonótica de estos microorganismos hacia los seres humanos en algunas regiones del país, b) mostrar la variabilidad genética presente en *G. intestinalis* y los eventos de intercambio genético intra e inter ensamblaje, por medio de la evaluación de nueve blancos genéticos y finalmente c) determinar la diversidad genética de *Blastocystis* y su variación intra subtipo, gracias al uso en conjunto del gen ribosomal *18s* y el gen metabólico *sdhA*. En conclusión, estos resultados muestran la gran necesidad de seguir evaluando las características genéticas de estos microorganismos, primordialmente en Colombia, para lograr establecer intervenciones de control y prevención más efectivas, y más aún, teniendo en cuenta, que la gran diversidad encontrada puede influir directamente en su prevalencia, dinámica de transmisión y virulencia. Quizás, en un futuro cercano, el uso de análisis genómicos de cepas aisladas a partir de muestras colombianas, puedan dar solución a algunos de los vacíos del conocimiento que permanecen en torno a estos microorganismos y sus características biológicas.

6. ESTADO DEL CONOCIMIENTO

Los parásitos intestinales suelen presentarse con frecuencia en la población humana, tanto en pacientes sintomáticos como asintomáticos y estar asociados a condiciones socioeconómicas particulares como poblaciones con bajos recursos económicos y hábitos de higiene y saneamiento ambiental inadecuado (1). Aproximadamente mil millones de personas en países en vía de desarrollo en todo el mundo se encuentran infectadas, asociadas, además de lo mencionado anteriormente, a la disminución de la productividad y a efectos negativos sobre el desarrollo cognitivo y la educación (2). En América Latina y el Caribe por lo menos 241 millones de personas viven en áreas donde al menos una de estas infecciones parasitarias es endémica (3).

Dentro de los parásitos intestinales, se encuentran varios grupos de organismos capaces de habitar las vías digestivas del ser humano, como son los protozoarios y los helmintos, siendo los protozoarios los que se encuentran con mayor frecuencia.

Los protozoarios se caracterizan por ser microorganismos eucariotas, unicelulares que se clasifican en diferentes grupos, a saber: 1) Los Apicomplexa quienes carecen de estructuras de locomoción y poseen un complejo apical, como *Cryptosporidium* 2) los Sarcomastigophora que involucran a los Sarcodinos o amebas y a los Mastigophora que poseen flagelos, como es el caso de *Giardia*, 3) otro grupo son los Ciliophora u organismos con cilios, y, 4) los Microspora o microsporidios caracterizados por la producción de esporas (4). Por otro lado, *Blastocystis,* otro microorganismo de importancia en el ambiente intestinal, ha sido clasificado dentro de los protistas, en el grupo Stramenopila*,* junto con las algas y los Oomycetos (5).

Dentro de este grupo de microorganismos existen algunos de gran importancia para el ser humano, por estar asociados a infecciones del tracto gastrointestinal. Uno de ellos es *G. intestinalis*, quien ha sido establecido como el protozoo encontrado con mayor frecuencia en infecciones intestinales sintomáticas en humanos (6) y otro que es *Blastocystis*, quien ha sido catalogado como el protozoo más frecuentemente encontrado en el intestino humano tanto en personas sintomáticas como asintomáticas (7, 8), siendo importante como indicador de desbalance de la microbiota intestinal y relacionado con alteraciones gastrointestinales (9, 10). A continuación, se describen las características de estos dos microorganismos.

6.1 *Giardia*

6.1.1 Historia y clasificación

Giardia fue descrita por primera vez en 1681 por Antonie van Leeuwenhoek cuando la observó en sus propias heces, sin embargo, fue hasta 1902 que Stiles la llamó *Giardia*

intestinalis. A partir de allí se han encontrado una serie de controversias con respecto al nombre debido a que se le ha llamado *G. duodenalis, G. lamblia* y *G. intestinalis*, los cuales actualmente se consideran sinónimos (11). Aquí se hará referencia a ella como *G. intestinalis.* Además de esta, otras especies han sido identificadas por medio de microscopia de luz diferenciándose morfológicamente como *G. agilis* en anfibios y *G. muris* en roedores (ver figura 1). Recientemente otras especies morfológicamente similares a *G. intestinalis* han sido determinadas por microscopía electrónica como *G. ardeae* en garzas, *G. psittaci* en aves, *G. microti* en roedores (11, 12) y *G. varani* encontrada en lagartos (13).

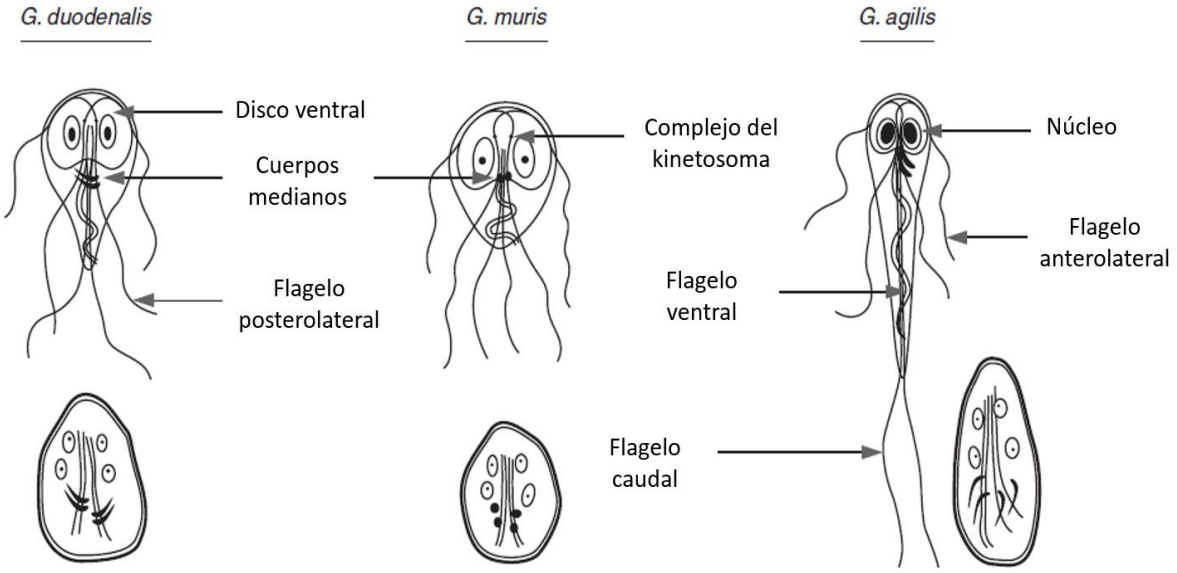


Figura 1. Diferenciación morfológica de trofozoítos y quistes de especies de *Giardia.* Tomado y modificado de (14).

Las primeras descripciones asumieron diferentes especies por hospedero, llevando a una sobreestimación de especies las cuales fueron basadas en características únicamente morfológicas. *Giardia* fue inicialmente clasificada en el grupo de las Diplomonas con organismos binucleados, dentro de los organismos ancestrales, debido a la ausencia de mitocondria, pero con el hallazgo de los mitosomas (un organelo pequeño, carente de genoma y rodeado por una doble membrana) su posición fue cambiando (15). Actualmente este protozoo eucariota unicelular se encuentra clasificado dentro del grupo Excavata, Clase: Fornicata, Orden: Diplomonadida y Familia: Giardiidae (16).

6.1.2 Biología

Giardia es el parásito más común causante de infecciones intestinales en humanos (17), pero a pesar de esto desde el año 2004 se catalogó a la giardiasis como una enfermedad desatendida (18). La giardiasis sintomática y asintomática suele ser frecuente,

principalmente en países en desarrollo (19), ya que *G. intestinalis* no solo afecta a los seres humanos, sino por el contrario también se ha encontrado en una gran variedad de animales que involucran animales domésticos (12) como perros, gatos, ganado y humanos y en animales silvestres (20), que adquieren la infección por el consumo de agua y alimentos contaminados (21). Es por esto, que la transmisión de este protozoario es considerada tanto zoonótica como zooantroponótica y aunque se desconoce la frecuencia de transmisión entre los diferentes hospederos (21) se sabe que existe un factor de riesgo potencial para una diseminación masiva.

Este parásito se caracteriza por tener cuatro pares de flagelos. Su ciclo de vida es directo, con dos estadios característicos: uno móvil denominado trofozoíto causante de la sintomatología y uno inmóvil llamado quiste, el cual es infeccioso una vez es eliminado en las heces (22) (ver figuras 2a y 2b). Normalmente en las heces duras o pastosas son eliminados los quistes, los cuales pueden entrar en contacto con fómites, aguas o alimentos y ser ingeridos nuevamente. Una vez son deglutidos y están cerca al intestino delgado y gracias al pH ácido del estómago se desenquistan y se convierten en el estadio móvil, el trofozoíto, el cual realiza el proceso reproductivo por medio de fisión binaria (23). Los trofozoítos se adhieren a los enterocitos y posteriormente por el tránsito intestinal caen en la luz del intestino, principalmente en el yeyuno, donde gracias a las sales biliares se convierten en quistes nuevamente y transitan hacia el colon para ser eliminados en las heces (24). En el caso de que el hospedero presente heces diarreicas es posible observar la eliminación de trofozoítos, aunque estos mueren inmediatamente cuando son expuestos al ambiente fuera del hospedero. Cuando se presentan manifestaciones clínicas, se pueden observar síntomas iniciales como malestar abdominal, náuseas y anorexia, posteriormente viene una fase aguda de aproximadamente 3 a 4 días caracterizada por heces diarreicas y fétidas, fiebre baja, distensión abdominal, flatulencia y ocasionalmente heces con sangre o moco. En casos crónicos se pueden presentar diarreas intermitentes, pérdida de peso, desnutrición y retardo en el crecimiento en el caso de niños infectados (24).

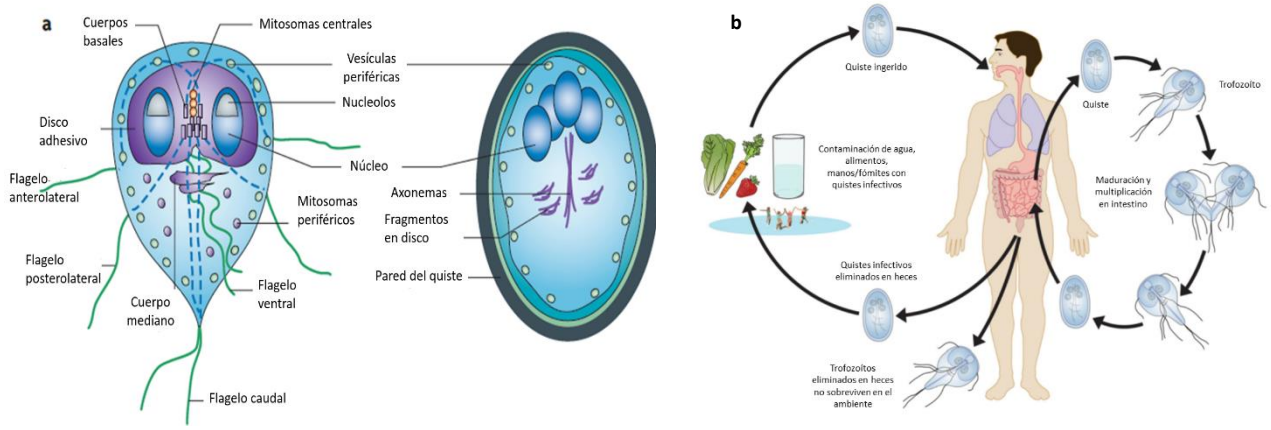


Figura 2. a. Características del trofozoíto (izquierda) y quiste (derecha) de *G. intestinalis* modificado de (18). b. Ciclo de vida de *G. intestinalis.* Modificado de (25).

El parásito sobrevive períodos largos en el ambiente en el estadio de quiste, siendo la dosis infecciosa de aproximadamente 10 quistes. Este parásito no genera una infección invasiva, ni es secretor de toxinas que puedan generar daños en los tejidos (18). Por el contrario se ha determinado que genera un daño mecánico sobre los enterocitos, lo que se ha asociado con algunas de las manifestaciones clínicas principales como síndrome de mala absorción intestinal y síndrome de colon irritable (18).

6.1.3 Epidemiología molecular

Estudios previos detectaron elevada heterogeneidad genética de *G. intestinalis* entre aislados de la misma y de diferentes especies de hospederos. Utilizando análisis de zimodemos se encontraron dos grupos mayores de aislamientos de *G. intestinalis*, el "Belga" y el "Polaco" (26-28), y 3 grupos basados en variaciones en antígenos de superficie (29) y en la *ssuARNr* (30). Luego, estudios con RFLPs, southern blot y la *ssuARNr* encontraron dos genotipos diferentes (31). El uso de RAPDs, mostró agrupaciones de aislamientos con tasas de crecimiento y metabolismo diferente, pero con patrones moleculares similares (11, 32). Posteriormente, el análisis de 50 enzimas diferentes, permitió agrupar aislamientos en 4 grupos genéticamente distintos (I-IV) (28), concordante con análisis por medio de RFLPs (26, 33-35) y PFGE para determinar re arreglos cromosómicos (36, 37) que igualmente permitieron establecer 4 grupos diferentes. Posteriormente fueron categorizados en los ensamblajes A (grupos I y II) y B (grupos III y IV), los cuales se establecieron por diferencias genéticas fijadas en 60% de los loci evaluados (38, 39) (ver figura 3 y tabla 1).

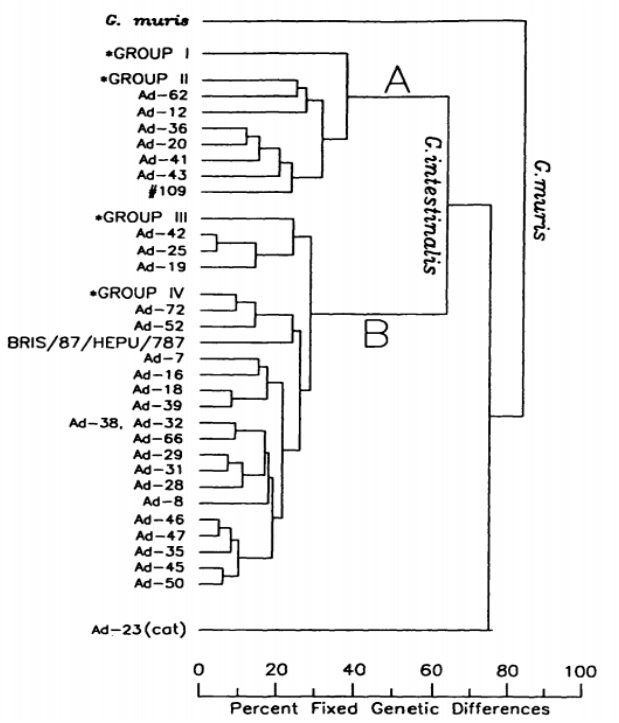


Figura 3. Fenograma construido a partir de electroforesis de alozimas, mostrando distancias genéticas entre dos ensamblajes (A y B) de *G. intestinalis.* Tomado de (38).

Tabla 1. Designación propuesta para genotipos de *G. intestinalis.* Modificado de (11).

Designación propuesta	**Grupo Nash**	**Ensamblaje Mayrhofer**	**Origen**	**Hospederos**
Genotipo A-I	1	A (Grupo 1)	Polaco	Humano, castor, gato, lemur, oveja, perro, becerro, caballo, cerdo y vaca
Genotipo A-2	2	A (Grupo 2)		Humano, castor
Genotipo B	3	B (Grupos 3 y 4)	Belga	Humano, castor, cobayo, perro, mono
		C		Perro
		D		Perro
		E (o A en ganado)		Vaca, oveja, cabra, cerdo
		F		Gato
		G		Roedor

Para diferenciar especies, se hicieron comparaciones cuantitativas utilizando secuencias del *ssuARNr* (*18S*) por ser una región altamente conservada, pero con regiones variables en los

extremos 5´ y 3´ que pueden ser utilizadas para diferenciar incluso ensamblajes dependiendo de los iniciadores que sean utilizados y también se utilizaron genes constitutivos como el de la *triosa fosfato isomerasa (tpi),* la *glutamato deshidrogenasa (gdh),* la *beta giardina (bg)* como genes de referencia para la genotipificación y la subgenotipificación, junto con otros menos frecuentes como el *factor de elongación 1 alfa (ef1alfa)* (11, 40). Estos diferentes loci, son utilizados según el nivel de diferenciación requerido, donde los más conservados como la *ssuARNr*, con una tasa de sustitución de 0,01, es usada para diferenciar especies y ensamblajes y los más variables como *tpi*, con una tasa de sustitución de 0,12, son usados para subtipificación. Genes como *gdh* y *bg* tienen tasas de sustitución intermedia por lo que pueden ser usados en diferentes aplicaciones (41). El diseño de iniciadores debe ser cuidadoso debido a que algunas regiones de estos marcadores son iguales generando tipificaciones erróneas de los ensamblajes y subtipos (41-43).

Estudios con los genes *ssuARNr (18S), tpi, gdh* y *bg* permitieron establecer tres grupos para *G. intestinalis* por Nash, el I, II y III, que, como se mencionó anteriormente, fueron denominados ensamblajes por Mayrhofer (11), donde el I y el II agrupan en el ensamblaje A y el III y posteriormente el IV dentro del ensamblaje B. El ensamblaje A y el B presentan un rango amplio de hospederos, donde el A es frecuente en bovinos, búfalos, ovejas, cabras, alpacas, cerdos, perros, gatos y caballos. El ensamblaje B es reportado en menor frecuencia en bovinos, ovejas y mascotas domésticas como caballos, perros, gatos y conejos (40). Por esta razón, Pijnacker y colaboradores (2016) sugirieron que el ensamblaje B debe ser de transmisión antroponótica y el ensamblaje A de transmisión zoonótica (44). Otros ensamblajes específicos de hospederos, con ocasional excepción (45-52) fueron determinados como el C y D ambos específicos de caninos, el ensamblaje E en ganado, el F en gatos, el G en roedores y el nuevo ensamblaje H en focas y gaviotas (53) (ver tabla 2).

Tabla 2. Ensamblajes actualmente reconocidos de *G. intestinalis*, su distribución de hospederos y taxonomía propuesta previamente. Adaptado de (54).

Ensamblaje	Hospedero	Nombre de especie propuesta
A	Humanos y otros primates, ganado, perros, gatos y algunas especies de mamíferos silvestres	*G. duodenalis*
B	Humanos y otros primates, perros, gatos y algunas especies de mamíferos silvestres	*G. entérica*
C	Perros y otros cánidos	*G. canis*
D	Perros y otros cánidos	
E	Ganado	*G. bovis*

F	Gatos	*G. cati*
G	Ratas	*G. simondi*
H	Mamíferos marinos (pinnípedos)	?

Teniendo en cuenta que los de origen Polaco son los del ensamblaje A y los Belgas del ensamblaje B, ambos relacionados con infecciones en humanos y castores, Adam y colaboradores (2001) propusieron una nueva designación de grupos así: genotipo A1 al grupo 1 de Nash, ensamblaje A de Mayrhofer, genotipo A2 al grupo 2 de Nash, ensamblaje A de Mayrhofer y genotipo B a los grupos 3 y 4 de Nash, ensamblaje B de Mayrhofer (ver tabla 1) (11). Sin embargo, esta propuesta hace evidente las dificultades en cuanto a la nomenclatura utilizada para la asignación de genotipos en *G. intestinalis*, generando confusiones y dificultades en estudios de tipificación.

Entre los genotipos A1 y A2 se encontró una divergencia en el gen *tpi* de tan solo el 1% en regiones codificantes y del 2% en regiones flanqueantes y entre el A1 y el B una divergencia del 19% en regiones codificantes y regiones flanqueantes muy distintas (11). Estos genotipos se han relacionado en algunos casos con diferencias biológicas importantes, por ejemplo, en el caso del grupo 3 (ensamblaje B) que se ha encontrado de mayor virulencia pero con crecimiento más lento en cultivo, lo que junto con sus diferencias genéticas, ha llevado a plantear la posibilidad de que el grupo 3 corresponda a una especie distinta (11). En el caso de otros ensamblajes encontrados, denominados ensamblajes C al G se han encontrado diferencias en secuencias de regiones codificantes y se ha determinado que en el caso de los que se encuentran en perros, tienen poco potencial de transmisión zoonótica. Algunos aislados de ganado han sido clasificados dentro del ensamblaje E, mientras que otros también de ganado, pertenecen al ensamblaje A genotipo 1, lo que representa gran importancia por ser potenciales generadores de infecciones en humanos (11).

Análisis moleculares para investigar la diversidad genética de *G. intestinalis,* han establecido hasta el momento 8 genotipos o ensamblajes (denominados de la A a la H), ampliamente distribuidos a nivel mundial. Los ensamblajes A y B son los más frecuentes, (ver figuras 4 y 5) (40), siendo el ensamblaje B el que presenta mayor frecuencia en humanos (54), lo que genera una gran relevancia, puesto que se ha asociado con casos de diarrea severa y prolongada, siendo catalogado como un ensamblaje de mayor virulencia (14, 40, 55-57). Estos resultados concuerdan con los hallazgos en Sur América, en Argentina donde se ha reportado el ensamblaje B en mayor frecuencia en comparación con el ensamblaje AII y algunas infecciones mixtas con B y AII (58). En Colombia, en muestras de heces de humanos, los subensamblajes AI, AII, BIII y BIV fueron encontrados (59-61), los subensamblajes AII, BIII y BIV se encontraron en heces de niños de las guarderías del Instituto Colombiano del

Bienestar Familiar (ICBF), junto con los ensamblajes C y D en muestras de perros (62). En Brasil, se encontró una frecuencia similar entre el ensamblaje A y B, con los subensamblajes AII, AIII, BIII y algunos genotipos nuevos de B que no han sido descritos previamente (63), mientras que en Perú, se encontró AI, AII y BIV, con predominancia de AI (64). Por otro lado, un estudio realizado en bovinos y ovinos en México, encontró el genotipo AI e infecciones mixtas de AI + BIII y de E + BIII (65), y en humanos se encontró mayor frecuencia del ensamblaje A y menor del B, con los subensamblajes AI y BIII (66). En Cuba se reportó una frecuencia similar entre el ensamblaje A y B (67).

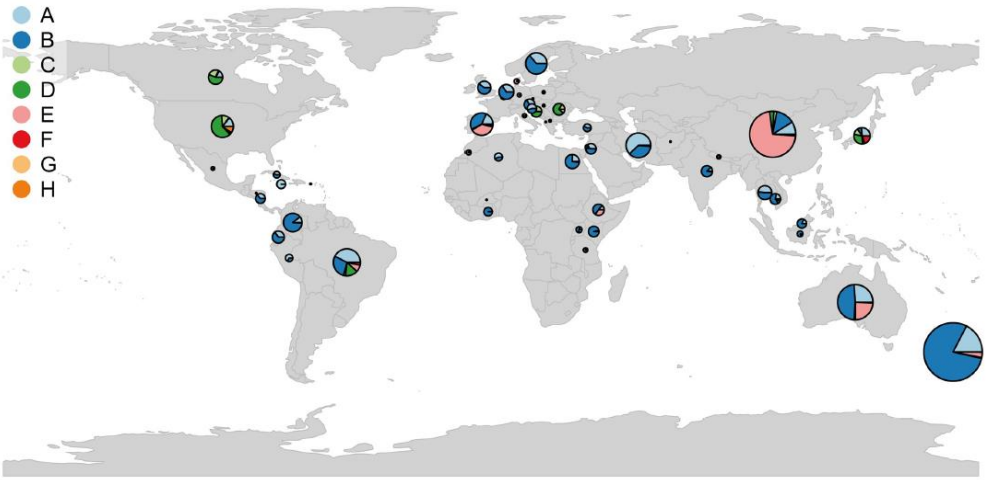


Figura 4. Distribución de ensamblajes de *G. intestinalis* detectados con el gen *gdh*. Tomado de (68).

Dentro de dichos ensamblajes, junto con la variación encontrada dentro de algunos de estos, se lograron establecer subensamblajes, como son los subensamblajes AI-AIII y BIII-BIV (ver figura 6) (14, 51, 53, 54, 69-71). Con el locus *gdh* se encontró una clara agrupación que define subensamblajes (I, II y III) dentro del ensamblaje A pero en el ensamblaje B estos no son claros aún (ver figura 6). En los ensamblajes C a G se encontró una alta especificidad de hospedero y con genes como el *tpi* no se observó el ensamblaje H (Figura 5), generando incongruencias entre las topologías obtenidas con distintos marcadores (40).

En 2011, Feng y Xiao, proponen nombrar a cada ensamblaje con una especie diferente asociándola principalmente al hospedero, llamando así al ensamblaje A: *G. duodenalis*, ensamblaje B: *G. enterica*, el C y el D: *G. canis,* el E: *G. bovis*, el F: *G. cati,* y el G: *G. simondi* (12, 40, 70) (ver tabla 2). Sin embargo, el reconocimiento de estas nuevas especies en la comunidad científica requiere de más estudios sobre la biología de cada una, morfología y genética. Además, en el caso de las especies que infectan caninos, sería importante tener en cuenta que involucra dos ensamblajes (el C y el D) cuya divergencia es similar a la de otros ensamblajes, y que con genes como *tpi* no forman un grupo monofilético (6, 40, 51).

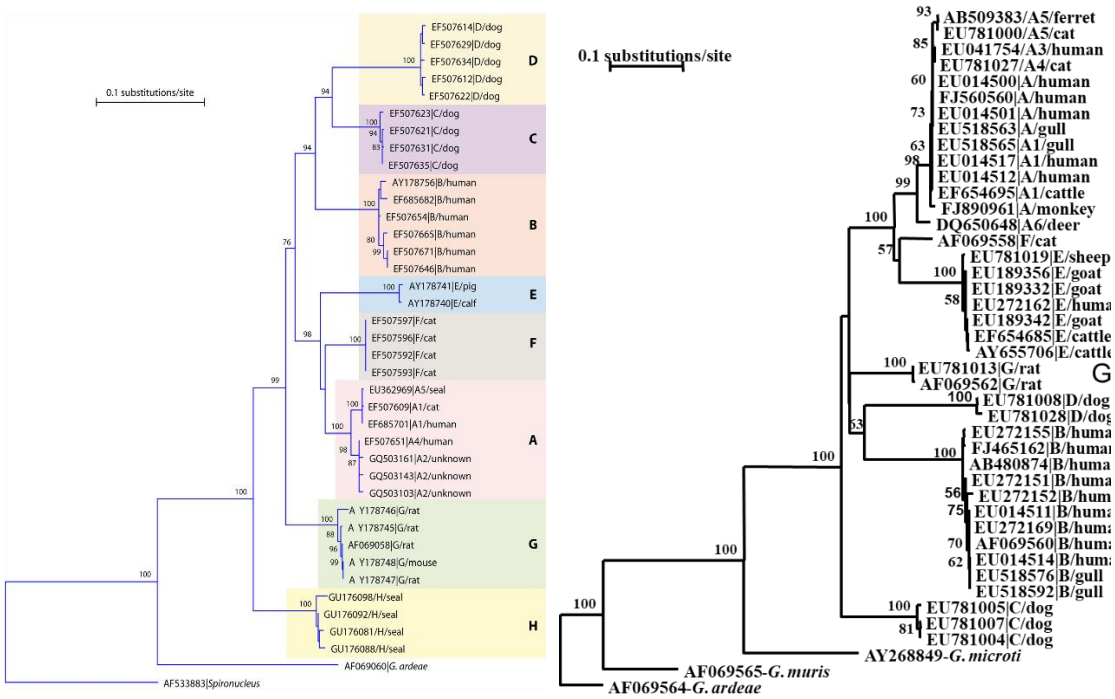


Figura 5. Relaciones filogenéticas entre ensamblajes de *G. intestinalis,* izquierda con el locus *gdh* y derecha con el locus *tpi*, evaluados con neighbor-joining. Tomado de (40).

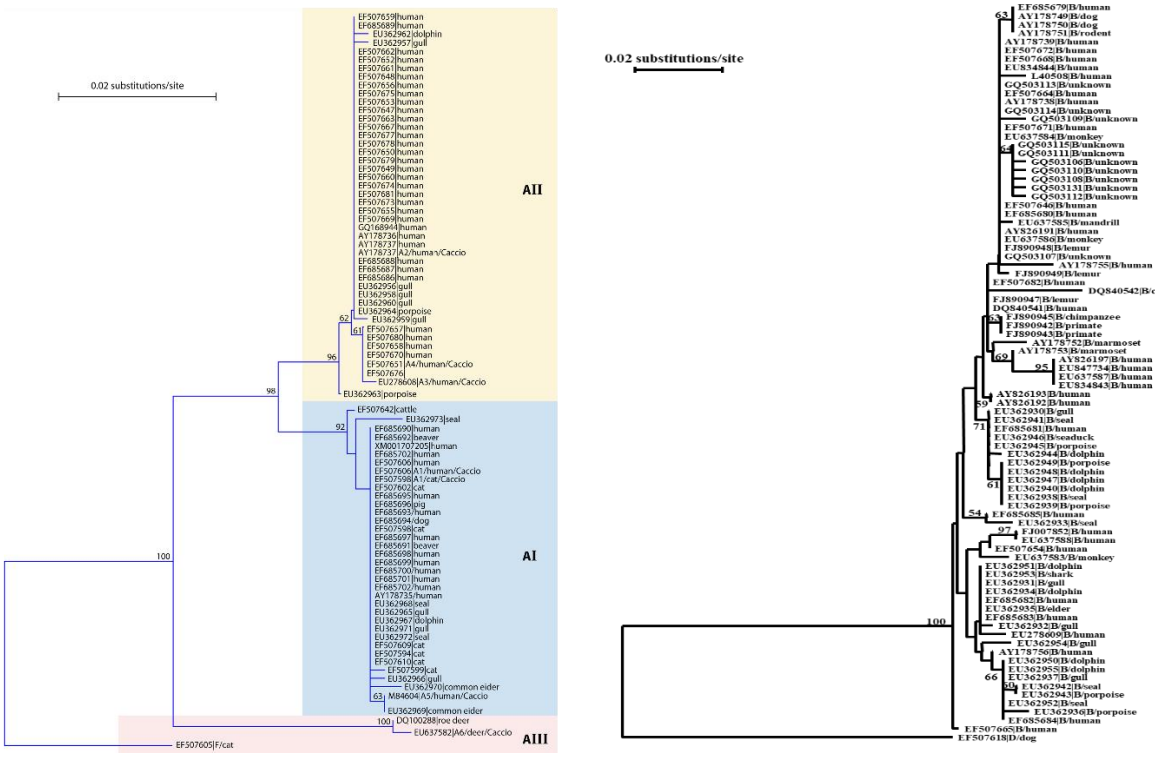


Figura 6. Relaciones filogenéticas entre subensamblajes A (izquierda) y entre subensamblajes B (derecha) de *G. intestinalis*, con el locus *gdh*, evaluados con neighbor-joining. Tomado de (40).

De igual manera cabe resaltar que los sub-ensamblajes BIII y BIV descritos por alozimas no están soportados por análisis de secuencias de ADN y que las secuencias conocidas no son suficientes para determinar polimorfismos específicos de secuencia de subgrupos (40). Hay varios subtipos en el ensamblaje E que no han sido claramente establecidos y existe insuficiencia de datos para evaluar la subestructura de ensamblajes C, D, F y G, por lo que es crucial realizar un análisis de secuencias multilocus que permita realizar una caracterización más robusta a nivel de ensamblaje y subensamblaje utilizando varios loci (55).

6.1.4 Diversidad y estructura genética

Análisis de genotipificación multilocus (MLG) se han realizado utilizando las secuencias de los 4 genes (*ssuARNr, bg, gdh* y *tpi*) de aislados de humanos y animales, mostrando congruencia únicamente para aislados del ensamblaje A, donde propusieron subtipos con 10 MLGs diferentes (AI-1, AI-2, AII-1, AII-2, AII-3, AII-4, AII-5, AII-6, AII-7, AIII-1) (6), junto con otros dos nuevos (AII-8 y AII-9) (71), pero en el caso del ensamblaje B, los resultados fueron inconsistentes (72-74). Para los ensamblajes B, C, D y E los electroferogramas suelen presentar dobles picos mientras que los ensamblajes A, F y G no, por lo que análisis con MLG parecen ser útiles solo para la tipificación del ensamblaje A (48).

Estudios de secuencias multilocus han sido reportados en aislamientos del subensamblaje A2, utilizando los genes de referencia *bg, tpi* y *gdh*, encontrando diferente número de SNPs asociados a mutaciones sinónimas, pero con análisis filogenéticos que muestran diferente historia evolutiva (6), sugiriendo posible recombinación, la cual cambia el patrón de herencia evidenciando el intercambio meiótico entre aislados A2 (75). De igual forma se evidenció en el estudio realizado por Ankarklev y colaboradores (2018), donde se evaluaron 6 genes nuevos junto con los tres de referencia, encontrando 7 subtipos y alelos idénticos en diferentes aislados y en combinaciones únicas, mostrando recombinación entre aislados del ensamblaje A y entre el ensamblaje A y E. Posiblemente coinfecciones entre AI y E y entre B y AII generaron eventos de recombinación por fusión de núcleos durante el enquistamiento (76).

Algunos análisis de poblaciones han utilizado redes de haplotipos y análisis moleculares de varianza con secuencias de los genes *tpi, gdh* y *bg* de diferentes fuentes de aislamiento, encontrando mayor variación en el gen *tpi* (77, 78) y al comparar dos poblaciones con el gen *gdh* y el *bg* hubo resultados diferentes sobre la diferenciación genética de los grupos (77). Dentro de los ensamblajes A, B y E, redes haplotípicas y análisis de diversidad genética del gen *tpi* mostraron mayor diversidad haplotípica y nucleotídica en el ensamblaje B,

seguida del E y en menor proporción en el A. Análisis de neutralidad con el indicador D de Tajima y el D de Fu y Li, mostraron expansión del tamaño poblacional, además de moderada diferenciación genética y elevado flujo de genes entre el ensamblaje B y E, es decir, dispersión de la misma población de *G. intestinalis* (78), posiblemente por cruzamiento meiótico reciente dentro del núcleo o entre núcleos generando homogenización de estos (79) y diferenciación genética entre poblaciones de Asia y Australia y América del ensamblaje A (78). Evidencia de baja heterocigosidad y recombinación tanto a nivel individual como intra ensamblaje entre ensamblajes A-G, fue observada con el uso de 7 loci, sugiriendo eventos meióticos entre ensamblajes con distancias genéticas grandes, lo que pone en duda que sean especies diferentes y por el contrario sugiere posibles mecanismos de adaptación del parásito que le permiten aumentar su dispersión y tal vez su virulencia (53). Otros estudios, encontraron 3 genotipos homocigotos para el ensamblaje A, recombinación individual y entre individuos del mismo ensamblaje, que junto con la baja heterocigosidad sugieren la posibilidad de procesos sexuales o meióticos (80). Lo anterior, concuerda con un estudio con muestras clínicas (humanos y animales – centros veterinarios) y muestras ambientales, donde construyeron redes haplotípicas con los genes *gdh, tpi* y *bg* y encontraron que aunque hay una alta diversidad de haplotipos hay una importante similitud genética, mostrando el potencial zoonótico de *G. intestinalis* (77) y que aunque existe esa alta variabilidad haplotípica, la mayoría de muestras provienen de pocos haplotipos, sugiriendo relación entre los grupos de los diferentes sitios de recolección (81). En Brasil, evidencia de baja divergencia genética fue detectada entre haplotipos del ensamblaje A idénticos a cepas europeas, mientras que elevada diversidad de haplotipos del ensamblaje B fue detectada (63).

6.1.5 Genómica y reproducción

El uso de técnicas que permiten secuenciar genomas completos ha permitido dilucidar características importantes con respecto al genoma de *G. intestinalis*. Hasta la fecha existen varios genomas publicados de *G. intestinalis,* de los cuales, el del ensamblage AI de la cepa WB, fue el primero en ser determinado, el cual está ensamblado por cromosomas, y junto con este, otros 6 genomas son considerados de referencia (82) (disponibles en GiardiaDB.org) y corresponden a los ensamblajes AII, BIV y E (41, 83-88), algunos con calidad mejorada gracias a la combinación de técnicas de segunda y tercera generación, que incluyen Illumina HiSeq y Oxford Nanopore (ONT) (89). Se ha logrado establecer que el genoma haploide de *Giardia* contiene un tamaño que oscila entre 10.7 a 12 Mpb, normalmente posee 5 cromosomas y se caracteriza por tener un genoma compacto (90). Se han encontrado diferentes grados de heterocigosidad entre los ensamblajes y subensamblajes, siendo el ensamblaje B (cepa GS) el que presenta una elevada heterocigosidad en secuencias alélicas, seguido del subensamblaje AII (cepa DH) con una menor heterocigosidad y aún menor en el subensamblaje AI y el ensamblaje E (76). Otros reportes con el subensamblaje AI (cepa WB-C6) mostraron tetraploidías con 2 núcleos

diploides, pero clones de WB y de otros ensamblajes mostraron aneuploidías entre los dos núcleos. El trofozoíto normalmente presenta dos núcleos generalmente diploides e idénticos genéticamente, caracterizados por presentar una baja heterocigosidad, lo que indica que alelos dentro de los núcleos, incluso dentro de núcleos opuestos son idénticos. Durante la división, los dos núcleos permanecen diferenciados, no se ha reportado apareamiento ni meiosis. Sin embargo, si *Giardia* es asexual debe acumular heterocigosidad alélica dentro y entre los núcleos generando divergencia entre estos, por lo tanto *Giardia* debe utilizar algún mecanismo para homogenizar los dos núcleos (75). Durante el enquistamiento se puede encontrar fusión de núcleos facilitando transferencia de material genético entre estos (91). Por eso la recombinación meiótica podría explicar esta homogeneidad por rearreglos y recombinación dentro y entre los núcleos (75). Cada núcleo es de linaje clonal y acumula mutaciones independientemente generando heterocigosidad alélica (71). Aunque algunos estudios han mostrado poca recombinación entre aislados de diferentes ensamblajes (92), contradictoriamente otros han inferido posible recombinación entre diferentes aislamientos de *Giardia* que junto con análisis de genómica comparativa han permitido demostrar recombinación entre ensamblajes, principalmente dentro del subensamblaje AII y transferencia de genes entre A y E (92). Una explicación a esto estuvo relacionada con posible reproducción sexual, que aunque no ha sido reportada como tal, si se han encontrado genes relacionados con procesos meióticos (93), por lo cual se ha sugerido recombinación somática entre los núcleos de un parásito (diplomixis) (ver figura 7) durante el enquistamiento donde podrían activarse algunos genes meióticos, sin recombinar entre parásitos (76, 91). También, es posible que la recombinación entre ensamblajes diferentes se deba a su fusión durante el enquistamiento, ya que en ratones infectados se ha visto la formación de microcolonias en el intestino, permitiendo el contacto de diferentes estadios móviles, posiblemente de diferente ensamblaje (94) y teniendo en cuenta que es posible la recombinación, en zonas de alta densidad parasitaria, en el hospedero, la diversidad genética puede ser mayor, ya que en regiones de alta endemicidad, existe una mayor probabilidad de tener infecciones mixtas (76).

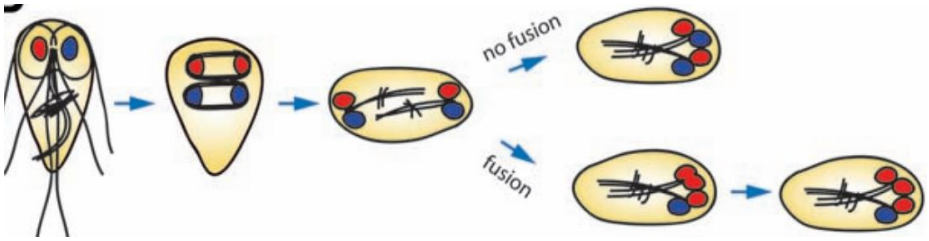


Figura 7. Modelo de intercambio de un plásmido durante el enquistamiento. El núcleo rojo del trofozoíto contiene el plásmido. Tomado de (91).

6.2 *Blastocystis*

6.2.1 Historia y clasificación

Inicialmente, en 1900 Alexeieff y Emile Brumpt clasificaron a *Blastocystis* como una levadura saprófita intestinal (95). En 1967, Zierdt y colaboradores, aunque mantuvieron su nombre, cambiaron su clasificación, ya que asociaron su morfología con la de un protista, puesto que presenta 1 o más núcleos, organelos como mitocondrias, aparato de Golgi, retículo endoplasmático, no crece en medios de hongos, es resistente a antifúngicos y susceptible a medicamentos antiprotozoarios (96). En 1996 Silberman y colaboradores por medio de análisis de la *ssuARNr* y el *factor de elongación 1 alfa*, lo clasificaron dentro del phylum Heterokontophyta (97), dentro del grupo de los Stramenopila, que hasta ese momento incluía protozoos heterótrofos y fotosintéticos (98), como algas, diatomeas, mohos mucilaginosos y Oomycetes. *Blastocystis,* se denominó como un protista botánico, por ser el único de este grupo capaz de infectar humanos (99). También lo han ubicado en un sexto reino, el Chromista (5, 100, 101).

6.2.2 Biología

Se caracteriza por ser un protista entérico, unicelular y anaerobio. Posee estructuras similares a mitocondrias y aunque carece de citocromo (102) posee hidrogenosomas involucrados en la biosíntesis de fosfolípidos que se acumulan en la vacuola, y le permiten hacer metabolismo de aminoácidos y el ciclo del ácido tricarboxílico (103). Esta estructura se ha relacionado con procesos apoptóticos, que sirven como un mecanismo para incrementar el número de células viables durante condiciones de estrés, cuando se reproduce *in vitro* por 7-12 horas en aerobiosis y en presencia de antiparasitarios (5, 104).

Con respecto a su morfología, se han encontrado 4 estadios presentes dentro de su ciclo de vida, la forma vacuolar, la granular, la ameboide y la de quiste (ver figura 8). La forma vacuolar es la típica en intestino y frecuente en heces. Esta forma, tiene una gran vacuola central que varía en tamaño y posee material compuesto de carbohidratos y lípidos (105), también es llamada cuerpo central, que es importante, no solo como organelo de almacenamiento, sino en la muerte celular programada. La forma granular, muestra diferentes gránulos ubicados en el cuerpo central, los cuales pueden ser metabólicos, reproductores y lipídicos. La forma ameboide, menos frecuente, presenta pseudópodos, aunque es inmóvil (99) y la forma de quiste, con un tamaño muy pequeño y una pared retráctil con una cobertura externa irregular. Los núcleos del quiste pueden variar de 1-4 (5, 7) y las formas pequeñas pueden tener entre uno y dos núcleos y las más grandes hasta cuatro (7).

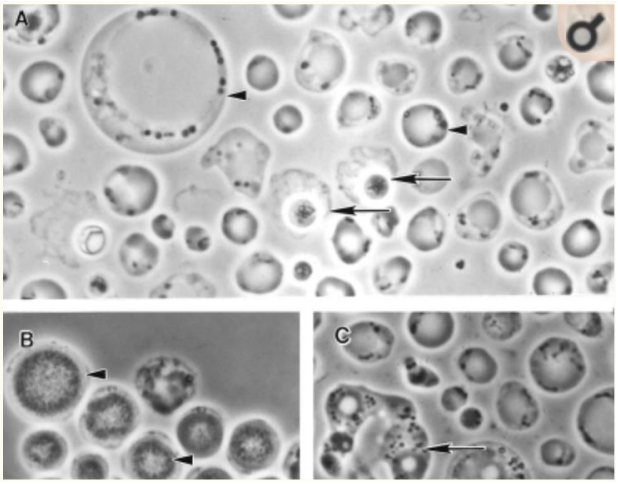


Figura 8. Morfología de *Blastocystis* ST4 con microscopio de contraste de fases. A) Forma vacuolar y de quiste, B) Forma granular y C) Forma ameboide. Tomado de (99).

La forma ameboide se ha visto aumentada en pacientes sintomáticos diarreicos, por lo que se sugiere que esta forma es la patogénica (106-108).

Con exactitud, el mecanismo de transmisión aún se desconoce, aunque al parecer ocurre principalmente por la forma de quiste que se encuentra en alimentos o aguas contaminadas con heces de portadores (109, 110). El ciclo de vida inicia cuando estos quistes son excretados en las heces de hospederos infectados y entran en contacto con agua o alimentos que son ingeridos por el hospedero (111, 112). Una vez llegan al intestino grueso, se desenquistan y liberan la forma vacuolar, esta se trasforma en la forma granular o en la ameboide y viceversa. La forma vacuolar se divide por fisión binaria, gemación o plasmotomía. La forma vacuolar se enquista en el lumen del intestino grueso para producir la forma de quiste que nuevamente es eliminada en las heces (5) (ver figura 9) . La forma vacuolar también se ha reportado como un posible estadio transmisor (100).

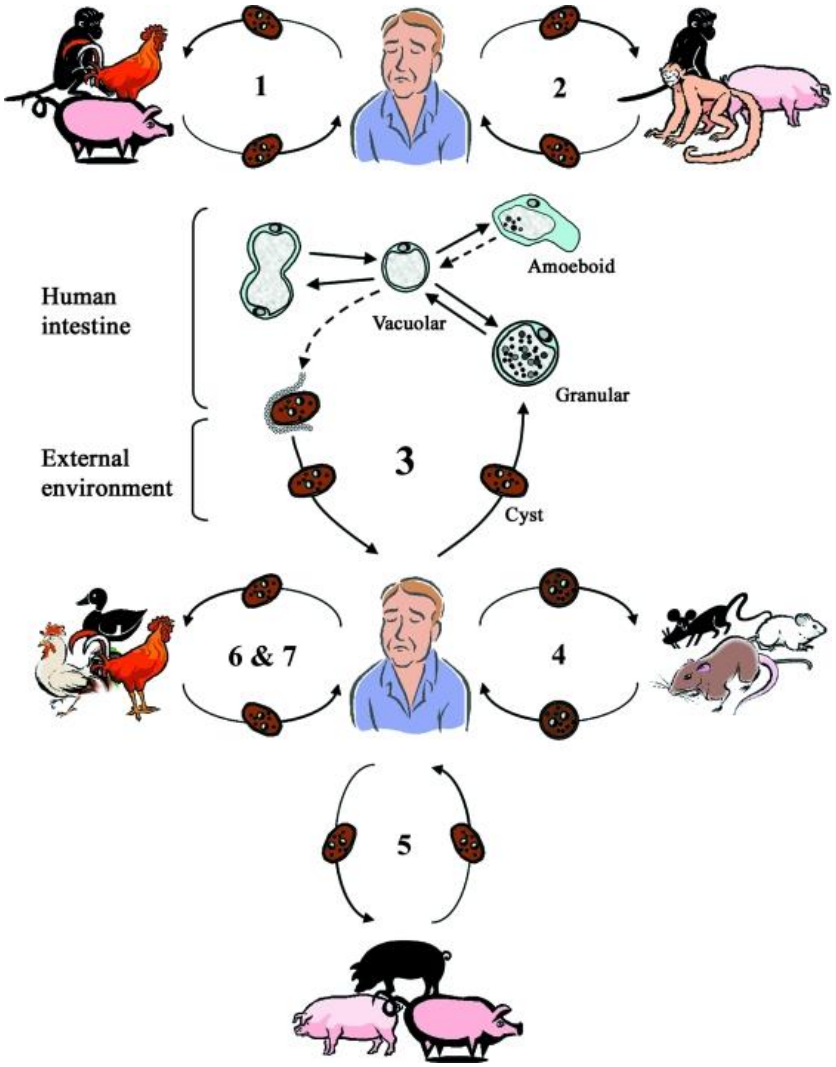


Figura 9. Ciclo de vida de *Blastocystis*, mostrando los subtipos 1 al 7. Tomado de (2008), (99).

La transmisión ocurre humano animal o animal humano, con una especificidad importante en el hospedero, además de un potencial patogénico en aislamientos relacionados con variaciones de secuencias de la *ssuARNr* (113).

Blastocystis presenta una distribución cosmopolita (100, 114) y coloniza alrededor de 1 millón de millones de personas, siendo considerado el protozoo eucariota más común en el intestino de humanos, (8), tanto así, que ha sido reportada una prevalencia del 100% en una población de niños en Senegal (115). Además, también ha sido ampliamente reportado en animales de granja, aves, roedores y reptiles (5). Entre los años 2003 y 2007 se determinó una prevalencia entre el 10 y 15% en individuos asintomáticos en países desarrollados y de 30 a 50% en los países en desarrollo (99, 116), siendo al parecer más común y severo en

pacientes inmunocomprometidos (117). Estudios realizados en Colombia han mostrado prevalencias importantes en regiones como Santander (25%), Santa Marta (62,6%), Bogotá (22,4%), Quindío (36,4%) y Cundinamarca (34,8%) (100) asociado a hospederos humanos, primates no humanos, cerdos, bovinos, aves, anfibios y menos frecuente en roedores, reptiles e insectos (118, 119).

En la actualidad se han reportado portadores asintomáticos y sintomáticos que no muestran diferencias en cuanto a los estadios del parásito y cargas parasitarias. Igualmente, no se conocen hasta el momento factores de virulencia, mecanismos de virulencia específicos o factores de riesgo que se relacionen específicamente con la sintomatología que pueda presentarse. Cuando se presentan síntomas, estos son principalmente intestinales, como dolor abdominal, constipación, diarrea, flatulencia, síndrome de colon irritable, disminución de peso y pérdida de sangre en heces (5, 100, 120). Datos recientes sugieren que *Blastocystis* es más común en individuos saludables, y podría ser indicador de contaminación fecal. Este organismo puede colonizar el intestino por diez años o más, por lo que puede ser encontrado como único agente parasitario en pacientes sintomáticos sin que necesariamente sea el agente causal inicial y sí se encuentre asociado a la exacerbación de los síntomas, siendo un patógeno emergente (121). No hay evidencia de proteínas que permitan la unión al epitelio, solo de hidrolasas consideradas como posibles factores de virulencia (122, 123). Por lo anterior, continúa siendo debatido si *Blastocystis* realmente actúa como un organismo patógeno, comensal u oportunista.

6.2.3 Epidemiología molecular

Aunque inicialmente los nombres de especies se relacionaron con su hospedero como *B. hominis* y *B. ratti*, en la actualidad y por consenso la nomenclatura se limita a especies de *Blastocystis*. Es por esto, que el género ha sido clasificado en varios subtipos o STs (113). Los STs son unidades discretas, sin variantes intermedias. Un nuevo ST debe divergir al menos en el 5% con respecto a los demás (121). No hay diferencia morfológica que pueda observarse o asociarse a un ST. Los STs han mostrado ser muy diversos, por lo que la subtipificación debe hacerse únicamente por secuenciación de una región de la *ssuARNr*, pese a que algunos autores han mostrado que el uso de un solo marcador podría acarrear conclusiones incorrectas para estudios evolutivos, puesto que el tiempo de divergencia puede no haber sido suficiente para fijar alelos con mutaciones (124) y estar mostrando una diversidad errónea.

Inicialmente, la diversidad genética de este microorganismo había sido demostrada por medio de amplificación azarosa de polimorfismos de ADN (RAPDs) y por medio de RFLPs de la *ssuARNr* (125). Algunos reportes mostraron que aislados de *Blastocystis* obtenidos a partir de muestras de cerdos de guinea y de pollos tenían patrones de ribodemos similares a los de aislados de humanos, sugiriendo transmisión entre animales y humanos en ambas direcciones (126, 127). Posteriormente se establecieron filogenias con el uso de la

secuencia completa de la *ssuARNr* (128) y el factor de elongación 1 alfa (129), donde se observaron los clados correspondientes a los STs.

A pesar de la existencia de genomas secuenciados, la mayoría de estudios se han enfocado en la tipificación de *Blastocystis,* basándose en el uso de un barcoding de ADN desarrollado para mostrar la relación entre un ST y una infección por *Blastocystis*, el cual solo requiere la secuenciación de un fragmento de 600 pb del gen de la *ssuARNr* en el extremo 5´, para identificar los subtipos y establecer la relación entre subtipos por análisis filogenéticos (130). En 2007, se propusieron 13 subtipos en humanos, animales y aves, donde los primeros 9 STs se encontraron en humanos (113, 131, 132). Existen 135 diferentes alelos de *18s* y se ha encontrado que cada ST tiene una asociación moderada a un hospedero (133). El ST3 es el más común y el más diverso (121) seguido del ST1, ST2 y ST4. Los ST5-ST9 son de origen zoonótico (59, 113, 131, 132). Infecciones mixtas han sido reportadas en China, Turquía, Francia, Dinamarca, Egipto, Brasil y Colombia (7).

Actualmente, se han encontrado 17 STs (133), de los cuales, en el continente americano, se han reportado a los ST1-ST9 y el ST12 en humanos y en animales a los ST1-ST8, ST10, ST14, y ST17 (134-136) (ver figura 10).

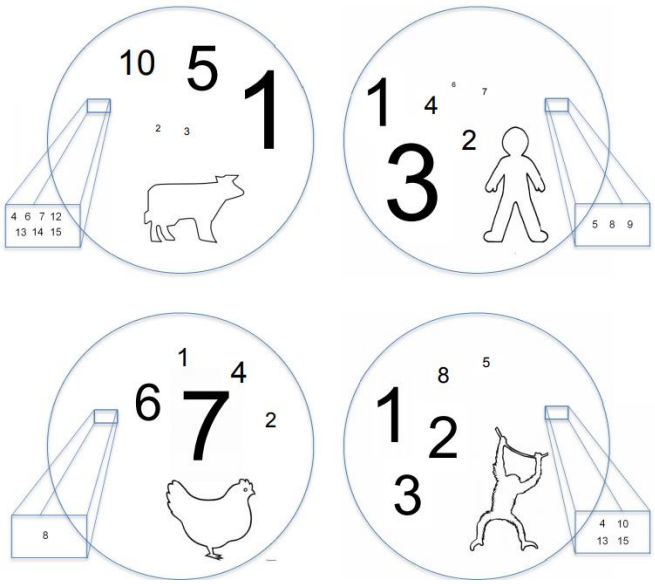


Figura 10. Rango de hospederos y prevalencia relativa de subtipos de *Blastocystis*. En los círculos, los números son los subtipos más comunes en el respectivo hospedero, con el tamaño del número proporcional a su prevalencia. Los números en los cuadros corresponden a los subtipos que constituyen menos del 5% del total de las muestras subtipificadas a la fecha. Como un indicador, las prevalencias de las figuras para los ST1-ST4 en humanos son 20.8%, 10.9%, 44.4% y 10.0% respectivamente. Tomado de (121).

Diferentes estudios realizados reportaron a los ST1, ST2 y ST3 como los más comunes en el continente americano (136). En humanos los ST1-ST9 varían en frecuencia, los ST1-ST4 son comunes en Europa, los ST1-ST3 son igualmente prevalentes en pacientes diarreicos y asintomáticos y el ST4 se ha encontrado en pacientes diarreicos y con síndrome de colon irritable en Europa (137). Los STs de *Blastocystis* en Suramérica eran poco conocidos, siendo los ST1-ST3 más comunes (7, 132). El ST2 y el ST3 se presentan en pacientes con urticaria, asociada principalmente a la forma ameboide del ST3 por adherencia al epitelio intestinal que cambia la homeóstasis y altera la respuesta inmune generando una respuesta inflamatoria (138). El ST4 es común en países como Suecia, Dinamarca y España (138) y aunque no había sido detectado en humanos en América, se había encontrado en un ciclo selvático en primates no humanos en Colombia. Posteriormente el ST4 se encontró en un grupo indígena en Colombia sugiriendo un origen étnico en Latinoamérica (131, 139).

Ramírez y colaboradores en el 2014, mostraron la relación entre los STs, los hospederos y la presencia de síntomas, encontrando que el ST1 se identificó en muestras de humanos asintomáticos y en ganado, el ST2 en humanos diarreicos y perros, el ST3 en ganado y humanos con síndrome de colon irritable, siendo el más frecuente tanto en pacientes asintomáticos como con diferentes sintomatologías, el ST4 en primates no humanos, el ST6 en aves y el ST8 en marsupiales (131). Existe una relación entre el ST y la presentación de síntomas (140). Sin embargo, en otro estudio no se encontró asociación estadísticamente significativa entre la infección por *Blastocystis* y los STs, ni con el género, la edad, los síntomas o la ubicación geográfica (134). Los ST5-ST8 son escasos en humanos (121) pero el ST7 se ha visto asociado a dolor abdominal y el ST6 se ha encontrado en cierta frecuencia en pacientes asintomáticos y en aves de Colombia (134). Dentro de los STs que se encuentran circulantes en Ecuador, Perú, Bolivia, Colombia, Brasil y Argentina (ver figura 11) se detectó mayor cantidad de alelos diferentes para los ST1-ST3 en muestras tanto de humanos, animales domésticos y del ciclo silvestre. Para el ST1 se encontraron los alelos 1, 2, 4, 81, 82 y 83, en el ST2 los alelos 9, 11, 12, 15, 67, 71 y 73, el ST3 alelos 34, 36, 38, 45, 49, 55, 134 y 128. Si se tiene presente el origen de las muestras, en algunos casos es posible plantear una probable transmisión por contaminación de aguas con heces de animales o posiblemente de humanos a animales. En el ST4 se identificaron los alelos 42 y 133, ST6 el alelo 122, ST7 el alelo 142 y en el ST8 el alelo 21 (114, 131, 134). Llama la atención el hallazgo de los alelos 42 y 133 en monos aulladores (131), puesto que el ST4 había sido reportado ser frecuente en Europa en pacientes con síndrome de colon irritable (141), además, el alelo 122 del ST6 fue detectado tanto en aves de Colombia como en humanos y el alelo 142 del ST7 (131, 134) que fue reportado previamente en algunos países de Europa, Asia Central y África en pacientes con síndrome de colon irritable (133). Los ST4, ST6 y ST7 posiblemente son nuevos en Colombia. Nuevos alelos fueron encontrados dentro de los ST5 -ST12 y un ST nuevo fue reportado, al parecer muy cercano al alelo 88 del ST1, pero un esquema MLST debe desarrollarse para lograr definir nuevos STs (114).

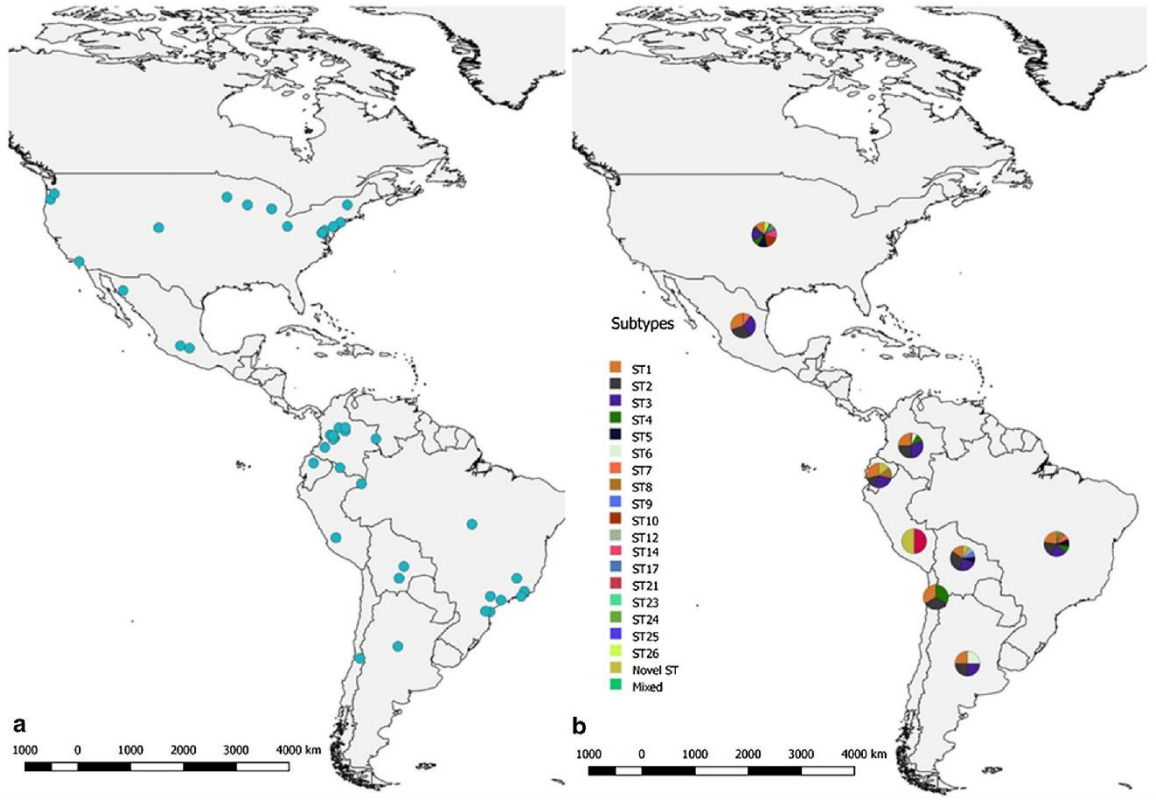


Figura 11. Distribución de *Blastocystis* en América. A) Distribución de *Blastocystis* en Norteamérica y Suramérica según los informes de muestras positivas por país. B) Distribución de subtipos de *Blastocystis* en los diferentes países donde se han tipificado las muestras. Tomado de (136).

6.2.4 Diversidad y estructura genética

Existen otros estudios, orientados al análisis de poblaciones de *Blastocystis* donde además del *ssuARNr*, incluyeron otros marcadores como el *espaciador interno transcrito* (*ITS)*. Uno de ellos, mostró que existe flujo genético entre poblaciones de América y Europa en el caso del ST1 y el ST2 (ver figura 12) (142) al igual que para el ST1 entre primates no humanos y humanos, con pocas mutaciones, pero por el contrario, en el ST2 las poblaciones en humanos fueron altamente diferenciadas de las encontradas en primates no humanos (143). En el caso del ST3, con secuencias del *ssuARNr*, se encontró gran diversidad haplotípica y muy baja diferenciación genética entre poblaciones de diferentes regiones geográficas en el mundo. Por otro lado, la diversidad intra subtipo para el ST3 usando marcadores genéticos del genoma del organelo similar a una mitocondria, junto con la del *ssuARNr* presentó relativa especificidad a hospederos humanos, posiblemente por transmisión entre humanos, mientras que con el MLST desarrollado para el ST4 se encontró que la variación intra subtipo fue diferente, ya que un clado fue casi clonal con diferentes hospederos, indicando posible transmisión zoonótica, llamando la atención que sea mucho más común en Europa y no en países donde la transmisión fecal oral es más probable (8).

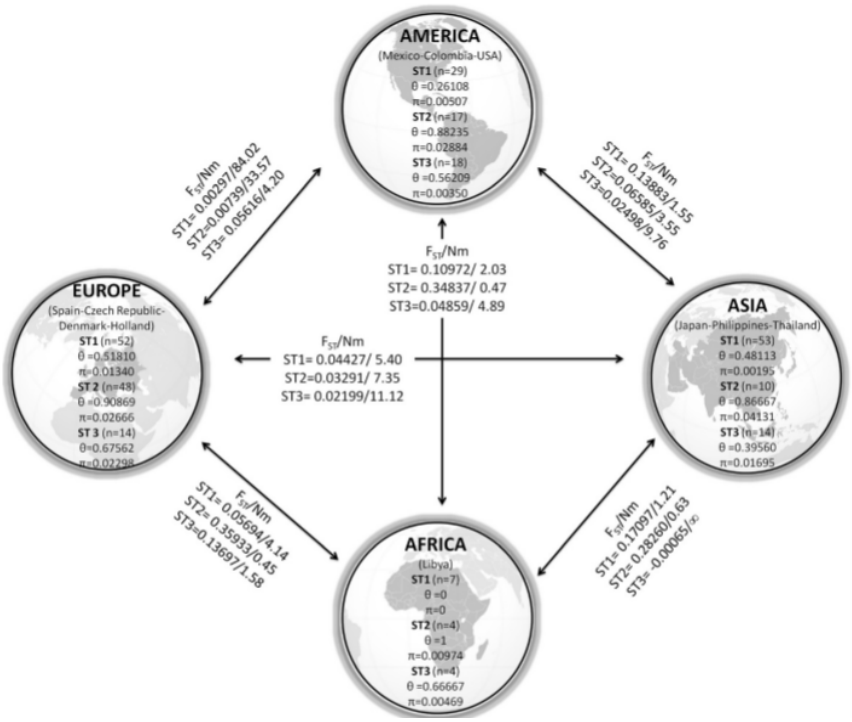


Figura 12. Valores de diversidad nucleotídica (π), índice de divergencia (θ), flujo de genes (Nm) e índice de diferenciación genética (FST) de STs de *Blastocystis* con el gen de la *ssuARNr* de acuerdo al área geográfica. Tomado de (35).

Otro marcador propuesto para discriminar subtipos o cepas de *Blastocystis*, fue basado en un gen metabólico, el gen de la *piruvato ferredoxin oxidoreductasa (PFOR)*, que ha mostrado mayor número de polimorfismos en comparación con el gen de la *ssuARNr*, generando agrupaciones en clados diferentes a STs, posiblemente por presiones selectivas diferentes (144).

6.2.5 Genómica en *Blastocystis*

Con el uso de la secuenciación de genomas, se confirmó la gran diversidad existente entre STs con respecto a su tamaño, contenido de guanina-citosina, número de intrones y contenido de genes (145). A la fecha, existen tres genomas secuenciados y anotados de buena calidad: el ST1 (145), el ST4 (146) y el ST7 (122). El ST1 (cepa ATCC 50177) secuenciado con la tecnologías de nueva generación (454 e Illumina) y el ST4 proveniente de un roedor de laboratorio de Singapur secuenciado con Illumina HiSeq (146), con tamaños aproximados de 16,5 Mb y 12,9 Mb, respectivamente y el ST7 aislado a partir de un paciente sintomático, secuenciado por Sanger, que tiene un tamaño de 18.8Mb, con aproximadamente 6000 genes, numerosos y pequeños intrones, con alrededor de 15 cromosomas (122). La divergencia entre proteínas homólogas es de una magnitud mayor,

en comparación a la observada entre especies dentro de otros géneros de protistas parásitos (ver figura 13). Esto abre la necesidad de estudiar la variación en el número y tipo de proteasas entre STs, para establecer la relación de estos con posible patogenicidad (161). Otros genomas nucleares ensamblados de menor calidad, los STs: ST2, ST3, ST4, ST6, ST8 y ST9, fueron liberados en bases de datos públicas en el 2015, pero no están anotados ni tienen predicción de genes (147). También, el genoma mitocondrial ha sido secuenciado en los ST1, ST4 y ST7 (148, 149).

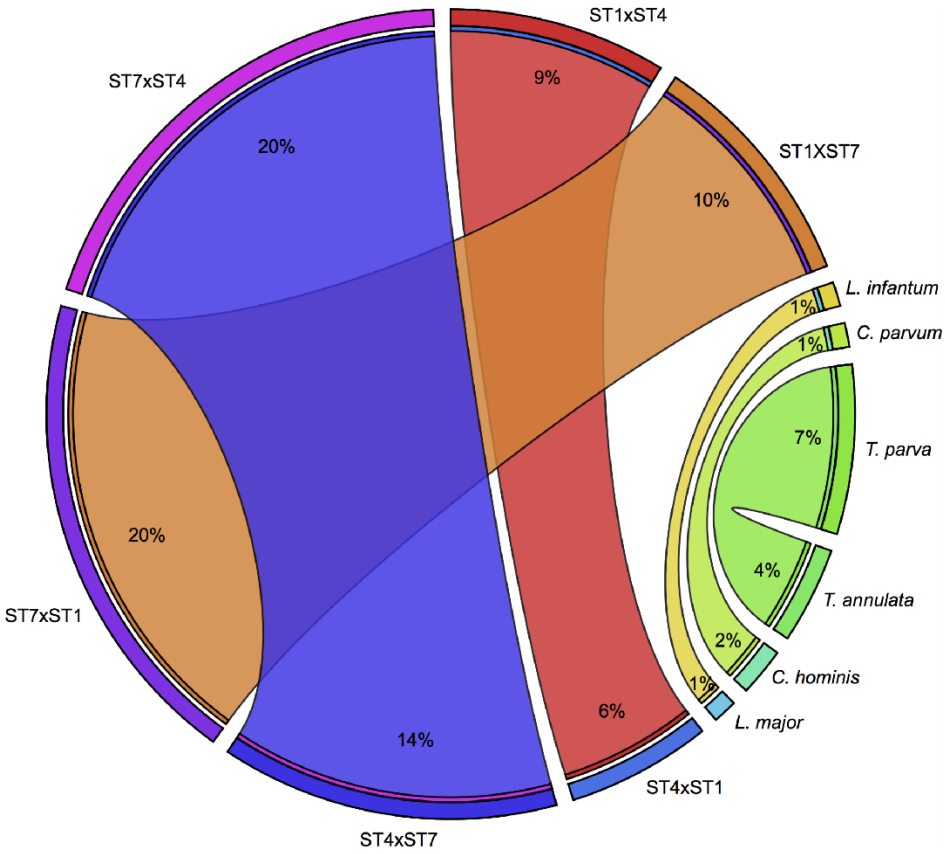


Figura 13. Comparación de genes únicos entre pares de subtipos (ST) de *Blastocystis* y otros patógenos. Se observa el porcentaje de un conjunto de genes codificantes para proteínas de cada organismo, el cual es único cuando se compara con otro grupo de genes codificantes para proteínas y viceversa(145).

7. OBJETIVOS

7.1. Objetivo General

Evaluar la diversidad y estructura genética de *Giardia intestinalis* y *Blastocystis,* mediante Análisis de Secuencias Multilocus (ASML).

7.2. Objetivos Específicos

1. Caracterizar los ensamblajes y subtipos de *G. intestinalis* y *Blastocystis* en muestras colectadas en diferentes regiones biogeográficas de Colombia.
2. Evaluar la diversidad intra especie y estructura genética de *G. intestinalis*, mediante el desarrollo y uso de nuevos marcadores moleculares.
3. Determinar la utilidad del gen *sdhA* para evaluar la diversidad y estructura genética de *Blastocystis,* en muestras humanas colectadas en cuatro departamentos de Colombia.

8. INTRODUCCIÓN A LOS CAPÍTULOS DEL ESTUDIO

Las enfermedades infecciosas son consideradas un gran reto a nivel mundial, puesto que se ha visto, que a pesar de los esfuerzos para disminuir las tasas de morbi-mortalidad en los seres humanos, aún existen falencias en las medidas preventivas y de control que repercuten en la continua transmisión de diferentes patógenos en la población humana, impidiendo el manejo de algunas enfermedades, principalmente en áreas endémicas (150). Dentro de dichas enfermedades se ha encontrado una frecuencia importante de las infecciones intestinales causadas por diferentes grupos microbianos. Las enfermedades infecciosas causadas por parásitos intestinales presentan una amplia distribución a nivel mundial. En el año 2001, alrededor de 3500 millones de personas se encontraban infectadas por protozoos y helmintos intestinales, de las cuales la población infantil mostró ser la más afectada (1). Principalmente por protistas que incluyen algunas especies del género *Blastocystis* (8), junto con protozoos como *Giardia, Cryptosporidium* y *Entamoeba,* quienes son los principales agentes protozoarios causales de cuadros diarreicos en humanos (2, 6, 151).

Por otra parte, técnicas moleculares para su caracterización han sido implementadas en muestras de diferentes regiones geográficas (6, 59, 71, 76, 138, 152), mostrando resultados que permiten hacer un estudio no sólo enfocado al diagnóstico sino también a profundizar más en las características genéticas de cada microorganismo, dado que en países en desarrollo donde la endemicidad es alta, la heterogeneidad de estos protozoos también lo es, posiblemente por la transmisión constante e intensa que puede ser tanto zoonótica como antroponótica (152, 153). Sin embargo, en el caso particular de *Blastocystis* y *G. intestinalis* el avance no ha sido significativo, puesto que los análisis se remiten normalmente a un solo marcador como el *18s* en el caso de *Blastocystis* (8, 130, 134, 141) y no se conoce cómo se afecta la variabilidad funcional y fenotípica con la variabilidad genética dentro de cada subtipo (ST). Para el caso de *G. intestinalis* se utilizan tan solo tres genes, el *gdh, tpi* y *bg*, los cuales suelen ser incongruentes entre sí (55) e insuficientes para resolver la asignación del ensamblaje B, debido a su carácter altamente polimórfico (71) y tener bajo poder de resolución para discriminar los subensamblajes del ensamblaje A, en algunos casos, haciendo difícil detectar eventos de recombinación (76). En cuanto a su estructura genética existen pocos estudios, asociados principalmente a los genes usualmente utilizados para tipificación.

Por lo anterior y teniendo en cuenta el vacío en el conocimiento que existe sobre las características genéticas, de diversidad y poblacionales de dichos microorganismos, principalmente en Colombia, y, que las condiciones sanitarias en varias regiones del país propenden por un aumento en la transmisión de estos, aumentando la posibilidad de recombinaciones o transferencia de material genético en zonas de alta endemicidad junto con la posible formación de cepas virulentas que amplíen su rango de hospederos

aumentando su dispersión, el presente estudio fue diseñado enfocado hacia tres etapas principalmente. La primera etapa (**Capítulo 1**), nos permitió evaluar muestras de heces humanas provenientes de cinco regiones naturales de Colombia e identificar los ensamblajes y subtipos de *G. intestinalis* y *Blastocystis*, respectivamente y las especies de *Cryptosporidium* y *Entamoeba*, circulantes en diferentes regiones del país. A su vez, nos permitió determinar algunos índices de diversidad genética entre las diferentes regiones evaluadas. La segunda etapa (**Capítulo 2**), se basó en el diseño de un esquema MLST (siglas en inglés para tipificación de secuencias multilocus) para *G. intestinalis*, donde se evaluaron 9 genes constitutivos para determinar la diversidad y estructura genética intra taxa, tanto *in silico*, como con muestras de ADN obtenidas a partir de heces humanas recolectadas en diferentes departamentos de Colombia y, en la última etapa (**Capítulo 3**), evaluamos un nuevo marcador genético, para la tipificación y evaluación de la diversidad y estructura genética de *Blastocystis.*

Capítulo 1: Para cumplir con la finalidad propuesta para esta etapa del estudio, se realizó la detección de ADN de *Blastocystis, G. intestinalis, Cryptosporidium* y del complejo *Entamoeba histolytica/dispar/moshkovskii* a partir de 649 muestras de heces humanas. Los resultados obtenidos en esta fase permitieron conocer el panorama general de estos microorganismos circulantes en las distintas regiones de Colombia evaluadas, sugiriendo que la región del Caribe es la que posiblemente presenta una mayor frecuencia para todos los microorganismos evaluados. *Blastocystis* mostró una mayor frecuencia, seguido de *G. intestinalis*. Además, se encontró que para *G. intestinalis* el ensamblaje más frecuente fue el B, y para *Blastocystis* los STs más comunes fueron los ST1-ST3, concordante con estudios previos (7, 132). Se reportó por primera vez en Colombia la presencia del ST9 de *Blastocystis* y el ensamblaje G de *G. intestinalis* en muestras humanas. Adicionalmente, detectamos gran diversidad genética en *Blastocystis* y *G. intestinalis* entre los departamentos evaluados, sugiriendo una mayor diversidad en el departamento de Córdoba. Sin embargo, un mayor número de muestras es requerido para confirmar estos resultados.

Por otra parte, y de forma paralela, se realizó una comparación entre las técnicas de detección, microscopía vs. PCR, encontrando una baja concordancia entre las pruebas utilizadas (kappa = 0,3551). Adicional a esto, por medio de microscopía y la PCR, se logró establecer la presencia de poliparasitismo, siendo *Blastocystis/G. intestinalis*, la coinfección más frecuente en una misma muestra. También se incluyeron muestras de mascotas de algunos individuos evaluados, mostrando el potencial zoonótico de transmisión que tienen estos microorganismos.

Como producto de este capítulo, se adjuntan los siguientes artículos científicos publicados en revistas indexadas internacionales en el primer cuartil (Q1) de Scimago:

1. Villamizar X, **Higuera A**, Herrera G, Vasquez-A LR, Buitron L, Muñoz LM, Gonzalez-C FE, Lopez MC, Giraldo JC, Ramírez JD. Molecular and descriptive epidemiology of

intestinal protozoan parasites of children and their pets in Cauca, Colombia: a cross-sectional study. BMC Infectious Diseases. 2019 Dec;19(1):1-1. doi: 10.1186/s12879-019-3810-0.

2. **Higuera A**, Villamizar X, Herrera G, Giraldo JC, Vasquez-A LR, Urbano P, Villalobos O, Tovar C, Ramírez JD. Molecular detection and genotyping of intestinal protozoa from different biogeographical regions of Colombia. PeerJ. 2020 Mar 9;8:e8554. doi: 10.7717/peerj.8554.

Capítulo 2: Luego de determinar los microorganismos que probablemente circulan en mayor frecuencia en algunas regiones del país, quisimos evaluar nuevos marcadores que nos permitieran profundizar sobre la diversidad de este parásito. Para esto, inicialmente se diseñaron iniciadores sobre siete loci diferentes correspondientes a genes codificantes para enzimas metabólicas, que, junto con los usualmente utilizados, *gdh* y *tpi*, fueron evaluados *in silico*, en 80 genomas disponibles en la base de datos pública del archivo de nucleótidos europeos (ENA). Dentro de los resultados obtenidos, se encontró que los nueve genes fueron de gran utilidad en la tipificación y determinación de la diversidad genética de *G. intestinalis* y por lo tanto pueden ser utilizados en la generación de un esquema MLST. Se encontró una mayor diversidad nucleotídica en el subensamblaje AII y expansión poblacional de cada ensamblaje analizado. La mayoría de las secuencias analizadas fueron concordantes entre los diferentes genes evaluados, mostrando que estos loci son buenos candidatos para establecer la variación intra taxa. Adicionalmente, se encontró evidencia de intercambio genético, muy probablemente por recombinación, principalmente entre los ensamblajes A y B, a pesar de presentar una elevada diferenciación genética entre todos los ensamblajes. Posteriormente, se determinó experimentalmente la utilidad de estos marcadores en ADN extraído a partir de un cultivo xénico de *G. intestinalis* y de muestras de heces provenientes de diferentes departamentos de Colombia. Dentro de los resultados obtenidos, se observaron clusters monofiléticos en las reconstrucciones filogenéticas generadas y con señales de reticulación indicadoras de posibles eventos de intercambio genético entre las muestras de Colombia, y que, junto con los índices de diversidad observados, mostraron una diversidad genética elevada de estas muestras evaluadas en comparación con los genomas analizados.

Como producto de esta fase, se publicó el siguiente artículo científico en la revista indexada internacional Pathogens que se encuentra en el primer cuartil (Q1) de Scimago:

1. **Higuera A**, Muñoz M, López MC, Reyes P, Urbano P, Villalobos O, Ramírez JD. Development of a multilocus sequence typing scheme for Giardia intestinalis. Genes. 2020 Jul;11(7):764. doi: 10.3390/genes11070764.

Capítulo 3: Posteriormente, y debido a que *Blastocystis* presenta una prevalencia tan elevada a nivel mundial incluyendo a Colombia y a que la mayoría de los estudios se centran únicamente en el uso del gen ribosomal *18s*, se evaluaron otras regiones del genoma de

Blastocystis que incluyeran genes metabólicos constitutivos para poder explorar su diversidad intra específica. Para esto, se diseñaron iniciadores sobre 12 genes, los cuales fueron posteriormente probados por PCR utilizando ADN de cultivo xénico y amplificados en muestras de ADN provenientes de heces humanas de diferentes departamentos de Colombia. Como resultado, se obtuvieron 117 secuencias para *Blastocystis* con el marcador para la *succinato deshidrogenasa subunidad A (sdhA)*, que mostraron una menor diversidad genética para *sdhA* en comparación con el gen ribosomal *18s*. Por otra parte, se realizaron inferencias filogenéticas que mostraron gran concordancia entre los STs asignados a las muestras al concatenar los dos genes y una relativa concordancia entre el concatenado y el gen *sdhA*, mostrando que, en conjunto, estos dos marcadores permiten observar las agrupaciones correspondientes a los STs asignados a las muestras evaluadas. Sorprendentemente, el marcador *sdhA* permitió observar variación intra ST y asociación de algunos de los subgrupos encontrados dentro de cada ST a la región geográfica de colección de las muestras. De igual manera se observó diferenciación genética entre los STs evaluados, tanto con el gen ribosomal *18s* como con *sdhA*.

Como producto de este capítulo, se sometió el siguiente artículo científico a la revista Helyion la cual es una revista indexada internacional y se encuentra en el primer cuartil (Q1) de Scimago:

1. **Higuera A**, Muñoz M, López MC, Reyes P, Urbano P, Villalobos O, Ramírez JD. Succinate dehydrogenase gene as a marker for studying *Blastocystis* genetic diversity.

8.1 CAPÍTULO 1

Características epidemiológicas de protozoos intestinales en Colombia.

Villamizar et al. BMC Infectious Diseases (2019) 19:190
https://doi.org/10.1186/s12879-019-3810-0

BMC Infectious Diseases

RESEARCH ARTICLE Open Access

Molecular and descriptive epidemiology of intestinal protozoan parasites of children and their pets in Cauca, Colombia: a cross-sectional study

Ximena Villamizar[1], Adriana Higuera[1], Giovanny Herrera[1], Luis Reinel Vasquez-A[2], Lorena Buitron[2], Lina Maria Muñoz[2], Fabiola E. Gonzalez-C[2], Myriam Consuelo Lopez[3], Julio Cesar Giraldo[4] and Juan David Ramírez[1*]

Abstract

Background: Parasitic infections, particularly those caused by protozoa, represent a considerable public health problem in developing countries. *Blastocystis*, *Giardia duodenalis*, *Cryptosporidium* spp. and the *Entamoeba* complex (*Entamoeba histolytica, Entamoeba dispar* and *Entamoeba moshkovskii)* are the most common etiological causes of intestinal parasitic infections.

Methods: We carried out a descriptive cross-sectional study in school-age children attending a daycare institution in commune eight of Popayán, Cauca (Southwest Colombia). A total of 266 fecal samples were collected (258 from children and eight from pets). *Blastocystis*, *G. duodenalis*, *Cryptosporidium* spp. and the *Entamoeba* complex were identified by microscopy, quantitative real-time PCR (qPCR) and conventional PCR. The concordance of qPCR and microscopy was assessed using the Kappa index. Molecular characterization was conducted to identify *Blastocystis* subtypes (18S), *G. duodenalis* assemblages (*tpi* and *gdh*) and *Cryptosporidium* species/subtypes (18S and GP60). Potential associations between intestinal parasitism and sociodemographic factors were examined using bivariate analyses.

Results: A total of 258 fecal samples from children were analyzed by microscopy and 255 samples were analyzed by qPCR. The prevalence of *Blastocystis* was between 25.19% (microscopy) and 39.22% (qPCR), that of *G. duodenalis* was between 8.14% (microscopy) and 10.59% (qPCR), that of *Cryptosporidium* spp. was estimated at 9.8% (qPCR), and that of the *Entamoeba* complex was between 0.39% (conventional PCR) and 0.78% (microscopy). The concordance between microscopy and qPCR was very low. *Blastocystis* ST1 (alleles 4, 8, and 80), ST2 (alleles 11, 12, and 15), ST3 (alleles 31, 34, 36, 38,57, and 151), and ST4 (alleles 42 and 91), *G. duodenalis* assemblages AII, BIII, BIV and D, *C. parvum* subtype IIa and *C. hominis* subtype IbA9G3R2 were identified. The only identified member of the *Entamoeba* complex corresponded to *E. histolytica*. No statistically significant association was identified between parasitic infection and any sociodemographic variable.

Conclusion: This study revealed the usefulness of molecular methods to depict the transmission dynamics of parasitic protozoa in southwest Colombia. The presence of some of these protozoa in domestic animals may be involved in their transmission.

Keywords: *Blastocystis*, *Giardia duodenalis*, *Cryptosporidium*, *Entamoeba histolytica/dispar/moshkovskii* complex, Zoonotic disease

* Correspondence: juand.ramirez@urosario.edu.co
[1]Grupo de Investigaciones Microbiológicas-UR (GIMUR), Programa de Biología, Facultad de Ciencias Naturales y Matemáticas, Universidad del Rosario, Bogotá, Colombia
Full list of author information is available at the end of the article

Background

Infections by Intestinal parasites occur worldwide, and their high prevalence rates represent a major public health problem. The major pathogens responsible for intestinal parasitic infections are protozoa. These microorganisms can cause significant morbidity in children as well as opportunistic infections in immunosuppressed patients [1]. Infection by intestinal parasites is generally associated with factors such as fecal contamination of soil and food, insufficient access to clean drinking water, lack of environmental sanitation, and vulnerable socioeconomic conditions. Given the socio-cultural features of developing countries, these regions tend to have the highest rates of infection by parasites [2]. Among the intestinal protozoa, *Blastocystis*, *Giardia duodenalis* (also known as *Giardia intestinalis* and *Giardia lamblia*), *Cryptosporidium* spp. and members of the *Entamoeba* complex (*Entamoeba histolytica, Entamoeba dispar and Entamoeba moshkovskii)* impose major burdens of diarrheal disease in children. The primary modes of protozoan parasite transmission are the fecal-oral route after direct or indirect contact with the infective forms (cysts/oocysts), human to human transmission, animal to human transmission, transmission by water, transmission through contaminated food, and airborne transmission (for *Cryptosporidium* spp. only) [3, 4].

Blastocystis are pleomorphic intestinal parasites commonly found in the gastrointestinal tracts of humans and both domestic and wild animals across the world [5]. Significant genetic diversity has been observed among the numerous *Blastocystis* isolates identified in humans and animals. *Blastocystis* can be grouped into subtypes with similar morphological characteristics: using the small subunit of ribosomal RNA (SSU rRNA), at least 17 subtypes (ST1 to ST17) and 151 different 18S alleles have been described [6–9]. ST1–ST8 and ST12 infect humans and animals (primates, pigs, cattle, rodents, and birds), ST9 infects only humans, and ST10, ST11, and ST13–ST17 have only been isolated from animals [10, 11]. In Colombia, the estimated prevalence of *Blastocystis* is 52.1% and studies have described the major circulating subtypes in human and animal populations (ST1–ST4 and ST6–ST8, of which ST1 and ST3 were found in humans and ST2 was found in both humans and dogs) [8, 11–13].

G. duodenalis is a single-celled flagellated parasite that infects the gastrointestinal tracts of humans and other mammals [14]. To date, eight genetic groups of *G. duodenalis* (assemblages A to H) have been identified [15]. Assemblages A (including AI and AII) and B (BIII and BIV) are responsible for most human infections and have also been identified in a wide range of mammals. The remaining assemblages show more restricted host ranges: assemblages C and D have been identified in canines, E in cattle, F in cats, G in rodents and H in seals and gulls [16]. In Colombia, the estimated prevalence of *G. duodenalis* in children is 15.4% and assemblages A and B have been detected with different frequencies depending on the population studied [13, 17–20].

Cryptosporidium spp. mainly infects the intestine and other extracellular spaces. Based on morphological, biological and molecular markers (SSU rRNA, HSP70, oocyst wall protein and the 60-kDa glycoprotein gp60, also known as gp40/15), at least 30 species and more than 70 genotypes have been identified. At least 20 species have been identified in humans and more than 90% of human infections are caused by *Cryptosporidium hominis* (anthroponotic) and *Cryptosporidium parvum* (zoonotic). Other species, including *Cryptosporidium meleagridis*, *Cryptosporidium canis*, *Cryptosporidium felis*, *Cryptosporidium ubiquitum*, and *Cryptosporidium cuniculus,* are less frequently detected in humans [21]. Ten subtypes of *C. hominis* (Ia–Ik) and 16 subtypes of *C. parvum* (IIa–IIp) have been described, with subtypes Ia, Ib, Id, and Ie of the former and subtypes IIa and IId of the latter having the highest prevalence worldwide [22]. In Colombia, the prevalence of this parasite is 0.5% and few studies have identified *C. parvum*, *C. hominis* or *C. viatorum* in humans [13].

The genus *Entamoeba* includes seven species: *E. histolytica*, *E. dispar*, *E. moshkovskii*, *Entamoeba bangladeshi*, *Entamoeba poleki*, *Entamoeba coli* and *Entamoeba hartmanni*. *E. histolytica*, *E. dispar* and *E. moshkovskii* are morphologically identical but genetically distinct; however, because the direct diagnostic methods currently in use do not permit their differentiation, they are typically reported as a complex. In humans the *Entamoeba* complex (*E. dispar*, *E. histolytica* and *E. moshkovskii*) is only differentiated by means of PCR. *E. dispar* is non-pathogenic while *E. histolytica* is pathogenic; the pathogenicity of *E. moshkovskii* is still controversial. In Colombia, all three species have been detected in asymptomatic children with a prevalence of approximately 15% [23].

Diagnosis of protozoan intestinal parasites is typically made using conventional methods such as microscopic examination of stool samples. This method is limited by its low specificity and sensitivity, which are related to the instability and rapid deterioration of some protozoan parasites outside the host [24]. Microscopic examination cannot distinguish between different species of *Cryposporidium*, between different assemblages of *Giardia*, between different subtypes of *Blastocystis* and between pathogenic and nonpathogenic species of *Entamoeba* [25]. For this reason, implementation of molecular techniques to diagnose circulating parasite species and subpopulations in vulnerable populations has been proposed to accurately measure parasite prevalence in endemic areas. Real-time PCR (qPCR) is an alternative technique that allows identification of parasitic DNA from preserved fecal matter and has a sensitivity between 80 and 100%, thus allowing more

sensitive detection of parasitic infection than microscopy [24, 26, 27]. However, few studies in Colombia have addressed the prevalence of these parasites or attempted to compare prevalence estimates using different diagnostic techniques.

Therefore, the aim of the present study was the epidemiological and molecular characterization of intestinal protozoa (*Blastocystis, G. duodenalis, Cryptosporidium* spp. and members of the *Entamoeba* complex) in children and dogs living in the commune eight of Popayán, Cauca, southwest Colombia. A secondary goal was to conduct an overall comparison of the diagnostic performance of microscopy versus qPCR.

Methods

Study population

School-aged children (age 12–54 months) attending a day-care institution located in commune eight in Popayán, Cauca, southwest Colombia and 8 samples derived from dogs, as the only pet used in the study were recruited. A total of 266 fecal samples were collected (258 samples from children and eight from their pets). All samples were used for identification of intestinal protozoa by conventional and molecular methods. For samples that were positive using molecular methods, detailed molecular characterization was performed.

Sociodemographic variables

At the time of providing informed consent, a structured survey was administered to collect information on the following variables: intestinal discomfort, socioeconomic stratum (In Colombia, the stratums are divided from 1 to 6 according to monthly income; stratums 1–2 are considered low-income, 3–4 middle-income and 5–6 high income), place of residence, age, sex, number of children in the house, monthly income, type of property, floor type, wall type, availability of public services, water quality, presence and number of pets, fecal elimination habits, hand washing habits and garbage storage/disposal procedures.

Detection of intestinal parasites

Microscopy

Fecal samples were split into two. The first half of the sample was fixed with SAF solution (sodium acetate, acetic acid and formaldehyde) for identification of intestinal parasites by direct examination (microscopy) in saline solution containing Lugol's iodine accompanied by the modified Ritichie-Frick concentration method and the Kato-Katz method as suggested by the World Health Organization (WHO) [28].

Real time PCR

DNA was extracted from the second half of the fecal sample using the Norgen Stool Extraction Kit, then The qPCRs were performed in 96 wells MicroAmp (Applied Biosystems), reactions in a total volume of 9 μL containing 3.5 μL of Taqman™ Mastermix (Roche), 1.0 μL of species-specific primers (10 μM) and primers of the internal amplification control (IAC) (10 μM), and 0.4 μL Taqman probes (5 μM) (*G. duodenalis, Blastocystis, Cryptosporidium*), 0.3 μL the water and 2.0 μL of DNA. The samples were processed by duplicate in an Applied Biosystems 7500 Fast equipment using default parameters of 40 cycles [13, 27, 29]. For the *Entamoeba* complex, conventional multiplex PCR was performed as previously reported [30]. We used DNA extracted from axenic cultures from *G. duodenalis, Blastocystis*, *E. histolytica, E. dispar, E. moshkovskii* and *C. hominis* as positive controls and fecal samples from patients from non-endemic regions that had previously tested negative for intestinal parasites by microscopy and qPCR as negative controls.

Genotyping for identifying *G. duodenalis* assemblages, *Blastocystis* subtypes and alleles, and *Cryptosporidium* species and subtypes

Genotyping was conducted for samples that were positive by qPCR for *G. duodenalis*, *Blastocystis*, and *Cryptosporidium* spp.. For identification of *Giardia* assemblages, these samples were subjected to conventional PCR using primers specific for the following molecular markers: (i) *gdh* (glutamate dehydrogenase) using primers GDHeF (5′-TCAA CGTYAAYCGYGGYTTCCGT-3′), GDHiF (5′-CAGTAC AACTCYGCTCTCGG-3′) and GDHiR (5′-GTTRTCCT TGCACATCTCC-3′) as reported elsewhere [31], and (ii) *tpi* (triose phosphate Isomerase) using primers AL3543 (5′-AAATIATGCCTGCTCGTCG-3′), AL3546 (5′-CAAA CCTTITCCGCAAACC-3′), AL3544 (5′-CCCTTCATCG GIGGTAACTT-3′), and AL3545 (5′-GTGGCCACCACIC CCGTGCC-3′) as reported elsewhere [32]. For identification of *Blastocystis* subtypes and alleles, SSU rRNA was amplified using primers RD5 (5′-ATCTGGTTGATCCT GTCCAG-3′) and BhRDr (5′-GAGTGCCTTTTTAACA ACAACG-3′) as previously described [33]. *Cryptosporidium* spp. were identified using direct sequencing of the SSU rRNA fragment using primers 18SF (5′-AGTG ACAAGAAATAACA ATACAGG3′) and 18SRv (5′-CCT GCTTTAAGCACTCTAATTTTC-3′) [34]. Subtyping of *C. hominis* and *C. parvum* was based on sequence analysis of gp60 genes. Each specimen was analyzed by the relevant method at least twice. Subtype families for *C. hominis* and *C. parvum* were determined based on sequence differences in the nonrepeat region of the gene. Within each subtype family, subtypes differed from one another, mostly in the number of serine-coding trinucleotide repeats (TCA, TCG, or TCT microsatellite) located in the 5′ region of the gene. The previously-established nomenclature system was used to differentiate subtypes within each subtype family as reported elsewhere [10].

Both strands of PCR products were sequenced using the dideoxy-terminal method for all genetic markers herein employed in a 310 Genetic Analyzer (Applied Biosystems). Sequences were edited and aligned using MEGA 7.0 [35] and compared with reference sequences using BLAST. For *Blastocystis*, a database established for allele and subtype confirmation was queried (http://pubmlst.org/blastocystis/). In the case of *G. duodenalis*, a phylogenetic reconstruction was performed using maximum likelihood methods with 1,000 bootstrap replicates. Reference sequences contained in GenBank with the following accession numbers were included: AI (M84604), AII (AY178737), BIII (AF069059), BIV (AY178739), C (U60982), D (U60986), E (AY178741), E (AB182127), F (AB569384), G (AF069058), G (AY178745), H (GU176089). The phylogenetic tree was rooted with *Giardia ardeae* (AF069060). For *Cryptosporidium,* sequences were compared with species and subtype control sequences harbored at CDC, Atlanta and kindly provided by Dr. Lihua Xiao.

Statistical analyses

Descriptive statistics were used to analyze data of interest. Variables were classified as categorical or continuous. Categorical variables were summarized by relative frequencies and their association with the presence of intestinal protozoa (*Blastocystis*, *G. duodenalis*, *Cryptosporidium* spp. and members of the *Entamoeba* complex) was assessed using chi-square tests. The 95% confidence intervals (CIs) were calculated for each of the associations. For continuous variables, the normality of the data was assessed using the Shapiro-Wilk test. Means and standard deviations were used to summarize normally-distributed variables, while medians and interquartile ranges were used for variables with non-normal distributions. The means and medians of each continuous variable were compared by t-tests or Mann Whitney U tests, depending on the fulfillment of the assumption of normality. All analyses were performed in STATA version 14.0 and values of $P < 0.05$ were considered statistically significant.

Concordance between microscopy and qPCR results was assessed by calculating the kappa index. A value of kappa close to one indicated that the results of both methods were concordant; and a value close to zero indicated that the methods were not concordant.

Results

Prevalence of intestinal parasitic infection

A total of 258 human fecal samples were analyzed by microscopy and 255 were analyzed by PCR (only for *Blastocystis*, *G. duodenalis*, *Cryptosporidium* spp. and *Entamoeba* complex). Three samples were not able to be analyzed by qPCR due to the low amount of the fecal sample that only allowed microscopy examination. Prevalence estimates by microscopy were 25.19% (65/258) for *Blastocystis*, 8.14% (21/258) for *G. duodenalis*, 0.78% (2/258) for members of the *Entamoeba* complex, 10.85% (28/258) for *E. coli*, 1.55% (4/258) for *Chilomastix* spp., 6.20% (16/258) for *Endolimax nana* and 0.38% (1/258) for *Entamoeba hartmani* (Fig. 1).

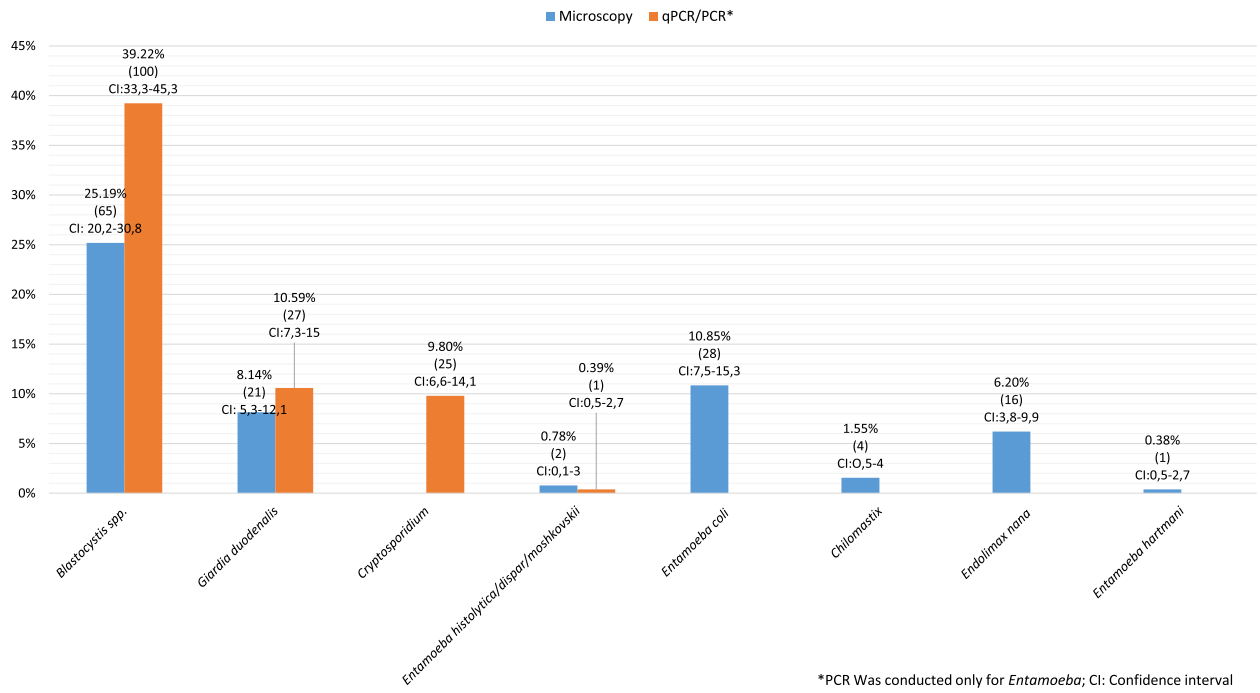


Fig. 1 Frequency of intestinal parasites detected by microscopy, qPCR and PCR in the samples analyzed

Cryptosporidium spp. could not be identified by microscopy due to logistical limitations. No helminths were detected across the samples. The presence of the most frequent intestinal protozoa parasites in Colombia was detected by qPCR. Prevalence estimates by qPCR were 39.22% (100/255) for *Blastocystis*, 10.59% (27/255) for *G. duodenalis*, 9.8% (25/255) for *Cryptosporidium* spp. and 0.9% (1/255) for *E. histolytica* (Fig. 1). Eight samples of pet feces were analyzed by qPCR and two samples showed evidence of infection by intestinal protozoa. *Blastocystis* was identified in one sample and in the other sample, *G. duodenalis* and *Cryptosporidium* spp. were detected; curiously, the owner of the *G. duodenalis* and *Cryptosporidium*-infected pet was not infected by either of these parasites, as was the case for the owner of the *Blastocystis*-infected pet.

Evaluation of polyparasitism

We found that among the stool samples analyzed by qPCR (for detection of *Blastocystis*, *Giardia*, and *Cryptosporidium* spp.) and conventional PCR (for detection of the *Entamoeba* complex), approximately 36% of individuals were infected by a single parasite, 11% were infected by two parasites and 1% were infected by three parasites (Fig. 2a). Using microscopy, we found that *E. coli*, *Chilomastix* spp., *E. nana* and *E. hartmani* were involved in polyparasitism (Fig. 2a). Using qPCR, we found that *Blastocystis/G. duodenalis* and *Blastocystis/Cryptosporidium* (Fig. 2b) coinfections occurred more frequently than *G. duodenalis/Cryptosporidium* coinfections ($P < 0.05$). Members of the *Entamoeba* complex were not involved in polyparasitism.

Comparison of the diagnostic performance of qPCR vs. microscopy

We analyzed the concordance between the qPCR and microscopy results and determined that qPCR was significantly more sensitive than microscopy for identification of *Blastocystis* (71.7% versus 56.1%, $P < 0.05$) and *G. duodenalis* (90% vs. 83.6%, $P < 0.05$). For samples testing positive both by microscopy and qPCR, *Blastocystis* was identified in 36.7% ($n = 44$) of samples by both techniques, in 16.7% ($n = 20$) of samples only by microscopy and in 46.7% ($n = 56$) of samples only by qPCR. The overall concordance between the two techniques was low, with a kappa index of 0.3551 (Fig. 3). *G. duodenalis* was identified in 46.8% ($n = 22$) of samples by both techniques, in 19.1% ($n = 9$) of samples only by microscopy and in 34% ($n = 16$) of samples only by qPCR. The kappa index was 0.3912 (Fig. 3).

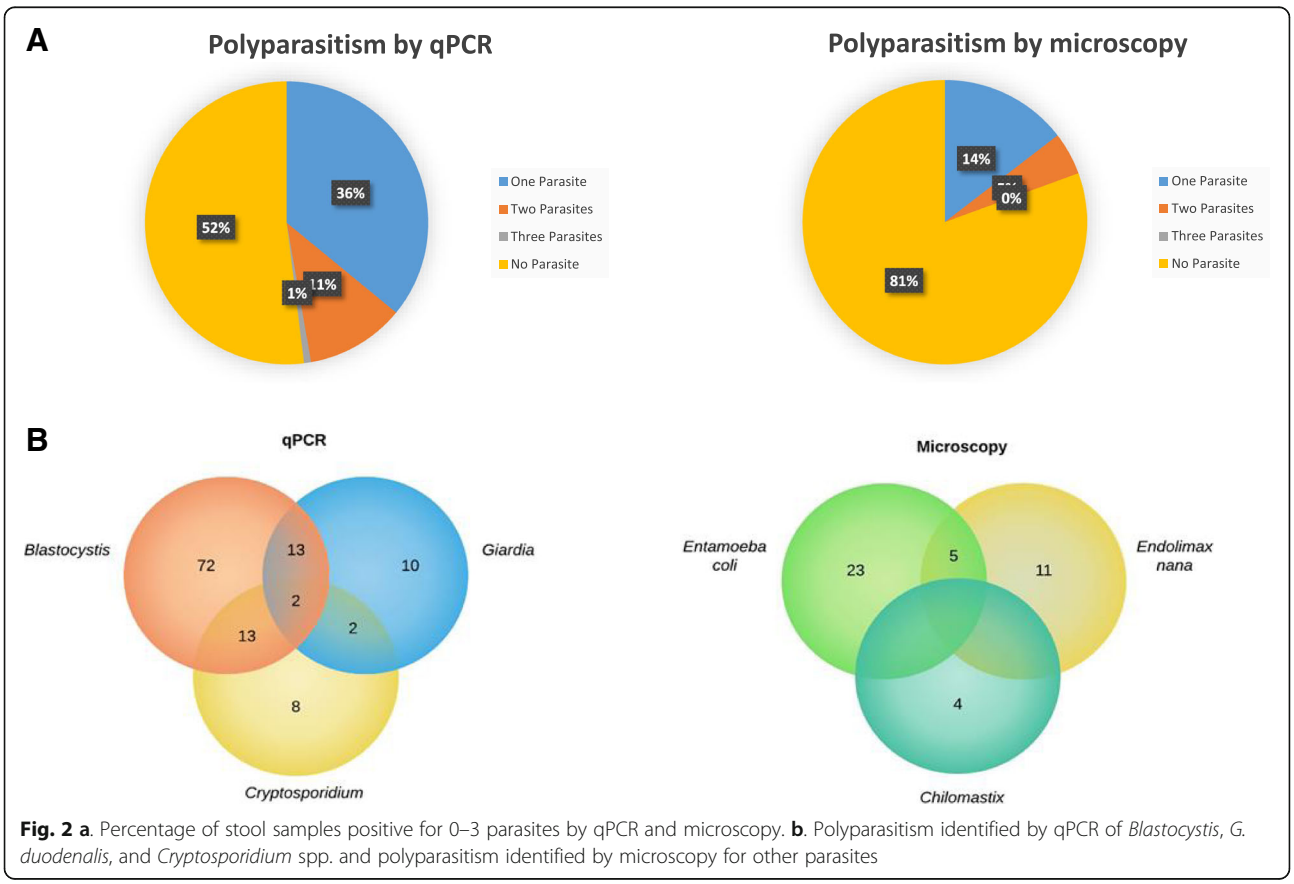


Fig. 2 a. Percentage of stool samples positive for 0–3 parasites by qPCR and microscopy. **b**. Polyparasitism identified by qPCR of *Blastocystis*, *G. duodenalis*, and *Cryptosporidium* spp. and polyparasitism identified by microscopy for other parasites

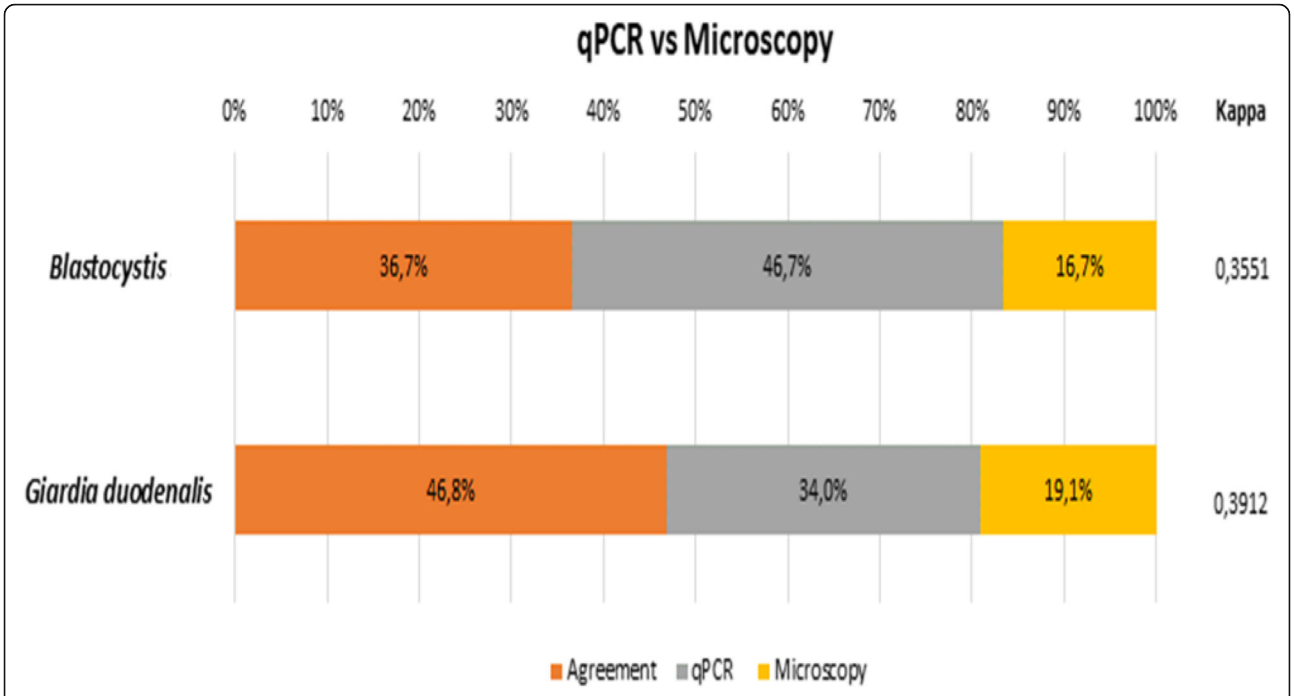


Fig. 3 Analysis of concordance as shown by the kappa index between qPCR and microscopy measurements for identification of *Blastocystis* and *G. duodenalis*

Blastocystis subtypes, *Giardia duodenalis* assemblages and *Cryptosporidium* species

The prevalence rates of *Blastocystis* subtypes, *G. duodenalis* assemblages and *Cryptosporidium* spp. were assessed by genotyping samples that were positive by qPCR. A total of 62 *Blastocystis* samples that were qPCR-positive were successfully subtyped. The most common subtypes were ST1 (38.7%, $n = 24$), ST2 (14.52%, $n = 9$), ST3 (43.55%, $n = 27$) and ST4 (3.22%, $n = 2$). In addition, the different alleles associated with each subtype were identified. For ST1, alleles 4, 8, and 80 were identified; for ST2, alleles 11, 12, and 15 were identified; for ST3, alleles 151, 31, 34, 36, 38, and 57 were identified; and for ST4, alleles 42 and 91 were identified. ST1 allele 4 had the highest frequency, while the most frequent subtype was ST3 (Fig. 4a). The single pet fecal sample testing positive for *Blastocystis* was identified as ST1, allele 4 (Fig. 4a).

Of the 27 samples testing positive for *G. duodenalis* by qPCR, only 12 could be genotyped using the *gdh* gene as no amplification products for Sanger sequencing were obtained from the remaining 15 samples. In the case of the *tpi* gene, multiple bands were observed after electrophoresis and were subsequently unable to be sequenced. The *G. duodenalis* assemblages identified were AII (8.3%, $n = 1$), BIII (50%, $n = 6$), BIV (33.3%, $n = 4$) and D (8.3%, $n = 1$) (Fig. 4b). The single dog fecal sample testing positive for *G. duodenalis* was genotyped as BIII. We also observed coinfection between *Blastocystis* STs and *G. duodenalis* in six samples. Coinfection occurred between the *G. duodenalis* BIII assemblage and *Blastocystis* ST1, ST2, and ST3; between the *G. duodenalis* BIV assemblage and *Blastocystis* subtypes ST2 and ST3; and between *G. duodenalis* assemblage D and *Blastocystis* subtype ST1.

Although 25 samples were qPCR-positive for *Cryptosporidium* spp., only four samples could be genotyped for gp60 and three samples for SSU-RNA. In three samples, *C. parvum* subtype IIa was identified and in one sample, *C. hominis* subtype IbA9G3R2 was identified. In the case of the single positive dog fecal sample, *C. parvum* subtype IIa was identified.

Association between sociodemographic variables and intestinal parasitism

Of the study population, similar proportions of participants were male ($n = 134$, 51.94%) and female ($n = 124$, 48.06%). The most common socioeconomic stratum of study participants was stratum one (80.62%), followed by stratum two (17.05%) and stratum three (2.33%). The vast majority (98.45%) of families had incomes lower than the minimum wage. However, 67.83% of families had their own homes, most with cement floors and brick walls, and 100% of the dwellings reported having an adequate system for sanitary elimination of excreta. Moreover, 99.61% of homes had access to water from treated aqueducts. The majority (88.76%) of children surveyed did not report gastrointestinal symptoms; only 6.2 and 3.49% reported diarrhea and abdominal pain, respectively. Characteristics of the study population are shown

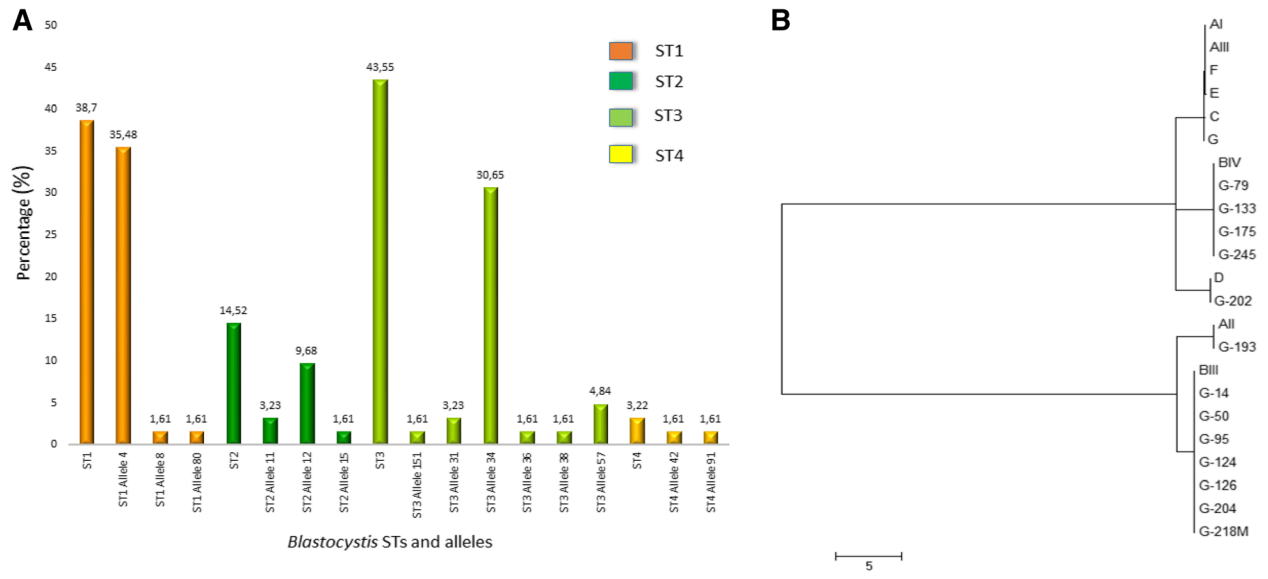


Fig. 4 a. Frequency of *Blastocystis* subtypes and alleles. **b**. Phylogenetic relationships of *Giardia duodenalis* inferred using neighbor-joining methods applied to *gdh* nucleotide sequences

in Table 1. No statistically significant associations were identified between any variable and the presence of intestinal parasites.

Discussion

Popayán, Cauca, a city in the southwest of Colombia, is blessed with significant forestry and mining resources. However, according to the last census, 22.2% of the population had unsatisfied basic needs, reflecting mainly a lack of housing and to a lesser extent, inadequate coverage of services such as education, health, drinking water and basic sanitation. These factors facilitate the development of multiple infections among the population, including intestinal parasites. Therefore, epidemiological studies such

Table 1 Sociodemographic variables studied under statistical analyzes

Variable	Category	*Blastocystis*			*Giardia duodenalis*			*Cryptosporidium*			*Entamoeba histolytica/ dispar/moshkovskii*		
		Positive	Negative	*P* value	Positive	Negative	*P* Value	Positive	Negative	*P* Value	Positive	Negative	*P* Value
Sex	Female	28	96	0,352	9	115	0,618	12	111	0,98	0	124	0,172
	Male	37	97		12	122		13	119		2	132	
Stratum	1	49	159	0,305	17	191	0,746	22	184	0,543	2	206	0,785
	2	15	29		4	40		3	40		0	44	
	3	1	5		0	6		0	6		0	6	
Type of population	Urban	64	188	0,626	21	231	0,461	24	225	0,567	1	251	0,000
	Rural	1	5		0	6		1	5		1	5	
Children per housing	1	34	113	0,802	12	135	0,709	17	129	0,791	0	147	0,448
	2	26	65		9	82		7	84		2	89	
	3	3	10		3	13		1	11		0	13	
	4	1	4		0	5		0	4		0	5	
	5	1	1		0	2		0	2		0	2	
Pets	Yes	17	63	0,328	5	75	0,457	8	70	0,872	1	79	0,560
	No	48	130		16	162		17	160		1	177	
Water quality	Treated	65	192	0,561	21	236	0,766	25	229	0,741	2	255	0,929
	Non-treated	0	1		0	1		0	1		0	1	

P-value: < 0.05

as the present one are imperative to understand the epidemiological and molecular features of the intestinal protozoa (*Blastocystis*, *G. duodenalis*, *Cryptosporidium* and *Entamoeba* spp.) affecting the infant population in developing countries.

The protozoa of interest were identified using both microscopy and qPCR. *Blastocystis* had the highest prevalence using both methods followed by *G. duodenalis* and *Cryptosporidium* spp. (detected only by qPCR) (Fig. 1). These findings are consistent with a national survey of parasitism, which showed that *Blastocystis*, *Giardia* and *Cryptosporidium* spp. were the most prevalent protozoa [36] and also with estimates from Latin America. Surprisingly, members of the *Entamoeba* complex, despite having the highest prevalence rates in the country, were detected at low frequency in our study regardless of the method of detection. When reviewing previous studies of intestinal parasites affecting the local population (Popayán, Cauca), we found no reports of the prevalence of members of the *Entamoeba* complex. Therefore, it is possible that the population studied here had a low prevalence of these parasites. However, the poor agreement between diagnostic methods (PCR and microscopy) for identification of *Entamoeba* complex members may be due to a high rate of false positives in microscopy because amoebic trophozoites can be easily confused with leukocytes (particularly macrophages that have phagocytosed red blood cells) and cysts of other amoebas [37].

Although the traditional method for diagnosis of gastrointestinal parasites is microscopy, it showed a low sensitivity for the identification of intestinal protozoa in our study compared with qPCR (Fig. 1). Likewise, comparison of these methods using the kappa index (for detection of *Blastocystis* and *G. duodenalis*.) showed a low concordance (Fig. 3). In the case of the *Entamoeba* complex, concordance could not be analyzed due to insufficient data, given the low prevalence of these parasites. Our concordance findings are similar to other studies, with a greater parasite prevalence detected by qPCR compared with microscopy [27, 38]. A study carried out by Sánchez et al. [13] among indigenous communities of the Amazon basin also showed low concordance between these methods. These results including evidence from Argentina and Ecuador support the use of molecular methods instead of microscopy for diagnosis of intestinal parasites [8] and for monitoring of patients post-treatment [39]. For logistical reasons, we did not attempt microscopic identification of *Cryptosporidium* spp. in the present study. It is important to note that this process is carried out using the Ziehl-Neelsen technique (modified acid-fast staining). This method has some disadvantages: it requires at least 50,000–500,000 oocysts per gram of fecal matter as well as significant operator expertise so as not to confuse the oocyst with other acid- and alcohol-resistant microorganisms of similar size such as *Cyclospora* or yeasts [40]. Further studies comparing microscopy with molecular methods should be conducted to establish the true prevalence of this protozoan because it has been underestimated in most developing countries. Because the immune systems of young children are not fully developed, they are more susceptible to infection and this could explain the high prevalence observed in our study. However, this parasite has been listed as one of the main etiological agents of diarrhea in children; this was demonstrated in a multicenter study conducted in Africa and Asia, where it was established as the second causative agent of diarrhea in children [41].

Polyparasitism was evaluated in the study population by qPCR and microscopy (Fig. 2). This polyparasitism may be caused by a variable immune response that may be influenced by nutritional status and repeated exposures to intestinal parasites [45]. Polyparasitism is the result of simultaneous infection with various helminths and intestinal protozoa, and is associated with ecological and environmental factors, different routes of infection and different exposures to the host [1]. Polyparasitism is very important for public health because it has a significant impact on general morbidity, nutritional status, immune reaction after treatment, and re-infection rates, causing an increase in the intensity of infection for most patients. Infection by multiple parasite species confers increased susceptibility to other infections [46], and qPCR allows greater sensitivity in the identification of polyparasitism, making it a useful tool for the evaluation of public health interventions [47]. In Colombia, studies of polyparasitism have been carried out in indigenous communities and found that both helminths and protozoa were involved [48].

We performed genotyping of the protozoa detected in our study and established that the circulating *Blastocystis* subtypes in the population were ST1, ST2, ST3, and ST4 (Fig. 4). The subtype with the highest prevalence was ST3. These results are consistent with two previous reports in Colombia, one study of nine localities in Colombia and another of indigenous communities of the Amazon region; both studies also identified the ST3 subtype as having the highest prevalence [12, 13]. Similarly, a study that included subtyped samples from several geographic regions around the world identified subtypes ST1 to ST9. Approximately 90% of the isolates belonged to ST1, ST2, ST3 and ST4 and ST3 caused most human infections worldwide [10]. We detected allele 4 of ST1 most frequently, as has been previously described in Colombia [8, 13]. Interestingly, we also detected ST4 allele 42, which has been previously reported in Colombia at low frequency, and allele 91, which has never been previously reported in the country. *Blastocystis* ST4 has a more restricted geographic distribution because

of its more recent colonization of humans; thus, alleles of this subtype are rarely detected in humans, and the 91 allele has been reported only in a study of Danish patients [49, 50]. By contrast, *Blastocystis* ST3 has been commonly identified from humans and non-human primates, and given its relatively high specificity for these hosts, infections are assumed to be caused by human-to-human transmission [7].

G. duodenalis assemblages A, B and D were identified in our study, and sub-assemblies AII, BIII, and BIV had higher prevalence than the BIV assemblage. The distribution of assemblages was not geographically limited, given that several studies carried out inside and outside the country revealed widely different distributions [16]. In Colombia, two studies conducted in specific areas of the country (one of clinical samples collected between 1997 and 2001 in the departments of Amazonas, Boyacá and Bogotá, and another conducted in the Amazon region) revealed that the most prevalent assemblage was A [13, 17]. However, other studies conducted using stool samples in central and Caribbean Colombia demonstrated a higher prevalence of assemblage B, similar to our study [18–20]. Our study population, similar to a study conducted in the central region of Colombia [20], was largely asymptomatic. In both studies, there was no association between symptoms and presence of the *G. duodenalis* B assemblage; other studies of children in Australia and Brazil presented similar findings [51–54]. These findings are contradicted by other studies, which suggested an association between severe diarrhea and this assemblage [55, 56].

Surprisingly, in our population we identified *G. duodenalis* assemblage D (Fig. 4b). This assemblage has been mainly detected in canines; however, a study conducted in German travelers identified assemblage D in two human samples originating in South Asia [57]. None of the canine fecal samples included in our study showed evidence of infection by *G. duodenalis* assemblage D. Therefore, the infection in this patient was likely to be transient. To verify this hypothesis, we would need to re-collect a fecal sample to establish the course of the infection. It would be beneficial to carry out additional studies in canines and humans to clarify our knowledge of the transmission dynamics of uncommon assemblages in humans and to evaluate the traceability of infections caused by them. Regarding *Cryptosporidium* spp., *C. parvum* and *C. hominis* (subtypes IIa and IbA9G3R2, respectively) were identified. These findings are consistent with the literature, as *C. parvum* and *C. hominis* are responsible for 90% of human infections. Likewise, *C. parvum* subtypes IIa and Ib are prevalent worldwide, and subtype IIa is considered to be predominant in humans and other animals [21, 22, 58].

In this study, the following protozoa were identified in stool samples from dogs: *Blastocystis* subtype ST1 allele 4, *G. duodenalis* assemblage BIII and *C. parvum* IIa. When comparing these results with those obtained from children, we found that the child living with the *Blastocystis*-infected dog was also infected by the same *Blastocystis* subtype and allele. By contrast, the child living with the *Giardia*- and *Cryptosporidium*-infected dog was not infected with either of these two protozoa. However, each of these parasites has been previously identified in both animals and humans, indicating that in our study population zoonotic transmission may have given rise to some of the infections [42–44].

In Colombia, several studies of intestinal parasites have included the sociodemographic variables listed here (Table 1). These studies have reported a heterogeneous distribution of associations. A study conducted in Calarcá found a higher prevalence of intestinal parasitism in children who do not achieve growth and development controls and a significant association between infection with *Blastocystis* and non-deworming of pets [59]. Another study carried out in the indigenous reservation in Nasa, Cauca, did not find statistical associations between sociodemographic conditions and parasitism. However, the authors described factors such as: low education of parents and low availability of aqueduct and sewerage that may play a role [60]. Finally, a study carried out in preschoolers and schoolchildren in Cajamarca found an association between parasitism and level of education [61]. In our study, we did not find any statistically significant associations with demographic variables.

Conclusions

The qPCR had better sensitivity for identification of *Blastocystis*, *G. duodenalis* and *Cryptosporidium* spp., which were detected at considerable frequency in the study population, with a higher prevalence of *Blastocystis*. The subtypes and alleles distributed in the population were determined and we identified an uncommon allele reported in humans (allele 91 of *Blastocystis* ST4). *G. duodenalis* assemblage D was identified in human feces, although this assemblage is typically found in canines. This may have represented possible zoonotic transmission because we detected these protozoa in canine feces. Our findings provide information for control entities about the distribution and transmission dynamics of intestinal parasites, which may help in implementation of strategies to reduce their prevalence in children.

Abbreviations

Blast: Basic Local Alignment Search Tool; Gdh: Glutamate dehydrogenase; Gp60: Glycoprotein 60; Mega: Molecular evolutionary genetics analysis; qPCR: Quantitative polymerase chain reaction; SSU rRNA: Small ribosomal ribonucleic acid subunit; ST: Sequence type; Tpi: Triose phosphate isomerase

Acknowledgements

We thank Edanz Group (www.edanzediting.com/ac) for editing a draft of this manuscript.

Funding
This work was funded by DEPARTAMENTO ADMINISTRATIVO DE CIENCIA Y TECNOLOGIA-COLCIENCIAS, project "Desarrollo de una estrategia y plataforma de Multilocus Sequence Typing (MLST) para la caracterización molecular de siete parásitos intestinales" code 122271250521 and contract number 130–2016 and by VICERECTORIA DE INVESTIGACIONES from Universidad del Cauca code 4204. The funders had no role in the design of the study, collection, analysis, interpretation of data and in writing the manuscript.

Availability of data and materials
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Authors' contributions
JDR reviewed the final version of the manuscript and supervised the project. GH performed the statistical analyses. LRV, MCL, AH, JCG, LB, LM and FG facilitated sample collection for the study, performed microscopy analyses and performed the DNA extraction. XV wrote the manuscript, analyzed the data and performed the molecular biology experiments. All authors have read and approved the final version of the manuscript.

Ethics approval and consent to participate
This project was a minimum risk investigation for the participants and was approved by the ethics committee of the department of internal medicine of the Universidad del Cauca (number VRI024/2016). Because the participants were minors, written consent and informed consent were obtained from parents or guardians prior to sample collection. In the case of pet's samples; the owners provided a written consent for sample collection.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note

Author details
[1]Grupo de Investigaciones Microbiológicas-UR (GIMUR), Programa de Biología, Facultad de Ciencias Naturales y Matemáticas, Universidad del Rosario, Bogotá, Colombia. [2]Centro de estudios en Microbiología y Parasitología (CEMPA), Departamento de Medicina Interna, Facultad de Ciencias de la Salud, Universidad del Cauca, Popayán, Colombia. [3]Departamento de Salud Pública, Facultad de Medicina, Universidad Nacional de Colombia, Bogotá, Colombia. [4]Programa de Biología, Universidad INCCA de Colombia, Bogotá, Colombia.

Received: 28 November 2018 Accepted: 12 February 2019
Published online: 26 February 2019

References

1. Harhay MO, Horton J, Olliaro PL. Epidemiology and control of human gastrointestinal parasites in children. Expert Rev Anti-Infect Ther. 2010;8(2):219–34.
2. Ortiz Pineda, Carolina; Lopez Paez, Myriam Consuelo; Rivas Muñoz, Favio Aurelio "Prevalencia de helmintos en la planta de aguas residuales del municipio El Rosal, Cundinamarca," Rev salud pública,vol. 14, no. 2, pp. 296–304, 2012.
3. Cacciò SM, Thompson RCA, McLauchlin J, Smith HV. Unravelling Cryptosporidium and Giardia epidemiology. Trends Parasitol. 2005;21(9):430–7.
4. Sponseller JK, Griffiths JK, Tzipori S. The evolution of respiratory cryptosporidiosis: evidence for transmission by inhalation. Clin Microbiol Rev. 2014;27(3):575–86.
5. Tan KSW. New insights on classification, identification, and clinical relevance of Blastocystis spp. Clin Microbiol Rev. 2008;21(4):639–65.
6. Noël C, Dufernez F. Molecular phylogenies of Blastocystis isolates from different hosts: implications for genetic diversity, identification of species, and zoonosis. J Clin Microbiol. 2005;43(1):348.
7. Stensvold CR, Alfellani M, Clark CG. Levels of genetic diversity vary dramatically between Blastocystis subtypes. Infect Genet Evol. 2012;12(2):263–73.
8. Ramírez JD, Sánchez LV, Bautista DC, Corredor AF, Flórez AC, Stensvold CR. Blastocystis subtypes detected in humans and animals from Colombia. Infect Genet Evol. 2014;22:223–8.
9. Parkar U, et al. Molecular characterization of Blastocystis isolates from zoo animals and their animal-keepers. Vet Parasitol. 2010;169(1–2):8–17.
10. 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(1):11–8.
11. Ramírez JD, et al. Geographic distribution of human Blastocystis subtypes in South America. Infect Genet Evol. 2016;41:32–5.
12. Ramírez JD, Flórez C, Olivera M, Bernal MC, Giraldo JC. Blastocystis subtyping and its association with intestinal parasites in children from different geographical regions of Colombia. PLoS One. 2017;12(2):1–13.
13. A. Sánchez et al., "Molecular epidemiology of Giardia, Blastocystis and Cryptosporidium among indigenous children from the Colombian Amazon basin," Front Microbiol, vol 8, no FEB, pp. 1–14, 2017.
14. Ryan U, Cacciò SM. Zoonotic potential of Giardia. Int J Parasitol. 2013; 43(12–13):943–56.
15. Andrews RH, Monis PT, Ey PL, Mayrhofer G. Comparison of the levels of intra-specific genetic variation within Giardia muris and Giardia intestinalis. Int J Parasitol. 1998;28(8):1179–85.
16. Sprong H, Cacciò SM, Van Der Giessen JWB. Identification of zoonotic genotypes of Giardia duodenalis. PLoS Negl Trop Dis. 2009;3(12):1–12.
17. Ravid Z, Duque S, Arévalo A, Nicholls RS, Wasserman M. Genetic diversity of Giardia intestinalis populations in Colombia. Biomedica. 2007;27(1):34–41.
18. Arroyo-Salgado B, Buelvas-Montes Y, Villalba-Vizcaíno V, Salomón-Arzuza O. Caracterización genética por reacción en cadena de la polimerasa de Giardia intestinalis en muestras de humanos y perros del Caribe colombiano. Enferm Infecc Microbiol Clin. 2014;32(7):424–7.
19. Rodriguez V, et al. Giardia duodenalis genotypes found in the Instituto Colombiano de Bienestar familiar day care centers and dogs in Ibague, Colombia. Biomedica. 2014;34(2):271–81.
20. Ramírez JD, et al. Molecular diagnosis and genotype analysis of Giardia duodenalis in asymptomatic children from a rural area in Central Colombia. Infect Genet Evol. 2015;32:208–13.
21. Azcona-Gutiérrez JM, et al. Molecular diversity and frequency of the diarrheagenic enteric protozoan Giardia duodenalis and Cryptosporidium spp. in a hospital setting in northern Spain. PLoS One. 2017;12(6):1–21.
22. Garcia-R JC, French N, Pita A, Velathanthiri N, Shrestha R, Hayman D. Local and global genetic diversity of protozoan parasites: spatial distribution of Cryptosporidium and Giardia genotypes. PLoS Negl Trop Dis. 2017;11(7):1–20.
23. López MC, et al. Molecular epidemiology of entamoeba: first description of Entamoeba moshkovskii in a rural area from Central Colombia. PLoS One. 2015;10(10):1–11.
24. Nino R, et al. Acta Tropica diagnosis of intestinal parasites in a rural community of Venezuela : advantages and disadvantages of using microscopy or RT-PCR. Acta Trop. 2017;167:64–70.
25. Haque R, et al. Prospective Case-Control Study of the Association between Common Enteric Protozoal Parasites and Diarrhea in Bangladesh. 2009; 22908:1191–7. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2883291/.
26. Cimino RO, et al. Identification of human intestinal parasites affecting an asymptomatic peri-urban Argentinian population using multi-parallel quantitative real-time polymerase chain reaction. Parasit Vectors. 2015:1–7.
27. Mejia R, et al. A novel, multi-parallel, real-time polymerase chain reaction approach for eight gastrointestinal parasites provides improved diagnostic capabilities to resource-limited at-risk populations. Am J Trop Med Hyg. 2013;88(6):1041–7.
28. Cristina I, et al. Evaluación de tres técnicas coproparasitoscópicas para el diagnóstico de geohelmintos intestinales. 26(1):15–24. http://www.scielo.org.co/pdf/iat/v26n1/v26n1a02.pdf.
29. Stensvold CR, Ahmed UN, Andersen LOB, Nielsen HV. Development and evaluation of a genus-specific, probe-based, internal-process-controlled real-time PCR assay for sensitive and specific detection of Blastocystis spp. J Clin Microbiol. 2012;50(6):1847–51.
30. Hamzah Z, Petmitr S, Mungthin M, Leelayoova S, Chavalitshewinkoon-Petmitr P. Differential detection of Entamoeba histolytica, Entamoeba dispar, and Entamoeba moshkovskii by a single-round PCR assay. J Clin Microbiol. 2006;44(9):3196–200.
31. Read CM, Monis PT, Thompson RCA. Discrimination of all genotypes of Giardia duodenalis at the glutamate dehydrogenase locus using PCR-RFLP. Infect Genet Evol. 2004;4(2):125–30.

32. Sulaiman IM, et al. Triosephosphate Isomerase Gene Characterization and Potential Zoonotic Transmission of Giardia duodenalis. 2003;9(11):15–7. https://wwwnc.cdc.gov/eid/article/9/11/03-0084_article.
33. Scicluna SM, Tawari B, Clark CG. DNA barcoding of Blastocystis. 2006;157:77–85. https://www.ncbi.nlm.nih.gov/pubmed/16431158.
34. Hunter PR, Hadfield SJ, Wilkinson D, Lake IR, Harrison FCD, Chalmers RM. Subtypes of Cryptosporidium parvum in humans and disease risk. 2007;13:1. https://www.ncbi.nlm.nih.gov/pubmed/17370519.
35. Mol Biol Evol. 2016 Jul;33(7):1870-4. https://doi.org/10.1093/molbev/msw054. Epub 2016 Mar 22.
36. Ministerio de Salud y Protección Social. U.D.a.E.N.D.P.I.E.P.E.M.U.D.A. El Ministerio. encuesta-nacional-de-parasitismo-2012-2014. 2015.
37. Iasmin D, Cavalcanti DA, Vital IR. Prevalencia de infección por Entamoeba histolytica en escuelas públicas de la ciudad de Maceió, Alagoas, Brasil prevalence of infection caused by Entamoeba histolytica in public schools of Maceio city, Alagoas, Brazil. Rev Cubana Med Trop. 2013;65(1):4–12.
38. Formenti F, et al. Molecular biology can change the classic laboratory approach for intestinal protozoan infections. 2017;8:1–6. https://www.ncbi.nlm.nih.gov/pubmed/29163458.
39. Boadi S, Polley SD, Kilburn S, Mills GA, Chiodini PL. A critical assessment of two real-time PCR assays targeting the (SSU) rRNA and gdh genes for the molecular identi fi cation of Giardia intestinalis in a clinical laboratory. no table 1. 2014. p. 811–816.
40. Tahvildar-biderouni F, Salehi N. Detection of Cryptosporidium infection by modified ziehl-neelsen and PCR methods in children with diarrheal samples in pediatric hospitals in Tehran. 2014;7(4):125–30. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4017569/.
41. Kotloff KL, et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the global enteric multicenter study, GEMS): a prospective, case-control study. Lancet. 2013;382(9888):209–22.
42. M. Adell-aledón et al., "Occurrence and molecular epidemiology of Giardia duodenalis infection in dog populations in eastern Spain," pp. 1–11, 2018.
43. J. Wang *et al.*, "Distribution and genetic diversity of Blastocystis subtypes in various mammal and bird species in northeastern China," pp. 1–7, 2018.
44. Gharieb RMA, Merwad AMA, Saleh AA, Abd El-Ghany AM. Molecular screening and genotyping. Vector Borne Zoonotic Dis. 2018;18(08):1–9.
45. Supali T, et al. Polyparasitism and its impact on the immune system. Int J Parasitol. 2010;40(10):1171–6.
46. R. Pullan and S. Brooker, "Europe PMC Funders Group The health impact of polyparasitism in humans: are we under- estimating the burden of parasitic diseases?," vol. 135, no. 7, pp. 783–794, 2009.
47. L. Meurs et al., "Diagnosing Polyparasitism in a high- prevalence setting in Beira , Mozambique: Detection of Intestinal Parasites in Fecal Samples by Microscopy and Real-Time PCR," pp. 1–18, 2017.
48. Fernández-niño JA, et al. Perfiles de poliparasitismo intestinal en una comunidad de la Amazonia colombiana. Biomédica. 2017;37(3):368–77.
49. Stensvold, Rune C, Nielsen HV. Comparison of microscopy and PCR for detection of intestinal parasites in Danish patients supports an incentive for molecular. J Clin Microbiol. 2012;50(2):540–1.
50. Stensvold CR, Christiansen DB, Elisabeth K, Olsen P, Nielsen HV. Short Report: Blastocystis sp . Subtype 4 is common in Danish Blastocystis -positive patients presenting with acute diarrhea. Am J Trop Med Hyg. 2011; 84(6):883–5.
51. Cian A, et al. Molecular epidemiology of Blastocystis sp. in various animal groups from two French zoos and evaluation of potential zoonotic risk. PLoS One. 2017;12(1):1–29.
52. Yoshikawa H, et al. Molecular characterization of Blastocystis isolates from children and rhesus monkeys in Kathmandu, Nepal. Vet Parasitol. 2009; 160(3–4):295–300.
53. Read C, Walters J, Robertson ID, Thompson RCA. Correlation between genotype of Giardia duodenalis and diarrhoea. Int J Parasitol. 2002;32(2): 229–31.
54. Kohli A, et al. Giardia duodenalis assemblage, clinical presentation and markers of intestinal inflammation in Brazilian children. 2008. p. 718–25.
55. Homan WL, Mank TG. Human giardiasis: genotype linked differences in clinical symptomatology. 2001;31:822–6. https://www.ncbi.nlm.nih.gov/pubmed/11403774.
56. Enrique L, et al. Molecular analysis of Giardia duodenalis isolates from symptomatic and asymptomatic children from La Habana, Cuba. Parasite Epidemiol. Control. 2017;2(3):105–13.
57. Broglia A, Weitzel T, Harms G, Cacció SM, Nöckler K. Molecular typing of Giardia duodenalis isolates from German travellers. Parasitol Res. 2013;10: 3449–56.
58. Wang R, et al. Cryptosporidium parvum IId family: clonal population and dispersal from Western Asia to other geographical regions. Sci Rep. 2014;4:1–5.
59. Londoño Á, Mejía S, Gómez-Marín J. Prevalencia y Factores de Riesgo Asociados a Parasitismo Intestinal en Preescolares de Zona Urbana en Calarcá, Colombia. Rev salud pública. 2009;11(1):72–81.
60. Gaviria LM, et al. Prevalencia de parasitosis intestinal, anemia y desnutrición en niños de un resguardo indígena Nasa, Cauca, Colombia. Rev Fac Nac Salud Pública, Medellín. 2015;35(3):390–9.
61. Rinaldo MDPJ. Parasitosis intestinal en preescolares y escolares atendidos en el centro médico EsSalud de Celendin, Cajamarca. Horiz. Med. 2016;16(3):35–42.

Molecular detection and genotyping of intestinal protozoa from different biogeographical regions of Colombia

Adriana Higuera[1], Ximena Villamizar[1], Giovanny Herrera[1], Julio Cesar Giraldo[2], Luis Reinel Vasquez-A[3], Plutarco Urbano[4], Oswaldo Villalobos[5], Catalina Tovar[6] and Juan David Ramírez[1]

[1] Grupo de Investigaciones Microbiológicas-UR (GIMUR), Departamento de Biología, Facultad de Ciencias Naturales, Universidad del Rosario, Bogota, Colombia
[2] Programa de Biología, Universidad INCCA de Colombia, Bogota, Colombia
[3] Centro de Estudios en Microbiología y Parasitología (CEMPA), Departamento de Medicina Interna, Facultad de Ciencias de la Salud, Universidad del Cauca, Popayan, Colombia
[4] Grupo de Investigaciones Biológicas de la Orinoquia, Unitrópico, Yopal, Colombia
[5] Hospital Local Santa María de Mompox, Programas Especiales (Lepra y TB), Mompox, Bolivar, Colombia
[6] Grupo de Enfermedades Tropicales y Resistencia Bacteriana, Universidad del Sinú, Monteria, Colombia

ABSTRACT

Background: Intestinal parasitic protozoa represent a serious problem of public health particularly in developing countries. Protozoa such as *Blastocystis, Giardia intestinalis*, *Entamoeba histolytica* and *Cryptosporidium* spp. are associated with diarrheal symptoms. In Colombia, there is little region-specific data on the frequency and circulating genotypes/species of these microorganisms. Therefore, the main objective of our study was to employ molecular detection and genotyping of *G. intestinalis* and *Blastocystis*, *Cryptosporidium* and *Entamoeba* spp. in samples from different biogeographical regions of Colombia.

Methods: We collected 649 human fecal samples from five biogeographical regions of Colombia: the Amazon, Andean, Caribbean, Orinoco and Pacific regions. *Blastocystis, G. intestinalis, Cryptosporidium* spp. and *Entamoeba* complex were detected by microscopy and conventional PCR. Molecular genotyping was conducted to identify *Blastocystis* subtypes (STs) (18s), *G. intestinalis* assemblages (triose phosphate isomerase and glutamate dehydrogenase) and *Cryptosporidium* species (18s). Genetic diversity indices were determined using dnasp.5.

Results: We detected *G. intestinalis* in 45.4% (n = 280) of samples, *Blastocystis* in 54.5% (n = 336) of samples, *Cryptosporidium* spp. in 7.3% (n = 45) of samples, *Entamoeba dispar* in 1.5% (n = 9) of samples, and *Entamoeba moshkovskii* in 0.32% (n = 2) of samples. *Blastocystis* STs 1–4, 8 and 9 and *G. intestinalis* assemblages AII, BIII, BIV, D and G were identified. The following *Cryptosporidium* species were identified: *C. hominis*, *C. parvum*, *C. bovis*, *C. andersoni*, *C. muris*, *C. ubiquitum* and *C. felis*. The Caribbean region had the highest frequency for each of the microorganisms evaluated (91.9% for *G. duodenalis*, 97.3% for *Blastocystis*, 10.8% for *Cryptosporidium* spp., 13.5% for *E. dispar* and 2.7% for *E. moshkovskii*). The Orinoco region had a high frequency of *Blastocystis* (97.2%) and the Andean region had a high frequency of *G. intestinalis* (69.4%). High and active transmission

Submitted 8 October 2019
Accepted 13 January 2020
Published 9 March 2020

Corresponding author
Juan David Ramírez,
juand.ramirez@urosario.edu.co

Academic editor
Erika Braga

Additional Information and Declarations can be found on page 18

DOI **10.7717/peerj.8554**

How to cite this article Higuera A, Villamizar X, Herrera G, Giraldo JC, Vasquez-A LR, Urbano P, Villalobos O, Tovar C, Ramírez JD. 2020. Molecular detection and genotyping of intestinal protozoa from different biogeographical regions of Colombia. **PeerJ 8:e8554 DOI 10.7717/peerj.8554**

was apparent in several regions of the country, implying that mechanisms for prevention and control of intestinal parasitosis in different parts of the country must be improved.

Subjects Microbiology, Parasitology, Infectious Diseases
Keywords *Giardia intestinalis*, *Cryptosporidium*, *Blastocystis*, *Entamoeba*, Molecular genotyping

INTRODUCTION

Infectious diseases are major public health challenges worldwide. Despite efforts to reduce human morbidity and mortality, shortcomings in prevention and control measures continue to impact the continued transmission of pathogens in the human population, preventing the management of some diseases in endemic areas (*Morens, Folkers & Fauci, 2004*). Infectious diseases caused by intestinal parasites have a wide distribution worldwide. In 2001, approximately 3,500 million people were infected by protozoa and intestinal helminths where the children were the most affected (*MinSalud, 2015*) by protozoal infections. Members of the genus *Blastocystis* are the most common eukaryotic microorganisms in the human and animal intestine (*Stensvold, Alfellani & Clark, 2012*), followed by *Giardia intestinalis* (synonyms: *G. duodenalis* and *G. lamblia*) and various *Cryptosporidium* and *Entamoeba* species (*Haque, 2007*). Together, these are the main protozoan causative agents of diarrheal disease in humans worldwide (*Caccio & Ryan, 2008*; *Haque, 2007*; *Jacobsen et al., 2007*).

Worldwide, approximately 200 million individuals are infected by *Giardia* species, while the frequency of *Cryptosporidium* infection ranges from 0.1% to 10% in developed and developing countries, respectively (*WHO, 2010*). The frequency of amebiasis, caused mainly by *Entamoeba histolytica*, is often reported as near 20%, but can vary greatly depending on the region and the techniques used to differentiate the *E. histolytica*/*dispar*/*moshkovskii* complex (*Silva et al., 2014*; *Tasawar, Kausar & Lashari, 2010*). The frequency of *Blastocystis* ranges between 0.5% and 24% in industrialized countries and between 30% and 76% in developing countries (*Wawrzyniak et al., 2013*). However, other studies have identified populations of children where the frequency of *Blastocystis* approaches 100% (*El Safadi et al., 2014*). In Colombia, the latest national survey by the Ministry of Health revealed that *Blastocystis* was the most commonly identified protozoa in human feces, with a nationwide frequency of 52%. *Blastocystis* were followed by *Entamoeba* (17%), *Giardia* (15%) and *Cryptosporidium* (0.5%) spp. (*MinSalud, 2015*).

Molecular tools have been developed to assess the genetic diversity of protozoan parasites at the intra-species level. In the case of *G. intestinalis*, eight genotypes or assemblages (A–H) have been identified and are distributed worldwide. Within these assemblages, sub-assemblages (AI–AIII and BIII–BIV) have been established (*Faria et al., 2017*; *Lasek-Nesselquist, Mark Welch & Sogin, 2010*; *Ryan & Cacciò, 2013*). In Latin America, similar frequencies of assemblages A and B were observed in Brazil (*Coronato Nunes et al., 2016*) and Cuba (*Pelayo et al., 2008*), while the frequencies of the AI, AII, AIII,

BIII and BIV sub-assemblages varied in Brazil, Argentina, Peru, Colombia and Mexico (*Coronato Nunes et al., 2016*; *Minvielle et al., 2008*; *Molina et al., 2011*; *Perez Cordon et al., 2008*; *Sánchez et al., 2017*; *Torres-Romero et al., 2014*). On the other hand, members of the genus *Blastocystis* can be classified into 17 subtypes (STs) (*Stensvold et al., 2007*) based on polymorphisms of 18S rDNA (*Scicluna, Tawari & Clark, 2006*). In humans, STs 1–3 are common in both Europe and South America (*Del Coco et al., 2017*; *Malheiros et al., 2011*; *Ramírez et al., 2016*; *Santin et al., 2011*), while ST4 is commonly found in Europe (*Stensvold et al., 2011*; *Wawrzyniak et al., 2013*) and was possibly associated with an enzootic cycle in nonhuman primates in Latin America (*Ramírez et al., 2014*; *Santin et al., 2011*). Approximately 20 different species have been identified within the genus *Cryptosporidium*, where *Cryptosporidium hominis* and *Cryptosporidium parvum* are the most common pathogens infecting humans (*Feng & Xiao, 2017*). Two markers (the small subunit of the ssuRNA and gp60) have been used to discriminate species and STs (*Khan, Shaik & Grigg, 2018*). Ten STs of *C. hominis* (Ia–Ik) and 16 STs of *C. parvum* (IIA–IIp) have been described (*Garcia-R et al., 2017*; *Xiao, 2010*). In Colombia, infections by *Cryptosporidium viatorum* (*Sánchez et al., 2017*) *Cryptosporidium galli* and *Cryptosporidium molnari* have been reported (*Sánchez et al., 2018*). Lastly, within the genus *Entamoeba*, the only pathogenic species is *E. histolytica*. However, the morphological similarities between the three species of the *E. histolytica*/*moshkovskii*/*dispar* complex (*Pritt & Clark, 2008*) make molecular tools required for species identification (*Ximénez et al., 2009*). A study in Colombia, in children under 16 years old, found a frequency of infections of 49.1%, being *E. dispar* the most frequently detected and *E. moshkovskii* also reported (*López et al., 2015*).

Colombia has a wide variety of climates and biogeographical regions classified according to epidemiological features. For instance, the biogeographical regions of Colombia are characterized by different climatic and ecosystem conditions, ranging from temperate zones to permanent snow in the mountain peaks. Moreover, the number of inhabitants and economic activities are increasing and the availability of resources are decreasing in some of these regions (*IDEAM et al., 2007*; *Barón, 2002*) affecting the ecological niches where some pathogens could be circulating. Also, variation in socioeconomic conditions may be associated with behavioral factors that drive the transmission of microorganisms through contact with animals from both urban and rural areas, as well as the consumption of food and water under inadequate sanitary conditions. All these features make Colombia a country where the transmission of intestinal microorganisms is very likely (*MinSalud, 2015*). For this reason, it is mandatory to establish intervention programs to know what protozoa are being transmitted, including their biological and molecular characteristics to improve control and prevention plans. Therefore, the main objective of this study was to conduct molecular detection and genotyping of *Giardia, Blastocystis, Cryptosporidium* and *Entamoeba* species from samples collected in different biogeographical regions of Colombia. We also compared the concordance results by PCR and microscopy in the analyzed samples.

METHODS

Ethics approval and consent to participate

This study was a minimum risk investigation for participants. Both the ethical standards of the Colombian Ministry of Health (Youth Code) and the Helsinki Declaration of 2013 were followed. The parents or guardians of minors participating in the study signed informed consent forms and gave their permission to obtain samples. This study was approved by the research ethics committee of the Universidad del Rosario (registered in Act No. 394 of the CEI-UR), the ethics committee of the Department of Internal Medicine of the Universidad del Cauca (number VRI024/2016), and the INCCA University of Colombia (number 237894).

Study area

Colombia is a country with significant geographical, ethnic, cultural and socioeconomic diversity. Based on climatic, territorial and ecosystem diversity, the country is subdivided into six natural regions: Insular, Caribbean, Amazon, Andean, Orinoco and Pacific. These regions are not precisely geographical but coincide with clusters that are the recipients of government budget funds. Except for the Insular region, all regions were included in this study.

The Caribbean region is located in the north of the country and includes seven departments. In the department of Córdoba, eight samples were obtained from Montería city and in the department of Bolívar, 30 samples were collected in the city of Mompós. This coastal region has some mountainous areas, contains both tropical and dry forest ecosystems, and is strongly influenced by the presence of bodies of water. The Amazon region has the smallest human population but the greatest diversity of flora and fauna. Its biome mainly comprises tropical forest and is characterized by warm weather and abundant precipitation. The human population is primarily indigenous. The Amazon region is located in southern Colombia and consists of six departments (*Rangel-Ch & Aguilar, 1995*). Fifty samples each were obtained from the departments of Guainía and Amazonas. The participating municipalities were Caño Conejo, Coco Nuevo, Coco Viejo and Puerto Inírida in Guainía and the cities of Leticia and Puerto Nariño in Amazonas.

The Andean region is the most populous in the country, housing half of the Colombian population. This region comprises the northern zone of the Andes and contains three mountain ranges that contribute to high climatic variability because of the different altitudes found within the region. The Andean region contains the departments of Antioquia, Boyacá and Cundinamarca; 50 samples were collected from each of these departments. The municipalities or cities contributing samples were Medellín, Río Negro, Amalfi, Bello, Caldas and El Santuario (Antioquia); Paipa, Villa de Leyva, Arcabuco, Tuta and Tunja (Boyacá) and the municipalities of Chaguaní, San Francisco, Fómeque, Soacha and the city of Bogotá (Cundinamarca). Other departments, including Quindío, Risaralda, Caldas and part of Tolima, make up the Coffee Axis. Another 50 stool samples were obtained from inhabitants of the municipalities of Calarcá, Armenia, Pereira, Córdoba and the Corregimiento Barcelona, all located within the Coffee Axis.

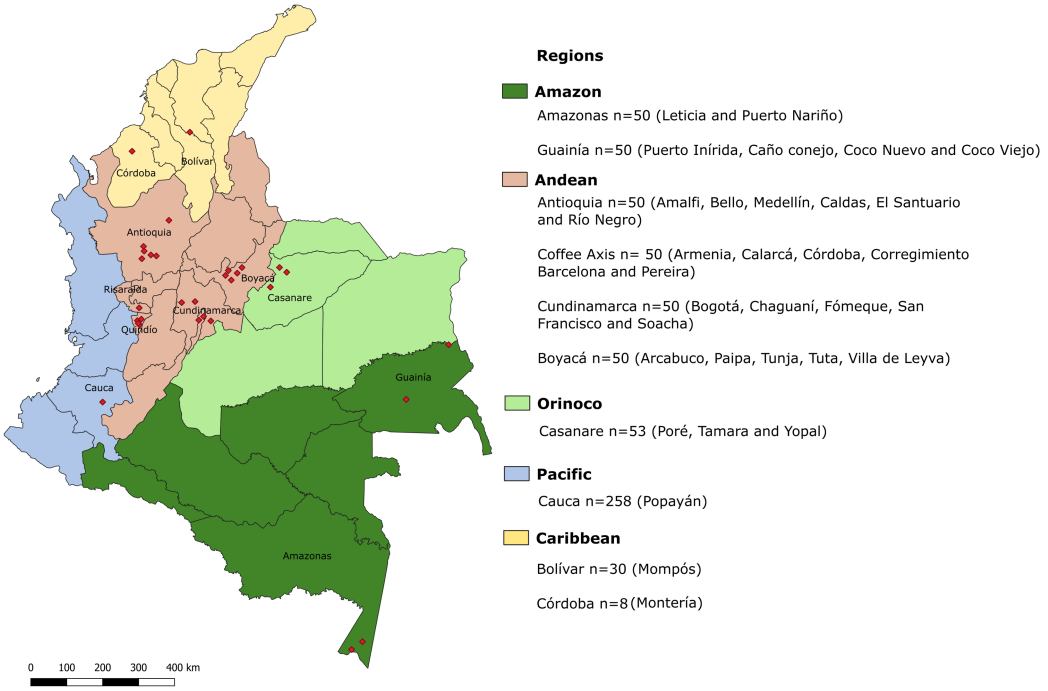


Figure 1 Geographic locations of regions in which samples were collected. Biogeographical regions of Colombia are indicated in colors. Each region is divided into departments. Red diamonds indicate the exact locations of sampling areas. In the legend, the departments sampled in each biogeographical region are indicated, along with the total number of samples for each department and the cities or municipalities from which the samples were obtained. Full-size DOI: 10.7717/peerj.8554/fig-1

The Orinoco contains a large number of rivers, warm ecosystems and tropical and subtropical forests. This region is sparsely populated and comprises four departments. One of these is Casanare, where we collected 53 samples from the municipalities of Poré, Yopal and Tamara.

Finally, the Pacific region, one of the wettest in the world, is characterized by tropical forest and high species diversity. Although it is naturally resource rich, the region has poor urban development and infrastructure. The Pacific region contains four departments, one of these is the Cauca department, where 258 samples were collected from inhabitants of commune 8 in the city of Popayán. These samples were collected within the study of Villamizar and colleagues (*Villamizar et al., 2019*), and are, in turn, part of the current study.

Study population

In total, 649 stool samples were collected from adults and children in different biogeographical regions of Colombia (Fig. 1). Convenience sampling was conducted to obtain samples from five different regions. The average age was 5 years (standard deviation: 6 years; range: 1–70 years). Microscopy is the gold standard used in Colombia to detect intestinal parasites, then most of the samples were evaluated by this diagnostic scheme following the protocol by Villamizar and collaborators (*Villamizar et al., 2019*), except for samples from Casanare, Bolívar and Córdoba, which could be only tested by

PCR, since the entire sample portion was preserved in ethanol 100%. The individuals included in the study lived in both rural and urban areas of different municipalities/cities. The percentages of samples obtained in each biogeographical region were: Amazon (15.4%, *n* = 100), Andean (30.8%, *n* = 200), Caribbean (5.9%, *n* = 38), Orinoquía (8.2%, *n* = 53) and Pacific (39.8%, *n* = 258). The majority (86%, *n* = 558) of samples were assessed for the presence of intestinal protozoa by microscopy. All samples were subjected to molecular detection of intestinal protozoa and in those that were positive, further molecular characterization was conducted.

DNA extraction

Prior to DNA extraction, approximately 300 µL of each sample was washed with sterile phosphate-buffered saline. Genomic DNA was extracted from stool samples using the Norgen Stool DNA Isolation Kit, Norgen Biotek Corp., following the recommendations of the manufacturer. During the lysis step, 10 µL of a recombinant plasmid, pZErO-2, was added (final concentration: 100 pg/µL). This plasmid contained the *Arabidopsis thaliana* aquaporin gene as an internal control for heterologous extrinsic amplification (*Duffy et al., 2013*).

Conventional PCR

Initially, an internal amplification control (IAC) PCR was performed to verify that there was no inhibition of this technique using stool samples as template. All samples were subjected to this amplification control, except those from Cauca (The samples were collected and extracted directly in Popayan and therefore IAC was not added to the sample). Overall, 358 (91.6%) samples were validated with a positive amplification for the IAC, and were subjected to PCR to detect *G. intestinalis*, *Blastocystis*, *Cryptosporidium* and *Entamoeba* complex DNA. On the other hand, 33 (8.4%) samples that were negative for the IAC were discarded. IAC PCRs and molecular detection for *Giardia*, *Blastocystis* and *Cryptosporidium* spp., were performed in a final volume of 9 µL containing 3.5 µL of GoTaq Green Master Mix (Promega), 2 µL of template DNA, and primers. For IAC PCRs, primers were used at a final concentration of 0.4 µM. Species-specific primers were used at a final concentration of 1 µM. The sequences of the primers used for IAC were: IACFw (5′-ACCGTCATGGAACAGCACGTA-3′) and IACRv (5′-CTCCC GCAACAAACCCTATAAAT-3′) (*Duffy et al., 2013*). For species-specific PCR, the primers used were: *G. intestinalis*, forward (5′-CATGCATGCCCGCTCA-3′) and reverse (5′-AGCGGTGTCCGGCTAGC-3′) (*Mejia et al., 2013*); *Blastocystis*, FWD F5 (5′-GGTCC GGTGAACACTTTGGATTT-3′) and R F2 (5′-CCTACGGAAACCTTGTTACGACTTC A-3′) (*Stensvold et al., 2012*); and *Cryptosporidium*, CcF18s (5′-GTTTTCATTAATC AAGAACGAAAGTTAGG-3′) and CcR18s (5′-GAGTAAGGAACAACCTCCAATCTC TAG-3′) (*Burnet et al., 2013*). The thermal cycling parameters were as follows: 95 °C for 5 min; 40 cycles of 95 °C for 15 s, 58 °C for 1 min and 72 °C for 30 s; 72 °C for 10 min. Each PCR was performed individually.

For the *Entamoeba* complex, a conventional multiplex PCR was performed using previously reported conditions and primers EntF (5′-ATGCACGAGAGCGAAAG

CAT-3′), EhR (5′-GATCTAGAAACAATGCTTCTCT-3′), EdR (5′-CACCACTTACTA TCCCTACC-3′) and EmR (5′-TGACCGGAGCCAGAGACAT-3′) (*Mahmoudi, Nazemalhosseini-Mojarad & Karanis, 2015*). Differentiation between *E. histolytica, E. dispar* and *E. moshkovskii* was based on the size of the amplicon using these primers. DNA extracted from axenic cultures of each protozoan provided by The University of Texas Medical Branch were used as positive controls.

Genotyping of *Giardia, Blastocystis* and *Cryptosporidium* spp.

Samples showing positive PCR amplification for *Giardia*, *Blastocystis* and *Cryptosporidium* spp. were subjected to conventional PCR to determine the assemblages, STs and species for each protozoan. For *Giardia*, two loci were amplified: glutamate dehydrogenase (*gdh*) and triose phosphate isomerase (*tpi*). To amplify *gdh*, primers GDHeF (5′-TCAA CGTYAAYCGYGGYTTCCGT-3′), GDHiF (5′-CAGTACAACTCYGCTCTCGG-3′) and GDHiR (5′-GTTRTCCTTGCACATCTCC-3′) were used (*Read, Monis & Thompson, 2004*). To amplify *tpi*, primers Al3543 (5′-AAATIATGCCTGCTCGTCG-3′), Al3546 (5′-CAAACCTTITCCGCAAACC-3′), Al3544 (5′-CCCTTCATCGGIGGTAAATT-3′) and Al3545 (5′-GTGGCCACCACICCC-3′) were used (*Sulaiman et al., 2003*). *Blastocystis* STs and alleles were determined by amplifying a region of ssuRNAr using primers BhRDr (5′-GAGCTTTTTAACTGCAACAACG-3′) and RD5 (5′-ATCTGGTTGA TCCTGCCAGT-3′) as reported previously (*Scicluna, Tawari & Clark, 2006*). To identify *Cryptosporidium* spp., a rRNA region was amplified using primers SSUrRNAF (5′-AGT GACAAGAAATAACAATACAGG-3′) and SSUrRNAR (5′-CCTGCTTTAAGCACTCT AATTTTC-3′) as described previously (*Hunter et al., 2007*).

Once all PCRs were performed, the size of each amplicon was assessed using 2% agarose gel electrophoresis followed by staining with SYBR Safe. Subsequently, each product was purified with ExoSAP-IT® following the manufacturer's recommendations. Both strands of each amplicon were sequenced using the Sanger method by Macrogen (Seoul, South Korea). Sequences were edited in MEGA 7.0 (*Kumar, Stecher & Tamura, 2016*) and compared with publicly available sequences using BLAST to verify that they corresponded to the expected taxonomic unit.

Blastocystis sequences were submitted to a database to identify STs and alleles (https://pubmlst.org/blastocystis/) (*Jolley & Maiden, 2010*). For *Giardia*, a multiple sequence alignment, including reference sequences for *gdh* and *tpi*, was performed using MUSCLE (*Edgar, 2004*) implemented in MEGA 7.0. A phylogenetic tree was constructed using maximum likelihood methods and 1,000 bootstrap replicates to determine assemblages. The accession numbers of the GenBank reference sequences used for *gdh* were: AI (M84604.1), AII (AY178737.1), AIII (EU637582.1), BIII (AF069059.1), BIV (AY178739.1), C (U60982.2), D (U60986.2), E (AY178741.1), F (AB569384.1), G (AF069058.2) and H (GU176089.1). Reference sequences contained in GenBank with the following accession numbers for *tpi* were used: AI (AF069556 .1), AII (AF069557.1), AIII (DQ650648.1), BIII (AF069561.1), BIV (AF069560.1), C (AF069563.1), D (DQ246216.1), E (AY228645. 1), F (AF069558.1) and G (EU781013.1). As outgroup, reference sequences of *G. ardeae* (AF069060.2) for *gdh* and *G. microti* (AY228649.1) for *tpi*

were used. *Cryptosporidium* species were determined by comparing target sequences with sequences available in GenBank.

Indices of genetic diversity

To assess the degree of DNA polymorphism, we constructed a multiple alignment of concatenated sequences for each of the loci evaluated for both *G. intestinalis* and *Blastocystis* using MAFFT v7. For the *gdh* and *tpi* loci of *G. intestinalis*, we analyzed 30 (295 sites including gaps) and 25 (465 sites including gaps) sequences, respectively. In case of *Blastocystis* with the 18s gene, we analyzed 114 (1,635 sites including gaps) sequences. All these sequences were used to calculate the indices of diversity (π and Θ), number of polymorphic (segregating) sites (S), number of haplotypes (h), and the haplotype diversity by department. DnaSP v5 software was used for these analyses.

Statistical analysis

Data were summarized using univariate statistics in Stata 14 (StataCorp, 2015, Stata Statistical Software: Release 14). Subsequently, Cohen's kappa indices were calculated to assess agreement between the results of microscopy and molecular techniques, both globally and for each of the parasites individually.

RESULTS

Sample description and detection of protozoa

The ages of individuals from which samples were collected ranged between 1 and 70 years (average, 4.8 years; standard deviation, 5.5 years). The largest number of samples (39.8%) were collected in the Pacific region (Department of Cauca), while 30.8% were collected in the Andean region (Departments of Antioquia, Boyacá, Cundinamarca and the Coffee Axis). The majority (74.9%) of samples came from rural areas.

Comparison of protozoan detection by microscopy and PCR

The majority of samples were positive by microscopy (68.3%) and molecular methods (71.2%). The frequency of positive samples by PCR (n = 616) *vs* microscopy (n = 649) was calculated for each protozoan: *G. intestinalis* (PCR 41.1% *vs* microscopy 24.5%), *Blastocystis* (PCR 49.0% *vs* microscopy 33.6%), *Cryptosporidium* (PCR 5.6% *vs* microscopy 27.3%), and the *Entamoeba* complex (PCR 22.9% *vs* microscopy 0.2%). The concordance between direct microscopy and by conventional PCR was analyzed both globally and for each protozoan. In all cases, a low concordance between detection techniques was observed, with kappa indices of 0.3807 for detection of all protozoa and 0.2699, 0.1478, 0.0149 and −0.0036 for *G. intestinalis*, *Blastocystis*, *Cryptosporidium* spp. and the *Entamoeba* complex, respectively.

Frequency of protozoa

In total, 558 samples were assessed by microscopy. Of these 558 samples, 25.4% were assessed for *G. intestinalis* (n = 142), 33.9% for *Blastocystis* (n = 189), 23.1% for members of the *Entamoeba* complex (n = 129), 27.2% for *Entamoeba coli* (n = 152), 15.9% for *Cryptosporidium* spp. (n = 89), 12.3% for *Cyclospora* spp. (n = 69), 5.7% for geohelminths

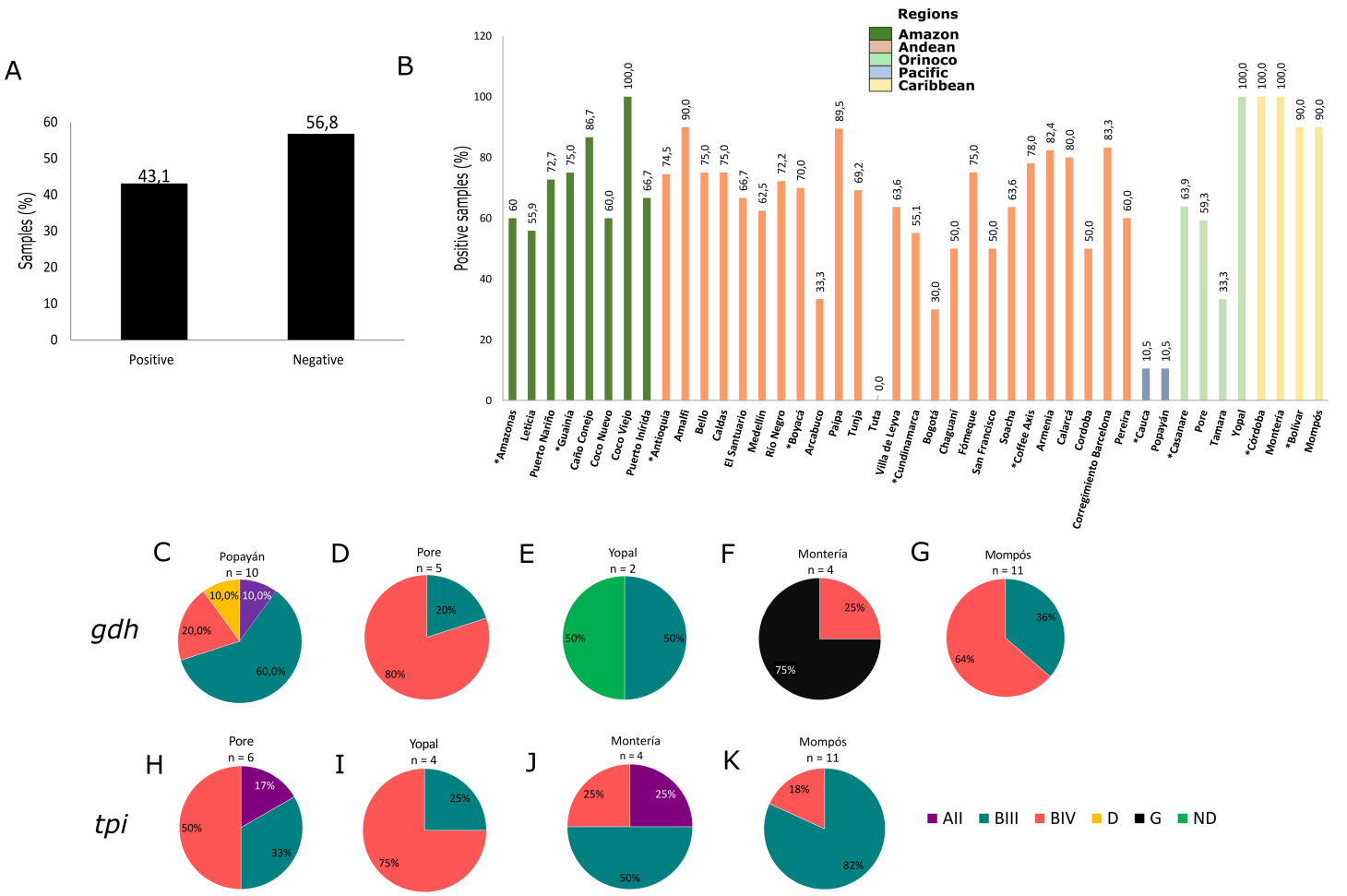


Figure 2 **Frequency and assemblages for *G. intestinalis*.** (A) Total percentage of positive and negative samples for *G. intestinalis*. (B) Frequency of positive samples for *G. intestinalis* by region. Departments indicated by * are highlighted. Colors indicate the biogeographical regions to which the sampled areas belong. (C–G) Frequencies of assemblages obtained in selected cities of some regions for the *gdh* locus. (H–K) Frequencies of assemblages obtained in selected cities of some regions for the *tpi* locus. Full-size DOI: 10.7717/peerj.8554/fig-2

such as *Strongyloides stercoralis* (n = 32), 5.0% for *Ascaris lumbricoides* (n = 28), 4.3% for *Trichuris trichiura* (n = 24) and 4.6% for *Uncinaria* (n = 26).

Giardia intestinalis

Using molecular detection by PCR, 43.1% of samples tested positive for *G. intestinalis* (Fig. 2A). The Caribbean region showed the highest frequency at 89.5% (95% CI [83.1–100.7]), followed by the Andean region (mainly the Department of Antioquia and the Coffee Axis) and the Amazon region (municipalities of Coco Viejo and Caño Conejo, Department of Guainía) in which the majority of sampled areas had a frequency greater than 60% (Table 1; Fig. 2B). Certain areas such as Yopal of Casanare and Paipa of Boyacá were distinguished by their extremely high *G. intestinalis* frequency (89.5% and 100%, respectively). From positive samples, assemblages were identified using the *gdh* marker for 33 samples as follows: AII (3.0%), BIII (36.3%), BIV (48.8%), D (3.0%) and G (9.1%). Sub-assemblage BIV was the most frequent, mainly in the municipality of Poré

Table 1 Prevalence of protozoa assessed by PCR in each biogeographical region.

Regions	*n*	*G. intestinalis*	*Blastocystis*	*Cryptosporidium* spp.	*E. dispar*	*E. moshkovskii*
Amazon	100	60.0%	49.0%	5.0%	0%	0%
		(*n* = 60)	(*n* = 49)	(*n* = 5)		
		95% CI [56.2–74.9]	95% CI [48.7–68.1]	95% CI [1.1–10.2]		
Andean	200	68.0%	58.0%	5.5%	0%	0%
		(*n* = 136)	(*n* = 116)	(*n* = 11)		
		95% CI [64.8–77.4]	95% CI [53.6–67.2]	95% CI [2.03–8.1]		
Caribbean	38	89.5%	94.7%	10.5%	13.1%	2.6%
		(*n* = 34)	(*n* = 36)	(*n* = 4)	(*n* = 5)	(*n* = 1)
		95% CI [83.1–100.69]	95% CI [92.1–102.5]	95% CI [0.8–20.8]	95% CI [2.5–24.5]	95% CI [−1.9 to 12.7]
Orinoco	53	43.4%	66.0%	0%	7.5%	0%
		(*n* = 23)	(*n* = 35)		(*n* = 4)	
		95% CI [29.9–56.1]	95% CI [61.5–70.5]		95% CI [−1.6 to 15.6]	
Pacific	258	10.5%	38.76%	9.70%	0%	0.39%
		(*n* = 27)	(*n* = 100)	(*n* = 25)		(*n* = 1)
		95% CI [6.8–14.3]	95% CI [33.2–45.2]	95% CI [6.2–13.4]		95% CI [−0.4 to 1.2]

Note:
n, number of samples; 95% CI, 95% confidence interval.

(Casanare), followed by sub-assemblage BIII with significant frequency in the cities of Yopal (Casanare) and Popayán (Cauca); the lattermost city had the greatest variety of assemblages. In the city of Monteria, a high frequency of assemblage G (75%) was observed. Twenty-five samples were genotyped using the *tpi* marker, and the AII, BIII and BIV assemblages were detected at frequencies of 8%, 56% and 36%, respectively. The highest frequency (82%) was observed for the BIII sub-assemblage in Mompós, followed by the BIV sub-assemblage (75%) in the city of Yopal (Figs. 2C–2K). From one sample collected in the city of Yopal, we found an inconsistency with the assigned assemblages using different markers. In the case of *gdh* marker, this sample clustered between AI and AII sub-assemblages, and could not be determined its assemblage with *gdh*, but with *tpi* marker this sample clustered with the BIV sub-assemblage.

Blastocystis

Of the samples analyzed, 51.8% were positive for *Blastocystis* by PCR (Fig. 3A). The Caribbean region showed the highest frequency of *Blastocystis* at 94.7% (95% CI [92.1–102.5]), followed by the Orinoco and the Andean regions (Table 1). In the Coffee Axis, the Corregimiento region of Barcelona, Amalfi in the Department of Antioquia and Caño Conejo in the department of Guainía, high frequency rates of 100%, 100% and 90% were observed, respectively (Fig. 3B). A total of 116 samples were subtyped from Popayán, Poré, Tamara, Yopal, Montería and Mompós. STs 1 (41.4%, *n* = 48), 2 (18.1%, *n* = 21), 3 (36.2%, *n* = 42), 4 (2.6%, *n* = 3), 8 (0.9%, *n* = 1) and 9 (0.9%, *n* = 1) were detected. The Poré (Casanare) and Popayán (Cauca) regions showed the greatest diversity of STs (Figs. 3C–3H). Greater allelic diversity was found within ST3 and the most frequently observed allele in Cauca was 34, followed by allele 57 and then other alleles including

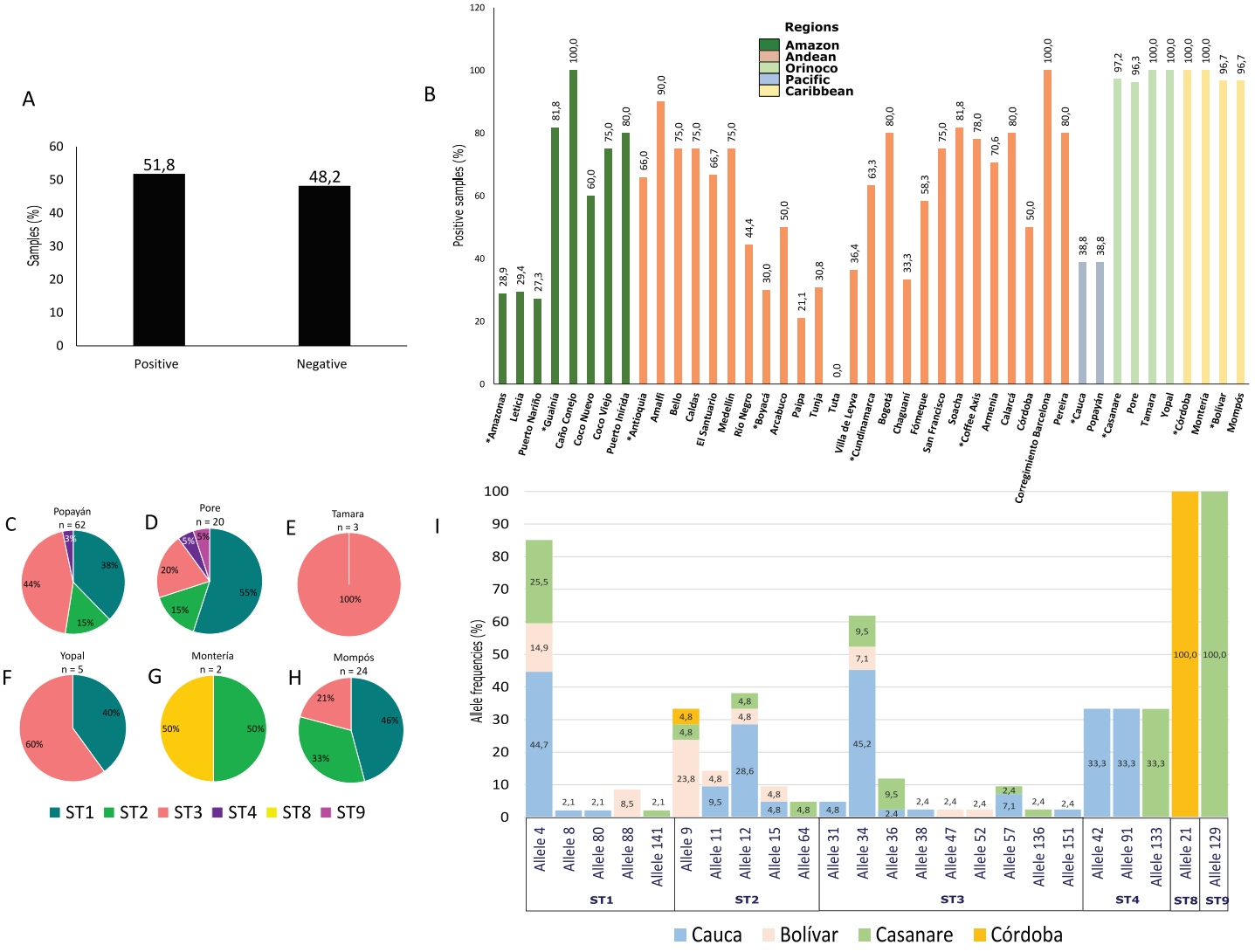


Figure 3 **Frequency, STs and alleles for *Blastocystis*.** (A) Total percentage of positive and negative samples for *Blastocystis*. (B) Frequency of positive samples for *Blastocystis* by region. Departments indicated by * are highlighted. Colors indicate the biogeographical regions to which the sampled areas belong. (C–H) Frequencies of STs obtained in six selected cities from Pacific, Orinoco and Caribbean regions for the 18s locus. The raw data is supplied. (I) Allele frequencies by ST and by department. Full-size DOI: 10.7717/peerj.8554/fig-3

31, 36, 38 and 151. For ST1, the most frequent allele was 4 in Cauca, Bolívar and Casanare, although alleles 8, 80, 88 and 141 were also observed. For ST2 the most frequent alleles were 9 in Bolívar and 12 in Cauca, although alleles 11, 15 and 64 were also detected. Noteworthy findings included the detection of alleles 42 and 91 in Cauca, ST4 allele 133 in Casanare, ST8 allele 21 in Córdoba and ST9 allele 129 in Casanare (Fig. 3I). The lattermost finding represents the first report of this allele in human samples from Colombia.

Diversity indices

Genetic diversity indices by department were calculated based on these alignments. For *G. intestinalis*, the number of segregating (polymorphic) sites (*S*) was 202 for *gdh* and

Table 2 Genetic diversity indices of *G. intestinalis* and *Blastocystis* by department.

	Locus	Parameter	Bolívar	Casanare	Córdoba	Cauca	Total
G. intestinalis	gdh	*n*	11	7	4	8	30
		π	0.08719	0.16324	0.31284	0.05862	0.18028
		Θ	0.14248	0.20964	0.34426	0.0736	0.3224
		S	95	121	141	47	202
		h	11	7	4	4	24
		Hd	0.891	1	1	0.75	0.977
		SD	0.039	0.076	0.177	0.139	0.017
	tpi	*n*	11	10	4	ND	25
		π	0.00689	0.04833	0.11351	ND	0.03922
		Θ	0.01228	0.08307	0.12426	ND	0.07113
		S	15	95	93	ND	105
		h	5	9	4	ND	17
		Hd	0.618	0.978	1	ND	0.903
		SD	0.164	0.054	0.177	ND	0.054
Blastocystis	18s	*n*	24	28	2	60	114
		π	0.27703	0.14334	0.48703	0.08725	0.71859
		Θ	0.35406	0.27501	0.48703	0.21444	0.56508
		S	209	84	169	6	19
		h	24	22	2	5	90
		Hd	0.989	0.966	1	0.219	0.993
		SD	0.012	0.024	0.5	0.07	0.03

Note:
n, number of sequences; π, nucleotide diversity; Θ, theta (per site) from eta; *S*, number of segregating sites; *h*, number of haplotypes; Hd, haplotype diversity; SD, standard deviation; ND, not determined.

105 for *tpi*, with haplotypic diversities of 0.977 and 0.903, respectively. The nucleotide diversity indices π and Θ, as well as haplotypic diversity (Hd), were high for the population in Córdoba for both loci and in Casanare for *tpi*. The lowest diversity in *gdh* was found among sequences from Cauca. Unfortunately, *tpi* sequences from Cauca showed electropherograms of poor quality and were not analyzed.

For *Blastocystis* sequences, the departments of Bolívar and Córdoba showed a greater number of polymorphic (segregating) sites (*S*). In particular, the Bolívar sequences showed the highest number of haplotypes (24), with a haplotypic diversity of 0.989 and higher nucleotide diversity indices compared with Casanare and Cauca. The latter had the lowest sequence diversity (Table 2).

Cryptosporidium and *Entamoeba* spp.

From the samples tested using PCR, 6.9% (*n* = 45) were positive for *Cryptosporidium* spp. Frequency was higher in the Caribbean (10.5%) and Pacific (9.7%) regions of Colombia (Table 1). We were able to identify species from 15 samples: 25% (*n* = 4) corresponded to *C. parvum*, 12.5% (*n* = 2) to *C. hominis*, 18.7% (*n* = 3) to *C. andersoni*, 25% (*n* = 4) to *C. muris* and 6.2% (*n* = 1) each to *C. ubiquitum* and *C. felis*. Samples where we identified *Cryptosporidium* spp. came from the cities of Popayán (*C. hominis* and

C. parvum), Leticia (*C. andersoni*), Paipa (*C. muris* and *C. andersoni*) and the Departments of Antioquia (*C muris*, *C. andersoni* and *C. ubiquitum*), Bolívar (*C. hominis* and *C. muris*) and Córdoba (*C. felis*). As for members of the *Entamoeba* complex, 1.4% (n = 9) of samples tested positive for *E. dispar* in the departments of Bolívar (Mompós) and Casanare (Poré and Yopal) and 0.5% (n = 3) tested positive for *E. moshkovskii* in the departments of Cauca (Popayán) and Bolívar (Mompós). No samples were positive for *E. histolytica* (Table 1). All sequences were deposited on GenBank under the accession numbers MN877659–MN877714.

DISCUSSION

Colombia is a privileged country with natural wealth, geographical variety and ecosystem diversity. However, the climatic conditions and location of the country, in addition to the unequal distribution of resources in different regions, give rise to some primarily rural areas with unfavorable socioeconomic conditions and inadequate sanitary conditions. These factors directly influence the transmission of parasitic diseases among the residents of a given region (*Ortiz, López & Rivas, 2012*). Another factor that plays a major role in transmission of infectious protozoa is age: children tend to be the most common hosts and adults are likely to be an important source of transmission to children (*Carvajal-Restrepo et al., 2019*). Age may be associated with susceptibility to infection due to age-dependent immunological conditions that favor colonization by protozoa as well as age-dependent malnutrition and behavioral factors that affect transmission (*Harhay, Horton & Olliaro, 2010*). Likewise, these factors can influence the transmission of helminths, which explains the finding of some of them in the samples evaluated by microscopy. It is important to clarify that although the objective of our study was not the detection of these geohelminths, we wanted to report them due to the great importance they have mainly in the child population, associated not only with immunological and malnutrition problems but also with growth and development (*Papier et al., 2014*). Our findings also support a high transmission rate of helminths in the country, which has severe implications in control programs across the country.

Our results showed a high frequency of intestinal protozoa present in different regions of the country. Using microscopic detection (with the exception of the Caribbean and Orinoco), we observed that the Andean region had the highest frequency of *G. intestinalis*, *Blastocystis*, members of the *Entamoeba* complex and *Cryptosporidium* spp. Using molecular tests, the region with the highest frequency of all protozoa evaluated was the Caribbean. There was a low concordance (kappa index = 0.38) between the two techniques evaluated (microscopy and PCR). However, it is not possible to assert that the Andean or Caribbean regions truly had higher frequency of these protozoa, as our study had an important selection bias: sampling was carried out at convenience, with a higher number of samples obtained in regions such as the Andean and Pacific regions. However, it is important to note that the Andean region concentrates the largest population in Colombia and the Pacific region is one of the rainiest areas in the world, with low economic conditions, inadequate health conditions and poor access to education which might to some extant explain our findings (*Barón, 2002*). Thus, future studies would be necessary to

collect a larger number of samples of comparable quantity in each region. Despite this, our results are in agreement with those of a survey conducted using microscopy by the Ministry of Health in 2015 (*MinSalud, 2015*). For example, using molecular tests, we observed that the areas with higher frequencies of these protozoa coincided with the Caribbean region and the Andean region in most cases, with the exception of *G. intestinalis*, because in the survey by the Ministry of Health, was observed at higher frequency in the Colombian Amazon.

In Colombia, most reports on protozoan pathogens have focused solely on microscopic detection (*Agudelo-Lopez et al., 2008*; *Carvajal-Restrepo et al., 2019*). Several studies have shown differences in detection rates using molecular tests, which allow identification of cryptic species and their genotypes in addition to detection (*Morgan et al., 1998*; *Stensvold et al., 2018*). Thus, there is clear value in using complementary techniques (*Beyhan & Taş Cengiz, 2017*; *Sri-Hidajati et al., 2018*; *Mateo et al., 2014*) for molecular epidemiological studies, which may help to better elucidate the transmission dynamics of microorganisms and to establish better prevention and control plans. Another advantage of using molecular techniques is their sensitivity in cases of polyparasitism (*Meurs et al., 2017*). Polyparasitism is an important factor in the transmission of parasitic diseases, and the presence of different infectious agents, including helminths and protozoa, may serve as an indicator of inadequate sanitary conditions, immune suppression, nutritional deficiencies and continual reinfection (*Supali et al., 2010*). In our study, 29.3% of samples evaluated were positive for both *Blastocystis* and *G. intestinalis*, 1.7% were positive for *Blastocystis*, *G. intestinalis* and *Cryptosporidium* spp., 3.4% were positive for *G. intestinalis* and *Cryptosporidium* spp., and 3.4% for *Blastocystis* and *Cryptosporidium* spp. The remaining co-infection combinations occurred at less than 1.4%. None of these combinations showed any geographical associations.

Few studies of these protozoan pathogens have been conducted in Colombia. Studies of samples from indigenous communities in the Amazon (*Sánchez et al., 2017*), from a rural region in La Vírgen (*Ramírez et al., 2015*), and from children in rural schools in the municipality of Apulo (*Hernández et al., 2019*), Cundinamarca, found *G. intestinalis* in human fecal samples. Sub-assemblages AI, AII, BIII and BIV and sub-assemblages AII, BIII and BIV were detected in the feces of children in nurseries of the Colombian Institute of Family Welfare. Assemblages C and D were detected in samples from dogs in Tolima (*Rodríguez et al., 2014*). These results are consistent with the detection of assemblages AII, BIII and BIV in the Orinoco, Pacific and Caribbean regions in our study, with the exception of the presence of assemblage D in the Pacific and assemblage G in the Caribbean (Figs. 2C–2K). As in other studies outside Colombia, we observed no restriction of assemblages to specific geographic regions (*Broglia et al., 2013*; *Feng & Xiao, 2011*). Assemblage D is typically associated with dogs, while assemblage G infects rodent including rats and mice (*Caccio & Ryan, 2008*). Thus, there is the potential for human infection by these assemblages in humans and they could potentially maintain an active cycle of transmission or generate transient infections, in humans (*Heyworth, 2016*). The association between these assemblages and the development of disease is not clear (*Sprong, Cacciò & Van Der Giessen, 2009*), but they may be acquired through the

consumption of untreated water in rural regions where potable drinking water systems are absent, allowing closer contact with animal feces and increasing the risk of zoonotic transmission (*Fantinatti et al., 2016*). In agreement with this, assemblage H was detected in a study of water supplied by treatment plants in Nariño (southwest Colombia) (*Sánchez et al., 2018*), suggesting that water or the feces of wild animals that have not been studied as possible reservoirs could explain the presence of these assemblages. It is also important to consider that in the Caribbean region consumption of exotic animals and animal products, including iguana eggs, small crocodiles, freshwater turtles and armadillos, is very common. The potential role of these foods in the transmission of infections is unknown.

Another protozoan detected with high frequency was *Blastocystis*, mainly in the Caribbean and Orinoco regions. Frequency rates were often above 80%, especially in some regions of the Amazon such as Caño Conejo, Puerto Inírida (Guainía), of the Andean region such as Amalfi (Antioquia), the city of Bogotá, Soacha (Cundinamarca) and Calarcá, Corregimiento Barcelona, Pereira (Coffee Axis) (Fig. 3B). These findings are in agreement with results obtained using microscopy in Colombia that showed significant frequency in Caribbean regions such as Santa Marta (62.6%) and in Andean areas such as Santander (25%), Bogotá (22.4%), Quindío (36.4%) and Cundinamarca (34.8%) (*Londono-Franco et al., 2014*). When performing subtyping of *Blastocystis*, we identified STs 1–4, 8 and 9, with STs 1–3 having the highest frequency as reported by Del Coco and collaborators in a review made in 2017 and another study in Brazil (*Del Coco et al., 2017*; *Malheiros et al., 2011*). The municipality of Poré, Casanare showed the greatest diversity of STs. Similarly, the lower proportion of ST4 observed in the Caribbean and in the Pacific coincides with previous reports suggesting that this subtype is of recent origin in humans from the Americas (*Stensvold et al., 2012*) and of ethnic origin in Colombia associated with the enzootic cycle (*Jiménez, Jaimes & Ramírez, 2019*; *Ramírez et al., 2014*; *Santin et al., 2011*). Surprisingly, one sample was positive for ST8 in the Caribbean and another for ST9 in Casanare in the Orinoco region; these STs are rarely detected in humans (*Stensvold & Clark, 2016*). This is the first report of ST9 in Colombia, previously, one study in Italy reported the presence of ST9 in samples from symptomatic humans (*Meloni et al., 2011*). However, more studies are required to evaluate the potential zoonotic origin of this ST and its relationship with the presence of symptoms (*Stensvold et al., 2009*). A previous study reported the presence of ST8 in Colombia in marsupial stool samples (*Ramírez et al., 2014*), while another study detected this ST in arboreal nonhuman primates in Asia and South America (*Alfellani et al., 2013*).
Few studies have reported the presence of this ST in humans, but it could apparently be involved in zoonotic transmission to humans (*Meloni et al., 2011*; *Stensvold et al., 2007*), where a great variety of animals could be involved in the transmission. This is because there is a great diversity of fauna and ecosystems in the country. For instance, in the Caribbean and Orinoco regions exist diverse ecosystems including savannah, mountainous forest, bodies of water, jungles and moorland (*Barón, 2002*; *IDEAM et al., 2007*; *Vergara, 2018*), these are exploited by each department to generate economic resources, and the presence there of nonhuman primates, rodents, birds and pigs infected with intestinal

protozoa could increase the risk of zoonotic transmission in rural areas as has been reported in the country and in Brazil (*Rondón et al., 2017*; *Valença-Barbosa et al., 2019*).

We also characterized the alleles of each of the STs. No geographical associations were observed for STs or alleles. Allele 4 of ST1 was detected in the regions of Cauca (44.7%), Casanare (25.5%) and Bolívar (14.9%) (Fig. 3I), and was the most frequently observed as previously reported (*Ramírez et al., 2014*; *Sánchez et al., 2017*). In addition, alleles 8, 80, 88 and 141 were found in ST1, alleles 9, 11, 12, 15 and 64 within ST2, and alleles 31, 34, 36 were detected in ST3. The presence of alleles 38, 47, 52, 57, 136 and 151 provided evidence of the great intra-subtype diversity present and mostly agreed with studies of STs circulating in Ecuador, Peru, Bolivia, Colombia, Brazil and Argentina in samples from humans, domestic animals and the enzootic cycle (*Ramirez et al., 2017*, *2014*, *2016*). For ST4, alleles 42, 91 and 133 were identified, which had previously been reported in Colombia; in particular, allele 91 that is possibly of European origin (*Ramírez et al., 2014*; *Stensvold et al., 2011*). For ST8 isolates reported in Colombia and Brazil (*Ramírez et al., 2016*), the 21 allele probably had a zoonotic origin. For ST9 we detected allele 129, of which there is no previous report in Colombia. The origin of this ST has not been established, and then it is not possible to make inferences about its transmission. As mentioned above, great diversity was present among the STs characterized for *Blastocystis*, and establishing the transmission dynamics for several of the STs detected at low frequency would be a useful task.

In addition to detecting protozoa and determining the frequencies of STs and assemblages, the genetic diversity among sequences of *G. intestinalis* and *Blastocystis* was evaluated by department and by marker. For *G. intestinalis*, diversity indexes were higher for assemblages from the Casanare and Córdoba departments for *gdh* (Table 2). However, the low number of sequences obtained from the Córdoba region means that the degree of diversity would need to be verified using more samples in a future study. For *Blastocystis*, greater diversity was found in the Caribbean region in the departments of Bolívar and Córdoba. This was expected since *Blastocystis* usually has high inter-subtype variability (*Stensvold et al., 2012*), as reported in a study of SSU DNAr genes conducted in Mexico, where the results showed similar diversity indices within each subtype, despite their different geographical regions and different inter-subtype indices (*Villegas-Gómez et al., 2016*), and a greater diversity between the STs of a control group compared with one associated with irritable bowel syndrome (*Vargas-Sanchez et al., 2015*). Like *G. intestinalis*, the number of sequences for Córdoba was very small, avoiding any strong conclusions regarding this population.

Finally, in the case of other less frequently detected protozoa such as *Cryptosporidium* and *Entamoeba* spp., we identified in some cases the species present in positive samples. For *Cryptosporidium* spp., *C. andersoni* was detected in the Amazon; *C. muris, C. ubiquitum* and *C. andersoni* were detected in the Andean region; *C. hominis*, *C. muris* and *C. felis* were detected in the Caribbean; and *C. hominis* and *C. parvum* were detected in the Pacific. These results agreed with previous studies (*Galván-Díaz, 2018*; *Sánchez et al., 2017*, *2018*), except for *C. ubiquitum* and *C. andersoni* which had not previously been reported in human feces in Colombia. These two species are associated with a wide variety

of animal hosts including domestic and wild ruminants, rodents, omnivores and primates (*Fayer, Santín & Macarisin, 2010*). The low host specificity of *C. ubiquitum* along with the shared habitats of different animals can contribute to its wide distribution and therefore to possible infections in humans, especially immunocompromised patients (*Fayer, Santín & Macarisin, 2010*). *C. parvum* and *C. hominis* are the species most frequently detected in humans. However, cattle and other domestic and wild animals infected with different species can have great importance in public health and the transmission of this parasite (*Ryan, Fayer & Xiao, 2014*). The detection of species associated with bovine hosts and cats infecting humans allows us to infer that the transmission of *Cryptosporidium* spp. in the regions evaluated is zoonotic and possibly also from human to human. Likewise, these findings suggest the great need to evaluate prevention and control measures for parasitic infections and the need to improve water sanitation infrastructure for human consumption.

For the *Entamoeba* complex, we only obtained samples positive for *E. dispar* in the departments of Bolívar and Casanare and samples positive for *E. moshkovskii* in Bolívar. This indicated probable orofecal transmission in the areas evaluated. Our findings are consistent with a study conducted by Lopez et al. showing a high frequency of *E. dispar* and *E. moshkovskii* in La Virgen, Cundinamarca, Colombia and low frequency of *E. histolytica* (*López et al., 2015*). Due to the low number of samples detected for *Cryptosporidium* and *Entamoeba* spp., it was not possible to establish any type of geographical associations for these parasites. For these microorganisms, the number of samples collected in each region should be expanded to establish with greater certainty the true frequencies of the circulating variants in the country.

CONCLUSIONS

In conclusion, ours is the first study to assess the frequency and genotypes of intestinal protozoa using sampling areas located in five biogeographical regions of Colombia. Our results showed frequent transmission of intestinal protozoa and high genetic diversity of *G. intestinalis* and *Blastocystis*, mainly in the Caribbean and the Andean regions. The sampled areas need to be expanded to establish the transmission rates and genetic characteristics of these microorganisms more accurately. Likewise, it is necessary to note that the results of the present work are an important contribution to explore the frequencies of these parasites in the country but new studies are required to obtain more representative information of the biogeographical regions, increasing the number of samples and including more cities/municipalities to be evaluated in order to establish the real frequency of these microorganisms in the different bioregions. Future studies could consider evaluating samples from different countries in South America as well, which would permit assessment of the frequency of intestinal parasites as well as their STs and assemblages at the continental level. The high frequency of *Blastocystis* and *G. intestinalis* in the samples analyzed likely represents active orofecal transmission involving different hosts in addition to humans, within the life cycles of these protozoa. This might put populations living in vulnerable socioeconomic conditions at risk and it is

therefore necessary to implement new strategies for control and prevention of these microorganisms.

ABBREVIATIONS

Blast	Basic Local Alignment Search Tool
gdh	glutamate dehydrogenase
Hd	haplotypic diversity
Mega	Molecular evolutionary genetics analysis
S	polymorphic (segregating) sites
SSUrRNA	Small subunit ribosomal ribonucleic acid
ST	subtype
tpi	triose phosphate isomerase

ACKNOWLEDGEMENTS

We thank Edanz Group for editing a draft of this manuscript.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

Adriana Higuera was financed by the Departamento Administrativo de Ciencia, Tecnología e Innovación (Colciencias) within the framework of the National Program for Promoting Research Training (sponsorship call 757). This work was funded by the Departamento Administrativo de Ciencia, Tecnología e Innovación (Colciencias) through project "Desarrollo de una estrategia y plataforma de Multilocus Sequence Typing (MLST) para la caracterización molecular de siete parásitos intestinales," code 122271250521 and contract number 130–2016. Funding was also received from the Vicerrectoria de Investigaciones of the Universidad del Cauca (code 4204). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Departamento Administrativo de Ciencia, Tecnología e Innovación (Colciencias): 122271250521 and 130–2016.
Vicerrectoria de Investigaciones of the Universidad del Cauca: 4204.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Adriana Higuera conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

- Ximena Villamizar performed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Giovanny Herrera performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Julio Cesar Giraldo performed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Luis Reinel Vasquez-A performed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Plutarco Urbano analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Oswaldo Villalobos analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Catalina Tovar analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Juan David Ramírez conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

Human Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

This study was a minimum risk investigation for participants. Both the ethical standards of the Colombian Ministry of Health (Youth Code) and the Helsinki Declaration of 2013 were followed. The parents or guardians of minors participating in the study signed informed consent forms and gave their permission to obtain samples. This study was approved by the research ethics committee of the Universidad del Rosario (registered in Act No. 394 of the CEI-UR), the ethics committee of the Department of Internal Medicine of the Universidad del Cauca (number VRI024/2016), and the INCCA University of Colombia (number 237894).

DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

The raw data is available as a Supplemental File. All sequences are available at GenBank: MN877659–MN877714.

Data Availability

The following information was supplied regarding data availability:

The raw data is available as a Supplemental File.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.8554#supplemental-information.

REFERENCES

Agudelo-Lopez S, Gómez-Rodríguez L, Coronado X, Orozco A, Valencia-Gutierrez CA, Restrepo-Betancur LF, Galvis-Gómez LA, Botero-Palacio LE. 2008. Prevalence of intestinal parasitism and associated factors in a village on the Colombian Atlantic Coast. *Revista Salud Publica (Bogota)* **10**:633–642.

Alfellani MA, Jacob AS, Perea NO, Krecek RC, Taner-Mulla D, Verweij JJ, Levecke B, Tannich E, Clark CG, Stensvold CR. 2013. Diversity and distribution of *Blastocystis* sp. subtypes in non-human primates. *Parasitology* **140(8)**:966–971 DOI 10.1017/S0031182013000255.

Barón JD. 2002. Las regiones económicas de Colombia: un análisis de clusters. *Centro de estudios económicos regionales, Banco de la República* **23**:40.

Beyhan YE, Taş Cengiz Z. 2017. Comparison of microscopy, ELISA, and real-time PCR for detection of *Giardia intestinalis* in human stool specimens. *Turkish Journal of Medical Sciences* **47**:1295–1299 DOI 10.3906/sag-1612-71.

Broglia A, Weitzel T, Harms G, Cacció SM, Nöckler K. 2013. Molecular typing of *Giardia duodenalis* isolates from German travellers. *Parasitology Research* **112(10)**:3449–3456 DOI 10.1007/s00436-013-3524-y.

Burnet JB, Ogorzaly L, Tissier A, Penny C, Cauchie HM. 2013. Novel quantitative TaqMan real-time PCR assays for detection of *Cryptosporidium* at the genus level and genotyping of major human and cattle-infecting species. *Journal of Applied Microbiology* **114(4)**:1211–1222 DOI 10.1111/jam.12103.

Caccio SM, Ryan U. 2008. Molecular epidemiology of *Giardiasis. Molecular and Biochemical Parasitology* **160(2)**:75–80 DOI 10.1016/j.molbiopara.2008.04.006.

Carvajal-Restrepo H, Orrego-Morales C, Vega-Orrego T, Arango-Arango S, Buitrago-Agudelo D, Maya-Betancourt M, Massaro-Ceballos M, Maya-Betancourt V, Restrepo-Álvarez L, Silva-Cáceres N, Suarez-Urquijo S, Massaro-Ceballos M, Romero-Montoya M, Sánchez-Jiménez M, Cardona-Castro N. 2019. Screening for intestinal parasites in adults from three different regions of Colombia. *Infectio* **23(1)**:33–38 DOI 10.22354/in.v23i1.753.

Coronato Nunes B, Pavan MG, Jaeger LH, Monteiro KJL, Xavier SCC, Monteiro FA, Bóia MN, Carvalho-Costa FA. 2016. Spatial and molecular epidemiology of *Giardia intestinalis* deep in the Amazon, Brazil. *PLOS ONE* **11(7)**:e0158805 DOI 10.1371/journal.pone.0158805.

Del Coco VF, Molina NB, Basualdo JA, Cordoba MA. 2017. *Blastocystis* spp.: advances, controversies and future challenges. *Revista Argentina de Microbiologia* **49(1)**:110–118 DOI 10.1016/j.ram.2016.08.004.

Duffy T, Cura CI, Ramirez JC, Abate T, Cayo NM, Parrado R, Bello ZD, Velazquez E, Muñoz-Calderon A, Juiz NA, Basile J, Garcia L, Riarte A, Nasser JR, Ocampo SB, Yadon ZE, Torrico F, De Noya BA, Ribeiro I, Schijman AG. 2013. Analytical performance of a multiplex real-time PCR assay using TaqMan probes for quantification of Trypanosoma cruzi satellite DNA in blood samples. *PLOS Neglected Tropical Diseases* **7(1)**:e2000 DOI 10.1371/journal.pntd.0002000.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32(5)**:1792–1797 DOI 10.1093/nar/gkh340.

El Safadi D, Gaayeb L, Meloni D, Cian A, Poirier P, Wawrzyniak I, Delbac F, Dabboussi F, Delhaes L, Seck M, Hamze M, Riveau G, Viscogliosi E. 2014. Children of Senegal River Basin show the highest prevalence of *Blastocystis* sp. ever observed worldwide. *BMC Infectious Diseases* **14(1)**:164 DOI 10.1186/1471-2334-14-164.

Fantinatti M, Bello AR, Fernandes O, Da-Cruz AM. 2016. Identification of *Giardia lamblia* assemblage E in humans points to a new anthropozoonotic cycle. *Journal of Infectious Diseases* **214(8)**:1256–1259 DOI 10.1093/infdis/jiw361.

Faria CP, Zanini GM, Dias GS, Da Silva S, Sousa MDC. 2017. New multilocus genotypes of *Giardia lamblia* human isolates. *Infection, Genetics and Evolution* **54**:128–137 DOI 10.1016/j.meegid.2017.06.028.

Fayer R, Santín M, Macarisin D. 2010. *Cryptosporidium ubiquitum* n. sp. in animals and humans. *Veterinary Parasitology* **172(1–2)**:23–32 DOI 10.1016/j.vetpar.2010.04.028.

Feng Y, Xiao L. 2011. Zoonotic potential and molecular epidemiology of *Giardia* species and *Giardiasis*. *Clinical Microbiology Reviews* **24(1)**:110–140 DOI 10.1128/CMR.00033-10.

Feng Y, Xiao L. 2017. Molecular epidemiology of Cryptosporidiosis in China. *Frontiers in Microbiology* **8**:1701 DOI 10.3389/fmicb.2017.01701.

Galván-Díaz AL. 2018. Cryptosporidiosis in Colombia: a systematic review. *Current Tropical Medicine Reports* **5(3)**:144–153 DOI 10.1007/s40475-018-0149-3.

Garcia-R JC, French N, Pita A, Velathanthiri N, Shrestha R, Hayman D. 2017. Local and global genetic diversity of protozoan parasites: spatial distribution of *Cryptosporidium* and *Giardia* genotypes. *PLOS Neglected Tropical Diseases* **11(7)**:e0005736 DOI 10.1371/journal.pntd.0005736.

Haque R. 2007. Human intestinal parasites. *Journal of Health, Population, and Nutrition* **25**:387–391.

Harhay MO, Horton J, Olliaro PL. 2010. Epidemiology and control of human gastrointestinal parasites in children. *Expert Review of Anti-Infective Therapy* **8(2)**:219–234 DOI 10.1586/eri.09.119.

Hernández PC, Morales L, Chaparro-Olaya J, Sarmiento D, Jaramillo JF, Ordoñez GA, Cortés F, Sánchez LK. 2019. Intestinal parasitic infections and associated factors in children of three rural schools in Colombia: a cross-sectional study. *PLOS ONE* **14(7)**:e0218681 DOI 10.1371/journal.pone.0218681.

Heyworth MF. 2016. *Giardia duodenalis* genetic assemblages and hosts. *Parasite* **23(4)**:13 DOI 10.1051/parasite/2016013.

Hunter PR, Hadfield SJ, Wilkinson D, Lake IR, Harrison FCD, Chalmers RM. 2007. Subtypes of *Cryptosporidium parvum* in humans and disease risk. *Emerging Infectious Diseases* **13(1)**:82–88 DOI 10.3201/eid1301.060481.

IDEAM, IGAC, IAvH, Invemar, Sinchi and IIAP. 2007. *Ecosistemas continentales, costeros y marinos de Colombia: IDEAM, IGAC, IAvH, Invemar, I. Sinchi, IIAP*. Bogotá: Imprenta Nacional de Colombia.

Jacobsen KH, Ribeiro PS, Quist BK, Rydbeck BV. 2007. Prevalence of intestinal parasites in young Quichua children in the highlands of rural ecuador. *Journal of Health, Population, and Nutrition* **25**:399–405.

Jiménez PA, Jaimes JE, Ramírez JD. 2019. A summary of *Blastocystis* subtypes in North and South America. *Parasites and Vectors* **12(1)**:376 DOI 10.1186/s13071-019-3641-2.

Jolley KA, Maiden MCJ. 2010. BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* **11(1)**:595 DOI 10.1186/1471-2105-11-595.

Khan A, Shaik JS, Grigg ME. 2018. Genomics and molecular epidemiology of *Cryptosporidium* species. *Acta Tropica* **184**:1–14 DOI 10.1016/j.actatropica.2017.10.023.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33(7)**:1870–1874 DOI 10.1093/molbev/msw054.

Lasek-Nesselquist E, Mark Welch D, Sogin M. 2010. The identification of a new *Giardia duodenalis* assemblage in marine vertebrates and a preliminary analysis of *G. duodenalis* population biology in marine systems. *International Journal for Parasitology* **40(9)**:1063–1074 DOI 10.1016/j.ijpara.2010.02.015.

Londono-Franco AL, Loaiza-Herrera J, Lora-Suarez FM, Gomez-Marin JE. 2014. *Blastocystis* sp. frequency and sources among children from 0 to 5 years of age attending public day care centers in Calarca. *Colombia Biomedica* **34**:218–227 DOI 10.1590/s0120-41572014000200008.

López MC, León CM, Fonseca J, Reyes P, Moncada L, Olivera MJ, Ramírez JD. 2015. Molecular epidemiology of *Entamoeba*: first description of *Entamoeba moshkovskii* in a rural area from Central Colombia. *PLOS ONE* **10(10)**:e0140302 DOI 10.1371/journal.pone.0140302.

Mahmoudi MR, Nazemalhosseini-Mojarad E, Karanis P. 2015. Genotyping of *Giardia lamblia* and *Entamoeba* spp. from river waters in Iran. *Parasitology Research* **114(12)**:4565–4570 DOI 10.1007/s00436-015-4702-x.

Malheiros AF, Stensvold CR, Clark CG, Braga GB, Shaw JJ. 2011. Short report: molecular characterization of *Blastocystis* obtained from members of the indigenous Tapirape ethnic group from the Brazilian Amazon region, Brazil. *American Journal of Tropical Medicine and Hygiene* **85(6)**:1050–1053 DOI 10.4269/ajtmh.2011.11-0481.

Mateo M, Mateo M, Montoya A, Bailo B, Saugar JM, Aguilera M, Fuentes I, Carmena D. 2014. Detection and molecular characterization of *Giardia duodenalis* in children attending day care centers in Majadahonda, Madrid, Central Spain. *Medicine (Baltimore)* **93(15)**:e75 DOI 10.1097/MD.0000000000000075.

Mejia R, Vicuña Y, Vaca M, Cooper PJ, Chico M, Nutman TB, Broncano N, Sandoval C. 2013. A novel, multi-parallel, real-time polymerase chain reaction approach for eight gastrointestinal parasites provides improved diagnostic capabilities to resource-limited at-risk populations. *American Journal of Tropical Medicine and Hygiene* **88(6)**:1041–1047 DOI 10.4269/ajtmh.12-0726.

Meloni D, Sanciu G, Poirier P, El Alaoui H, Chabé M, Delhaes L, Dei-Cas E, Delbac F, Luigi Fiori P, Di Cave D, Viscogliosi E. 2011. Molecular subtyping of *Blastocystis* sp. isolates from symptomatic patients in Italy. *Parasitology Research* **109(3)**:613–619 DOI 10.1007/s00436-011-2294-7.

Meurs L, Polderman AM, Vinkeles Melchers NVS, Brienen EAT, Verweij JJ, Groosjohan B, Mendes F, Mechendura M, Hepp DH, Langenberg MCC, Edelenbosch R, Polman K, Van Lieshout L. 2017. Diagnosing polyparasitism in a high-prevalence setting in Beira, Mozambique: detection of intestinal parasites in fecal samples by microscopy and real-time PCR. *PLOS Neglected and Tropical Diseases* **11(1)**:e0005310 DOI 10.1371/journal.pntd.0005310.

MinSalud. 2015. Encuesta nacional de parasitismo intestinal en población escolar Colombia. *Available at https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/ET/encuesta-nacional-de-parasitismo-2012-2014.pdf*.

Minvielle MC, Molina NB, Polverino D, Basualdo JA. 2008. First genotyping of *Giardia lamblia* from human and animal feces in Argentina, South America. *Memórias do Instituto Oswaldo Cruz* **103(1)**:98–103 DOI 10.1590/S0074-02762008000100015.

Molina N, Minvielle M, Grenovero S, Salomon C, Basualdo J. 2011. High prevalences of infection with *Giardia intestinalis* genotype B among children in urban and rural areas of Argentina.

Annals of Tropical Medicine & Parasitology **105(4)**:299–309 DOI 10.1179/136485911X12987676649665.

Morens DM, Folkers GK, Fauci AS. 2004. The challenge of emerging and re-emerging infectious diseases. *Nature* **430(6996)**:242–249 DOI 10.1038/nature02759.

Morgan UM, Pallant L, Dwyer BW, Forbes DA, Rich G, Thompson RCA. 1998. Comparison of PCR and microscopy for detection of *Cryptosporidium parvum* in human fecal specimens: clinical trial. *Journal of Clinical Microbiology* **36(4)**:995–998 DOI 10.1128/JCM.36.4.995-998.1998.

Ortiz C, López MC, Rivas FA. 2012. Prevalencia de helmintos en la planta de aguas residuales del municipio El Rosal, Cundinamarca. *Revista de Salud Publica (Bogota)* **14(2)**:296–304 DOI 10.1590/s0124-00642012000200010.

Papier K, Williams GM, Luceres-Catubig R, Ahmed F, Olveda RM, McManus DP, Ross AG. 2014. Childhood malnutrition and parasitic helminth interactions. *Clinical Infectious Diseases* **59(2)**:234–243 DOI 10.1093/cid/ciu211.

Pelayo L, Nuñez FA, Rojas L, Furuseth Hansen E, Gjerde B, Wilke H, Mulder B, Robertson L. 2008. *Giardia* infections in Cuban children: the genotypes circulating in a rural population. *Annals of Tropical Medicine & Parasitology* **102(7)**:585–595 DOI 10.1179/136485908X355247.

Perez Cordon G, Cordova Paz Soldan O, Vargas Vasquez F, Velasco Soto JR, Sempere Bordes L, Sanchez Moreno M, Rosales MJ. 2008. Prevalence of enteroparasites and genotyping of *Giardia lamblia* in Peruvian children. *Parasitology Research* **103(2)**:459–465 DOI 10.1007/s00436-008-1007-3.

Pritt BS, Clark CG. 2008. Amebiasis. *Mayo Clinic Proceedings* **83(10)**:1154–1160 DOI 10.4065/83.10.1154.

Ramirez JD, Florez C, Olivera M, Bernal MC, Giraldo JC. 2017. *Blastocystis* subtyping and its association with intestinal parasites in children from different geographical regions of Colombia. *PLOS ONE* **12(2)**:e0172586 DOI 10.1371/journal.pone.0172586.

Ramírez JD, Sanchez LV, Bautista DC, Corredor AF, Florez AC, Stensvold CR. 2014. Blastocystis subtypes detected in humans and animals from Colombia. *Infection, Genetics and Evolution* **22**:223–228 DOI 10.1016/j.meegid.2013.07.020.

Ramírez JD, Heredia RD, Hernández C, León CM, Moncada LI, Reyes P, Pinilla AE, Lopez MC. 2015. Molecular diagnosis and genotype analysis of *Giardia duodenalis* in asymptomatic children from a rural area in central Colombia. *Infection, Genetics and Evolution* **32**:208–213 DOI 10.1016/j.meegid.2015.03.015.

Ramírez JD, Sánchez A, Hernández C, Flórez C, Bernal MC, Giraldo JC, Reyes P, López MC, García L, Cooper PJ, Vicuña Y, Mongi F, Casero RD. 2016. Geographic distribution of human *Blastocystis* subtypes in South America. *Infection, Genetics and Evolution* **41**:32–35 DOI 10.1016/j.meegid.2016.03.017.

Rangel-Ch O, Aguilar M. 1995. Una aproximación sobre la diversidad climática en las regiones naturales de Colombia. *Available at http://www.uneditorial.net/pdf/TomoI.pdf#page=15*.

Read CM, Monis PT, Thompson RA. 2004. Discrimination of all genotypes of *Giardia duodenalis* at the glutamate dehydrogenase locus using PCR-RFLP. *Infection, Genetics and Evolution* **4(2)**:125–130 DOI 10.1016/j.meegid.2004.02.001.

Rodríguez V, Espinosa O, Carranza JC, Duque S, Arévalo A, Clavijo JA, Urrea DA, Vallejo GA. 2014. Genotipos de *Giardia duodenalis* en muestras de niños de las guarderías del Instituto Colombiano de Bienestar Familiar y de perros en Ibagué. *Colombia Biomédica* **34**:271–281 DOI 10.7705/biomedica.v34i2.1713.

Rondón S, Ortiz M, León C, Galvis N, Link A, González C. 2017. Seasonality, richness and prevalence of intestinal parasites of three neotropical primates (*Alouatta seniculus*, *Ateles hybridus* and *Cebus versicolor*) in a fragmented forest in Colombia. *International Journal for Parasitology: Parasites and Wildlife* **6(3)**:202–208 DOI 10.1016/j.ijppaw.2017.07.006.

Ryan U, Cacciò SM. 2013. Zoonotic potential of Giardia. *International Journal for Parasitology* **43(12–13)**:943–956 DOI 10.1016/j.ijpara.2013.06.001.

Ryan U, Fayer R, Xiao L. 2014. *Cryptosporidium* species in humans and animals: current understanding and research needs. *Parasitology* **141(13)**:1667–1685 DOI 10.1017/S0031182014001085.

Sánchez A, Munoz M, Gomez N, Tabares J, Segura L, Salazar A, Restrepo C, Ruíz M, Reyes P, Qian Y, Xiao L, López MC, Ramirez JD. 2017. Molecular epidemiology of *Giardia*, *Blastocystis* and *Cryptosporidium* among Indigenous children from the Colombian Amazon Basin. *Frontiers in Microbiology* **8(Pt. 1)**:248 DOI 10.3389/fmicb.2017.00248.

Sánchez C, López MC, Galeano LA, Qvarnstrom Y, Houghton K, Ramírez JD. 2018. Molecular detection and genotyping of pathogenic protozoan parasites in raw and treated water samples from southwest Colombia. *Parasites & Vectors* **11(1)**:563 DOI 10.1186/s13071-018-3147-3.

Santin M, Gomez-Munoz MT, Solano-Aguilar G, Fayer R. 2011. Development of a new PCR protocol to detect and subtype *Blastocystis* spp. from humans and animals. *Parasitology Research* **109(1)**:205–212 DOI 10.1007/s00436-010-2244-9.

Scicluna SM, Tawari B, Clark CG. 2006. DNA barcoding of *Blastocystis*. *Protist* **157(1)**:77–85 DOI 10.1016/j.protis.2005.12.001.

Silva MTN, Santana JV, Bragagnoli G, Da Nóbrega Marinho AM, Malagueño E. 2014. Prevalence of *Entamoeba histolytica*/*Entamoeba dispar* in the city of Campina Grande, in northeastern Brazil. *Revista do Instituto de Medicina Tropical de São Paulo* **56(5)**:451–454 DOI 10.1590/S0036-46652014000500015.

Sprong H, Cacciò SM, Van Der Giessen JWB, on behalf of the ZOOPNET network and partners. 2009. Identification of zoonotic genotypes of *Giardia duodenalis*. *PLOS Neglected Tropical Diseases* **3(12)**:e558 DOI 10.1371/journal.pntd.0000558.

Sri-Hidajati BS, Basuki S, Pusarawati S, Kusmartisnawati K, Arwati H, Tantular I, Fardah A, Darma A, Handajani R, Soedarmo SM. 2018. Comparison of multiplexing single round PCR and microscopy in diagnosis of amoebiasis. *African Journal of Infectious Diseases* **12(1S)**:120–126 DOI 10.21010/ajid.v12i1S.18.

Stensvold CR, Ahmed UN, Andersen LO, Nielsen HV. 2012. Development and evaluation of a genus-specific, probe-based, internal-process-controlled real-time PCR assay for sensitive and specific detection of *Blastocystis* spp. *Journal of Clinical Microbiology* **50(6)**:1847–1851 DOI 10.1128/JCM.00007-12.

Stensvold CR, Alfellani M, Clark CG. 2012. Levels of genetic diversity vary dramatically between *Blastocystis* subtypes. *Infection, Genetics and Evolution* **12(2)**:263–273 DOI 10.1016/j.meegid.2011.11.002.

Stensvold CR, Clark CG. 2016. Current status of *Blastocystis*: a personal view. *Parasitology International* **65(6)**:763–771 DOI 10.1016/j.parint.2016.05.015.

Stensvold CR, Lewis HC, Hammerum AM, Porsbo LJ, Nielsen SS, Olsen KEP, Arendrup MC, Nielsen HV, Mølbak K. 2009. *Blastocystis*: unravelling potential risk factors and clinical significance of a common but neglected parasite. *Epidemiology and Infection* **137(11)**:1655–1663 DOI 10.1017/S0950268809002672.

Stensvold CR, Olsen KEP, Christiansen DB, Nielsen HV. 2011. *Blastocystis* sp. subtype 4 is common in Danish *Blastocystis*-Positive patients presenting with acute diarrhea. *American Journal of Tropical Medicine and Hygiene* **84(6)**:883–885 DOI 10.4269/ajtmh.2011.11-0005.

Stensvold CR, Suresh GK, Tan KSW, Thompson RCA, Traub RJ, Viscogliosi E, Yoshikawa H, Clark CG. 2007. Terminology for *Blastocystis* subtypes—a consensus. *Trends in Parasitology* **23(3)**:93–96 DOI 10.1016/j.pt.2007.01.004.

Stensvold CR, Winiecka-Krusnell J, Lier T, Lebbad M. 2018. Evaluation of a PCR method for detection of *Entamoeba polecki*, with an overview of its molecular epidemiology. *Journal of Clinical Microbiology* **56(5)**:e00154-18 DOI 10.1128/JCM.00154-18.

Sulaiman IM, Fayer R, Bern C, Gilman RH, Trout JM, Schantz PM, Das P, Lal AA, Xiao L. 2003. Triosephosphate isomerase gene characterization and potential zoonotic transmission of *Giardia duodenalis*. *Emerging Infectious Diseases* **9(11)**:1444–1452 DOI 10.3201/eid0911.030084.

Supali T, Verweij JJ, Wiria AE, Djuardi Y, Hamid F, Kaisar MMM, Wammes LJ, Van Lieshout L, Luty AJF, Sartono E, Yazdanbakhsh M. 2010. Polyparasitism and its impact on the immune system. *International Journal for Parasitology* **40(10)**:1171–1176 DOI 10.1016/j.ijpara.2010.05.003.

Tasawar Z, Kausar S, Lashari MH. 2010. Prevalence of *Entamoeba histolytica* in humans. *Pakistan Journal of Pharmaceutical Sciences* **23**:344–348.

Torres-Romero JC, De Jesus Euan-Canto A, Benito-González N, Padilla-Montaño N, Huchin-Chan C, Lara-Riegos J, Cedillo-Rivera R. 2014. Intestinal parasites and genotyping of *Giardia duodenalis* in children: first report of genotype B in isolates from human clinical samples in Mexico. *Memórias do Instituto Oswaldo Cruz* **109(3)**:388–390 DOI 10.1590/0074-0276140507.

Valença-Barbosa C, Do Bomfim TCB, Teixeira BR, Gentile R, Da Costa Neto SF, Magalhães BSN, De Almeida Balthazar D, Da Silva FA, Biot R, D'Avila Levy CM, Santos HLC. 2019. Molecular epidemiology of *Blastocystis* isolated from animals in the state of Rio de Janeiro, Brazil. *PLOS ONE* **14(1)**:e0210740 DOI 10.1371/journal.pone.0210740.

Vargas-Sanchez G-B, Romero-Valdovinos M, Ramirez-Guerrero C, Vargas-Hernandez I, Ramirez-Miranda ME, Martinez-Ocaña J, Valadez A, Ximenez C, Lopez-Escamilla E, Hernandez-Campos ME, Villalobos G, Martinez-Hernandez F, Maravilla P, Escobar-Gutiérrez A. 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.

Vergara FJ. 2018. *Nueva Geografía de Colombia, Vol 1: Escrita por Regiones Naturales (Classic Reprint)*. London: FB&C Limited, 1178.

Villamizar X, Higuera A, Herrera G, Vasquez-A LR, Buitron L, Muñoz LM, Gonzalez-C FE, Lopez MC, Giraldo JC, Ramírez JD. 2019. Molecular and descriptive epidemiology of intestinal protozoan parasites of children and their pets in Cauca, Colombia: a cross-sectional study. *BMC Infectious Diseases* **19(1)**:190 DOI 10.1186/s12879-019-3810-0.

Villegas-Gómez I, Martínez-Hernández F, Urrea-Quezada A, González-Díaz M, Durazo M, Hernández J, Orozco-Mosqueda GE, Villalobos G, Maravilla P, Valenzuela O. 2016. Comparison of the genetic variability of *Blastocystis* subtypes between human carriers from two contrasting climatic regions of México. *Infection, Genetics and Evolution* **44**:334–340 DOI 10.1016/j.meegid.2016.07.036.

Wawrzyniak I, Poirier P, Viscogliosi E, Dionigia M, Texier C, Delbac F, Alaoui HE. 2013. *Blastocystis*, an unrecognized parasite: an overview of pathogenesis and diagnosis. *Therapeutic Advances in Infectious Disease* **1(5)**:167–178 DOI 10.1177/2049936113504754.

WHO. 2010. Guía de atención integral en VIH/SIDA. *Available at http://www.who.int/hiv/pub/guidelines/ecuador_art.pdf?ua=1*.

Xiao L. 2010. Molecular epidemiology of *Cryptosporidiosis*: an update. *Experimental Parasitology* **124(1)**:80–89 DOI 10.1016/j.exppara.2009.03.018.

Ximénez C, Morán P, Rojas L, Valadez A, Gómez A. 2009. Reassessment of the epidemiology of amebiasis: state of the art. *Infection, Genetics and Evolution* **9(6)**:1023–1032 DOI 10.1016/j.meegid.2009.06.008.

8.2 CAPÍTULO 2

Análisis de la diversidad y estructura genética de *Giardia intestinalis*.

genes

MDPI

Article

Development of a Multilocus Sequence Typing Scheme for *Giardia intestinalis*

Adriana Higuera [1], Marina Muñoz [1], Myriam Consuelo López [2], Patricia Reyes [2], Plutarco Urbano [3], Oswaldo Villalobos [4] and Juan David Ramírez [1,*]

1 Grupo de Investigaciones Microbiológicas-UR (GIMUR), Departamento de Biología, Facultad de Ciencias Naturales, Universidad del Rosario, Bogotá 111211, Colombia; adriana.higuera@urosario.edu.co (A.H.); claudiamarina23@gmail.com (M.M.)
2 Departamento de Salud Pública, Universidad Nacional de Colombia, Bogotá 111321, Colombia; mlopezp@unal.edu.co (M.C.L.); preyes@unal.edu.co (P.R.)
3 Grupo de Investigaciones Biológicas de la Orinoquia, Unitrópico, Yopal 8500, Colombia; plurbanus@gmail.com
4 Hospital Local Santa María de Mompox, Programas Especiales (Lepra y TB), Mompox, Bolívar 132560, Colombia; o.villalobosr@gmail.com
* Correspondence: juand.ramirez@urosario.edu.co

Received: 14 June 2020; Accepted: 28 June 2020; Published: 8 July 2020

Abstract: *Giardia intestinalis* is an intestinal protozoan most commonly found in humans. It has been grouped into 8 assemblages (A-H). Markers such as the glutamate dehydrogenase gene, triose phosphate isomerase and beta-giardin (β-giardin) have been widely used for genotyping. In addition, different genetic targets have been proposed as a valuable alternative to assess diversity and genetics of this microorganism. Thus, our objective was to evaluate new markers for the study of the diversity and intra-taxa genetic structure of *G intestinalis* in silico and in DNA obtained from stool samples. We analysed nine constitutive genes in 80 complete genome sequences and in a group of 24 stool samples from Colombia. Allelic diversity was evaluated by locus and for the concatenated sequence of nine loci that could discriminate up to 53 alleles. Phylogenetic reconstructions allowed us to identify AI, AII and B assemblages. We found evidence of intra- and inter-assemblage recombination events. Population structure analysis showed genetic differentiation among the assemblages analysed.

Keywords: *G intestinalis*; genetic diversity; genetic structure; recombination

1. Introduction

Giardia intestinalis (synonym *G lamblia*, *G duodenalis*), a single-celled eukaryotic protozoan, is the most common cause of parasitic diarrhoea in humans worldwide [1]. It infects approximately 2% of adults and between 6% and 8% of children in developed countries. About 33% of people have had giardiasis in developing countries [2]. Transmission of this protozoan is considered both zoonotic and zooanthroponotic since it is present in domestic [3] and wild animals [4]. The frequency of transmission among hosts is unclear [5]. Still, a risk for massive spread is known to exist.

Different molecular tools [6–8] and genetic markers [9,10] with different mutation rates [11], have been used to evaluate the inter- and intra-specific variation of *G intestinalis* [12,13], based mainly on 3 loci, β-giardin, triose phosphate isomerase and glutamate dehydrogenase [14,15]. These loci supported the identification of eight assemblages, termed A through H. These assemblages can be host-specific [16–19] and have allowed determination of assemblages A and B as most frequent [13], with assemblage B being most common in humans [20]. Assemblage B is also associated with more severe and prolonged disease and is considered the most virulent [3,13,21–23]. Sub-assemblages, such as AI, AII, AIII, BIII and BIV [24,25], have also been established using the above loci. However, despite

their utility, these typing markers have produced contradictory results [14] or low resolution [24] when identifying assemblages in some samples.

Some authors have proposed the study of new genes [26], which, when added to typically used markers, might better elucidate intra-specific diversity, along with nucleotide heterozygosity, allelic divergence and even recombination processes and inter/intra-genetic exchange [24,26–29]. This possibility is studied through single nucleotide polymorphisms and phylogenetic analyses and even comparative genomics. Such analyses indicate that sexual or meiotic processes may promote the generation of more virulent strains or expand their host range [26]. Additionally, exploring other genetic markers will allow characterisation of sub-assemblages not clearly established in assemblage E and perhaps others, and provide needed information on the substructures of assemblages C, D, F and G [21].

Evaluation of additional regions of the genome of *G intestinalis* is needed to identify new markers for understanding its diversity and evolution. Such markers should possess sufficient discriminatory power to establish groupings related to epidemiological factors. Thus, investigation of new markers should focus on detection and typing, and allow additional inference on reproduction, evolution, zoonotic potential and population structure [30–32]. Some studies show that multilocus sequences are useful for identifying species, genera and populations, characterising isolates with conserved genes with low variation, and thus establishing allelic profiles in study populations [33]. Initially, this tool was widely used for bacteria [33,34], and subsequently has been implemented with diploid eukaryotes [35–38] and fungi [39–41]. This is because, despite the availability of complete genome analysis, the Multilocus sequence typing (MLST) approach is more accessible and economic, together with the selection of suitable markers, it is possible to generate high-resolution information for analysis of genetic diversity and evolution, without the bias of complete information on the genome that includes regions that are not informative or exposed to different selection pressures, which could be useful in other types of studies.

Few studies on *G intestinalis* are available that sought to address additional genetic markers. Yet, generating multilocus analyses is essential for understanding the genetic characteristics of circulating strains in different geographical regions and monitoring their evolution and adaptation. Such analysis will encourage the design of strategies to decrease infection incidence [21]. In the present study, we evaluated different coding loci for constitutive enzymes involved in metabolic pathways, such as glycolysis and the Tricarboxylic acid cycle (TCA), focused on identifying genotypic characteristics of *G intestinalis* tested in publicly available whole-genome sequences (WGS) and subsequently analyse these markers in DNA from stool samples from some regions of Colombia. Our proposed new markers are capable of elucidating diversity, population structure and possible recombination events between and within *G intestinalis* assemblages.

2. Results

2.1. Analysis of New Loci Using WGS Data

2.1.1. Genetic Diversity of Housekeeping Genes

Sequences of each assemblage, AI (WB), AII (DH), B (GS and GS_B) and E (P15) (Supplementary Materials, Table S1) for various genes were aligned, and conserved regions among assemblages used for primer design (Supplementary Materials, Table S2). Initially, eight genes were chosen, based on established criteria (see methods below). We used target genes and associated primers to evaluate *Pyrophosphate-fructose 6-phosphate 1-phosphotransferase alpha subunit (PFP-ALPHA1), Fructose-bisphosphate aldolase (FBA), Phosphoglycerate kinase (PGK), enolase, Acetyl-CoA synthetase (ACS), NADP-dependent malic enzyme (NADP-ME), Serine palmitoyltransferase 2 (SPT), Glutamate dehydrogenase (GDH)* and *Triose phosphate isomerase (TPI)* genes in 80 available WGS from *G intestinalis*. (Supplementary Materials, Table S3). We mapped short reads from sequences of interest and obtained consensus sequences for each gene using Short read sequence typing 2 (SRST2) [42]. The SRST2 output file did not report the

GPI gene in any genome, and it was excluded from the study. Thus, nine genes were used in the investigation of diversity.

The initial analysis of molecular characteristics of conserved genes, showed interesting differences among loci and the concatenated sequence. Multiple alignment of concatenated sequences with the nine loci over a length of 11,978 bp identified 2,651 polymorphic sites. The highest number haplotypes and haplotypic diversity were h:34 and 0.842, respectively. *NADP-ME* and *SPT* genes displayed the highest values for both nucleotide diversity (Pi: 0.107 and 0.108. respectively), Theta (per site) from Eta (0.060 and 0.059. respectively) and numbers of segregating polymorphic sites (464 and 458, respectively). The *SPT* showed the fewest haplotypes (h: 11) and haplotypic diversity (Hd: 0.686). In contrast, genes such as *GDH*, commonly used to type *G intestinalis*, showed a nucleotide diversity value, Pi of 0.058 and a Theta (per site) from Eta of 0.034, both being the lowest among analysed loci (Table 1).

Table 1. Diversity indices obtained for the evaluated loci.

	Marker	No. of Nucleotide Sites	No. of Sequences	Total Number of Sites	S	h	Hd (SD)	Pi (SD)	Theta	Tajima's D test	Rm
WGS data consensus sequences	Concatenated	11978	85	11496	2651	34	0.842 (0.036)	0.086 (0.003)	0.0488	2.676 *	191
	ACS	2190	85	1857	394	19	0.81 (0.032)	0.078 (0.002)	0.0451	2.550 *	30
	Enolase	1338	85	1252	267	13	0.761 (0.036)	0.071 (0.003)	0.0455	1.983	16
	FBA	972	85	937	180	12	0.762 (0.03)	0.07 (0.002)	0.0423	2.242 *	11
	PFP-ALHA1	1650	85	1635	392	15	0.756 (0.034)	0.092 (0.003)	0.05	2.911 **	23
	PGK	1230	85	1204	272	16	0.799 (0.033)	0.091 (0.002)	0.047	3.169 **	22
	GDH	1386	85	1345	224	16	0.781 (0.036)	0.058 (0.002)	0.034	2.362 *	17
	NADP-ME	1689	85	1625	464	15	0.791 (0.034)	0.107 (0.004)	0.06	2.641 *	42
	SPT	1665	85	1641	458	11	0.686 (0.031)	0.108 (0.003)	0.059	2.896 **	24
	TPI	774	85	774	194	14	0.74 (0.034)	0.095 (0.003)	0.053	2.718 **	24
Sequences obtained from stool samples	*ACS*	562	101	297	278	26	0.831 (0.025)	0.21 (0.027)	0.326	–1.201	38
	Enolase	428	97	350	162	19	0.756 (0.029)	0.077 (0.005)	0.108	–0.978	26
	GDH	365	108	227	192	21	0.716 (0.039)	0.091 (0.013)	0.249	–2.118 **	26
	SPT	487	91	175	134	10	0.579 (0.031)	0.114 (0.013)	0.221	–1.639	15
	TPI	450	106	448	133	21	0.793 (0.026)	0.109 (0.002)	0.061	2.531 *	52

Total number of sites (excluding sites with gaps/missing data); S: Number of polymorphic (segregating) sites; h: Number of Haplotypes; Hd: Haplotype (gene) diversity; Pi: Nucleotide diversity; SD: Standard Deviation; Theta (per site) from Eta; Rm: Minimum number of recombination events. * Statistical significance: $p < 0.05$; ** Statistical significance: $p < 0.01$.

Assemblage diversity indices were also calculated. The concatenated alignments AI and B assemblages showed low nucleotide diversity compared to the AII assemblage, for which relatively higher values were obtained for most loci. In contrast, Hd among these assemblages, was slightly lower for AI. Further, genes *TPI* and *PGK* showed a value of zero for AI and AII assemblages for both nucleotide diversity and haplotypic diversity (Supplementary Materials, Tables S4–S6). Notably, positive results were observed for the evolutionary divergence parameter, Tajima D [43], for all loci, most of them statistically significant ($p < 0.05$; $p < 0.01$). However, when analysing assemblages, results for all loci in AI and B assemblages were negative (Supplementary Materials, Tables S4 and S6).

Finally, we evaluated the utility of loci, including numbers of polymorphisms, typing efficiency (TE) and discriminatory power (DP), using MLSTest software (CONICET, Salta, Argentina) [44]. Numbers of possible alleles found among loci used and their combinations, showed, for example, that the combination of all nine loci could identify up to 53 alleles. With six loci, up to 51 different alleles were detected, and with a single locus, particularly *ACS*, a minimum of 11 different alleles (Supplementary Materials, Table S7). We also compared TE and DP among all loci, finding that the *GDH* locus displays the highest TE. The highest DP, 0.885, for all loci was somewhat above the DP for *ACS*, 0.815 (Table 2).

Table 2. Calculation of typing efficiency and discriminatory power of evaluated loci.

Name	*ACS*	*Enolase*	*FBA*	*PFP-ALPHA1*	*PGK*	*GDH*	*NADP-ME*	*SPT*	*TPI*	All Loci
Number of Alleles	21	15	14	15	18	20	15	11	14	53
Number of Polymorphisms	727	353	215	407	298	265	528	482	194	3469
Typing Efficiency	0.029	0.042	0.065	0.037	0.06	0.075	0.028	0.023	0.072	0.041
DP (95% Confidence Interval)	0.815 (0.748–0.881)	0.768 (0.693–0.843)	0.775 (0.715–0.835)	0.756 (0.688–0.824)	0.801 (0.732–0.869)	0.796 (0.72–0.873)	0.791 (0.722–0.861)	0.686 (0.623–0.749)	0.74 (0.671–0.81)	0.885 (0.816–0.955)

2.1.2. Phylogenetic Analysis and Recombination Signals

The phylogenetic inferences constructed from sequences from concatenated genomic data and by gene, identified three main clusters, corresponding to *G intestinalis* assemblages most commonly found in humans, the AI, AII and B (Figure 1A). The concatenated sequences of SRR3177757 and SRR3177873 genomes did not group within any established assemblage and are termed ND (not defined). Notably, comparing the position of these genomes in the phylogenetic tree, they coincide in grouping form, but the SRR3177757 genome shows evidence of inter-assemblage recombination, specifically among AI, AII and E. The SRR3177873 sequence, though showing reticulation signals, is located farther from these assemblages (Figure 1B).

Phylogenies constructed with the *GDH* locus, in particular, demonstrate all three assemblages, along with evidence of intra-assemblage AII reticulation (Supplementary Materials, Figure S1). In contrast, trees generated with other loci, showed inconsistencies in tree topology due to locations of some evaluated genomes (Supplementary Materials, Figures S2–S9). For example, *FBA*, *NADP-ME* and *TPI* loci could not establish SRR3177751 and SRR3177919 genomes with certainty in an assemblage, and only a small recombination signal was observed in the *FBA* gene (Supplementary Materials, Figures S4, S7 and S9). Further, *enolase*, *PFP-ALPHA* and *PGK* loci did not clearly group all sequences with any assemblage but did group the sequences in the phylogenetic network (Supplementary Materials, Figures S3, S5 and S6).

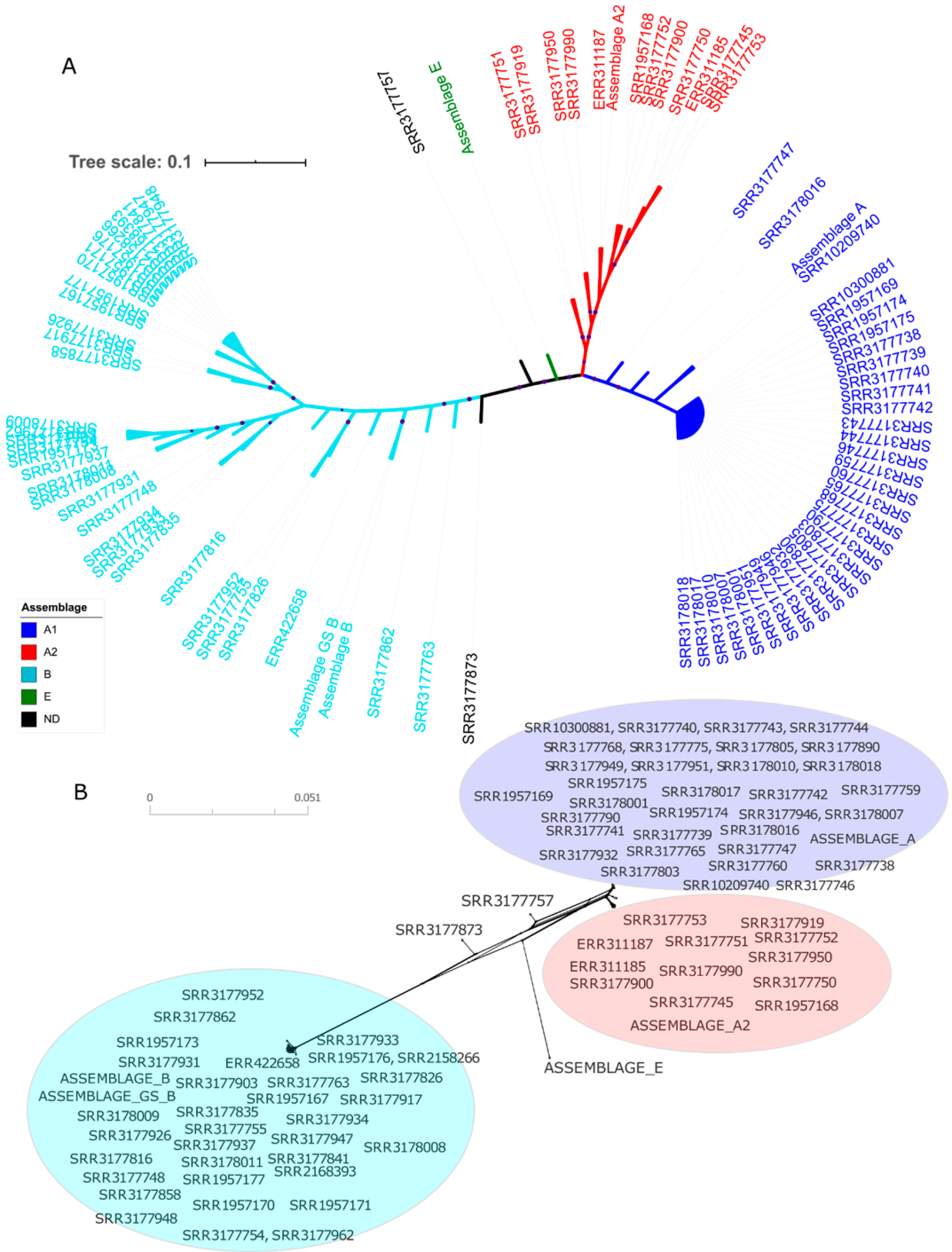


Figure 1. Phylogenetic reconstructions with sequences obtained from WGS data. Phylogenetic inferences were constructed from concatenated sequences of the nine selected genes. (**A**) Phylogenetic tree constructed from the alignment of concatenated sequences of all the genes evaluated. The tree was constructed with FastTree [45] software and visualized with ITOL [46] software. The maximum likelihood (ML) method was used under the Jukes Cantor nucleotide evolution model, with a Bootstrap of 1000 repetitions. A Bootstrap value greater than 90% is represented with a purple circle above each node. The colours indicate the assemblage to which the evaluated sequences belong (Blue: AI, Red: AII, Turquoise: B, Green: E and ND: not defined). (**B**) A phylogenetic network, using Splitstree software [47], was built with the Neighbornet algorithm; the colours correspond to assemblages. The access numbers of the genomes used are indicated.

Based on the comparison of clusters formed with concatenated sequences vs. clusters of each gene, *enolase* and *PFP-ALPHA1* genes presented the greatest number of inconsistencies in tree topology (Supplementary Materials, Figures S3 and S5). Still, the main clusters determined by maximum likelihood (ML) both for concatenated sequences (Figure 1A) and for each gene (Supplementary Materials, Figures S1A–S9A) were consistent with clusters found in phylogenetic networks using Neighbornet algorithm in Splitstree software [47] (Figure 1B and Supplementary Materials, Figures S1B–S9B). These findings support the presence of three established assemblages.

In general, crosslinking signals were observed mainly between A and E assemblages and within the AII assemblage, indicating possible recombination events. We compared this evidence of recombination with calculated indices of minimum numbers of recombination events (Rm), for all evaluated alignments, using Dnasp software [48]. The highest Rm value was found with concatenated sequences, followed by *NADP-ME* and *ACS* genes (191, 42 and 30, respectively) (Table 1). This result is consistent with phylogenetic networks, except for the case of *NADP-ME* (Supplementary Materials, Figure S7). Further, we obtained 11 different recombinants in a search for recombination sites among concatenated sequences. The most frequent was in the SRR3177873 sequence, with breakpoints at different positions of the alignment, depending on parental sequences (Supplementary Materials, Figure S10).

After phylogenetic topologies were generated and inter- and intra-assemblage crosslinks identified, a second analysis was developed in STRUCTURE [49,50]. K = 4 populations were established a priori. Clear signs of admixture between pre-established populations were observed (Figure 2). Populations are distinct, yet admixture is observed between assemblage E and the other assemblages, possibly by genetic interchange. Next, we used RDP4 software [51], to identify possible recombination events and identify their origin. We found 17 unique recombination events. Eight were detected by at least four different methods, providing further support for the presence of the event. Both detection by each method, and recombinant genomes and their possible parents from the concatenated sequences are presented (Table 3). Recombinant genomes, SRR3177757 and SRR3177873, were detected by at least five different methods. Recombination score through alignment (by position) for all detected events was calculated (Supplementary Materials, Figure S10).

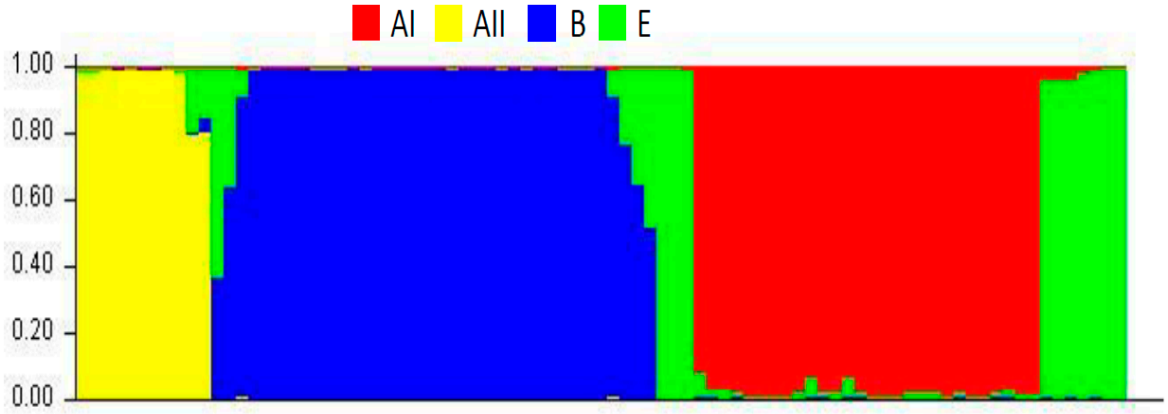


Figure 2. Analysis of STRUCTURE with the genome sequences. K = 4 populations were established a priori, corresponding to assemblages AI, AII, B and E. Colours indicate respective assemblages.

Table 3. Detection of recombination events.

Event No.	Found in	Recombinants	Major Parent	Minor Parent	Detection Methods						
					RDP	GENE CONV	BootScan	MaxiChi	Chimaera	SiScan	3Seq
1	1	SRR3177873	SRR317790	SRR3177948	+	+	+	+	+	+	+
2	1	SRR3177873	SRR3177900	ERR422658	+	+	+	+	+	+	+
3	2	Assemblage_A2	SRR317799	SRR3177750	-	+	-	+	-	+	+
4	1	ERR422658	SRR3177816	Unknown	-	+	-	+	+	+	+
5	1	SRR3177873	SRR3177900	ERR422658	-	+	-	-	-	-	-
6	1	SRR3177873	SRR317790	Unknown	-	+	+	-	-	-	-
7	1	SRR3177751	SRR317791	Unknown	-	-	-	+	-	-	-
8	4	SRR3177750	SRR317790	Unknown	+	+	+	+	+	+	+
9	2	SRR3177950	SRR317790	SRR3177919	-	+	-	+	+	-	-
10	1	SRR3177873	SRR3177950	ERR422658	-	+	-	-	-	-	-
11	1	SRR3177757	SRR102097	ERR422658	-	+	+	+	+	+	+
12	1	SRR3177757	SRR102097	SRR3177816	-	+	-	+	+	+	+
13	10	SRR3177948	SRR3178011	SRR3177926	-	-	-	+	+	-	+
14	1	SRR3177862	Assemblage_B	Unknown	-	-	-	+	+	-	+
15	1	SRR3177931	Unknown	SRR3178011	-	-	-	+	-	-	+
16	5	SRR3177952	SRR3178011	SRR3177926	-	-	-	+	-	-	+
17	1	SRR3177763	Unknown	SRR3177926	+	-	-	+	+	+	+

Finally, an allelic plot was constructed, using the classification of each genome in relation to the assemblage determined in the phylogenetic trees (bootstrap > 80), both by gene and concatenated sequence. (Figure 3). The presence of different colours in the same genome indicates inconsistencies in phylogenetic tree topology and subsequently in assemblage assignment. Such findings are consistent with possible recombination signals observed in different analyses. For example, SRR3177873 sequences are grouped in the AII sub-assemblage by *ACS*, *enolase*, *FBA*, *GDH*, *NADP-ME* and *TPI*, but in assemblage B with *PFP-ALPHA1*, *PGK* and *SPT*. No assemblage could be assigned with the concatenated sequence. Both SRR317799 and SRR3177751 showed inconsistencies between AII and AI assemblages. Some genes in the assemblage were indeterminate.

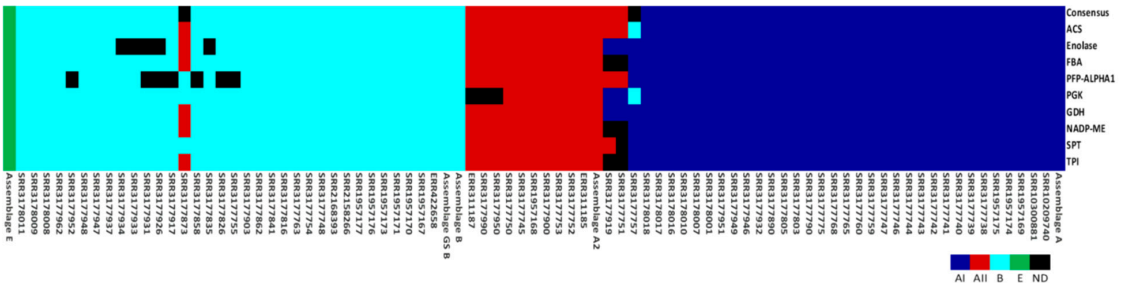


Figure 3. Allele plot constructed for with the alignment of the concatenated genomic sequences and for each gene. Colours indicate AI, AII, B, E, and ND (Not determined) assemblages. Each column corresponds to a genome and each row is a genetic marker. The assemblages were assigned depending on clusters obtained in the phylogenetic tree of each gene.

2.1.3. Population Structure

Statistics used for genetic differentiation between populations are shown in Table 4. Established populations correspond to AI, AII and B assemblages. We evaluated concatenation and individual sequences by locus. We found higher values for the Gst statistic for all loci when comparing AI vs. AII and AI vs. B assemblages with respect to AII vs. B assemblages. For example, we saw noticeably lower value for concatenated sequences when evaluating AII vs. B assemblages (0.038); Gst for the other two comparisons were 0.344 (AI vs. AII) and 0.301 (AI vs. B). In contrast, Kxy, the average ratio of nucleotide differences between populations, and Dxy, the number of average nucleotide substitutions between populations, were greatly increased for both concatenated and all loci, and between the AI vs. B and AII assemblages vs. B. Fst indices were relatively high (Fst > 0.25) [52] for all cases, indicating a structure with elevated genetic differentiation between populations, in this case, the assemblages.

Table 4. Genetic differentiation among populations with concatenated sequences and each locus.

	Assemblages		Hs	Ks	Kxy	Gst	DeltaSt	GammaSt	Nst	Fst	Dxy	Da
Concatenated	AI	AII	0.353	1.616	13.844	0.344	0.004	0.760	0.809	0.808	0.012	0.009
	AI	B	0.536	7.910	193.217	0.301	0.078	0.923	0.962	0.961	0.162	0.156
	AII	B	0.909	12.207	194.079	0.038	0.061	0.951	0.951	0.948	0.163	0.155
ACS	AI	AII	0.300	3.952	36.542	0.421	0.006	0.774	0.822	0.822	0.018	0.015
	AI	B	0.435	3.732	298.743	0.396	0.075	0.976	0.989	0.988	0.151	0.149
	AII	B	0.757	8.637	301.443	0.115	0.060	0.935	0.971	0.967	0.152	0.147
Enolase	AI	AII	0.192	0.689	12.086	0.533	0.003	0.858	0.947	0.947	0.010	0.009
	AI	B	0.306	1.244	177.760	0.516	0.069	0.986	0.993	0.993	0.141	0.140
	AII	B	0.561	1.532	178.279	0.241	0.055	0.979	0.994	0.993	0.142	0.141
FBA	AI	AII	0.083	0.426	10.610	0.799	0.004	0.905	0.932	0.932	0.011	0.011
	AI	B	0.349	4.646	124.082	0.485	0.064	0.930	0.963	0.963	0.132	0.128
	AII	B	0.526	7.123	127.157	0.307	0.049	0.869	0.960	0.959	0.136	0.130
PFP-ALPHAI	AI	AII	0.213	0.736	16.048	0.526	0.004	0.896	0.919	0.918	0.010	0.009
	AI	B	0.176	1.496	295.903	0.695	0.089	0.990	0.995	0.994	0.181	0.180
	AII	B	0.424	3.025	293.796	0.333	0.077	0.977	0.991	0.990	0.180	0.178
PGK	AI	AII	0.138	0.795	9.667	0.675	0.002	0.780	0.950	0.950	0.008	0.008
	AI	B	0.467	2.942	216.423	0.366	0.088	0.974	0.988	0.987	0.178	0.176
	AII	B	0.651	3.995	217.676	0.233	0.052	0.942	0.990	0.989	0.179	0.177
GDH	AI	AII	0.251	1.342	18.435	0.443	0.004	0.829	0.876	0.875	0.014	0.012
	AI	B	0.402	1.580	152.200	0.423	0.056	0.980	0.990	0.989	0.113	0.112
	AII	B	0.724	3.149	150.617	0.127	0.042	0.949	0.979	0.977	0.112	0.109
NADP-ME	AI	AII	0.181	2.074	15.462	0.572	0.003	0.703	0.745	0.745	0.010	0.007
	AI	B	0.397	5.064	340.229	0.432	0.103	0.972	0.987	0.985	0.209	0.206
	AII	B	0.702	9.445	342.710	0.159	0.078	0.933	0.978	0.974	0.211	0.205
SPT	AI	AII	0.167	1.702	22.629	0.603	0.005	0.827	0.858	0.858	0.014	0.012
	AI	B	0.110	0.423	350.307	0.801	0.107	0.998	0.999	0.999	0.213	0.213
	AII	B	0.246	2.125	349.750	0.493	0.079	0.985	0.991	0.990	0.213	0.211
TPI	AI	AII	0.137	0.178	6.500	0.641	0.003	0.935	0.947	0.946	0.008	0.008
	AI	B	0.245	0.701	146.086	0.607	0.094	0.991	0.996	0.995	0.189	0.188
	AII	B	0.513	1.222	144.529	0.257	0.071	0.979	0.994	0.993	0.187	0.185

Hs: haplotype-based statistic; Ks: statistic based on nucleotide sequences; Kxy: average proportion of nucleotide differences between populations; Gst: genetic differentiation index based on the frequency of haplotypes; Dxy: average number of nucleotide substitutions per site between populations; Da: net nucleotide substitutions per site between populations.

2.2. *Analysis of New Loci Using* G intestinalis *in Stool Samples*

2.2.1. Stool Samples

We randomly chose 24 samples positive for *G intestinalis*, collected in different regions of Colombia, from Amazonas [53], Casanare, Bolívar, and Córdoba, as reported elsewhere [54].

2.2.2. Amplification of New Loci in DNA Samples

We experimentally evaluated primers designed on loci analysed in silico. Initially, we tested with DNA extracted from *G intestinalis* axenic culture. All primers adequately amplified corresponding regions with each locus. A single band of the expected size was obtained for each marker (Supplementary Materials, Table S1). We then tested primers for each locus with a small set (n = 24) of positive samples. Of the 24 samples, 95.8% (n = 23) amplified *GDH*; 83.3% (n = 20) amplified *TPI*; 66.7% (n = 16), *ACS*; 29.2% (n = 7) *SPT*, and 50.0% (n = 12), *enolase*. Many other samples amplified the target genes, and a band of the expected size was evident. However, other bands of different sizes were also observed. Also, concentrations of products obtained after the polymerase chain reaction (PCR) were low in some cases, showing bands so thin that it was impossible to obtain results from these sequences or no amplification occurred. Poor quality electropherograms were obtained for the *NADP-ME* gene, and sequences for this gene were ignored. Finally, five genes (*GDH*, *TPI*, *ACS*, *SPT* and *enolase*) were evaluated with DNA from the 24 stool samples. At least three markers were amplified in each sample,

and all five loci were amplified in a few samples. (Supplementary Materials, Table S8). These results were not included in the MLST analysis.

2.2.3. Genetic Diversity and Phylogenetic Reconstruction

Diversity indices for *ACS*, *SPT*, *enolase*, *GDH* and *TPI* loci were calculated with sequences obtained from stool samples, together with consensus sequences from the in silico analysis (Table 1). The *ACS* locus has a higher number of polymorphic sites (S = 278), compared to other loci, and the *TPI* locus the lowest number (S = 133), even though it displays the largest number of sites analysed. Also, nucleotide diversity, Pi, for the *ACS* gene was 0.210, followed by *SPT* and *TPI* loci with estimates of 0.114 and 0.109, respectively. Similarly, the *ACS* locus presented the highest value for Theta (per site) from Eta. Thus, locus that shows the greatest diversity is *ACS*. It is important to highlight that calculation of D Tajima as a neutrality test produced almost universally negative values. Not all values statistically significant, but still may indicate population expansion, at least for those loci, D Tajima for the *TPI* locus was positive, suggesting balancing selection. Another index calculated for this data set, corresponding to Rm, indicates a high value (Rm = 52) for the *TPI* locus compared to other target genes (Table 1).

Subsequently, we prepared a phylogenetic reconstruction by gene that included the stool samples. *G intestinalis* from stool samples form a cluster different from other sequences, except for the *TPI* gene. Three clusters corresponding to assemblages previously identified in silico analysis are observed (Figure 4). However, within the cluster corresponding to assemblage B, consensus sequences from the WGS data are closely grouped, and stool samples sequences are somewhat more distant. These observations suggest genetic differences that become more evident when other genes are included.

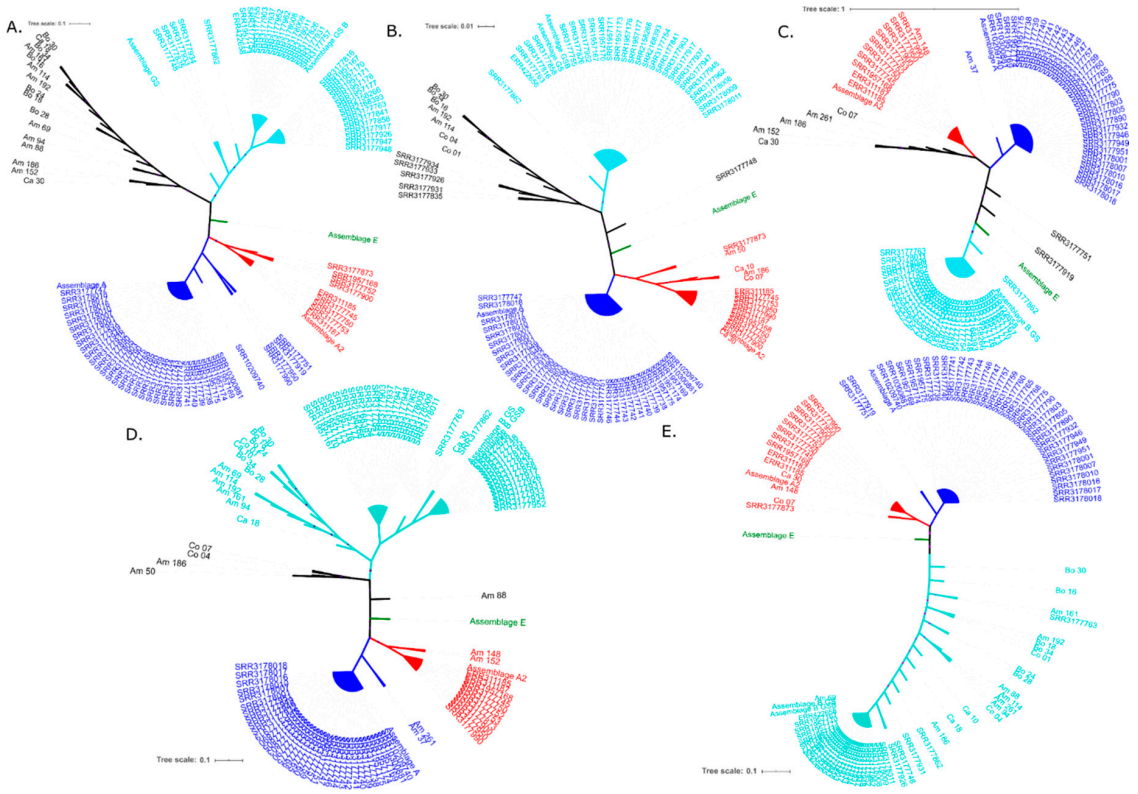


Figure 4. Gene trees including sequences obtained from human faeces as well as sequences extracted from public genomes. (**A**) *ACS*. (**B**) *Enolase*. (**C**) *SPT*. (**D**) *GDH*. (**E**) *TPI*. Phylogenetic inferences were constructed using maximum likelihood (ML) under the Jukes Cantor nucleotide evolution model, with 1000 bootstrap iterations. The tree was constructed with FastTree [45] software and visualized

with ITOL [46] software. A Bootstrap value greater than 80% is represented with a purple circle above each node. The colours indicate the assemblage to which the evaluated sequences belong (Blue: AI. Red: AII. Turquoise: B. Green: E and ND: not defined). Access numbers of the genomes and origins of stool samples are indicated (AM: Amazonas. BO: Bolívar. CA: Casanare CO: Córdoba).

To explore possible explanations for the topology change in the obtained phylogenies, we constructed phylogenetic networks. Possible reticulation events are observed in the phylogenetic networks between DNA sequences derived from stool samples, both intra and inter-assemblage. These events are primarily associated with *ACS*, *GDH* and, to a lesser extent, *SPT* loci (Figure 5). *TPI* gene sequences could not discriminate between clusters of the AI and AII sub-assemblages, in contrast to other markers.

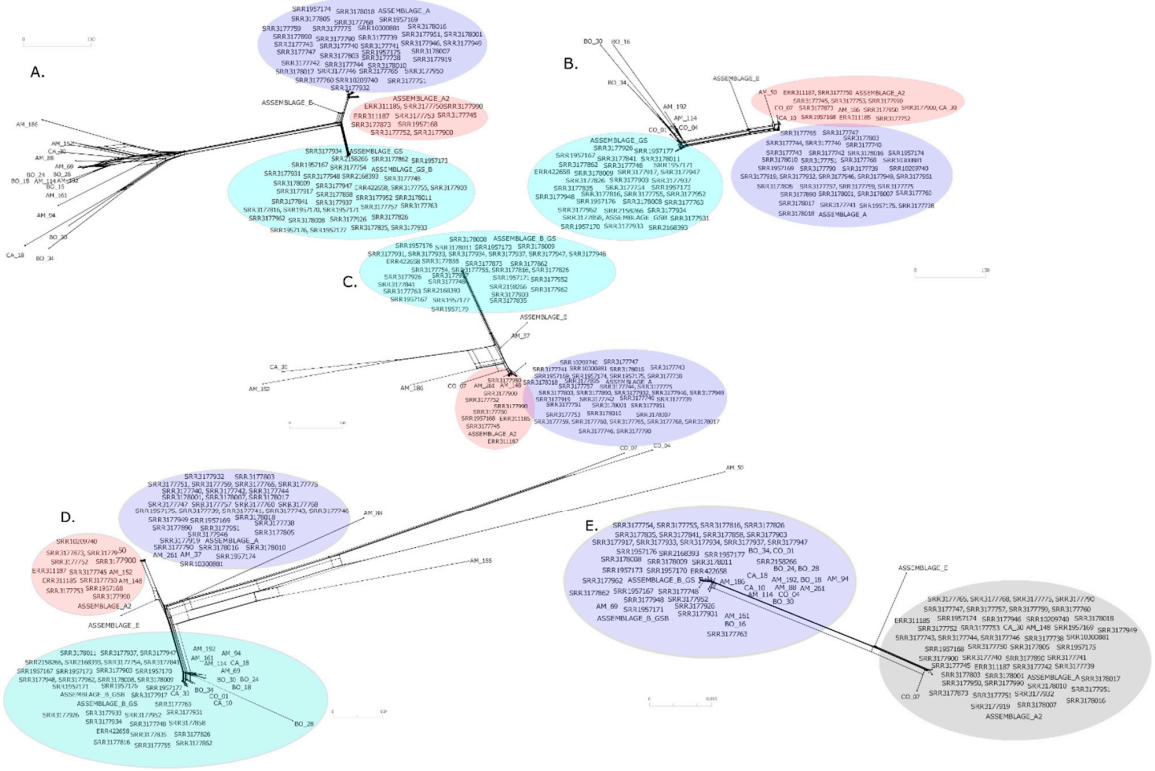


Figure 5. Phylogenetic network with SplitsTree [47] software, with sequences obtained from human faeces. (**A**) *ACS*. (**B**) *Enolase*. (**C**) *SPT*. (**D**) *GDH*. (**E**) *TPI*. Networks were built with the Neighbornet algorithm. Colours correspond to the assemblages (Blue: AI. Red: AII. Turquoise: B. Green: E. and ND: not defined). For the TPI gene, Assemblage A is highlighted in grey because AI and AII sub-assemblages are not discriminated. Access numbers of public genomes and origins of stool samples are indicated (AM: Amazonas. BO: Bolívar. CA: Casanare CO: Córdoba).

3. Discussion

The genes identified in the present study are useful for analysing calculated TE and DP. TE, an indicator of grouping of members with common characteristics, showed that proposed target genes are adequate to detect assemblages commonly found in human samples. DP, which allows differentiation of individuals belonging to different groups [44], showed that those genes are sufficient for identifying individuals that are slightly divergent from assemblages A and B. Considering these two parameters, implementing a typing scheme based on several loci is crucial. Markers must be adequate to assign an isolate to a "sequence type" or ST and powerful enough to differentiate one sequence type from another, but without discriminating to the point where each sequence becomes a different sequence type [55]. The latter could overestimate diversity and generate several STs. These results would make it difficult to establish phylogenetic relationships with epidemiological

factors, as observed with *Candida albicans* [40]. Further, each ST corresponds to a relatively recent lineage, reflecting changes in the accessory genome, for example, a gene acquired parasexually [55]. Thus, we consider the inclusion of these new loci evaluated in this study useful for both typing and for studying the divergence within and between *Giardia* assemblages.

For *G intestinalis*, no database built using a MLST scheme is available. Most studies, not surprisingly, use the same typing genes. The use of multiple loci to evaluate genetic characteristics of a microorganism has great advantages, such as ease of accessibility, basic bioinformatic requirements, and exponential enrichment of freely accessible databases [39]. Genomic data for *G intestinalis* are available [56] and diversity studies have acquired information on regions of the genome that allow typing with sufficient DP [57]. In particular, we found that a combination of six different loci allows detection of up to 51 different alleles (Supplementary Materials, Table S7). Further, a combination of nine loci shows high DP followed by *ACS* and *PGK* markers (Table 2), though a greater number of polymorphisms were found with *ACS* and *NADP-ME*. However, no delimited genotypes or subgroups were observed among any established assemblages, consistent with previously reported MLGs using genes typically used for typing [58] and the previously reported MLST [24]. This background highlights the need for additional molecular targets to fully characterise the genetics of *G intestinalis*. Genes in other regions of the genome will provide a comprehensive understanding of genetic diversity and genotypes.

Interestingly, in addition to diversity found in public sequences used for the MLST, haplotypic diversity was increased by including stool samples in the calculation of diversity indices. This result indicates that field samples are diverse in comparison to genomes evaluated in silico, as observed at the phylogenetic level. The finding likely reflects multiple sources of parasite transmission in the areas where samples were collected. Transmission may be affected by socioeconomic conditions in sampled populations [53], high dynamism of metropolitan areas, and high contact rates with different infected hosts, including symptomatic and asymptomatic patients, with different *G intestinalis* assemblages.

Evolutionary history was evaluated using neutrality analysis with the Tajima D test. We obtained positive D Tajima values for all loci analysing all genome sequences as a single population. From this approach, we can infer that the frequency of new alleles is low and that the population may be under balancing selection and contracting (Table 1). However, the assemblage analysis for all loci of the AI assemblage and the majority of loci of the assemblage B showed that the population is expanding (D Tajima < 0) (Supplementary Materials, Table S4; Table 1), which may be due to oral-faecal transmission that facilitates spread of cysts to new hosts [30]. Movement between hosts is crucial for gene flow and spread of rare alleles [59]. An expanding population was also inferred from results using stool samples. The use of constitutive loci allowed us to establish substantial diversity in each assemblage and population in general, despite the expectation that genes used are well conserved. These genes also allowed us to elucidate the evolutionary history of loci and the concatenation of all loci. Most values were significant, indicating that mutations may affect microorganism function and respond to selective pressure, as seen in studies using bacteria under an MLST scheme [60–62]. However, confirmation of this hypothesis will require increasing the number of individuals evaluated per assemblage to identify all evolutionary trends in *Giardia*.

Using phylogenetic analysis, we were able in most cases assign an assemblage for each consensus sequence in the in silico analysis, using loci independently and the concatenate of all loci. However, some sequences showed variations and did not clearly group with an assemblage. In some instances, clustering changed depending on the evaluated gene, showing confusion in delimiting criteria for assemblages (Figures 1 and 3 and Supplementary Materials, Figures S1–S9). Our numbers of sequences for the in silico analysis was limited to 80, and the number of inconsistencies represents an important finding that must be considered when evaluating established assemblages. Sequences inclusion of stool samples increased the number of phylogenetic inconsistencies and diversity (Figure 4 and Table 1). Also, we found several subgroups within clusters, mainly for assemblage B, that likewise vary depending on the evaluated gene. Thus, substantial diversity exists within the assemblage. Such diversity

could be related to a greater virulence [13] and evasion of host immunity [63] or may reflect the high heterozygosity found in this assemblage, how has been proposed in different studies of pathogenesis. An increase in the number of sequences for all assemblages, including non-A/B/E assemblages, would add certainty to intra-assemblage diversity. Our findings suggest reconsidering the classification of *G intestinalis* exclusively in these assemblages, because of inconsistencies within clusters, high diversity, and alternative assemblage/cluster assignments depending on the analysed locus.

When comparing ML trees results with phylogenetic networks, each assemblage shows important divergence with respect to AII. Such divergences in notable for A and B, and to a lesser extent, AI. Further, we observed reticulation signals and possible genetic exchange among assemblages, mainly between A and E, and for AI and AII (Figures 1B and 5 and Supplementary Materials, Figures S1B–S9B). Inconsistencies in topologies of phylogenetic trees were previously reported [19]. These results suggest genetic exchanges, between isolates, and also with other microorganisms such as bacteria [56]. Recombination events evaluated with genomic data [29] are proposed, consistent with sequence analyses of several loci, mainly *GDH*, *TPI*, *β-giardin* and *small subunit ribosomal ribonucleic acid (ssurRNA)* [13]. Other genes with greater variability have also been proposed [24], based initially on changes in topologies of phylogenetic analyses [24,26,33,64]. Results obtained have opened debate on taxonomy and cell division processes in *G intestinalis* [65], considering that, as a member of the diplomonads, it is typically asexual [66]. Also, changes in phylogenetic topology may be attributable to inadequate sampling, limited divergence, hybridisation, cryptic speciation with undocumented phenotypic differences, and incomplete lineage sorting (ILS) [67,68].

Further, admixture was observed among assemblages with the contribution of E assemblage alleles to AI, AII and B assemblages (Figure 2). Mixtures are mainly due to recombination events based on results with RDP (Table 3). Still, other genetic contributions should not be ruled out. The allelic plot shows, for some of sequences, no agreement between assigned assemblage and evaluated markers (Figure 3). No assemblage could be established for some genes with sufficient support (bootstrap > 80). These genes may belong in assemblages not evaluated in our study or may be the product of events that influence evolutionary dynamics of populations studied. Adaptive traits may be transferred that promote divergence due to events such as recombination, introgression or hybridisation [69]. Population structure statistics, such as the Fst were consistent with divergence among assemblages observed in phylogenetic networks (Figure 1B). Values higher than 0.8 were found for genetic structure among populations. We did not evaluate structure by geographic distances, and our results agree with studies based on haplotypic networks with the *TPI* gene from different continents. Genetic differentiation is reported between assemblage A and populations in Asia, Australia, and America. Moderate genetic differentiation is also seen with comparison using assemblages B and E. The latter case shows the dispersion of the same population of *G intestinalis* [30]. We consider our results robust in support of the utility of proposed new loci to type *G intestinalis*, and for investigating diversity, evolution, genetic structure and plausible genetic exchange events.

4. Materials and Methods

4.1. Selection of New Genetic Markers and Design of Primers

Using sequenced, curated, and annotated genomes of *G intestinalis* in the EuPathDB database: The Eukaryotic Pathogen Genomics Resource (GiardiaDB) (https://giardiadb.org/) (Supplementary Materials, Table S1), we searched for genes encoding constitutive enzymes that participate in metabolic processes, mainly glycolysis, alcoholic fermentation and TCA. Not all enzymes in the latter two are present in *G intestinalis*. These proteins are highly conserved in eukaryotes, and most enzymes are reported for this microorganism [70]. Other enzymes in the sphingolipid biosynthesis pathway were also evaluated [71]. Specifically, we selected the following genetic targets from the glycolysis cycle: *Glucose-6-phosphate isomerase (GPI), Pyrophosphate-fructose 6-phosphate 1-phosphotransferase alpha subunit (PFP-ALPHA1), Fructose-bisphosphate aldolase (FBA), Phosphoglycerate kinase (PGK)* and *enolase*. We chose

Acetyl-CoA synthetase (ACS) from the end of the alcoholic fermentation process, *NADP-dependent malic enzyme (NADP-ME)* from the TCA cycle and *Serine palmitoyltransferase 2 (SPT)* from lipid synthesis (Supplementary Materials, Table S1).

For bioinformatic design of primers, CDS sequences of orthologous and synthetic genes for these enzymes were downloaded for AI (isolate WB), AII (isolate DH), B (isolates GS and GS_B) and E (isolate P15) assemblages of eight genes (Supplementary Materials, Table S1). We aligned sequences of assemblages for each gene, using MUSCLE [72] implemented in MEGA 7.0. (Pennsylvania State University, PA, USA) [73]. We focused on identifying conserved regions between assemblages. Primers were designed using the Primer-BLAST tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) and were chosen considering: (1) primers were located in conserved regions for all assemblages, (2) the final product had a size between 300 and 700 bp, (3) selected markers were single-copy genes in the *G intestinalis* genome, and (4) information on dimers, formation of forks, melting temperature (Tm), per cent GC, size of initiator, size of the amplified region, and specificity were available. Single-copy genes were needed to estimate intra- and inter-assemblage allelic diversity, avoiding bias by recombination between copies of the genes. The fourth criterion was verified using Basic Local Alignment Search Tool (BLASTn).

Two approaches were used to verify single-copy genes. First, the Ortholog Groups of Protein Sequences (OrthoMCL DB) database, available at https://orthomcl.org/, was queried. The corresponding group of orthologs was searched for each protein to identify any alternate AA sequences. In the second, we used the CD-HIT comparative analysis algorithm, a tool for grouping and comparing biological sequences [74]. Sequences of each gene and each assemblage, previously downloaded from giardiadb.org in fasta format, were evaluated (Supplementary Materials, Table S1) and compared with transcript sequences available in NCBI. We looked numbers of sequences found for each gene with an identity ≥ 90% and Kmer = 2. Numbers of sequences found using these criteria within the same cluster were considered as possible variations in the number of gene copies.

4.2. In Silico Evaluation of Ten Genetic Markers

A total of 10 loci were evaluated in 130 *G intestinalis* genomes available online. Among these loci, we included genes commonly used for typing of *G intestinalis*, glutamate dehydrogenase (GDH) and triose phosphate isomerase (TPI). Genes were used to evaluate genetic diversity, recombination, typing capacity, and discrimination of *G intestinalis* in whole-genome sequencing (WGS) data available in the public database of The European Nucleotide Archive (ENA) (Supplementary Materials, Table S3).

We verified downloaded genomes using Kraken 2 software [75], which assesses DNA short-read sequences with a database of genomes from eukaryotes. One advantage of this approach is high sensitivity and speed, along with Kmers used for alignment to classify reads at different taxonomic levels [75]. We used hits greater than 80% as a cut-off point for data from short reads, obtaining 80 verified genomes for *G intestinalis*. All other genomes were discarded (Supplementary Materials, Table S3).

Subsequently, we used the Short Read Sequencing Typing 2 (SRST2) tool [42] to extract genes of interest from the 80 selected WGS. This tool maps reads of each genome (sets of reads in fastq format) on a database of reference alleles in fasta format (the ten selected genes), to detect the presence of a gene or locus, and identify the allele that best matches the locus among all allelic sequences used for reference. The reference allele database was constructed using the sequences of each gene for each assemblage downloaded from EupathDB (UGA, Athens, GA, USA). Consensus sequences for each gene were obtained each public genome, together the STs were assigned, based on alleles that best matched each locus. The software did not yield consensus sequences for the GPI gene and it was eliminated from subsequent analyses.

4.3. Utility of Selected Loci for Typing

Initially, consensus sequences of each gene and their concatenation were aligned using the multiple sequence alignment programme MAFFT v7 (Suita, Osaka, Japan) [76]. Subsequently, MLSTest software (CONICET, Salta, Argentina) was used to calculate numbers of alleles, typing efficiency (TE) and discriminatory power (DP) with 95% CI [44]. Alignments of consensus sequences from WGS data for each of nine markers were included as input data. An optimisation scheme was used to show optimal numbers of loci with different possible combinations and allelic profiles.

4.4. Phylogenetic Inferences and Recombination Signals

Phylogenetic trees were constructed from alignment of consensus sequences from WGS data (Supplementary Materials, Table S3). Subsequently, a phylogenetic tree was constructed by gene and with concatenation of all genes, using maximum likelihood under the Jukes Cantor nucleotide evolution model. The analysis used 1000 bootstrap replicates in FastTree 2.1 [45]. Each cluster was defined with bootstrap values ≥ 80.0%. Visualisation and editing used the online tool, Interactive Tree Of Life V4 (http://itol.embl.de) [46]. Additionally, phylogenetic networks were constructed to detect recombination signals between evaluated genes. The analysis used the SplitsTree5 programme [47], with the Neighbornet algorithm and 1000 iterations.

Once assemblages were established for each consensus sequence with each loci, an allelic plot schema was constructed with concatenated sequences and by gene. Each assemblage was represented in a different colour, the AI in blue, the AII in red, the B in turquoise, the E in green, and sequences that did not correspond to any assemblage were left in black. The scheme compares assemblages assigned to each consensus sequence with each gene, so that assemblage agreement among loci is represented. This scheme represents alleles found, such that the number of colours in the allelic plot represents the number of clusters discriminated by each marker, as reported in other studies [77].

To verify the existence of recombination events, we performed an additional analysis with the Recombination Detection Programme version 4 (RDP4) [51], using the alignment of concatenated sequences for the nine genes. RDP, GENECONV, BOOTSCAN, MaxChi Square (MaxChi), CHIMAERA, SISCAN and 3SEQ [78] were used, and recombination events described by multiple methods represent more robust results. To predict genetic admixture signals, we used the STRUCTURE 2.3.4 programme [49]. The number of established populations was K = 4, based on the four assemblages evaluated (AI, AII, B and E). We used 600,000 iterations of the Markov chain Monte Carlo algorithm with a Length of Burn-in Period of 60,000 iterations.

4.5. Indices of Diversity and Genetic Structure

All sequences in WGS data were used to calculate diversity indices, for each gene and concatenated sequence. Indices were also calculated with sequences grouped by assemblage. Input data in the DnaSP v.5 [48] programme (http://www.ub.edu/dnasp) were sequences aligned for each marker. Indices used included nucleotide diversity (Pi)—the average number of nucleotide differences per site between a pair of DNA sequences; Theta (per site) from the total number of mutations (Eta); numbers of polymorphic (segregating) sites (S); numbers of haplotypes (h); and haplotypic diversity (Hd). The latter index indicates the probability that two random haplotypes are different. Tajima D was calculated to determine if sequences evaluated reflected neutral variation or were involved in a selection process. This index indicates a balancing selection for positive values and a purifying selection for negative values [43]. The minimum number of recombination events (Rm) was also estimated. Some indices, such as haplotypic diversity and nucleotide diversity, are reported with their respective standard deviation.

Separate statistics for genetic differentiation among assemblages for each gene and concatenated sequence were also calculated. Assemblages that could not be defined for some sequences in some genes were assessed, as was the concatenated sequence from the E assemblage. Genetic differences

were estimated using statistics based on haplotypes (Hs), nucleotide sequences (Ks), and several others that reflect gene flow from nucleotide sequences, including Wright's F (Fst), Delta ST, Gamma ST, and Nst. Average number of nucleotide differences in pairs (Kxy), nucleotide substitutions per site (Dxy), net nucleotide substitutions per site (Da), and gene flow from haplotypes (Gst) were then calculated. DnaSP v.5 software [48] was used for the analysis (http://www.ub.edu/dnasp).

4.6. Assays from Human Stool Samples

4.6.1. Ethical Statement

All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of the National University of Colombia (002-012-15 February 12, 2015) and the ethics committee of the Universidad del Rosario (registered in Act No. 394 of the CEI-UR). This project was conducted under the contract number RGE131 of access to genetic resources granted by the "Ministerio de Medio ambiente y Desarrollo sostenible".

4.6.2. Study Population, Detection and Typing of *G intestinalis*

Convenience sampling of human faeces samples was performed in the departments of Amazonas, Bolívar, Casanare, and Córdoba in Colombia. The collection, extraction and typing of samples positive for *G intestinalis* was performed as described in Sánchez et al. [53] for samples from the Amazon and in Higuera et al. [54] for samples from Córdoba, Bolívar and Casanare. A small set of 24 DNAs from *G intestinalis* positive samples were taken at random to evaluate markers assessed in silico. The number of samples by location were: 12 from Amazonas, six from Bolívar, three from Casanare and three from Córdoba.

4.6.3. DNA Marker Assay of Stool Samples

We amplified each locus from DNA extracted from *G intestinalis* axenic cultures. PCR was performed in a final volume of 25 µL, containing 2 µL of quenched DNA, 12.5 µL of Go Taq Master Mix Green (Promega) (cat. No. M7122) (Madison, WIS, USA). at a final concentration of 1X and primers at a concentration of 1 µM each. Primers used for PCR are shown in Table 1, along with their respective expected band sizes. Thermal profile conditions for all loci were 95 °C for 5 min, 40 cycles of 95 °C for 1 min, 62 °C for 1 min, 72 °C for 1 min, and 10 min at 72 °C of final elongation. After verifying that all markers worked with control DNA, each *G intestinalis* positive stool sample was amplified using the above conditions. All PCR products were verified by observation on 2% agarose gel, stained with SYBR Safe, Thermo Fisher Scientific (cat No. S33102) (MA, CA, USA). Each PCR product was purified with ExoSAP-IT®, Affymetrix ™ (cat. No. 15513687) (Göteborg, Sweden) following the manufacturer's recommendations. Both chains of each product were sequenced with the Sanger method. Sequences were edited in MEGA 7.0 [73] to extract the fragment of interest. Once the sequence was cleaned, it was compared with publicly available sequences using the BLAST algorithm to verify that the fragment corresponded to the expected taxonomic unit.

4.6.4. Phylogenetic Reconstructions and Diversity Indices from Stool Samples

Trees and phylogenetic networks were constructed with sequences obtained for each locus. Consensus sequences from the SRST2 tool output were concatenated with sequences obtained from human faeces of Colombian origin. This process used the same procedures described above for phylogenetic reconstructions. Diversity indices were also calculated by gene with the Dnasp v.5 programme [48].

5. Conclusions

We highlight loci useful of an MLST scheme for typing of *G intestinalis*. These loci can also be used as alternatives to and supplements for genomic studies of *Giardia* diversity. We demonstrate intra-taxa diversity and show both genetic structures for established assemblages, and admixture among populations due to genetic exchange, apparently by recombination among individuals. Analysis of proposed loci should extend to future studies that include genomic sequences of additional assemblages to define their diversity and population structure. Further, future studies might focus on increasing the number of samples to evaluate markers on a large scale and extend sampling and analysis to other hosts and water sources that may be sources/reservoirs of infection. Such investigations will help elucidate transmission dynamics of the pathogen. Finally, future studies should assess a broader set of stool samples using nested PCR to examine the usefulness of genetic markers developed in the present study.

Availability of Data and Materials: Sequences data that support the findings of this study were deposited in GenBank with accession codes: MN877659–MN877686, MN877687 and MN877710. The accession numbers for ACS, Enolase and SPT are MT499125–MT499159.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4425/11/7/764/s1. Figure S1: Phylogenetic reconstructions for the GDH locus from public data. (A). Phylogenetic tree built from concatenation of all genes evaluated. (B). Splits tree built with the Neighbornet algorithm; Figure S2: Phylogenetic reconstructions for the ACS locus from public data. (A). Phylogenetic tree built from concatenation of all genes evaluated. (B). Splits tree was built with the Neighbornet algorithm; Figure S3: Phylogenetic reconstructions for the enolase locus from public data. (A). Phylogenetic tree built from concatenation of all genes evaluated. (B). Splits tree built with the Neighbornet algorithm; Figure S4: Phylogenetic reconstructions for the FBA locus from public data. (A). Phylogenetic tree built from concatenation of all genes evaluated. (B). Splits tree built with the Neighbornet algorithm; Figure S5: Phylogenetic reconstructions for the PFP-ALPHA1 locus from public data. (A). Phylogenetic tree built from concatenation of all genes evaluated. (B). Splits tree built with the Neighbornet algorithm; Figure S6: Phylogenetic reconstructions for the PGK locus from public data. (A). Phylogenetic tree built from concatenation of all genes evaluated. (B). Splits tree built with the Neighbornet algorithm; Figure S7: Phylogenetic reconstructions for the NADP-ME locus from public data. (A). Phylogenetic tree built from concatenation of all genes evaluated. (B). Splits tree built with the Neighbornet algorithm; Figure S8: Phylogenetic reconstructions for the SPT locus from public data. (A). Phylogenetic tree built from concatenation of all genes evaluated. (B). Splits tree built with the Neighbornet algorithm; Figure S9: Phylogenetic reconstructions for the TPI locus from public data. (A). Phylogenetic tree built from concatenation of all genes evaluated. (B). Splits tree built with the Neighbornet algorithm; Figure S10: Graph of recombination score through alignment for all detected events. Events 1-17 are displayed. Table S1. Identification of GiardiaDB sequences and EC numbers of enzymes used as targets for the design of primers; Table S2. Sequences of primers designed for each gene evaluated; Table S3. Whole-genome sequencing information (WGS) data genomes available in the public database of The European Nucleotide Archive (ENA); Table S4. Diversity indices obtained for nine loci evaluated from the AI assemblage; Table S5. Diversity indices obtained for nine loci evaluated from the AII assemblage; Table S6. Diversity indices obtained for nine loci evaluated from assemblage B; Table S7. Analysis of scheme optimisation and the optimum number of loci; Table S8. Amplified loci for the 24 DNA samples evaluated.

Author Contributions: Conceptualisation. J.D.R. and A.H.; methodology. A.H.; software. M.M.; validation. A.H. and M.M.; formal analysis. M.M. and A.H.; investigation. A.H.; resources. J.D.R., M.C.L., P.R., P.U., O.V.; data curation. J.D.R.; writing—original draft preparation. A.H.; writing—review and editing. J.D.R.; visualisation. A.H.; supervision. J.D.R.; project administration. J.D.R.; funding acquisition. J.D.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Departamento Administrativo de Ciencia, Tecnología e Innovación (Colciencias) through project "Desarrollo de una estrategia y plataforma de Multilocus Sequence Typing (MLST) para la caracterización molecular de siete parásitos intestinales," code 122271250521 and contract number 130–2016 granted to Universidad del Rosario and Universidad Nacional de Colombia. The APC was funded by Dirección de Investigación e Innovación from Universidad del Rosario. Adriana Higuera was financed by the Departamento Administrativo de Ciencia, Tecnología e Innovación (Colciencias) within the framework of the National Program for Promoting Research Training (sponsorship call 757).

Acknowledgments: We thank Angie Johana Sánchez for her help in the molecular detection of *G intestinalis* in the faecal samples from Amazonas. We thank Dirección de Investigación e Innovación from Universidad del Rosario for providing the english editing service of this manuscript. We also thank the High Computing Cluster (HPC) from Universidad del Rosario for the service provided to conduct bioinformatic analyses.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; the collection of samples, analyses, or interpretation of data; the writing of the manuscript, or the decision to publish the results.

References

1. Haque, R. Human Intestinal Parasites. *J. Health Popul. Nutr.* **2007**, *25*, 387–391.
2. CDC, Centers for Disease Control and Prevention. Parasites—Giardia. Available online: https://www.cdc.gov/parasites/giardia/infection-sources.html (accessed on 11 May 2020).
3. Thompson, R.C.; Monis, P.T. Variation in Giardia: Implications for taxonomy and epidemiology. *Adv. Parasitol.* **2004**, *58*, 69–137. [CrossRef]
4. Appelbee, A.J.; Thompson, R.C.; Olson, M.E. Giardia and Cryptosporidium in mammalian wildlife—Current status and future needs. *Trends Parasitol.* **2005**, *21*, 370–376. [CrossRef]
5. Thompson, R.C.; Ash, A. Molecular epidemiology of Giardia and Cryptosporidium infections. *Infect. Genet. Evol.* **2016**, *40*, 315–323. [CrossRef]
6. Abaza, S.M.; Sullivan, J.J.; Visvesvara, G.S. Isoenzyme profiles of four strains of Giardia lamblia and their infectivity to jirds. *Am. J. Trop. Med. Hyg.* **1991**, *44*, 63–68. [CrossRef]
7. Andrews, R.H.; Adams, M.; Boreham, P.F.; Mayrhofer, G.; Meloni, B.P. Giardia intestinalis: Electrophoretic evidence for a species complex. *Int. J. Parasitol.* **1989**, *19*, 183–190. [CrossRef]
8. Homan, W.L.; van Enckevort, F.H.; Limper, L.; van Eys, G.J.; Schoone, G.J.; Kasprzak, W.; Majewska, A.C.; van Knapen, F. Comparison of Giardia isolates from different laboratories by isoenzyme analysis and recombinant DNA probes. *Parasitol. Res.* **1992**, *78*, 316–323. [CrossRef]
9. Nash, T. Surface antigen variability and variation in Giardia lamblia. *Parasitol. Today* **1992**, *8*, 229–234. [CrossRef]
10. Weiss, J.B.; van Keulen, H.; Nash, T.E. Classification of subgroups of Giardia lamblia based upon ribosomal RNA gene sequence using the polymerase chain reaction. *Mol. Biochem. Parasitol.* **1992**, *54*, 73–86. [CrossRef]
11. Wielinga, C.M.; Thompson, R.C. Comparative evaluation of Giardia duodenalis sequence data. *Parasitology* **2007**, *134*, 1795–1821. [CrossRef]
12. Adam, R.D. Biology of Giardia lamblia. *Clin. Microbiol. Rev.* **2001**, *14*, 447–475. [CrossRef] [PubMed]
13. Feng, Y.; Xiao, L. Zoonotic potential and molecular epidemiology of Giardia species and giardiasis. *Clin. Microbiol. Rev.* **2011**, *24*, 110–140. [CrossRef]
14. Faria, C.P.; Zanini, G.M.; Dias, G.S.; da Silva, S.; Sousa, M.D.C. New multilocus genotypes of Giardia lamblia human isolates. *Infect. Genet. Evol.* **2017**, *54*, 128–137. [CrossRef]
15. Read, C.M.; Monis, P.T.; Thompson, R.A. Discrimination of all genotypes of Giardia duodenalis at the glutamate dehydrogenase locus using PCR-RFLP. *Infect. Genet. Evol.* **2004**, *4*, 125–130. [CrossRef]
16. Armson, A.; Yang, R.; Thompson, J.; Johnson, J.; Reid, S.; Ryan, U.M. Giardia genotypes in pigs in Western Australia: Prevalence and association with diarrhea. *Exp. Parasitol.* **2009**, *121*, 381–383. [CrossRef]
17. Lebbad, M.; Mattsson, J.G.; Christensson, B.; Ljungström, B.; Backhans, A.; Andersson, J.O.; Svärd, S.G. From mouse to moose: Multilocus genotyping of Giardia isolates from various animal species. *Vet. Parasitol.* **2010**, *168*, 231–239. [CrossRef]
18. Palmer, C.S.; Traub, R.J.; Robertson, I.D.; Devlin, G.; Rees, R.; Thompson, R.A. Determining the zoonotic significance of Giardia and Cryptosporidium in Australian dogs and cats. *Vet. Parasitol.* **2008**, *154*, 142–147. [CrossRef]
19. Sprong, H.; Cacciò, S.M.; van der Giessen, J.W.B.; ZOOPNET Network and Partners. Identification of Zoonotic Genotypes of Giardia duodenalis. *PLoS Negl. Trop. Dis.* **2009**, *3*, e558. [CrossRef]
20. Ryan, U.; Cacciò, S.M. Zoonotic potential of Giardia. *Int. J. Parasitol.* **2013**, *43*, 943–956. [CrossRef]
21. Lee, M.F.; Auer, H.; Lindo, J.F.; Walochnik, J. Multilocus sequence analysis of Giardia spp. isolated from patients with diarrhea in Austria. *Parasitol. Res.* **2017**, *116*, 477–481. [CrossRef] [PubMed]
22. Plutzer, J.; Ongerth, J.; Karanis, P. Giardia taxonomy, phylogeny and epidemiology: Facts and open questions. *Int. J. Hyg. Environ. Health* **2010**, *213*, 321–333. [CrossRef] [PubMed]

23. Wang, H.; Zhao, G.; Chen, G.; Jian, F.; Zhang, S.; Feng, C.; Wang, R.; Jinfeng, Z.; Dong, H.; Hua, J.; et al. Multilocus genotyping of Giardia duodenalis in dairy cattle in Henan, China. *PLoS ONE* **2014**, *9*, e100453. [CrossRef]
24. Ankarklev, J.; Lebbad, M.; Einarsson, E.; Franzén, O.; Ahola, H.; Troell, K.; Svärd, S.G. A novel high-resolution multilocus sequence typing of Giardia intestinalis Assemblage a isolates reveals zoonotic transmission, clonal outbreaks and recombination. *Infect. Genet. Evol.* **2018**, *60*, 7–16. [CrossRef]
25. Monis, P.T.; Andrews, R.H.; Mayrhofer, G.; Ey, P.L. Molecular systematics of the parasitic protozoan Giardia intestinalis. *Mol. Biol. Evol.* **1999**, *16*, 1135–1144. [CrossRef]
26. Lasek-Nesselquist, E.; Mark Welch, D.M.; Sogin, M.L. The identification of a new Giardia duodenalis assemblage in marine vertebrates and a preliminary analysis of G. duodenalis population biology in marine systems. *Int. J. Parasitol.* **2010**, *40*, 1063–1074. [CrossRef]
27. Poxleitner, M.K.; Carpenter, M.L.; Mancuso, J.J.; Wang, C.J.; Dawson, S.C.; Cande, W.Z. Evidence for karyogamy and exchange of genetic material in the binucleate intestinal parasite Giardia intestinalis. *Science* **2008**, *319*, 1530–1533. [CrossRef]
28. Teodorovic, S.; Braverman, J.M.; Elmendorf, H.G. Unusually Low Levels of Genetic Variation among Giardia lamblia Isolates. *Eukaryot. Cell* **2007**, *6*, 1421–1430. [CrossRef]
29. Xu, F.; Jerlström-Hultqvist, J.; Andersson, J.O. Genome-wide analyses of recombination suggest that Giardia intestinalis assemblages represent different species. *Mol. Biol. Evol.* **2012**, *29*, 2895–2898. [CrossRef]
30. Choy, S.H.; Mahdy, M.A.K.; Al-Mekhlafi, H.M.; Low, V.L.; Surin, J. Population expansion and gene flow in Giardia duodenalis as revealed by triosephosphate isomerase gene. *Parasites Vectors* **2015**, *8*, 454. [CrossRef]
31. Durigan, M.; Ciampi-Guillardi, M.; Rodrigues, R.C.A.; Greinert-Goulart, J.A.; Siqueira-Castro, I.C.V.; Leal, D.A.G.; Yamashiro, S.; Bonatti, T.R.; Zucchi, M.I.; Franco, R.M.; et al. Population genetic analysis of Giardia duodenalis: Genetic diversity and haplotype sharing between clinical and environmental sources. *MicrobiologyOpen* **2017**, *6*, e00424. [CrossRef]
32. Cooper, M.A.; Adam, R.D.; Worobey, M.; Sterling, C.R. Population genetics provides evidence for recombination in Giardia. *Curr. Biol.* **2007**, *17*, 1984–1988. [CrossRef] [PubMed]
33. Urwin, R.; Maiden, M.C. Multi-locus sequence typing: A tool for global epidemiology. *Trends Microbiol.* **2003**, *11*, 479–487. [CrossRef]
34. Sullivan, C.B.; Diggle, M.A.; Clarke, S.C. Multilocus sequence typing: Data analysis in clinical microbiology and public Health. *Mol. Biotechnol.* **2005**, *29*, 245–254. [CrossRef]
35. Diosque, P.; Tomasini, N.; Lauthier, J.J.; Messenger, L.A.; Monje Rumi, M.M.; Ragone, P.G.; Alberti-D'Amato, A.M.; Pérez Brandán, C.; Barnabé, C.; Tibayrenc, M.; et al. Optimised multilocus sequence typing (MLST) scheme for Trypanosoma cruzi. *PLoS Negl. Trop. Dis.* **2014**, *8*, e3117. [CrossRef]
36. Herrera, G.; Hernández, C.; Ayala, M.S.; Flórez, C.; Teherán, A.A.; Ramírez, J.D. Evaluation of a Multilocus Sequence Typing (MLST) scheme for Leishmania (Viannia) braziliensis and Leishmania (Viannia) panamensis in Colombia. *Parasites Vectors* **2017**, *10*, 236. [CrossRef]
37. Ramírez, J.D.; Tapia-Calle, G.; Guhl, F. Genetic structure of Trypanosoma cruzi in Colombia revealed by a High-throughput Nuclear Multilocus Sequence Typing (nMLST) approach. *BMC Genet.* **2013**, *14*, 96. [CrossRef]
38. Yeo, M.; Mauricio, I.L.; Messenger, L.A.; Lewis, M.D.; Llewellyn, M.S.; Acosta, N.; Bhattacharyya, T.; Diosque, P.; Carrasco, H.J.; Miles, M.A. Multilocus Sequence Typing (MLST) for Lineage Assignment and High Resolution Diversity Studies in Trypanosoma cruzi. *PLoS Negl. Trop. Dis.* **2011**, *5*, e1049. [CrossRef]
39. Muñoz, M.; Camargo, M.; Ramírez, J.D. Estimating the Intra-taxa diversity, population genetic structure, and evolutionary pathways of Cryptococcus neoformans and Cryptococcus gattii. *Front. Genet.* **2018**, *9*, 148. [CrossRef]
40. Muñoz, M.; Wintaco, M.; Muñoz, S.; Ramírez, J.D. Dissecting the Heterogeneous Population Genetic Structure of Candida albicans: Limitations and Constraints of the Multilocus Sequence Typing Scheme. *Front. Microbiol.* **2019**, *10*, 1052. [CrossRef]
41. Bougnoux, M.E.; Morand, S.; d'Enfert, C. Usefulness of multilocus sequence typing for characterisation of clinical isolates of Candida albicans. *J. Clin. Microbiol.* **2002**, *40*, 1290–1297. [CrossRef] [PubMed]
42. Inouye, M.; Dashnow, H.; Raven, L.A.; Schultz, M.B.; Pope, B.J.; Tomita, T.; Zobel, J.; Holt, K.E. SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med.* **2014**, *6*, 90. [CrossRef]

43. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **1989**, *123*, 585–595.
44. Tomasini, N.; Lauthier, J.J.; Llewellyn, M.S.; Diosque, P. MLSTest: Novel software for multilocus sequence data analysis in eukaryotic organisms. *Infect. Genet. Evol.* **2013**, *20*, 188–196. [CrossRef] [PubMed]
45. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **2009**, *26*, 1641–1650. [CrossRef] [PubMed]
46. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, W256–W259. [CrossRef] [PubMed]
47. Huson, D.H.; Bryant, D. Application of phylogenetic networks in evolutionary studies. *Mol. Boil. Evol.* **2006**, *23*, 254–267. [CrossRef]
48. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [CrossRef]
49. Hubisz, M.J.; Falush, D.; Stephens, M.; Pritchard, J.K. Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Resour.* **2009**, *9*, 1322–1332. [CrossRef]
50. Rieux, A.; Halkett, F.; De Lapeyre de Bellaire, L.; Zapater, M.F.; Rousset, F.; Ravigné, V.; Carlier, J. Inferences on pathogenic fungus population structures from microsatellite data: New insights from spatial genetics approaches. *Mol. Ecol.* **2011**, *20*, 1661–1674. [CrossRef]
51. Martin, D.P.; Murrell, B.; Golden, M.; Khoosal, A.; Muhire, B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol.* **2015**, *1*, vev003. [CrossRef]
52. Wright, S. *Evolution and the Genetics of Populations: A Treatise in Four Volumes: Vol. 4: Variability within and among Natural Populations*; University of Chicago Press: Chicago, IL, USA, 1978.
53. Sánchez, A.; Muñoz, M.; Gómez, N.; Tabares, J.; Segura, L.; Salazar, Á.; Restrepo, C.; Ruíz, M.; Reyes, P.; Qian, Y.; et al. Molecular epidemiology of Giardia, Blastocystis and Cryptosporidium among indigenous children from the Colombian Amazon Basin. *Front. Microbiol.* **2017**, *8*, 248. [CrossRef] [PubMed]
54. Higuera, A.; Villamizar, X.; Herrera, G.; Giraldo, J.C.; Vasquez-A, L.R.; Urbano, P.; Villalobos, O.; Tovar, C.; Ramírez, J.D. Molecular detection and genotyping of intestinal protozoa from different biogeographical regions of Colombia. *PeerJ* **2020**, *8*, e8554. [CrossRef] [PubMed]
55. Turner, K.M.; Feil, E.J. The secret life of the multilocus sequence type. *Int. J. Antimicrob. Agents* **2007**, *29*, 129–135. [CrossRef] [PubMed]
56. Jerlström-Hultqvist, J.; Ankarklev, J.; Svärd, S.G. Is human giardiasis caused by two different Giardia species? *Gut Microbes* **2010**, *1*, 379–382. [CrossRef]
57. Ankarklev, J.; Franzén, O.; Peirasmaki, D.; Jerlström-Hultqvist, J.; Lebbad, M.; Andersson, J.; Svärd, S.G. Comparative genomic analyses of freshly isolated Giardia intestinalis assemblage A isolates. *BMC Genom.* **2015**, *16*, 697. [CrossRef]
58. Cacciò, S.M.; Beck, R.; Lalle, M.; Marinculic, A.; Pozio, E. Multilocus genotyping of Giardia duodenalis reveals striking differences between assemblages A and B. *Int. J. Parasitol.* **2008**, *38*, 1523–1531. [CrossRef]
59. Blouin, M.S.; Yowell, C.A.; Courtney, C.H.; Dame, J.B. Host movement and the genetic structure of populations of parasitic nematodes. *Genetics* **1995**, *141*, 1007–1014.
60. Miyoshi-Akiyama, T.; Hayakawa, K.; Ohmagari, N.; Shimojima, M.; Kirikae, T. Multilocus sequence typing (MLST) for characterisation of Enterobacter cloacae. *PLoS ONE* **2013**, *8*. [CrossRef]
61. Chassain, B.; Lemée, L.; Didi, J.; Thiberge, J.M.; Brisse, S.; Pons, J.L.; Pestel-Caron, M. Multilocus sequence typing analysis of Staphylococcus lugdunensis implies a clonal population structure. *J. Clin. Microbiol.* **2012**, *50*, 3003–3009. [CrossRef]
62. Pérez-Losada, M.; Arenas, M.; Castro-Nallar, E. Multilocus sequence typing of pathogens. In *Book Genetics and Evolution of Infectious Disease*, 2nd ed.; Elsevier: Amsterdam, The Netherlands, 2017; pp. 383–404. [CrossRef]
63. Minetti, C.; Lamden, K.; Durband, C.; Cheesbrough, J.; Fox, A.; Wastling, J.M. Determination of Giardia duodenalis assemblages and multilocus genotypes in patients with sporadic giardiasis from England. *Parasites Vectors* **2015**, *8*, 444. [CrossRef]
64. Cacciò, S.M.; Ryan, U. Molecular epidemiology of giardiasis. *Mol. Biochem. Parasitol.* **2008**, *160*, 75–80. [CrossRef] [PubMed]
65. Carpenter, M.L.; Assaf, Z.J.; Gourguechon, S.; Cande, W.Z. Nuclear inheritance and genetic exchange without meiosis in the binucleate parasite Giardia intestinalis. *J. Cell Sci.* **2012**, *125*, 2523–2532. [CrossRef] [PubMed]
66. Birky, C.W., Jr. Sex: Is Giardia doing it in the dark? *Curr. Biol.* **2005**, *15*, R56–R58. [CrossRef] [PubMed]

67. Hart, M.W.; Sunday, J. Things fall apart: Biological species form unconnected parsimony networks. *Biol. Lett.* **2007**, *3*, 509–512. [CrossRef] [PubMed]
68. Tarcz, S.; Przyboś, E.; Surmacz, M. An assessment of haplotype variation in ribosomal and mitochondrial DNA fragments suggests incomplete lineage sorting in some species of the Paramecium aurelia complex (Ciliophora, Protozoa). *Mol. Phylogenet. Evol.* **2013**, *67*, 255–265. [CrossRef] [PubMed]
69. Detwiler, J.T.; Criscione, C.D. An infectious topic in reticulate evolution: Introgression and hybridisation in animal parasites. *Genes* **2010**, *1*, 102–123. [CrossRef]
70. Han, J.; Collins, L.J. Reconstruction of sugar metabolic pathways of Giardia lamblia. *Int. J. Proteom.* **2012**, *2012*, 980829. [CrossRef]
71. Hernández, Y.; Shpak, M.; Duarte, T.T.; Mendez, T.L.; Maldonado, R.A.; Roychowdhury, S.; Rodrigues, M.L.; Das, S. Novel role of sphingolipid synthesis genes in regulating giardial encystation. *Infect. Immun.* **2008**, *76*, 2939–2949. [CrossRef]
72. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [CrossRef]
73. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef]
74. Huang, Y.; Niu, B.; Gao, Y.; Fu, L.; Li, W. CD-HIT suite: A web server for clustering and comparing biological sequences. *Bioinformatics* **2010**, *26*, 680–682. [CrossRef] [PubMed]
75. Wood, D.E.; Lu, J.; Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **2019**, *20*, 257. [CrossRef] [PubMed]
76. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef] [PubMed]
77. Patiño, L.H.; Camargo, M.; Muñoz, M.; Ríos-Chaparro, D.I.; Patarroyo, M.A.; Ramírez, J.D. Unveiling the multilocus sequence typing (MLST) schemes and core genome phylogenies for genotyping Chlamydia trachomatis. *Front. Microbiol.* **2018**, *9*, 1854. [CrossRef]
78. Martin, D.P.; Williamson, C.; Posada, D. RDP2: Recombination detection and analysis from sequence alignments. *Bioinformatics* **2005**, *21*, 260–262. [CrossRef]

8.3 CAPÍTULO 3

Análisis de la diversidad y estructura genética de *Blastocystis*.

Heliyon

Succinate dehydrogenase gene as a marker for studying Blastocystis genetic diversity

--Manuscript Draft--

Manuscript Number:	
Article Type:	Original Research Article
Section/Category:	Microbiology
Keywords:	Blastocystis; ssu rRNA; succinate dehydrogenase subunit A; genetic diversity; genetic population structure
Manuscript Classifications:	110.130: Evolutionary Biology; 110.190: Microbiology; 110.280: Molecular Biology; 110.350: Zoology; 110.400: Epidemiology
Corresponding Author:	Juan David Ramírez Universidad Del Rosario COLOMBIA
First Author:	Adriana Higuera
Order of Authors:	Adriana Higuera
	Marina Muñóz
	Myriam Consuelo López
	Patricia Reyes
	Plutarco Urbano
	Oswaldo Villalobos
	Juan David Ramírez
Abstract:	Blastocystis has been reported as the most common eukaryotic microorganism residing in the intestines of both humans and animals, with a prevalence of up to 100% in some populations. Since this is a cryptic species, sequence polymorphisms are the single strategy to analyses its genetic diversity, being traditionally used the analysis of ssu rRNA gene sequence to determine alleles and subtypes (STs) for this species. This multicopy gene with non-identical copies has shown high diversity among different STs, making necessary to explore other genes to assess intraespecific diversity. This study evaluated the use of a novel genetic marker, succinate dehydrogenase (SDHA), for the typing and evaluation of the genetic diversity and genetic population structure of Blastocystis. In total, 375 human fecal samples were collected and subjected to PCR, subtyped using the ssu rRNA marker, and then the SDHA gene was amplified via PCR for 117 samples. We found some incongruences between tree topologies for both molecular markers. However, the clustering by ST previously stablished for Blastocystis was congruent in the concatenated sequence. SDHA showed lower reticulation signals and better intra ST clustering ability. Clusters with geographical associations were observed intra ST. The genetic diversity was lower in the marker evaluated compared to that of the ssu rRNA gene (nucleotide diversity = 0.03344 and 0.16986, respectively) and the sequences analyzed showed population expansion with genetic differentiation principally among STs. The ssu rRNA gene was useful to explore interspecific diversity but together with the SDHA gene the resolution power to evaluate intra ST diversity was higher. These results showed the potential of the SDHA marker for studying the intra ST genetic diversity of Blastocystis related with geographical location and the diversity inter ST using the concatenated sequences.
Suggested Reviewers:	Liliana Rojas lhily@yahoo.com
	Gustavo Vallejo gvallejo@ut.edu.co
	Graham Clark graham.clark@lshtm.ac.uk

Opposed Reviewers:	Rune Stensvold run@ssi.dk Conflict of interests

Bogotá D.C., August 3rd, 2020

Editor
Heliyon

Dear Editor,

Blastocystis is a cosmopolitan parasite that has been considered the most common eukaryote of the human intestine worldwide. To date, there is still controversies regarding the pathogenicity of this microorganisms where some authors suggest it is a commensal but others a parasite. The molecular epidemiology of *Blastocystis* is complex and it is accepted for the scientific community to use the ssuRNAr for the discrimination of 17 subtypes (STs) of this protozoan. However, it is well known in other microorganisms that one sole marker is not enough for an accurate understanding of the genetic diversity and population structure of parasites. Therefore, the continuous search for novel genetic markers are needed in terms of unveiling the biology, mechanisms of reproduction and evolution of parasitic protozoa as is the case of *Blastocystis.*

By means of this letter we intend to submit the manuscript entitled "**Succinate dehydrogenase (SDH) gene as a marker for estimating *Blastocystis* genetic diversity"** where we developed primers that amplified the SDH marker in fecal samples from different regions of Colombia. We found discrimination in the four main groups (using SDH sequences) of the samples evaluated, without geographical association, or STs. The genetic diversity was lower in the marker evaluated compared to that of the ssuRNAr gene (nucleotide diversity = 0.03143 and 0.15918, respectively) and the sequences analyzed showed a greater haplotypic diversity and population expansion without genetic differentiation. These results showed the potential of the SDH marker for studying the genetic diversity and structure of *Blastocystis* populations.

We believe this contribution might be an interesting manuscript for the wide audience HELIYON attracts.

My sincere regards,

Juan David Ramírez González, Ph.D.
Corresponding author
Faculty of Natural Sciences and Mathematics
Chief
Grupo de Investigaciones Microbiológicas – UR (GIMUR)
Universidad el Rosario

Succinate dehydrogenase gene as a marker for studying *Blastocystis* genetic diversity

Adriana Higuera[1], Marina Muñóz[1], Myriam Consuelo López[2], Patricia Reyes[2], Plutarco Urbano[3], Oswaldo Villalobos[4], Juan David Ramírez[1*]

[1]Grupo de Investigaciones Microbiológicas-UR (GIMUR), Departamento de Biología, Facultad de Ciencias Naturales, Universidad del Rosario, Bogotá, Colombia

[2] Departamento de Salud Pública Universidad Nacional de Colombia, Bogotá, Colombia

[3]Grupo de Investigaciones Biológicas de la Orinoquia, Unitrópico, Yopal, Colombia

[4]Hospital Local Santa María de Mompox, Programas Especiales (Lepra y TB), Mompox, Bolívar, Colombia

[*]Correspondence:

Juan David Ramírez

[1]Grupo de Investigaciones Microbiológicas-UR (GIMUR), Departamento de Biología, Facultad de Ciencias Naturales, Universidad del Rosario, Bogotá, Colombia

juand.ramirez@urosario.edu.co

Running title: **Succinate dehydrogenase gene for subtyping *Blastocystis***

Abstract

Blastocystis has been reported as the most common eukaryotic microorganism residing in the intestines of both humans and animals, with a prevalence of up to 100% in some populations. Since this is a cryptic species, sequence polymorphisms are the single strategy to analyses its genetic diversity, being traditionally used the analysis of *ssu rRNA* gene sequence to determine alleles and subtypes (STs) for this species. This multicopy gene with non-identical copies has shown high diversity among different STs, making necessary to explore other genes to assess intraespecific diversity. This study evaluated the use of a novel genetic marker, succinate dehydrogenase (*SDHA*), for the typing and evaluation of the genetic diversity and genetic population structure of *Blastocystis*. In total, 375 human fecal samples were collected and subjected to PCR, subtyped using the *ssu rRNA* marker, and then the *SDHA* gene was amplified via PCR for 117 samples. We found some incongruences between tree topologies for both molecular markers. However, the clustering by ST previously stablished for *Blastocystis* was congruent in the concatenated sequence. *SDHA* showed lower reticulation signals and better intra ST clustering ability. Clusters with geographical associations were observed intra ST. The genetic diversity was lower in the marker evaluated compared to that of the *ssu rRNA* gene (nucleotide diversity = 0.03344 and 0.16986, respectively) and the sequences analyzed showed population expansion with genetic differentiation principally among STs. The *ssu rRNA* gene was useful to explore interspecific diversity but together with the *SDHA* gene the resolution power to evaluate intra ST diversity was higher. These results showed the potential of the *SDHA* marker for studying the intra ST genetic diversity of *Blastocystis* related with geographical location and the diversity inter ST using the concatenated sequences.

Keywords: *Blastocystis*, *ssu rRNA*, succinate dehydrogenase subunit A, genetic diversity, genetic population structure

Background

Blastocystis spp., are anaerobic intestinal protists belonging to the phylum Heterokontophyta [1] of the Stramenopila group, which includes heterotrophic and photosynthetic protozoa [2]. *Blastocystis* has a cosmopolitan distribution [3, 4] and is the most common eukaryotic protozoan in the human intestine [5], with prevalence's up to 100% in a population of children in Senegal [6]. However, the role of this parasite at the intestinal level is still a matter of contention as it is present in both asymptomatic [7] and symptomatic patients, in the latter associated with inflammatory bowel disease (IBD), irritable bowel syndrome (IBS) [8] and chronic or acute urticarial lesions [9]. In some studies, it has been suggested that the potential pathogenic of *Blastocystis* could be related with the subtypes or STs [10-12], however, this association remains in debate. Also, the STs have been related to its geographical distribution [4,12,13] and a relative specificity to different hosts [14]. For all these, it is important to increase studies in molecular epidemiology in areas where the STs are circulating, to determine the impact of *Blastocystis* diversity and associate it with both biological and clinical factors. Initially, the genetic diversity of this microorganism had been demonstrated through a variety of different techniques [15-17]; however, it was not until the establishment of phylogeny with the use of the complete small subunit ribosomal RNA (*ssu rRNA*) sequence [18] and elongation factor 1 alpha [19], that the clusters corresponding to the STs were observed. Subsequently, with the use of genomic sequencing, that diversity and the differences between some STs with respect to their size, guanine-cytosine content, number of introns, and gene content were elucidated [20]. Currently, there are sequenced genomes for ST1,

ST4 and ST7 [20-22] and some drafts genomes for ST2, ST3, ST6, ST8 and ST9 available on GenBank, with sizes ranging between 12 -18 Mb.

Due to the existence of genomic sequencing and multiple improvements in the ease and access of new generation sequencing techniques today, the sequencing of just a few genes is now feasible. Therefore, DNA barcoding has been developed for *Blastocystis* and is frequently used by the scientific community to identify STs with the *ssu rRNA* gen [23]. With this unique marker, the genus had been classified into 17 STs [15,24], but currently 28 STs have been reported [25].The use of this marker to determine STs has revealed a high diversity among them [5, 25], so it has been proposed that each ST could correspond to a different species [26]. However, due to the increasing number of STs found, it would be important to add other markers with different resolution power to explore the diversity in other genetic targets and conjunction with *ssu rRNA*, solving the variation both inter and intraspecific , even more considering the heterogeneity of *ssu rRNA* which was reported in the ST7 strain B isolate [27], whit 17 non-identical copies [22, 27]. A few additional markers have also been evaluated, such as the pyruvate ferredoxin oxidoreductase (*PFOR*) gene, which allowed finding three clades with different ST samples, a lower nucleotide diversity and haplotype polymorphism for clade III [28]. Other study used the internal transcribed spacer *ITS* which revealed novel variants intra ST1 and a high flow gene among different countries of Europe and America [29]. However, many aspects of *Blastocystis* remain unknown, so it is necessary to develop additional markers that allow for typing of the genus and determining some aspects about their biology, evolutionary history, reproduction mechanisms, possible recombination, and population genetic events.

Recently, studies of the genome for *Blastocystis* demonstrated the existence of synteny between ST1, ST4, and ST7 [30,31], with the presence of genes corresponding to complex I and II of the electron transport chain and an absence of III and IV [32], where they also determined 4 subunits, encoded by nuclear DNA, of the mitochondrial respiratory chain complex II that could work via succinate dehydrogenase (SDH) or fumarate reductase [22]. These pathways reportedly reverse reactions in other protozoa, such as *Trypanosoma cruzi* [33]. In eukaryotes, such as yeasts, SDH is a key enzyme that catalyzes the passage of succinate to fumarate during the tricarboxylic acid (TCA) cycle [34], it basically has 4 subunits called A, B, C and D. Subunit A is strongly associated with the mitochondrial inner membrane and covalently bound to flavin adenine dinucleotide (FAD), B contains three Fe-S groups and C and D which are integrals membrane proteins [35]. Although SDH is a mitochondrial protein, it is encoded by a nuclear gene and part of its structure is conserved among eukaryotes. Because nuclear genes are less affected by deleterious mutations associated with asexual reproduction [36] and due to the importance of this enzyme, we wanted to establish if this gene could be informative and could be used as an additional marker in the subtyping and/or discrimination of *Blastocystis* subtypes creating a better understanding of the genetic diversity and population structure of this stramenopile. Therefore, the main objective of this study was to determine the usefulness of the succinate dehydrogenase gene of the subunit A (*SDHA*) as a possible additional marker in the subtyping, study of the genetic diversity and population genetic structure of *Blastocystis*.

Materials and methods

Selection of new genetic markers

A search was made of the genes corresponding to the constitutive enzymes, primarily those involved with the glycolysis cycle and a few from the Krebs cycle, using the genome available for ST7 in the *Blastocystis* Genome Browser (http://www.genoscope.cns.fr/). Subsequently, the larger sequences were downloaded into a FASTA format. The genes included in the search were *Triosephosphate isomerase (TPI), biphosphate aldolase (BPA), glucokinase (GK), Glucose 6 phosphate isomerase (GPI), Hexokinase (HK), Phosphofructokinase (PFK), Glyceraldehyde-3-phosphate dehydrogenase (GAPDH), Phosphoglycerate kinase (PGK), Phosphoglycerate mutase (PGM), Enolase (enol-1), Pyruvate kinase (PK)* and the subunits A and B de *Succinate dehydrogenase (SDHA, SDHB*).

Subsequently, these genes were searched in the annotation report for ST1 and ST4 available at GenBank. Then, an alignment was performed using MUSCLE [37] implemented in MEGA 7.0 [38] between the sequences downloaded from the annotated genomes and from each of the previously downloaded ST7 genes. All sequences of ST1, ST4 and ST7 were left at the same length. Once the sequences for ST1, ST4 and ST7 were aligned for each gene, conserved regions were determined, suitable for the design of primers.

For the particular case of the *SDH* gene, in the Genome Browser database we found sequences for subunits A and B, so we downloaded a sequence of 1,865 nt from subunit A and 855 nt for subunit B. The other *SDH* subunits were not specified in the database for ST7, so they were not taken into account. Conserved regions were not observed when aligning the sequences of the *SDH* subunit B of the three STs, therefore this gene was discarded for the design of primers. Then, we just used the subunit A (*SDHA*).

The resultant sequences were then used as a template for the design of the specific primers. These primers were designed using the Primer-BLAST tool. Several criteria were taken into account for each pair of primers such as dimer formation, fork formation, melting temperature (Tm), percentage of GC, size of each initiator, size of the amplified region, and specificity. The latter was verified through BLASTn, which ensured that a specific hit with 100% identity for *Blastocystis* was obtained. The amplification conditions for each pair of primers were standardized using the *Blastocystis* xenic culture DNA ST3.

Additionally, the Ortholog Groups of Protein Sequences (OrthoMCL DB) database available at https://orthomcl.org/ was used to explore the copy number of each of the genes, where a single copy was found for the *SDHA*, *GPI*, *PFK* and *PGM* genes.

PCR of control DNA with newly designed primers

Once the primers designed for each marker were obtained, a PCR with control DNA from the *Blastocystis* xenic culture ST3 was performed to verify the size of the amplified products. The reactions were performed in a final volume of 12.5 µL, 2 µL of DNA, 6.25 µl of Go Taq Master Mix Green (Promega) (cat. No. M7122) at a final concentration of 1X and at a final concentration for each primer of 1 µM. The thermal cycling parameters were as follows 95 ° C for 5 min, 40 cycles of 95 ° C for 1 min, 59 ° C for 1 min, 72 ° C for 1 min, and 72 ° C for 10 min. Each amplification reaction was observed in 2% agarose gels and stained with SYBR Safe, Thermo Fisher Scientific (cat No. S33102) verifying the presence of a single band and the expected size.

Subsequently, 4 of the 12 markers evaluated were amplified with the control DNA, creating unique bands with an expected size, including *TPI* (550 bp), *PFK* (437 bp), *PK* (479 bp), and *SDHA* (514 bp). The sequences of the primers that worked for each locus, with the control DNA, were the following: *TPI* (Fw 5´ GCGTTCACAGAACCTCCGTA 3´; Rv 5´ CCTCCAACCTGAACAGCGAT3´), *PFK* (Fw 5´ TACCACTTCGTGCGCTTGAT3´; 5´ACGCAGGACACGATGAACTT3´), *PK* (Fw 5´ CGTCAGATCACCGTCGGAAA3´; Rv 5´ ACCAAGATCGTATGCACGCT3´) and *SDHA* (Fw5´ GTCGATCCATCGCTTCCACT3´; Rv 5´ CAGTCCGCCCATGTTGTAGT3´).

These four genes were chosen to evaluate them in the *Blastocystis* positive DNA samples.

Study population

A convenience sampling was performed in the departments of Amazonas, Bolívar, Casanare, and Córdoba in Colombia (Figure 1). In total, 375 samples of human feces were collected. From the municipality of Puerto Nariño in the department of Amazonas, samples (75.7%, n = 284) were collected from 3 indigenous rural settlements, as described by Sánchez and collaborators [39]. Samples were also collected from the urban area of Mompós in the department of Bolívar (8%, n = 30), the municipalities of Poré, Tamara and Yopal in the department of Casanare (14.1%, n = 53), and samples from Montería in the department of Córdoba (2.1%, n = 8) as depicted in Figure 1. The samples were collected in plastic containers with 70% ethanol (ratio 1: 4 feces: alcohol) and kept refrigerated for subsequent DNA extraction.

DNA Extraction, *Blastocystis* DNA detection, and subtyping

Prior to DNA extraction, each sample was washed with 1x sterile PBS. Genomic DNA from the stool samples was obtained using the commercial kit, Stool DNA Isolation kit

from Norgen (cat. No 27600), following the manufacturer's instructions. DNA samples were stored at -30° C until use. Initially, the samples were subjected to PCR to detect *Blastocystis.* Samples from the department of Amazonas were processed for the detection of *Blastocystis* according to the protocol reported by Sanchez et al. [39]. Conventional PCR was performed for the detection of *Blastocystis* of the samples obtained from the departments of Bolívar, Casanare, and Córdoba. PCR for detecting *Blastocystis*, was performed in a final volume of 9 µL containing 3.5 µL of GoTaq Green Master Mix (Promega) (cat. No M7122), 2 µL of the template DNA, and the primers. Species-specific primers were used at a final concentration of 1 µM. The sequences of the primers used for *Blastocystis* were FWD F5 (5´-GGTCCGGTGAACACTTTGGATTT-3´) and R F2 (5´-CCTACGGAAACCTTGTTACGACTTCA-3´) [40]. The thermal cycling parameters were as follows 95°C for 5 min; 40 cycles of 95°C for 15 s, 58°C for 1 min., 72°C for 30 s and then 72°C for 10 min. The expected size of the amplified fragments was 119 bp. [5].

Conventional PCR was performed in order to determine the subtypes and alleles of the samples that were positive for *Blastocystis*. Thus, through the amplification of a 600pb region of the 5 'end of the *ssu rRNA* using primers BhRDr (5´-GAGCTTTTTAACTGCAACAACG-3´) and RD5 (5´-ATCTGGTTGATCCTGCCAGT-3´) as reported previously [23]. Once all of the PCR subtyping were performed, the size of each amplicon was assessed using 2% agarose gel electrophoresis followed by staining with SYBR Safe, Thermo Fisher Scientific (cat No. S33102). Subsequently, each product was purified with ExoSAP-IT®, Affymetrix™ (cat. No 15513687) following the manufacturer's recommendations. Both strands of each product amplified were sequenced using the Sanger

method by Macrogene (Korea). Sequences were edited in MEGA 7.0 [38] and submitted to a database to identify the subtypes and alleles (https://pubmlst.org/blastocystis/).

PCR in the DNA testing of fecal samples with designed primers

The primers that produced the best PCR results with DNA control (*TPI, PK, PFK* and *SDHA*), were chosen to be evaluated in the DNA testing of the human fecal samples. The conditions for these PCR were the same as those used with the DNA control.

Then, each of the products of each marker was purified with ExoSAP-IT®, Affymetrix™ (cat. No 15513687) following the manufacturer's recommendations. Each product amplified were sequenced using the paired end Sanger method by Macrogene (Korea). Sequences were edited in the MEGA 7.0 [38].

However, after PCR, sequencing and editing processes of sequences, suitable results were only obtained for the *SDHA* marker. With the other markers, the intensity of the band was very weak, and in some cases, there was no amplification, or the sequencing results were not successful.

Phylogenetic reconstructions and determination of the haplotype networks

The edited sequences for *ssu rRNA* and the *SDHA* markers in the FASTA format were aligned using the multiple sequence alignment program MAFFT v7 [41]. Subsequently, different phylogenetic trees using each marker and the concatenated sequences of both markers were constructed. Also, phylogenetic trees by ST with *SDHA* and ssur RNA were constructed. The trees were run with the maximum likelihood method under the model of nucleotide evolution GTR+F, and 1000 bootstrap replicates using IQtree [42]. The visualization and edition of the phylogenetic trees was carried out with the online tool

Interactive Tree Of Life V32 [43]. The ST7 was used as the outgroup. In addition, phylogenetic networks were constructed to detect reticulation signals between samples evaluated by gene and their concatenation. The analysis used the SplitsTree5 program [44] with the Neighbornet algorithm and 1,000 iterations.

Also we use https://microreact.org/showcase to visualize the relation between geographical regions with the phylogenetic trees by ST constructed with IQtree software.

Furthermore, to determine the number of haplotypes in the population for the *SDHA* gene, a fasta sequence matrix was constructed for the haplotype network analysis using prior alignment with MAFFT v7 [41]. This alignment of the fasta format was imported into the Geneious Prime software (available at https://www.geneious.com/) to export a file with .phy extension and subsequently submitted to the Network 5.0 program [45] (available on http://www.fluxus-engineering.com/sharenet.htm) to build a haplotype network by geographical region based on the median-joining model with 1000 iterations. Analysis of the haplotype networks allowed for the determination of the intraspecific relationships between the different haplotypes and the mutational positions generated between them.

Evaluation of Discrimination power and Typing Efficiency

The sequences of each gene and their concatenation were aligned using the multiple sequence alignment programme MAFFT v7 software [41]. Then, these alignments were included as input data to MLSTest software, to evaluate the discriminatory power (DP) with 95% CI and typing efficiency (TE) of *SDHA* and *ssu rRNA* genes. TE is an indicator of grouping of members with common characteristics and DP allows differentiation of individuals belonging to different groups [46].

Calculation of diversity and the genetic differentiation indices

To determine the number of polymorphisms, present in the sequences of the *ssu rRNA* and the *SDHA* genes, the previous alignment of each locus was used with the MAFFT v7 software [41]. The sequences used in this study were grouped into three populations, determined by the departments of origin of the samples Amazon, Bolívar, and Casanare. We analyzed 92 sequences in total for the determination of both loci, considering populations by departments and STs. A total number of sites (including gaps) at 376 for the *SDHA* locus and 1314 for the *ssu rRNA*. All sequences, except 1 sequence from Córdoba, were used to calculate the diversity indices such as nucleotide diversity (Pi) and Theta (per site) from the total number of mutations (Eta), number of polymorphic (segregating) sites (S), number of haplotypes (h) and haplotypic diversity (Hd).

In addition, in order to determine if the sequences evaluated presented a neutral evolution or were involved in a selection process, the Tajima D was calculated, which indicate a balancing selection with a positive value and purifying selection if the value is negative [47]. Finally, statistics of the genetic differentiation between populations (departments and STs) for the *SDHA* and *ssu rRNA* loci were calculated. The department of Córdoba was excluded since only one sequence had been obtained for this group. Peer genetic differences were estimated for the populations, calculating the Wright's statistic F (Fst). Then, the average number of nucleotide differences in pairs (Kxy), nucleotide substitutions per site (Dxy), net nucleotide substitutions per site (Da), and gene flow from the haplotypes (Gst). DnaSP v.5 software was used for the analysis (available at http://www.ub.edu/dnasp).

Results

Newly designed primers with control DNA

Four of twelve new loci evaluated for *Blastocystis* were amplified with a single band of the expected size when we tested them on the control DNA of *Blastocystis* ST3 from a xenic culture. The other 8 loci were discarded because of different reasons, such as, they did not amplify, the size band was incorrect, some faint bands or multiple bands in some cases were showed, preventing their use for the purpose of this study. Due to this, only four loci (*TPI, PPK*, *PK* and *SDHA*) were evaluated in the DNA of faecal samples.

Sample description and detection of *Blastocystis* DNA

The average age of the human population was 8.4 years (SD: 7.5 years, range: 1-70). The largest number of samples collected was in the department of Amazonas (75.7%; n=261), then 20.3% (n=76) for San Juan del Socó, 8.3% (n=31) Nuevo Paraíso, 11.2% (n=42) Villa Andrea, and 29.9% (n=112) Puerto Nariño. From Casanare, 14.1% (n=53) of the total samples were collected, which corresponded to the following municipalities Poré (9.6%; n=36), Tamara (1.6%; n=6), and Yopal (2.9%; n=11). Moreover, for Bolívar (Mompós) and also Córdoba (Montería), 8.0% (n=30) and 2.1% (n=8) were collected, respectively (Figure 1).

Regarding the detection of *Blastocystis* DNA by PCR with the *ssu rRNA* marker [5], a total of 86.6% [95% CI: 88.1–94.5, n = 305] of the samples were positive. From the total samples, the percentages of positive samples for each department were Amazonas (66.5%, n = 234), Casanare (9.9%, n = 35), Bolívar, Mompós (8.2%, n = 29) and Córdoba, Montería (1.9%, n = 7). The percentages of positive samples by municipality in the department of

Amazonas were San Juan del Socó (18.1%, n = 68), Nuevo Paraíso (7.7%, n = 29), Villa Andrea (10.1%, n = 38), and Puerto Nariño (26.4%, n = 99) and, the percentages of positive samples by municipality in the department of Casanare were Poré (6.9%, n = 26), Yopal (1.6%, n = 6), and Tamara (0.8%, n = 3).

Blastocystis subtyping (*ssu rRNA*) and *SDHA* amplification

A total of 77.4% (n = 236) of the positive samples for *Blastocystis* DNA were subtyped and the following STs identified ST3 (27.5%, n = 84), ST1 (26.2%, n = 80) and ST2 (23.6%, n = 72). The 69 samples that were positive for *Blastocystis* could not be subtyped because the quality of the electropherogram obtained after sequencing was not optimal and the sequence to establish the ST could not be obtained. On the other hand, in some of the analyzed sequences the presence of mixed infections was evident, where the STs in low proportion could not be established by the sequencing technique used.

Initially, we tested DNA primers for four loci in all the positive samples (n = 305). A total of 8.5% (n =26), 19.01% (n =58), 12.8% (n = 39) and 41.3% (n = 126) samples were amplified with primers for *TPI, PPK*, *PK* and *SDHA*, respectively. In the case of *TPI, PPK* and *PK* the amount of sequences obtained was too low and the quality of their electropherograms were not appropriated to get a consensus sequence, so we could not analyze them. By contrast, in the case of *SDHA,* we obtained clean sequences for each sample. The sequences were deposited under the accession numbers MT072325 to MT072444. The quantity of samples amplified for this locus, by department, were Bolívar (15%, n = 18), Casanare (19.6%, n = 23), Córdoba (0.8%, n = 1), and Amazonas (64.1%, n = 75). In order to compare *ssu* rRNA and *SDHA* genes, only those samples where the amplification of both genes could be obtained were used, so 9 samples were discarded.

Genetic diversity and differentiation

The genetic diversity indexes were calculated using all the sequences obtained for each loci by department and by ST, with the exception of Córdoba (only 1 sample). Comparing the results of haplotype diversity (Hd) and nucleotide diversity (Pi) by department (Figure 2a, b) and by ST (Figure 2c,d) with each loci, greater Pi values were found in the *ssu rRNA* locus, principally in the Bolívar department (Figure 2b), in contrast to *SDHA* locus which showed similar diversity indexes in all departments (Figure 2a). Similar results were obtained by ST, where the highest nucleotide diversity was shown in ST1 of *ssu rRNA* gene and the lower haplotype diversity value showed was in ST1 of *SDHA* gene (Figure 2c, d). In general, *ssu rRNA* gene showed higher diversity in comparison with *SDHA* (Supplementary material Table S1). The number of polymorphic sites obtained for the *ssu rRNA* was 193 segregating sites and the number of haplotypes 56, unlike the *SDHA* with 126 sites and 36 haplotypes. It should also be noted that, with respect to the sampled departments, Casanare and Bolívar showed a larger Theta per site in the *ssu rRNA* locus. Also, a greater number of haplotypes in Amazonas was observed with both genes, being higher with *SDHA* locus. Conversely, in the *SDHA* locus, nucleotide diversity indices were very similar between the three Colombian departments and the ST3 showed higher nucleotide and haplotype diversity for this locus (Figure. 2 a and 2 c; Supplementary material Table S1). Minimum number of recombination event was calculated for both genes by geographical region and ST. In all cases *ssu rRNA* showed to be more diverse (Supplementary material Table S1).

Then, we applied the D Tajima test [47] to determine if there were any selective processes occurring in the sequences. We observed that a statistically significant negative value when

evaluating the sequences of the three departments and ST for each of the loci. This indicated a process of population expansion with a high frequency circulating among the rare alleles (Supplementary material Table S1).

Phylogenetic reconstructions

Figures 3 a- c depicts the phylogenetic trees constructed for the concatenated (*ssu rRNA* and *SDHA* sequences), *ssu rRNA* and *SDHA* genes, respectively, where the colors indicate the ST1, ST2 and ST3 that were determined for the samples and the ST7 used as outgroup. In comparison, the concatenated and *ssu rRNA* trees showed similar topology between them with little changes and incongruences among some clusters, for instance, inside of the ST1 cluster we could identify changes in the clustering of samples as Bol_26, Bol_23 and Cor_01 from ST2 and Cas_03 from ST. Besides, we observed some subgroups clearly established inside of each ST. The *SDHA* tree, exhibited 4 clusters with bootstrap supported greater than 80%, with a relative evidence of clustering associated with the STs where it is possible to detect the clusters of ST1, ST2 and ST3. The topology obtained to concatenated, *ssu rRNA* and *SDHA* genes are similar to those reported to *Cryptococcus* [48]. Concatenated and *ssu rRNA* presents signatures of evolution by gradualism where big changes result from many cumulative small changes [49]. Figures 3a and 3b show this profile. In the case of *SDHA* gene (Figure 3c), phylogenetic reconstruction coincides with the principle of punctuated equilibrium (abrupt and rapid changes that give rise to well-differentiated clusters) [49].

SDHA tree had some incongruences of clustering in comparison with the other trees (Figure 3). For these inconsistences, we wanted to determine some signals of reticulation for the concatenated and each individual gene (Figures 3d – f). Higher signals of reticulation were

evidenced in the case of the concatenated and the *ssu rRNA* gene and lower in the *SDHA* gene. In the other hand, we wanted to verify if the *SDHA* marker could resolve differentiation intra ST. For this purpose, phylogenetic trees for each ST with the sequences of *SDHA* were constructed. In comparison with *ssu rRNA* trees by ST, at least three subgroups were detected inside each *SDHA* phylogenetic tree with enough support, indicating diversity intra ST with this genetic target and highlighting the resolution power of this marker (Figure 4). Then, using Micro React tool (https://microreact.org/showcase) some relation between each marker and geographical location was checked (Figures 5a-c). In the ST1, it was clear the cluster of Amazonas department and other cluster conformed by Bolívar and Casanare sequences. In the case of ST3 the sequences from Amazonas could be splitted in two clusters and an additional cluster was conformed for Bolívar and Casanare. In the case of ST2, a cluster is observed for Amazonas sequences and another mixed.

In addition, we use the MLSTest to evaluate the discrimination power (DP) and typing efficiency (TE) [46] of *SDHA* and *ssu rRNA* genes. TE is an indicator of grouping of members with common characteristics. In our case, the value of TE for *SDHA* was higher (TE=0.257) than *ssu rRNA* (TE= 0.052) and concatenated (TE=0.029), showing that *SDHA* gene has better grouping capacity which was observed intra STs. DP, which allows differentiation of individuals belonging to different groups showed values between 0.904 – 1. So, both markers can discriminate individuals from different STs with a confidence interval range 0.865 - 1.

Haplotypes network and genetic differentiation

Through the sequences obtained for the *SDHA* gene, haplotype networks were constructed under the medium-joining model [45], in order to understand the genetic variability of the

circulating strains of *Blastocystis* in the areas of the sampled departments. A total of 36 haplotypes were found, of which a group of them were observed with a small number of mutations among them, which were all from the department of Amazonas. In contrast, it was possible to observe some samples with a very high number of mutations compared to other haplotypes. For example, those referred to as CA_017, BO_024, and AM_094, corresponded to samples from Casanare, Bolívar, and Amazonas, respectively (Figure 5d), according to the great diversity present in the genus *Blastocystis*. In addition, some shared haplotypes were observed between the collection regions. The distribution of haplotypes using the ST of each of the samples was observed. In this case, it was possible to observe that the samples that corresponded to ST1, presented a greater number of haplotypes, which had diverged by a varied number of mutational changes, despite coming from different geographical regions. With respect to the samples classified as ST2 and ST3, no groups of clearly associated haplotypes were observed (Figure 5d).

Finally, the genetic differentiation was determined considering each department and STs like different populations. Fst and Gst statistics, gave values between -0.00818 to 0.10163 and -0.00501 to 0.03104, respectively, showing no genetic differentiation among Bolívar and Casanare but in case of Amazonas we found moderated differentiation between Amazonas and Casanare and Amazonas and Bolívar with *SDHA* marker [50]. In case of *ssu rRNA* and *SDHA* by ST, the genetic differentiation was evident between the populations evaluated (Table 1). The number of average nucleotide differences between populations (Kxy) was much higher for the *ssu rRNA* marker, the higher value being 49.72857 and the number of average nucleotide substitutions between populations (Dxy) was higher between

Casanare and Bolívar with the *ssu rRNA* gene, 0.24023, followed by ST1 and ST3 with a value of 0.20658 (Table 1).

Table 1. Genetic differentiation among populations with *SDHA* and *ssu rRNA* locus.

Locus	Population 1	Population 2	Kxy	Gst	Fst	Dxy	Da
SDHA	Amazonas	Bolívar	12.29333	0.01140	0.03612	0.03396	0.00123
	Amazonas	Casanare	13.13600	0.03104	0.09839	0.03629	0.00357
	Bolívar	Casanare	12.97857	-0.00501	-0.00818	0.03585	-0.00029
ssu rRNA	Amazonas	Bolivar	42.44667	0.02318	0.10163	0.20506	0.02084
	Amazonas	Casanare	27.81500	0.01439	0.04528	0.13437	0.00608
	Bolivar	Casanare	49.72857	0.00318	0.03660	0.24023	0.00879
SDHA	ST1	ST2	10.86800	0.07731	0.19569	0.03002	0.00587
	ST1	ST3	14.28797	0.04534	0.15387	0.03947	0.00607
	ST2	ST3	14.20798	0.03763	0.14072	0.03925	0.00552
ssu rRNA	ST1	ST2	35.25926	0.07572	0.34376	0.17033	0.05855
	ST1	ST3	42.76134	0.08018	0.45593	0.20658	0.09418
	ST2	ST3	35.99288	0.05890	0.44264	0.17388	0.07697

Kxy: Average proportion of nucleotide differences between populations.
Gst: Genetic differentiation index based on the frequency of haplotypes.
Fst: estimate gene flow from nucleotide sequences.
Dxy: The average number of nucleotide substitutions per site between populations.
Da: The number of net nucleotide substitutions per site between populations.

Discussion

Blastocystis globally has been typed into different STs worldwide due to the use of barcoding that amplified 600 bp of the *ssu rRNA* gene, which could easily discriminate between STs [23], showing a high diversity within the genus. This DNA region is interesting because is expected to be highly conserved [27], instead, exhibits relative high diversity within *Blastocystis* STs. It could indicate that the divergence time might not have

been sufficient to fix the alleles with their mutations [51] showing an overestimated diversity. In this case, the need to colonize a greater host diversity even with strong competition in the intestinal environment could be the reason to be changing. But, due to the unknown role of this microorganism inside the gut, it has been proposed to explore additional markers with different evolutionary rate in order to elucidate what is happening in these populations. A few authors, have reported the need for use of other genes or non-coding regions of the genome, such as the *internal transcribed spacers* (*ITS*)[29], the use of single copy markers from MROs, which allowed for the detection of coinfections even in the same ST [27] or the use of the *PFOR* gene that generated clades that were different from the STs, because they were subject to different selective pressures that showed an evolutionary history different from that of the ssu rRNA gene [28].

In the current study, the subtyped samples were amplified with primers designed for the *SDHA* gene. Interestingly, the *SDHA* locus did not show high genetic diversity compared to the *ssu rRNA* locus (Figure 2, Supplementary material Table S1) and showed a grouping in the phylogenetic reconstruction that was relatively associated with the STs and the collection areas (Figures 3 and 5). This suggested that this gene was probably under different evolutionary forces, which were similar to the *PFOR* gene reported in the Mexican samples [28]. Also, it was important to consider that ribosomal genes were highly utilized in eukaryotes and required a large production of ribosomes in times of massive growth, that could encode hundreds of copies of their transcriptional units [52], leading to recombination events and potentially greater diversity compared to the constitutive metabolic genes, which did not require high variation in order to maintain their function. The latter was characterized by a particular genomic composition with respect to their

structure, composition, and conservation of sequences and expression [53]. For example, stabilizing selection for the conservation of metabolic functioning could eventually present recombination and had sufficient discrimination power without being subject to diversified selection [54]. However, significantly conserved genes would not be very informative for identifying variants successfully [54]. In this sense, we wanted to explore reticulation signals between the samples with both markers. The results obtained showed a greater number of reticulation signals in the concatenated and the *ssu rRNA* gene compared to *SDHA* (Figure 3d -f) and agree with the index determined for recombination, which was higher for the *ssu rRNA* gene (Supplementary material Table S1). Future studies should consider including more metabolic genes to understand the true genetic diversity of *Blastocystis.* It is also important to highlight that in some samples mixed infections with the *ssu rRNA* marker were evidenced and that therefore it is possible that STs are present in a lower proportion, which could be amplified with the *SDHA* marker and could influence the phylogenetic inconsistencies shown between both markers (Figure 3). This occurred with three ST2 samples that formed a cluster within ST1, because we obtained an electropherogram with multiple signals in the same position indicating mixed infection, but it was not possible to detect the other sequences present in the samples. The determination of the ST presented in a sample must be resolved using other sequencing techniques as deep sequencing [25].

Phylogenetic reconstructions revealed different topology. In comparison of both genes we propose differences in evolutionary patterns (gradualism in the case of *ssu rRNA* and punctuated evolution in the case of *SDHA* gene) as was reported in *Cryptococcus* [48]. Besides, we could establish at least tree subgroups inside each ST using *SDHA* gene and to

elucidate capability of *ssu rRNA* to discriminate but *SDHA* to typing (Figure 4). Moreover, we found some clusters related with the geographical distribution of Amazonas and other cluster with Bolívar and Casanare mixed (Figure 5).

Besides, to detect intraespecific variation, it is necessary to increase the phylogenetic resolution using other markers with lower diversity. For this reason, we propose in addition to using *ssu rRNA* to include other markers as *SDHA* which let to obtain clusters with enough support of bootstrap and detect the intra and inter ST variation. These results agree with the TE [46] value obtained to *SDHA*, which showed more capability for this gene to cluster some samples and visualize subgroups inside each ST. However, it is important to evaluate this marker in other STs different from those studied in this study.

On the other hand, several studies have tried to relate this diversity to the different symptoms [55-57], hosts [7, 13], and even socioeconomic factors [39, 58]. However, these aspects have not yet been fully explicated, and this parasite is associated with many unknown factors regarding the biology and evolution of this microorganism and even the role of this protist in different hosts. In this study, we found a relationship between STs and geographic distribution by department, where the use of the two markers let us to observe a grouping corresponding to the department of Amazonas and another that includes samples from both Bolívar and Casanare. Geographically, Bolívar and Casanare are closer and there is a greater movement of humans between these two regions, so the possibility of transmission between these two areas is higher, instead the samples from the Amazon form a separate cluster, because it corresponds to samples from indigenous communities These communities are located near to forest regions and the possibility of genetic exchange between microorganisms is less. Also, within ST3, we were able to observe two separate

groups within the Amazon region, showing different populations of this microorganism in this region (Figure 4 and 5).

Haplotype networks built with *SDHA* sequences in the current study exhibited great diversity of the haplotypes distributed in the different geographic regions and among the STs. Some haplotypes were found to be associated specifically in the sampled geographic regions, principally in the Amazonas, but some shared haplotypes were observed by region and by ST (Figure 5d and Supplementary material Figure S1). This could be explained by different reasons, such as inadequate sampling (low numbers of samples collected), limited divergence, hybridization, cryptic speciation, and incomplete lineage sorting [59-61].

Interestingly, when the ST analysis was conducted, groups of haplotypes that were associated with ST1 and ST2 were observed, while the haplotypes associated with ST3 were seen at a greater distribution in the sampled regions (Supplementary material Figure S1). The latter could be in accordance with other reports made, where the ST3 was one of the subtypes that had the greatest diversity [5, 62, 63] and in turn, coincided with the fact that this subtype was the most variable compared to the others since it had more than 50 different alleles reported in the database for ST determination (https://pubmlst.org).

Nevertheless, despite the genetic diversity found and the great geographical distances between some of the sampled departments, no evidence of genetic differentiation for *SDHA* was found between Amazonas - Bolívar and Bolívar – Casanare and for *ssu rRNA* gene, just moderate genetic differentiation between Amazonas – Bolívar was found. High genetic differentiation between STs with either of the two loci evaluated was observed. Furthermore, estimators used to make comparisons between populations, such as the Kxy, Dxy, and Da showed more nucleotide substitutions and differences in the *ssu rRNA* gen

when we compare by geographical regions and STs (Table 1). The D Tajima test [47] was based on the comparison of the number of differences using pairs of nucleotide sequences and the number of segregating sites, where the negative values indicated the high frequency of rare alleles and was a signal of population expansion with both markers evaluated.

This results are similar to that observed in the protozoa such as *G. duodenalis* [61] and some of the coding genes for the variant surface antigens in *Plasmodium vivax* [64]. This expansion might be due to a possible selection that has occurred in the *Blastocystis* genetic pool, which varied depending on the mutational rate or the recombination rate [65]. Thus, it is necessary to evaluate the reproductive mechanisms of this microorganism, and the possibility of recombination events that tend to maintain the high variation observed, as has happened in protozoa such as *T. cruzi* [66] and *Leishmania* [67] and has been proposed for *G. duodenalis* [68, 69].

In the case of *Blastocystis*, the evidence of expansion in the samples analyzed might be related to the wide range of hosts presented by this protist, where the capacity for infection in mammals and other animals might have influenced the great diversity, and effectively increased the population size [70]. This population expansion likely occurred due to the ease of fecal transmission through contaminated food and water where the cysts were dispersed to new hosts moving easily throughout the Colombian territory, as has been reported for *G. duodenalis* [61] and for some nematodes, where it has been proposed that host movement is key to gene flow and scattering of these rare alleles [71]. Future studies are necessary to verify these dynamics of transmission and population genetics. Moreover, studies using whole genomes of *Blastocystis* will allow us to understand if recombination is occurring.

The *SDHA* marker could be considered as a possible candidate for the discrimination of groups presented within a population of this microorganism. However, this marker should be evaluated with a larger number of samples and hosts (including animals) as well as more geographic regions in Colombia and South America. Similarly, it could be very useful, along with other markers such as *ssu rRNA*, in the analysis of the genetic diversity and population structure of *Blastocystis*, since in our case it allowed us to show selective pressure forces leading to the expansion process that could be explained by the high number of hosts capable of becoming infected with this microorganism. However, more markers are required to obtain a more robust analysis of what has happened within the populations of this protozoan.

Some limitations of our study include the use of only one additional genetic marker, which could bias these assumptions, the low number of samples that we could amplify because we did not culture the samples and amplify directly from DNA extracted from stool samples, so the amount of DNA available from genes with just one copy or lower number of copies make difficult the amplification. Besides the concentration of this microorganism could be low. It is necessary to increase the number of samples evaluated from different regions to verify our assumptions. It is necessary to extend this study to other STs, since we just could test the ST1, ST2 and ST3.

Conclusions

The *SDHA* marker could eventually be used together with *ssu rRNA* marker, for typing and to evaluate the evolution, diversity, and population structure of *Blastocystis*. However, it was necessary to explore more regions of the genome that allowed for the development of

new markers to study this microorganism and to elucidate the incongruity still present around the biology of this protist.

List of abbreviations

ssu rRNA: Small subunit ribosomal ribonucleic acid

SDHA: Succinate dehydrogenase subunit A

STs: Subtypes

MROs: Mitochondrion-related organelles

PFOR: Pyruvate ferredoxin oxidoreductase

TCA: Tricarboxylic acid

FAD: Flavin adenine dinucleotide

ML: Maximum likelihood

TPI: Triosephosphate isomerase

BPA: Biphosphate aldolase

GK: Glucokinase

GPI: Glucose 6 phosphate isomerase

HK: Hexokinase

PFK: Phosphofructokinase

GAPDH: Glyceraldehyde-3-phosphate dehydrogenase

PGK: Phosphoglycerate kinase

PGM: Phosphoglycerate mutase

PK: Pyruvate kinase

Declarations

Ethics approval and consent to participate

This study was approved by the ethics committee of the National University of Colombia (002-012-15 February 12, 2015) and the ethics committee of the Universidad del Rosario (registered in Act No. 394 of the CEI-UR). The patients approved and signed the written informed consent. This project was conducted under the contract number RGE131 of access to genetic resources granted by the "Ministerio de Medio ambiente y Desarrollo sostenible" from Colombia.

Consent for publication

Not applicable.

Availability of data and materials

Sequences data that support the findings of this study have been deposited in GenBank with the accession codes: MT072325 to MT072444.

Competing interests

The authors declare that they have no competing interests

Funding

The first author was financed by the Departamento Administrativo de Ciencia, Tecnología e Innovación (Colciencias) within the framework of the National Program for Promoting Research Training (sponsorship call 757). This work was funded by the Departamento Administrativo de

Ciencia, Tecnología e Innovación (Colciencias) through project "Desarrollo de una estrategia y plataforma de Multilocus Sequence Typing (MLST) para la caracterización molecular de siete parásitos intestinales", code 122271250521 and contract number 130–2016. The funders had no role in the design of the study, in collection, analysis, interpretation of data, or in writing the manuscript

Author´s contributions

AH: carried out the extraction, amplification and analysis processes of the evaluated samples, and was a major contributor in writing the manuscript, MM: analysis processes of the evaluated samples, data analyses and revision of the manuscript, MCL: collected the Amazonas samples, PR: carried out the Amazonas samples collection, PU: carried out the collection of samples from Casanare, OV: carried out the collection of samples from Bolívar, JDR: contributed to data analysis and revision of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank Angie Johana Sanchez for her help in the molecular detection of *Blastocystis* in the fecal samples from Amazonas.

Figure legends

Figure 1. Geographic locations of departments in which samples were collected.

The departments where the sample collection was conducted are highlighted in color. The red dots indicate the exact location of the sampled municipalities. In the table a zoom of the department of the Amazon is observed, for a better observation of the sampled sites. In the legend on the right, the municipalities of each department and the number of samples taken are indicated. The map was created using ArcGIS version 10.7.

Figure 2. Nucleotide and Haplotype diversity in *ssu rRNA* and *SDHA* markers.

Nucleotide and haplotypic diversity values for both loci are shown by geographic region and by ST. a.) *SDHA* by geographical region, b.) *ssu rRNA* by geographical region, c.) *SDHA* by ST, b.) *ssu rRNA* by ST.

Figure 3. Phylogenetic reconstruction with the *ssu rRNA* and *SDHA* markers.

The evolutionary history was inferred using the maximum likelihood (ML) method based on the GTR+F model with 1000 bootstrap replicates. The pink dot on each node represents the bootstrap support >80%. The initials of the departments where the samples come from are indicated at the tips of the branches, followed by the sample code (Cas: Casanare, Ama: Amazonas, Bol: Bolívar, Cor: Córdoba) a.) Phylogenetic trees built from a.) the concatenated sequences obtained with the *ssu rRNA* and *SDHA* marker, b.) *ssu rRNA* marker and c.) *SDHA* marker. The corresponding colors are indicated in the legend with the ST determined for each sequence. ST7 was used as outgroup.

A phylogenetic network, using Splitstree software, was built with the Neighbornet algorithm. d.) the concatenated sequences obtained with the *ssu rRNA* and *SDHA* marker, e.) *ssu rRNA* marker and f.) *SDHA* marker.

Figure 4. Phylogenetic reconstruction with the *ssu rRNA* and *SDHA* markers by ST.

The evolutionary history was inferred using the maximum likelihood (ML) method based on the GTR+F model with 1000 bootstrap replicates. The pink dot on each node represents the bootstrap support >80%. The initials of the departments where the samples come from are indicated at the tips of the branches, followed by the sample code (Cas: Casanare, Ama: Amazonas, Bol: Bolívar, Cor: Córdoba). Phylogenetic trees built from the sequences

obtained with the *ssu rRNA* by ST a.) ST1, b.) ST2, c.) ST3. Phylogenetic trees built from the sequences obtained with the *SDHA* ST d.) ST1, e.) ST2, f.) ST3. The corresponding colors are indicated in the legend with the ST determined for each sequence. ST7 was used as outgroup.

Figure 5. Relationships of geographical regions with clusters and haplotypic network.

The relationships between geographic location and phylogenetic trees was visualized with https://microreact.org/showcase. These trees correspond with those constructed with IQtree. The location of each department in Colombia is shown. The size of the circles represents the frequency of the samples depending on each ST. Each phylogenetic tree is showing the clusters found inside each ST. a.) ST1, b.) ST2, c.) ST3. The network shows the genetic variability present among the sequences analyzed. d.) The length of the lines is in concordance with the amount of changes to generate a new haplotype. mv: mediate vector. mv, is a hypothesized (often ancestral) sequence which is required to connect existing sequences within the network with maximum parsimony. Without the median vector, there would be no shortest connection between the data set's sequences). The colors represent the regions where the samples were collected, which are indicated in the legend.

Supplementary material Figure S1. The network shows the genetic variability present among the sequences analyzed. The length of the lines is in concordance with the amount of changes to generate a new haplotype. mv: mediate vector. mv, is a hypothesized (often ancestral) sequence which is required to connect existing sequences within the network with maximum parsimony. Without the median vector, there would be no shortest connection between the data set's sequences). The colors represent the STs determined for each sample, which are indicated in the legend.

References

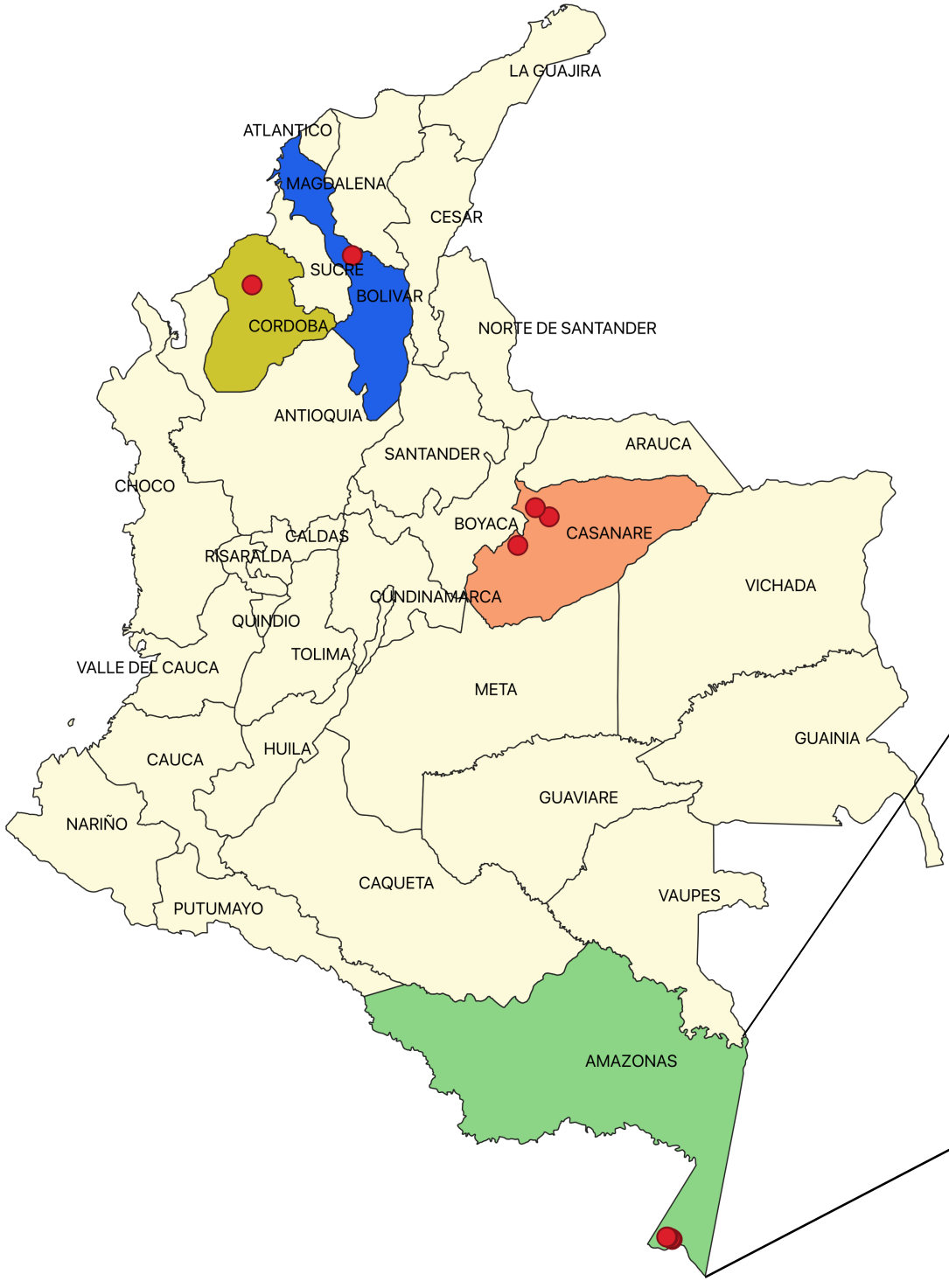
1. Silberman JD, Sogin ML, Leipe DD, Clark CG. Human parasite finds taxonomic home. Nature. 1996;380(6573):398. DOI: 10.1038/380398a0.
2. Arisue N, Hashimoto T, Yoshikawa H, Nakamura Y, Nakamura G, Nakamura F, et al. Phylogenetic position of Blastocystis hominis and of stramenopiles inferred from multiple molecular sequence data. J Eukaryot Microbiol. 2002;49(1):42-53. DOI: 10.1111/j.1550-7408.2002.tb00339.x.
3. Londono-Franco AL, Loaiza-Herrera J, Lora-Suarez FM, Gomez-Marin JE. [Blastocystis sp. frequency and sources among children from 0 to 5 years of age attending public day care centers in Calarca, Colombia]. Biomedica. 2014;34(2):218-27. DOI: 10.1590/S0120-41572014000200008.
4. Ramírez JD, Sánchez A, Hernández C, Flórez C, Bernal MC, Giraldo JC, et al. Geographic distribution of human Blastocystis subtypes in South America. Infect Genet Evol. 2016;41:32-5. doi: 10.1016/j.meegid.2016.03.017.
5. Stensvold CR, Alfellani M, Clark CG. Levels of genetic diversity vary dramatically between Blastocystis subtypes. Infect Genet Evol. 2012;12(2):263-73. doi: 10.1016/j.meegid.2011.11.002.
6. 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;14:164. doi: 10.1186/1471-2334-14-164.
7. 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;1(5):167-78. doi: 10.1177/2049936113504754.
8. Tan KSW, Mirza H, Teo JDW, Wu B, Macary PA. Current Views on the Clinical Relevance of Blastocystis spp.Curr Infect Dis Rep. 2010;12(1):28-35. doi: 10.1007/s11908-009-0073-8.
9. Verma R, Delfanian K. Blastocystis hominis associated acute urticaria. Am J Med Sci. 2013; 346(1):80-1. https://doi.org/10.1097/MAJ.0b013e3182801478.
10. Roberts T, Stark D, Harkness J, Ellis J. Subtype distribution of Blastocystis isolates identified in a Sydney population and pathogenic potential of Blastocystis. Eur J Clin Microbiol Infect Dis. 2013;32(3):335-43. doi: 10.1007/s10096-012-1746-z.
11. Vogelberg C, Stensvold CR, Monecke S, Ditzen A, Stopsack K, Heinrich-Gräfe U, et al. Blastocystis sp. subtype 4 is common in Danish Blastocystis-positive patients presenting with acute diarrhea. Parasitol Int. 2010;59(3):469-71. doi: 10.1016/j.parint.2010.03.009.
12. Stensvold CR, Christiansen DB, Olsen KE, Nielsen HV. Blastocystis sp. Subtype 4 is Common in Danish Blastocystis-Positive Patients Presenting with Acute Diarrhea. Am J Trop Med Hyg. 2011;84(6):883-5. doi: 10.4269/ajtmh.2011.11-0005.
13. Ramirez JD, Sanchez LV, Bautista DC, Corredor AF, Florez AC, Stensvold CR. Blastocystis subtypes detected in humans and animals from Colombia. Infect Genet Evol. 2014;22:223-8. doi: 10.1016/j.meegid.2013.07.020

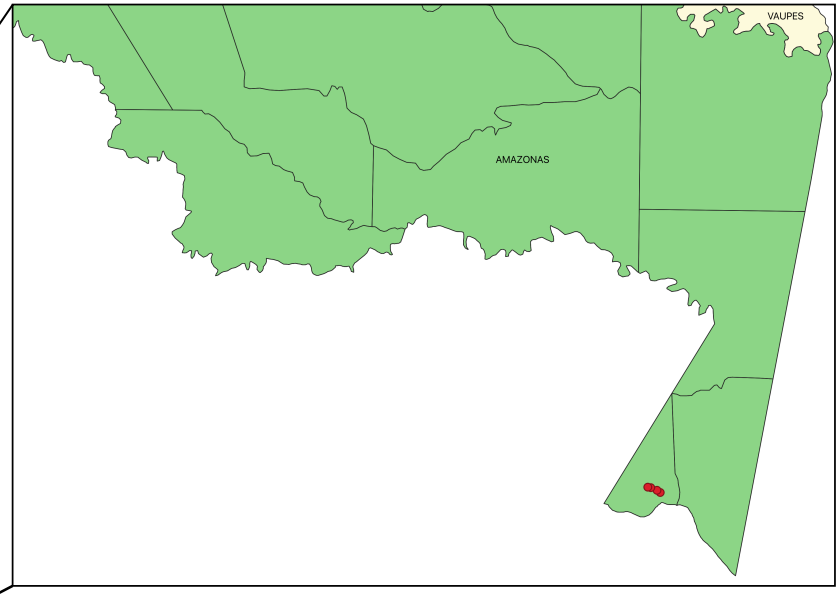
14. Tan KSW. New Insights on Classification, Identification, and Clinical Relevance of Blastocystis spp. Clin Microbiol Rev. 2008;21(4):639-65. doi: 10.1128/CMR.00022-08.
15. Abe N, Wu Z, Yoshikawa H. Molecular characterization of Blastocystis isolates from primates. Vet Parasitol. 2003;113(3-4):321-5. DOI: 10.1016/s0304-4017(03)00081-5.
16. Clark CG. Extensive genetic diversity in Blastocystis hominis. Mol Biochem Parasitol. 1997;87(1):79-83. DOI: 10.1016/s0166-6851(97)00046-7.
17. Yoshikawa H, Nagono I, Yap EH, Singh M, Takahashi Y. DNA polymorphism revealed by arbitrary primers polymerase chain reaction among Blastocystis strains isolated from humans, a chicken, and a reptile. J Eukaryot Microbiol. 1996;43(2):127-30. DOI: 10.1111/j.1550-7408.1996.tb04492.x.
18. Arisue N, Hashimoto T, Yoshikawa H. Sequence heterogeneity of the small subunit ribosomal RNA genes among blastocystis isolates. Parasitology. 2003;126(Pt 1):1-9. DOI: 10.1017/s0031182002002640.
19. Ho LC, Armiugam A, Jeyaseelan K, Yap EH, Singh M. Blastocystis elongation factor-1alpha: genomic organization, taxonomy and phylogenetic relationships. Parasitology. 2000;121 (Pt 2):135-44. DOI: 10.1017/s0031182099006113.
20. Gentekaki E, Curtis BA, Stairs CW, Klimes V, Elias M, Salas-Leiva DE, et al. Extreme genome diversity in the hyper-prevalent parasitic eukaryote Blastocystis. PLoS Biol. 2017;15(9):e2003769. doi: 10.1371/journal.pbio.2003769.
21. Wawrzyniak I, Courtine D, Osman M, Hubans-Pierlot C, Cian A, Nourrisson C, et al. Draft genome sequence of the intestinal parasite Blastocystis subtype 4-isolate WR1. Genom Data. 2015;4:22-3. doi: 10.1016/j.gdata.2015.01.009.
22. 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(3):R29. doi: 10.1186/gb-2011-12-3-r29.
23. Scicluna SM, Tawari B, Clark CG. DNA barcoding of blastocystis. Protist. 2006;157(1):77-85. DOI: 10.1016/j.protis.2005.12.001.
24. Stensvold CR, Suresh GK, Tan KS, Thompson RC, Traub RJ, Viscogliosi E, et al. Terminology for Blastocystis subtypes--a consensus. Trends Parasitol. 2007;23(3):93-6. DOI: 10.1016/j.pt.2007.01.004.
25. Maloney JG, Molokin A, da 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 May 1;9:e00138. doi: 10.1016/j.parepi.2020.e00138.
26. Clark CG, van der Giezen M, Alfellani MA, Stensvold CR. Recent developments in Blastocystis research. Adv Parasitol. 2013;82:1-32. doi: 10.1016/B978-0-12-407706-5.00001-0.
27. Poirier P, Meloni D, Nourrisson C, Wawrzyniak I, Viscogliosi E, Livrelli V, et al. Molecular subtyping of Blastocystis spp. using a new rDNA marker from the mitochondria-like organelle genome. Parasitology. 2014 Apr 1;141(5):670. doi: 10.1017/S0031182013001996.
28. 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(1):564. doi: 10.1186/s13071-018-3141-9.

29. Villalobos G, Orozco-Mosqueda GE, Lopez-Perez M, Lopez-Escamilla E, Cordoba-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. doi: 10.1186/s13071-014-0461-2.
30. Perez-Brocal V, Clark CG. Analysis of two genomes from the mitochondrion-like organelle of the intestinal parasite Blastocystis: complete sequences, gene content, and genome organization. Mol Biol Evol. 2008;25(11):2475-82. doi: 10.1093/molbev/msn193.
31. Wawrzyniak I, Roussel M, Diogon M, Couloux A, Texier C, Tan KS, et al. Complete circular DNA in the mitochondria-like organelles of Blastocystis hominis. Int J Parasitol. 2008;38(12):1377-82. doi: 10.1016/j.ijpara.2008.06.001.
32. Stechmann A, Hamblin K, Pérez-Brocal V, Gaston D, Richmond GS, van der Giezen M, et al. Organelles in Blastocystis that blur the distinction between mitochondria and hydrogenosomes. Curr Biol. 2008;18(8):580-5. doi: 10.1016/j.cub.2008.03.037.
33. Christmas PB, Turrens JF. Separation of NADH-fumarate reductase and succinate dehydrogenase activities in Trypanosoma cruzi. FEMS Microbiol Lett. 2000;183(2):225-8. DOI: 10.1111/j.1574-6968.2000.tb08962.x.
34. Kregiel D. Succinate dehydrogenase of Saccharomyces cerevisiae–the unique enzyme of TCA cycle–current knowledge and new perspectives. 2012. In: Dehydrogenases [Internet]. [211-34]. Available from: https://www.intechopen.com/books/dehydrogenases/succinate-dehydrogenase-of-saccharomyces-cerevisiae-the-unique-enzyme-of-tca-cycle-current-knowledge.
35. Huang S, Millar AH. Succinate dehydrogenase: the complex roles of a simple enzyme. Curr Opin Plant Biol. 2013;16(3):344-9. doi: 10.1016/j.pbi.2013.02.007.
36. Berg OG, Kurland CG. Why mitochondrial genes are most often found in nuclei. Mol Biol Evol. 2000 Jun 1;17(6):951-61. https://doi.org/10.1093/oxfordjournals.molbev.a026376
37. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792-7. DOI: 10.1093/nar/gkh340. https://doi.org/10.1093/nar/gkh340.
38. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33(7):1870-4. doi: 10.1093/molbev/msw054.
39. Sanchez A, Munoz M, Gomez N, Tabares J, Segura L, Salazar A, et al. Molecular Epidemiology of Giardia, Blastocystis and Cryptosporidium among Indigenous Children from the Colombian Amazon Basin. Front Microbiol. 2017;8:248. doi: 10.3389/fmicb.2017.00248.
40. Stensvold CR, Ahmed UN, Andersen LO, Nielsen HV. Development and evaluation of a genus-specific, probe-based, internal-process-controlled real-time PCR assay for sensitive and specific detection of Blastocystis spp. J Clin Microbiol. 2012;50(6):1847-51. doi: 10.1128/JCM.00007-12.
41. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772-80. doi: 10.1093/molbev/mst010.
42. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R. IQ-TREE 2: New models and efficient methods for phylogenetic

inference in the genomic era. Mol Biol Evol. 2020 May 1;37(5):1530-4. https://doi.org/10.1093/molbev/msaa015.
43. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016;44(W1):W242-5. doi: 10.1093/nar/gkw290.
44. Huson, DH, Bryant, D. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol. 2006; 23(2):254-267, doi: 10.1093/molbev/msj030.
45. Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 1999;16(1):37-48. doi: 10.1093/oxfordjournals.molbev.a026036.
46. Tomasini N, Lauthier JJ, Llewellyn MS, Diosque P. MLSTest: novel software for multi-locus sequence data analysis in eukaryotic organisms. Infect Genet Evol. 2013 Dec 1;20:188-96. doi: 10.1016/j.meegid.2013.08.029.
47. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989;123(3):585-95.
48. Muñoz M, Camargo M, Ramírez JD. Estimating the intra-taxa diversity, population genetic structure, and evolutionary pathways of Cryptococcus neoformans and Cryptococcus gattii. Front Genet. 2018;24(9):148. https://doi.org/10.3389/fgene.2018.00148.
49. Singh B. Evolutionary biology: concepts of punctuated equilibrium, concerted evolution and coevolution. J. Sci. Res. 2014;58:15-26.
50. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. Evolution. 1984;38(6):1358-70. doi: 10.1111/j.1558-5646.1984.tb05657.x.
51. Anderson TJ. The dangers of using single locus markers in parasite epidemiology: Ascaris as a case study. Trends Parasitol. 2001;17(4):183-8. DOI: 10.1016/s1471-4922(00)01944-9.
52. Eickbush TH, Eickbush DG. Finely orchestrated movements: evolution of the ribosomal RNA genes. Genetics. 2007;175(2):477-85. DOI: 10.1534/genetics.107.071399.
53. Wei K, Zhang T, Ma L. Divergent and convergent evolution of housekeeping genes in human-pig lineage. PeerJ. 2018;6:e4840. doi: 10.7717/peerj.4840.
54. Maiden M. Multilocus sequence typing of bacteria. Annu Rev Microbiol. 2006;60:561-88. DOI: 10.1146/annurev.micro.59.030804.121325.
55. Mohamed RT, El-Bali MA, Mohamed AA, Abdel-Fatah MA, El-Malky MA, Mowafy NM, et al. Subtyping of Blastocystis sp. isolated from symptomatic and asymptomatic individuals in Makkah, Saudi Arabia. Parasit Vectors. 2017;10(1):174. doi: 10.1186/s13071-017-2114-8.
56. Ramirez JD, Florez C, Olivera M, Bernal MC, Giraldo JC. Blastocystis subtyping and its association with intestinal parasites in children from different geographical regions of Colombia. PLoS One. 2017;12(2):e0172586. doi: 10.1371/journal.pone.0172586.
57. Zulfa F, Sari IP, Kurniawan A. Association of Blastocystis subtypes with diarrhea in children. IOP Publishing. Journal of Physics: Conference Series; 2017. p. 012031. DOI: 10.1088/1742-6596/884/1/012031.
58. Villamizar X, Higuera A, Herrera G, Vasquez-A LR, Buitron L, Muñoz LM, et al. Molecular and descriptive epidemiology of intestinal protozoan parasites of children and

their pets in Cauca, Colombia: a cross-sectional study. BMC Infect Dis. 2019;19(1):190. doi: 10.1186/s12879-019-3810-0.
59. Hart MW, Sunday J. Things fall apart: biological species form unconnected parsimony networks. Biol Lett. 2007;3(5):509-12. doi: 10.1098/rsbl.2007.0307.
60. Tarcz S, Przyboś E, Surmacz M. An assessment of haplotype variation in ribosomal and mitochondrial DNA fragments suggests incomplete lineage sorting in some species of the Paramecium aurelia complex (Ciliophora, Protozoa). Mol Phylogenet Evol. 2013;67(1):255-65. doi: 10.1016/j.ympev.2013.01.016.
61. Choy SH, Mahdy MAK, Al-Mekhlafi HM, Low VL, Surin J. Population expansion and gene flow in Giardia duodenalis as revealed by triosephosphate isomerase gene. Parasit Vectors. 2015;8:454. doi: 10.1186/s13071-015-1084-y.
62. Meloni D, Poirier P, Mantini C, Noël C, Gantois N, Wawrzyniak I, et al. Mixed human intra- and inter-subtype infections with the parasite Blastocystis sp. Parasitol Int. 2012;61(4):719-22. doi: 10.1016/j.parint.2012.05.012.
63. Rojas-Velázquez L, Morán P, Serrano-Vázquez A, Fernández LD, Pérez-Juárez H, Poot-Hernández AC, 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. doi: 10.1155/2018/3916263.
64. Son UH, Dinzouna-Boutamba SD, Lee S, Yun HS, Kim JY, Joo SY, et al. Diversity of *vir* Genes in *Plasmodium vivax* from Endemic Regions in the Republic of Korea: an Initial Evaluation. Korean J Parasitol. 2017;55(2):149-58. doi: 10.3347/kjp.2017.55.2.149.
65. Hudson RR, Kaplan NL. Deleterious background selection with recombination. Genetics. 1995;141(4):1605-17.
66. Lewis M, Llewellyn M, Yeo M, Messenger L, Miles M. Experimental and natural recombination in Trypanosoma cruzi. 2017. In: *American Trypanosomiasis Chagas Disease* [Internet]. Elsevier; [455-73]. doi:10.1016/B978-0-12-801029-7.00020-4.
67. Inbar E, Shaik J, Iantorno SA, Romano A, Nzelu CO, Owens K, et al. Whole genome sequencing of experimental hybrids supports meiosis-like sexual recombination in Leishmania. PLoS Genet. 2019;15(5):e1008042. doi: 10.1371/journal.pgen.1008042.
68. Cacciò SM, Sprong H. Giardia duodenalis: genetic recombination and its implications for taxonomy and molecular epidemiology. Exp Parasitol. 2010;124(1):107-12. doi: 10.1016/j.exppara.2009.02.007.
69. Cooper MA, Adam RD, Worobey M, Sterling CR. Population genetics provides evidence for recombination in Giardia. Curr Biol. 2007;17(22):1984-8. DOI: 10.1016/j.cub.2007.10.020.
70. Barrett LG, Thrall PH, Burdon JJ, Linde CC. Life history determines genetic structure and evolutionary potential of host-parasite interactions. Trends Ecol Evol. 2008;23(12):678-85. doi: 10.1016/j.tree.2008.06.017.
71. Blouin MS, Yowell CA, Courtney CH, Dame JB. Host movement and the genetic structure of populations of parasitic nematodes. Genetics. 1995;141(3):1007-14.





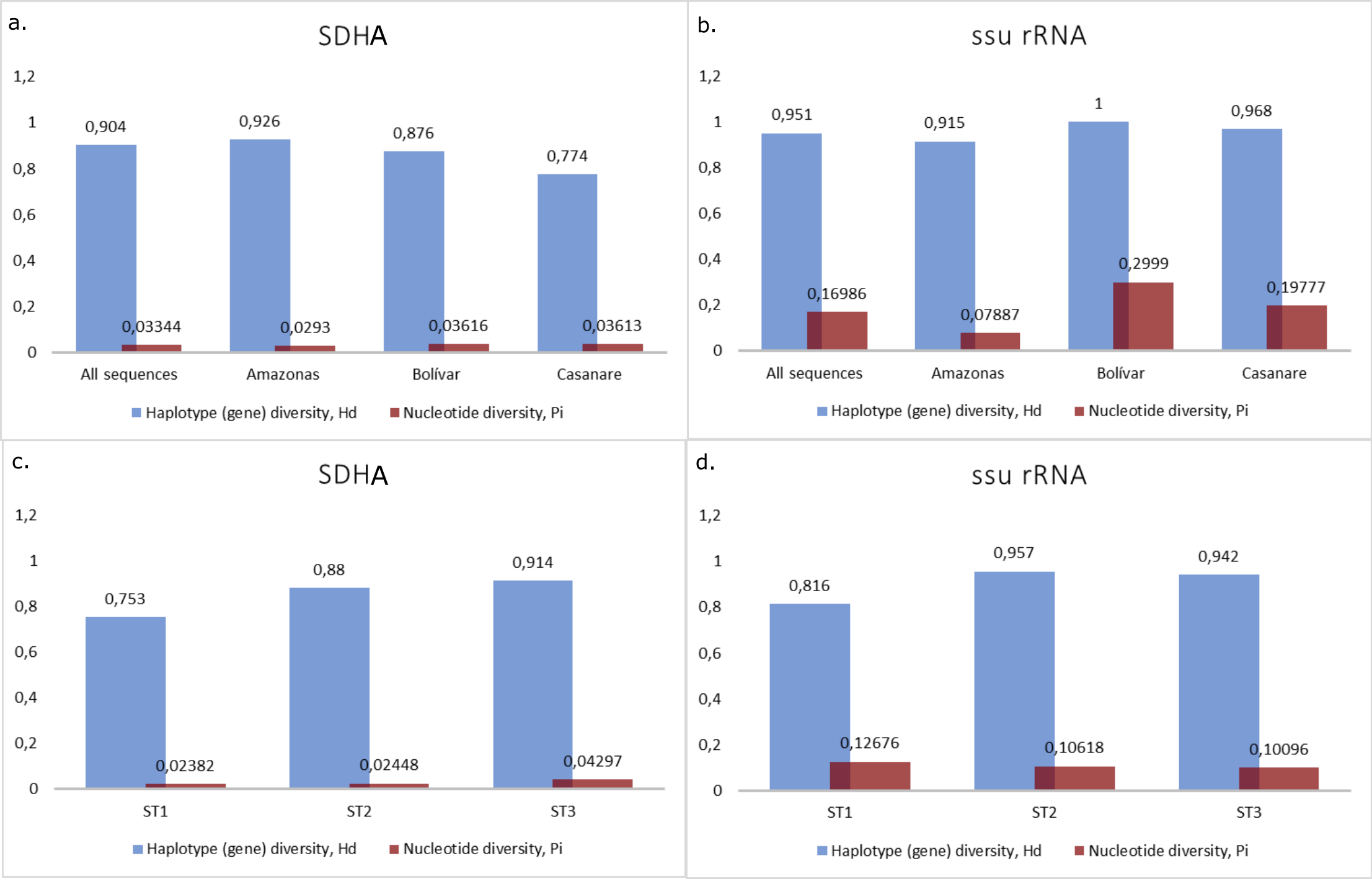
Departments

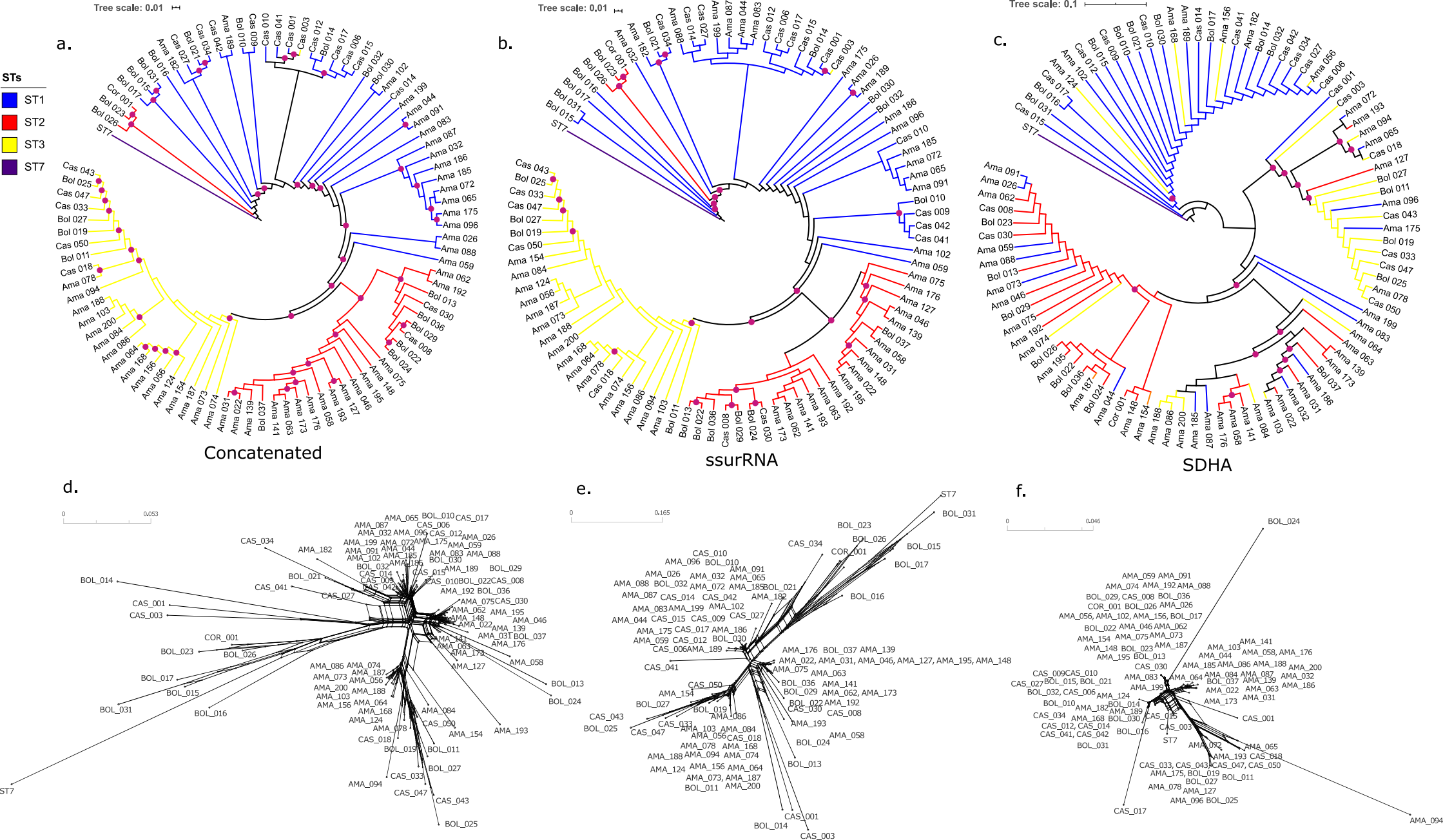
Amazonas
n = 284 (San Juan del Socó, Nuevo Paraíso, Villa Andrea and Puerto Nariño)

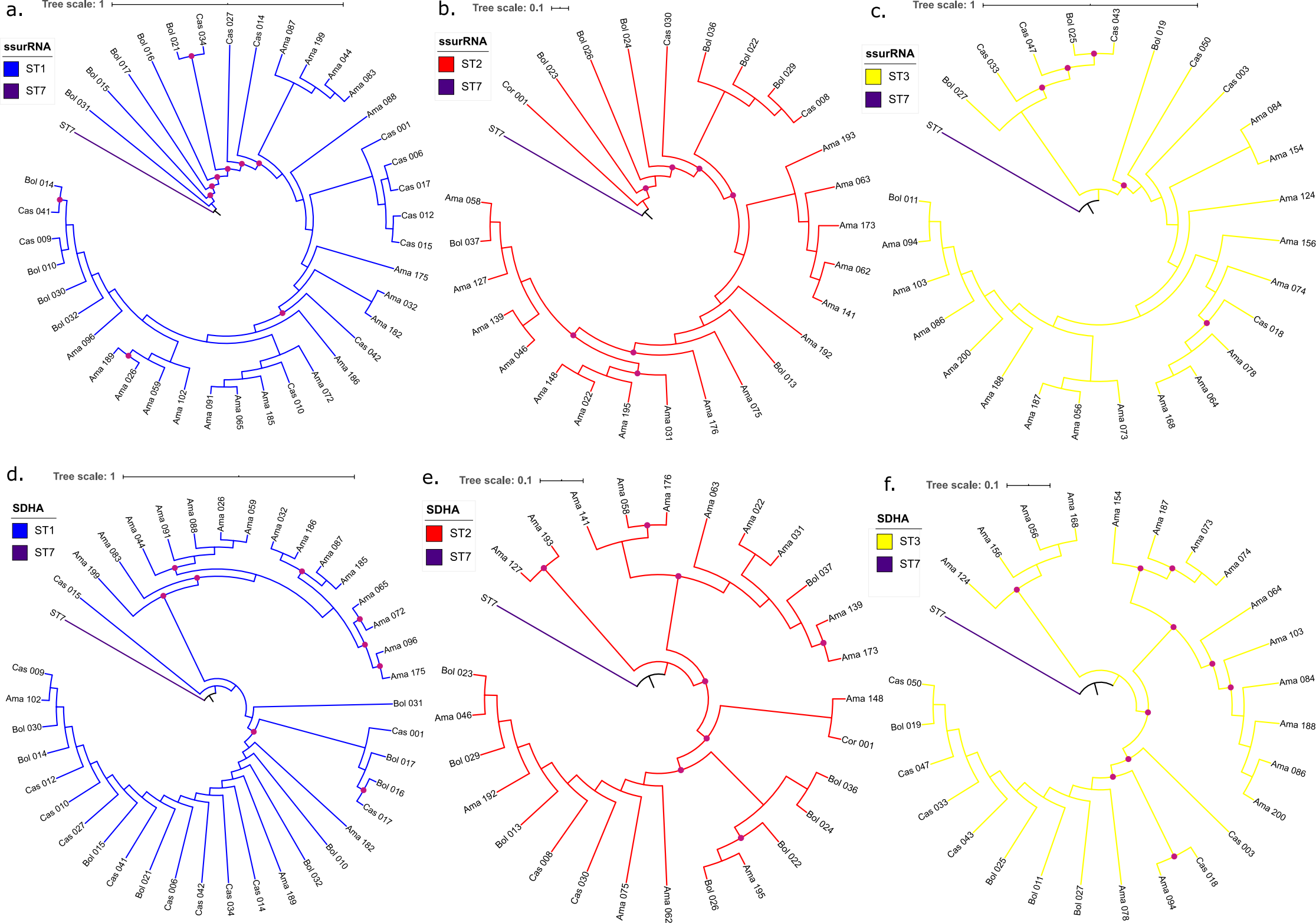
Casanare
n =53 (Poré, Tamara and Yopal)

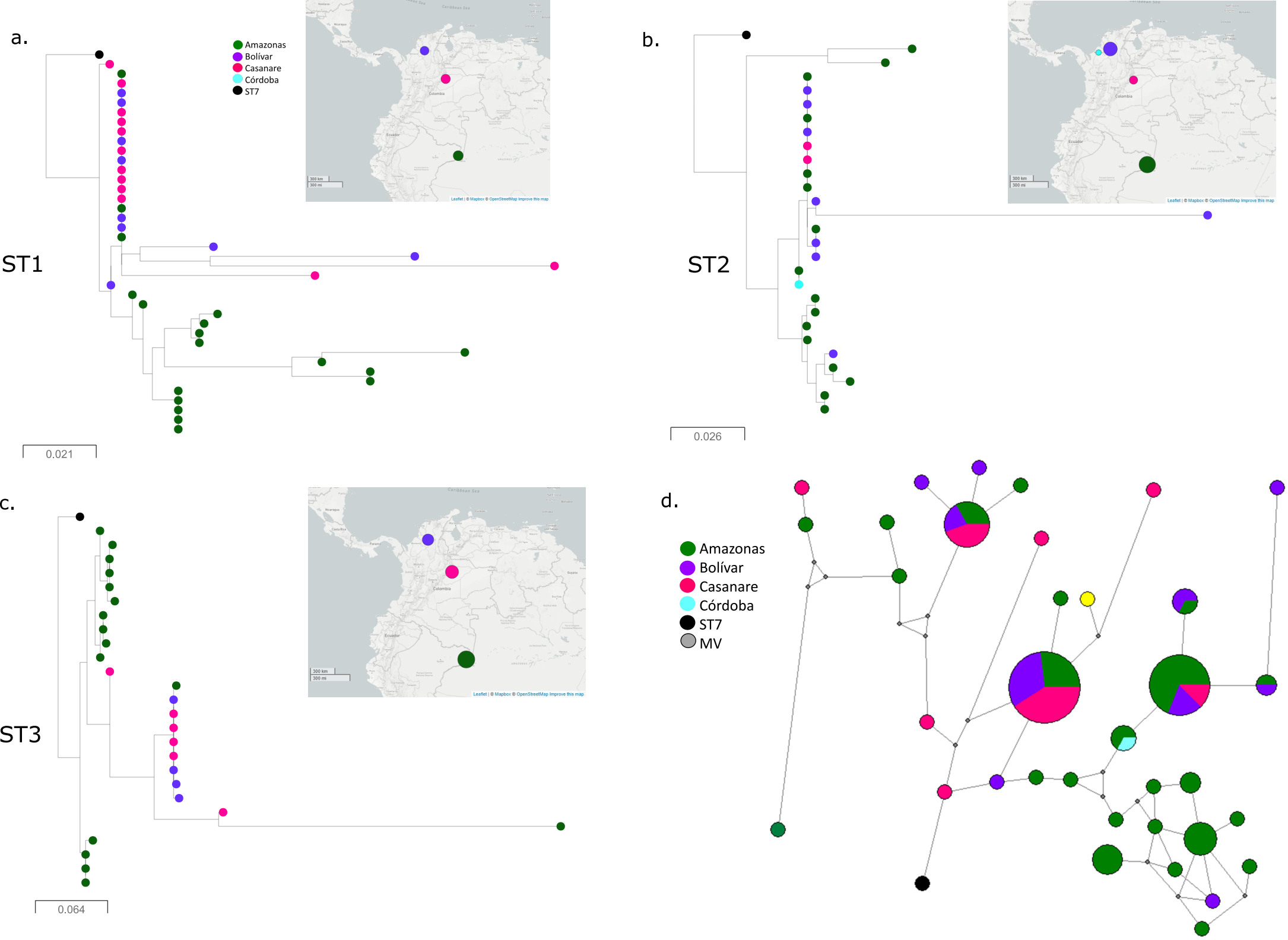
Bolívar
n = 30 (Mompós)

Córdoba
n = 8 (Montería)


a.
SDHA
1,2
1
0,8
0,6
0,4
0,2
0
0,904
0,926
0,876
0,774
0,03344
0,0293
0,03616
0,03613
All sequences
Amazonas
Bolívar
Casanare
Haplotype (gene) diversity, Hd
Nucleotide diversity, Pi
b.
ssu rRNA
1,2
1
0,8
0,6
0,4
0,2
0
0,951
0,915
1
0,968
0,16986
0,07887
0,2999
0,19777
All sequences
Amazonas
Bolívar
Casanare
Haplotype (gene) diversity, Hd
Nucleotide diversity, Pi
c.
SDHA
1,2
1
0,8
0,6
0,4
0,2
0
0,753
0,88
0,914
0,02382
0,02448
0,04297
ST1
ST2
ST3
Haplotype (gene) diversity, Hd
Nucleotide diversity, Pi
d.
ssu rRNA
1,2
1
0,8
0,6
0,4
0,2
0
0,816
0,957
0,942
0,12676
0,10618
0,10096
ST1
ST2
ST3
Haplotype (gene) diversity, Hd
Nucleotide diversity, Pi


a.
Tree scale: 0.01
STs
ST1
ST2
ST3
ST7
Concatenated
b.
Tree scale: 0.01
ssurRNA
c.
Tree scale: 0.1
SDHA
d.
e.
f.


a.
Tree scale: 1
ssurRNA
ST1
ST7
Bol 021
Cas 034
Cas 027
Cas 014
Ama 087
Ama 199
Ama 044
Ama 083
Bol 016
Bol 017
Bol 015
Bol 031
ST7
Ama 088
Cas 001
Cas 006
Cas 017
Cas 012
Cas 015
Ama 175
Ama 032
Ama 182
Cas 042
Ama 186
Ama 072
Cas 010
Ama 185
Ama 065
Ama 091
Ama 102
Ama 059
Ama 026
Ama 189
Ama 096
Bol 032
Bol 030
Bol 010
Cas 009
Cas 041
Bol 014
b.
Tree scale: 0.1
ssurRNA
ST2
ST7
Bol 026
Bol 024
Cas 030
Bol 036
Bol 022
Bol 029
Cas 008
Bol 023
Cor 001
ST7
Ama 193
Ama 063
Ama 173
Ama 062
Ama 141
Ama 192
Bol 013
Ama 075
Ama 176
Ama 031
Ama 195
Ama 022
Ama 148
Ama 046
Ama 139
Ama 127
Bol 037
Ama 058
c.
Tree scale: 1
ssurRNA
ST3
ST7
Bol 025
Cas 043
Cas 047
Cas 033
Bol 019
Cas 050
Bol 027
Cas 003
ST7
Ama 084
Ama 154
Ama 124
Bol 011
Ama 156
Ama 094
Ama 074
Ama 103
Cas 018
Ama 086
Ama 078
Ama 200
Ama 064
Ama 188
Ama 168
Ama 187
Ama 056
Ama 073
d.
Tree scale: 1
SDHA
ST1
ST7
Ama 088
Ama 026
Ama 059
Ama 032
Ama 186
Ama 087
Ama 185
Ama 091
Ama 044
Ama 083
Ama 199
Cas 015
ST7
Ama 065
Ama 072
Ama 096
Ama 175
Bol 031
Cas 001
Bol 017
Bol 016
Cas 017
Ama 182
Bol 010
Bol 032
Ama 189
Cas 014
Cas 034
Cas 042
Bol 021
Cas 006
Cas 041
Bol 015
Cas 027
Cas 010
Cas 012
Bol 014
Bol 030
Ama 102
Cas 009
e.
Tree scale: 0.1
SDHA
ST2
ST7
Ama 176
Ama 141
Ama 058
Ama 063
Ama 193
Ama 022
Ama 127
Ama 031
ST7
Bol 037
Ama 139
Ama 173
Bol 023
Ama 148
Ama 046
Cor 001
Bol 029
Bol 036
Ama 192
Bol 024
Bol 013
Bol 022
Cas 008
Ama 195
Cas 030
Bol 026
Ama 075
Ama 062
f.
Tree scale: 0.1
SDHA
ST3
ST7
Ama 168
Ama 056
Ama 154
Ama 187
Ama 073
Ama 156
Ama 074
Ama 124
ST7
Ama 064
Ama 103
Cas 050
Ama 084
Bol 019
Ama 188
Cas 047
Ama 086
Cas 033
Ama 200
Cas 043
Cas 003
Bol 025
Ama 094
Bol 011
Cas 018
Bol 027
Ama 078


a.
Amazonas
Bolívar
Casanare
Córdoba
ST7
ST1
0.021
b.
ST2
0.026
c.
ST3
0.064
d.
Amazonas
Bolívar
Casanare
Córdoba
ST7
MV

10. OTROS PRODUCTOS

10.1 Certificados de presentaciones en congresos.

ICOPA 2018
14TH INTERNATIONAL CONGRESS OF PARASITOLOGY

19SUN–24FRI AUGUST 2018
EXCO, DAEGU, KOREA | www.icopa2018.org

This certifies that the following abstract was presented as poster presentation

Development of molecular markers to depict the population genetic structure of four intestinal parasitic protozoa

By

Higuera et al., **Universidad del Rosario, Colombia**

Prof. Min-Ho Choi
Chair of Scientific Committee
Seoul National University College of Medicine, Korea

Prof. Tai-Soon Yong
Chair of Local Organizing Committee
Yonsei University College of Medicine, Korea

Jong-Yil Chai
President
Seoul National University College of Medicine, Korea

Bogotá D.C., January 10th, 2018

Doctor
Adriana Higuera
Colombia

Dear Dr. Higuera,

By means of this letter, I hereby certify that the following abstract was presented as poster during The 2nd International *Blastocystis* Conference held in Bogotá, Colombia at October 11-12th in 2018 under the auspices of The Faculty of Natural Sciences and Mathematics, the Direction of Investigations and Innovation from Universidad del Rosario, Annar Diagnostica S.A.S. and Genproducts S.A.S.

Title: Development of molecular markers to study the population genetic structure of Blastocystis
Authors: Adriana Marcela Higuera, Juan David Ramírez

My kindest regards

JUAN DAVID RAMÍREZ, Ph.D.
2IBC Coordinator
Faculty of Natural Sciences and Mathematics
Universidad del Rosario

Universidad del Rosario
2nd International *Blastocystis* Conference
Sede Quinta Mutis: Carrera 24 No. 63C-69
Tel: (571) 2970200 Ext. 4033
Bogotá, Colombia
2ibc@urosario.edu.co
www.urosario.edu.co

XXV CONGRESO LATINOAMERICANO DE PARASITOLOGÍA
V CONGRESO DE LA ASOC. PANAMEÑA DE MICROBIOLOGÍA Y PARASITOLOGÍA | XVII CONGRESO MEDICINA UP XXII JORNADA CIENTIFICA

La Federación Latinoamericana de Parasitología (FLAP)
y la Asociación Panameña de Microbiología y Parasitología (APMYP) :

CERTIFICAN QUE:

"Adriana Higuera, Ximena Villamizar, Giovanny Herrera, Julio César Giraldo, Luis Reinel Vasquez-A, Plutarco Urbano, Osvaldo Villalobos, Catalina Tovar, Juan David Ramírez, "

presentó el Cartel titulado:

"Detección molecular y genotipificación de protozoos intestinales de diferentes regiones biogeográficas de Colombia"

Sesión:Sesión de Carteles S2 Día: 2019/11/26

Dra. María Mireya Muñoz
Presidenta XXV Congreso FLAP 2019

Dr. Azael Saldaña
Presidente del Comité Científico

Dado en la Ciudad de Panamá, República de Panamá del 24 al 27 de noviembre de 2019.

XVII Congreso Colombiano de
Parasitología y Medicina Tropical
En busca de la eliminación de eventos transmisibles

Certifica que:

Adriana Marcela Higuera G,lvez

Participó en el:

XVII Congreso Colombiano de Parasitología y Medicina Tropical

Que se realizó en la ciudad de Santiago de Cali,
del 4 al 6 de diciembre de 2019

Clara B. Ocampo, MSc, PhD
Presidente

Juan David Ramírez González, PhD
Vicepresidente

10.2 Certificado curso internacional EupathDB.

Certificate of Participation

This award is presented to

Adriana Higuera Gelvez

for participation in the

2018 EuPathDB Eukaryotic Pathogen Database Resources http://EuPathDB.org Workshop

at the University of Georgia

June 17 – 21, 2018.

Jessica Kissinger

10.3 Certificado de realización de pasantía internacional.

DALHOUSIE UNIVERSITY

Biochemistry & Molecular Biology

October 22, 2019

Programa Doctoral de Becas
Universidad del Rosario
Bogotá, Colombia

To Whom It May Concern:

I am writing to verify that **Adriana Marcela Higuera Gélvez** has successfully completed her PhD research internship in my laboratory in the Department of Biochemistry and Molecular Biology at Dalhousie University, Halifax, Canada from May to October, 2019.

If you require any further information, please don't hesitate to contact me.

Sincerely,

Andrew J. Roger, PhD, FRS(C)
Professor, Department of Biochemistry and Molecular Biology
Canada Research Chair in Comparative Genomics & Evolutionary Bioinformatics
Dalhousie University
Halifax, Nova Scotia, Canada

Tel: 1-902-494-2620 (office)
Fax: 1-902-494-1355
Email: aroger@dal.ca

10.4 Certificados otros cursos

eventbrite

Order #1034281277

Introduction to Linux

Academic

Dalhousie University, Killam Library, Room B-400

Wednesday, 18 September 2019 from 2:30 PM to 4:00 PM (ADT)

Free Order

Order Information

Order #1034281277. Ordered by Adriana Marcela Higuera Gélvez on 30 August 2019 11:50 AM

Name

Adriana Marcela Higuera Gélvez

10342812771345843947001

eventbrite

Order #1034287751

Introduction to Shell Scripting

Academic

Dalhousie University, Killam Library, Room B-400

Thursday, 19 September 2019 from 1:00 PM to 2:30 PM (ADT)

Free Order

Order Information

Order #1034287751. Ordered by Adriana Marcela Higuera Gélvez on 30 August 2019 11:54 AM

Name

Adriana Marcela Higuera Gélvez

10342877511345853841001

eventbrite

Order #1034291647

Job Scheduling With Slurm

Academic

Dalhousie University, Killam Library, Room B-400

Thursday, 19 September 2019 from 2:30 PM to 4:00 PM (ADT)

Free Order

Order Information

Order #1034291647. Ordered by Adriana Marcela Higuera Gélvez on 30 August 2019 11:56 AM

Name

Adriana Marcela Higuera Gélvez

10342916471345859791001

11. REFERENCIAS BIBLIOGRÁFICAS

1. MinSalud. Encuesta nacional de parasitismo intestinal en población escolar Colombia, 2012 – 2014. 2015 [Available from: https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/ET/encuesta-nacional-de-parasitismo-2012-2014.pdf.
2. Haque R. Human Intestinal Parasites. Journal of Health, Population, and Nutrition. 2007;25(4):387-91.
3. WHO. Research priorities for helminth infections. World Health Organ Tech Rep Ser. 2012(972):xv-xvii, 1-174, back cover.
4. Cox FEG. Taxonomy and Classification of Human Parasitic Protozoa and Helminths. American Society for Microbiology. 2011.
5. Parija SC, Jeremiah S. Blastocystis: Taxonomy, biology and virulence. Trop Parasitol. 2013;3(1):17-25.
6. Caccio SM, Ryan U. Molecular epidemiology of giardiasis. Mol Biochem Parasitol. 2008;160(2):75-80.
7. Del Coco VF, Molina NB, Basualdo JA, Cordoba MA. [Blastocystis spp.: Advances, controversies and future challenges]. Rev Argent Microbiol. 2017;49(1):110-8.
8. Stensvold CR, Alfellani M, Clark CG. Levels of genetic diversity vary dramatically between Blastocystis subtypes. Infect Genet Evol. 2012;12(2):263-73.
9. Nourrisson C, Scanzi J, Pereira B, NkoudMongo C, Wawrzyniak I, Cian A, et al. Blastocystis is associated with decrease of fecal microbiota protective bacteria: comparative analysis between patients with irritable bowel syndrome and control subjects. PLoS One. 2014;9(11):e111868.
10. Nagel R, Cuttell L, Stensvold CR, Mills PC, Bielefeldt-Ohmann H, Traub RJ. Blastocystis subtypes in symptomatic and asymptomatic family members and pets and response to therapy. Intern Med J. 2012;42(11):1187-95.
11. Adam RD. Biology of Giardia lamblia. Clin Microbiol Rev. 2001;14(3):447-75.
12. Thompson RC, Monis PT. Variation in Giardia: implications for taxonomy and epidemiology. Adv Parasitol. 2004;58:69-137.
13. Upton SJ, Zien CA. Description of a Giardia varani-like flagellate from a water monitor, Varanus salvator, from Malaysia. J Parasitol. 1997;83(5):970-1.
14. Thompson RCA, Monis P. Variation in Giardia: Implications for Taxonomy and Epidemiology2004. 69-137 p.
15. Simpson AG, Inagaki Y, Roger AJ. Comprehensive multigene phylogenies of excavate protists reveal the evolutionary positions of "primitive" eukaryotes. Mol Biol Evol. 2006;23(3):615-25.
16. Cepicka I. Fornicata 2008 [Available from: http://tolweb.org/Fornicata/121182.
17. Yason JA, Rivera WL. Genotyping of Giardia duodenalis isolates among residents of slum area in Manila, Philippines. Parasitol Res. 2007;101(3):681-7.
18. Ankarklev J, Jerlstrom-Hultqvist J, Ringqvist E, Troell K, Svard SG. Behind the smile: cell biology and disease mechanisms of Giardia species. Nat Rev Microbiol. 2010;8(6):413-22.
19. Hellard ME, Sinclair MI, Hogg GG, Fairley CK. Prevalence of enteric pathogens among community based asymptomatic individuals. J Gastroenterol Hepatol. 2000;15(3):290-3.
20. Appelbee AJ, Thompson RC, Olson ME. Giardia and Cryptosporidium in mammalian wildlife--current status and future needs. Trends Parasitol. 2005;21(8):370-6.
21. Thompson RC, Ash A. Molecular epidemiology of Giardia and Cryptosporidium infections. Infect Genet Evol. 2016;40:315-23.
22. Huang DB, White AC. An updated review on Cryptosporidium and Giardia. Gastroenterol Clin North Am. 2006;35(2):291-314, viii.

23. Shen HE, Cao L, Li J, Tian XF, Yang ZH, Wang Y, et al. Visualization of chromosomes in the binucleate intestinal parasite Giardia lamblia. Parasitol Res. 2011;109(5):1439-45.
24. Lv S, Tian L-G, Liu Q, Qian M-B, Fu Q, Steinmann P, et al. Water-Related Parasitic Diseases in China. International Journal of Environmental Research and Public Health. 2013;10(5):1977-2016.
25. Esch KJ, Petersen CA. Transmission and epidemiology of zoonotic protozoal diseases of companion animals. Clin Microbiol Rev. 2013;26(1):58-85.
26. Homan WL, van Enckevort FH, Limper L, van Eys GJ, Schoone GJ, Kasprzak W, et al. Comparison of Giardia isolates from different laboratories by isoenzyme analysis and recombinant DNA probes. Parasitol Res. 1992;78(4):316-23.
27. Abaza SM, Sullivan JJ, Visvesvara GS. Isoenzyme profiles of four strains of Giardia lamblia and their infectivity to jirds. Am J Trop Med Hyg. 1991;44(1):63-8.
28. Andrews RH, Adams M, Boreham PF, Mayrhofer G, Meloni BP. Giardia intestinalis: electrophoretic evidence for a species complex. Int J Parasitol. 1989;19(2):183-90.
29. Nash T. Surface antigen variability and variation in Giardia lamblia. Parasitology Today. 1992;8(7):229-34.
30. Weiss JB, van Keulen H, Nash TE. Classification of subgroups of Giardia lamblia based upon ribosomal RNA gene sequence using the polymerase chain reaction. Mol Biochem Parasitol. 1992;54(1):73-86.
31. van Keulen H, Homan WL, Erlandsen SL, Jarroll EL. A three nucleotide signature sequence in small subunit rRNA divides human Giardia in two different genotypes. J Eukaryot Microbiol. 1995;42(4):392-4.
32. Morgan UM, Constantine CC, Greene WK, Thompson RCA. RAPD (random amplified polymorphic DNA) analysis of Giardia DNA and correlation with isoenzyme data. Transactions of the Royal Society of Tropical Medicine and Hygiene. 1993;87(6):702-5.
33. Nash TE, McCutchan T, Keister D, Dame JB, Conrad JD, Gillin FD. Restriction-endonuclease analysis of DNA from 15 Giardia isolates obtained from humans and animals. J Infect Dis. 1985;152(1):64-73.
34. De Jonckheere JF, Majewska AC, Kasprzak W. Giardia isolates from primates and rodents display the same molecular polymorphism as human isolates. Mol Biochem Parasitol. 1990;39(1):23-9.
35. Ey PL, Andrews RH, Mayrhofer G. Differentiation of major genotypes of Giardia intestinalis by polymerase chain reaction analysis of a gene encoding a trophozoite surface antigen. Parasitology. 1993;106 (Pt 4):347-56.
36. Adam RD, Nash TE, Wellems TE. The Giardia lamblia trophozoite contains sets of closely related chromosomes. Nucleic Acids Res. 1988;16(10):4555-67.
37. Campbell SR, van Keulen H, Erlandsen SL, Senturia JB, Jarroll EL. Giardia sp.: comparison of electrophoretic karyotypes. Experimental parasitology. 1990;71(4):470-82.
38. Mayrhofer G, Andrews RH, Ey PL, Chilton NB. Division of Giardia isolates from humans into two genetically distinct assemblages by electrophoretic analysis of enzymes encoded at 27 loci and comparison with Giardia muris. Parasitology. 1995;111 (Pt 1):11-7.
39. Monis PT, Mayrhofer G, Andrews RH, Homan WL, Limper L, Ey PL. Molecular genetic analysis of Giardia intestinalis isolates at the glutamate dehydrogenase locus. Parasitology. 1996;112 (Pt 1):1-12.
40. Feng Y, Xiao L. Zoonotic potential and molecular epidemiology of Giardia species and giardiasis. Clin Microbiol Rev. 2011;24(1):110-40.
41. M Wielinga C, C A Thompson R. Comparative evaluation of Giardia duodenalis sequence data2007. 1795-821 p.

42. Hopkins RM, Meloni BP, Groth DM, Wetherall JD, Reynoldson JA, Thompson RA. Ribosomal RNA sequencing reveals differences between the genotypes of Giardia isolates recovered from humans and dogs living in the same locality. The Journal of parasitology. 1997:44-51.
43. Berrilli F, Di Cave D, De Liberato C, Franco A, Scaramozzino P, Orecchia P. Genotype characterisation of Giardia duodenalis isolates from domestic and farm animals by SSU-rRNA gene sequencing. Veterinary Parasitology. 2004;122(3):193-9.
44. Pijnacker R, Mughini-Gras L, Heusinkveld M, Roelfsema J, van Pelt W, Kortbeek T. Different risk factors for infection with Giardia lamblia assemblages A and B in children attending day-care centres. Eur J Clin Microbiol Infect Dis. 2016;35(12):2005-13.
45. Armson A, Yang R, Thompson J, Johnson J, Reid S, Ryan UM. Giardia genotypes in pigs in Western Australia: Prevalence and association with diarrhea. Experimental Parasitology. 2009;121(4):381-3.
46. Foronda P, Bargues M, Abreu Acosta N, Periago M, Valero M, Valladares B, et al. Identification of genotypes of Giardia intestinalis of human isolates in Egypt2008. 1177-81 p.
47. Gelanew T, Lalle M, Hailu A, Pozio E, Cacciò SM. Molecular characterization of human isolates of Giardia duodenalis from Ethiopia. Acta Tropica. 2007;102(2):92-9.
48. Lebbad M, Mattsson JG, Christensson B, Ljungström B, Backhans A, Andersson JO, et al. From mouse to moose: Multilocus genotyping of Giardia isolates from various animal species. Veterinary Parasitology. 2010;168(3):231-9.
49. Palmer CS, Traub RJ, Robertson ID, Devlin G, Rees R, Thompson RA. Determining the zoonotic significance of Giardia and Cryptosporidium in Australian dogs and cats. Veterinary parasitology. 2008;154(1):142-7.
50. Read CM, Monis PT, Thompson RA. Discrimination of all genotypes of Giardia duodenalis at the glutamate dehydrogenase locus using PCR-RFLP. Infection, Genetics and Evolution. 2004;4(2):125-30.
51. Sprong H, Cacciò SM, van der Giessen JWB, on behalf of the Zn, partners. Identification of Zoonotic Genotypes of Giardia duodenalis. PLoS Neglected Tropical Diseases. 2009;3(12):e558.
52. Traub RJ, Inpankaew T, Reid SA, Sutthikornchai C, Sukthana Y, Robertson ID, et al. Transmission cycles of Giardia duodenalis in dogs and humans in Temple communities in Bangkok—a critical evaluation of its prevalence using three diagnostic tests in the field in the absence of a gold standard. Acta tropica. 2009;111(2):125-32.
53. Lasek-Nesselquist E, Mark Welch D, Sogin M. The identification of a new Giardia duodenalis assemblage in marine vertebrates and a preliminary analysis of G. duodenalis population biology in marine systems2010. 1063-74 p.
54. Ryan U, Cacciò SM. Zoonotic potential of Giardia. International Journal for Parasitology. 2013;43(12):943-56.
55. Lee MF, Auer H, Lindo JF, Walochnik J. Multilocus sequence analysis of Giardia spp. isolated from patients with diarrhea in Austria. Parasitol Res. 2017;116(2):477-81.
56. Plutzer J, Ongerth J, Karanis P. Giardia taxonomy, phylogeny and epidemiology: Facts and open questions. Int J Hyg Environ Health. 2010;213(5):321-33.
57. Wang H, Zhao G, Chen G, Jian F, Zhang S, Feng C, et al. Multilocus genotyping of Giardia duodenalis in dairy cattle in Henan, China. PLoS One. 2014;9(6):e100453.
58. Molina N, Minvielle M, Grenovero S, Salomon C, Basualdo J. High prevalences of infection with Giardia intestinalis genotype B among children in urban and rural areas of Argentina. Ann Trop Med Parasitol. 2011;105(4):299-309.
59. Sanchez A, Munoz M, Gomez N, Tabares J, Segura L, Salazar A, et al. Molecular Epidemiology of Giardia, Blastocystis and Cryptosporidium among Indigenous Children from the Colombian Amazon Basin. Front Microbiol. 2017;8:248.

60. Ramirez JD, Heredia RD, Hernandez C, Leon CM, Moncada LI, Reyes P, et al. Molecular diagnosis and genotype analysis of Giardia duodenalis in asymptomatic children from a rural area in central Colombia. Infect Genet Evol. 2015;32:208-13.
61. Hernández PC, Morales L, Chaparro-Olaya J, Sarmiento D, Jaramillo JF, Ordoñez GA, et al. Intestinal parasitic infections and associated factors in children of three rural schools in Colombia. A cross-sectional study. PLoS One. 2019;14(7):e0218681.
62. Rodríguez V, Espinosa O, Carranza JC, Duque S, Arévalo A, Clavijo JA, et al. Genotipos de Giardia duodenalis en muestras de niños de las guarderías del Instituto Colombiano de Bienestar Familiar y de perros en Ibagué, Colombia. Biomédica. 2014;34(2).
63. Coronato Nunes B, Pavan MG, Jaeger LH, Monteiro KJ, Xavier SC, Monteiro FA, et al. Spatial and Molecular Epidemiology of Giardia intestinalis Deep in the Amazon, Brazil. PLoS One. 2016;11(7):e0158805.
64. Perez Cordon G, Cordova Paz Soldan O, Vargas Vasquez F, Velasco Soto JR, Sempere Bordes L, Sanchez Moreno M, et al. Prevalence of enteroparasites and genotyping of Giardia lamblia in Peruvian children. Parasitol Res. 2008;103(2):459-65.
65. Otero-Negrete JJ, Ibarra-Velarde F, Martínez-Gordillo MN, Ponce-Macotela M. Prevalencia de Giardia intestinalis y predominio de genotipos zoonóticos en ovinos y bovinos de traspatio de cinco estados de la República Mexicana. Veterinaria México. 2011;42:219-26.
66. Torres-Romero JC, Euan-Canto AdJ, Benito-González N, Padilla-Montaño N, Huchin-Chan C, Lara-Riegos J, et al. Intestinal parasites and genotyping of Giardia duodenalis in children: first report of genotype B in isolates from human clinical samples in Mexico. Memórias do Instituto Oswaldo Cruz. 2014;109(3):388-90.
67. Pelayo L, Nunez FA, Rojas L, Furuseth Hansen E, Gjerde B, Wilke H, et al. Giardia infections in Cuban children: the genotypes circulating in a rural population. Ann Trop Med Parasitol. 2008;102(7):585-95.
68. Garcia RJ, French N, Pita A, Velathanthiri N, Shrestha R, Hayman D. Local and global genetic diversity of protozoan parasites: Spatial distribution of Cryptosporidium and Giardia genotypes. PLoS Negl Trop Dis. 2017;11(7):e0005736.
69. Monis PT, Andrews RH, Mayrhofer G, Ey PL. Molecular systematics of the parasitic protozoan Giardia intestinalis. Mol Biol Evol. 1999;16(9):1135-44.
70. Monis PT, Caccio SM, Thompson RCA. Variation in Giardia: towards a taxonomic revision of the genus. Trends in Parasitology. 2009;25(2):93-100.
71. Faria CP, Zanini GM, Dias GS, da Silva S, Sousa MDC. New multilocus genotypes of Giardia lamblia human isolates. Infect Genet Evol. 2017;54:128-37.
72. Caccio SM, Beck R, Lalle M, Marinculic A, Pozio E. Multilocus genotyping of Giardia duodenalis reveals striking differences between assemblages A and B. Int J Parasitol. 2008;38(13):1523-31.
73. Traub R, Monis P, Robertson I, Irwin P, Mencke N, Thompson RCA. Epidemiology and molecular evidence support the zoonotic transmission of Giardia among humans and dogs living in the same community2004. 53-62 p.
74. Geurden T, Levecke B, Cacciò S, Visser A, De Groote G, Casaert S, et al. Multilocus genotyping of Cryptosporidium and Giardia in non-outbreak related cases of diarrhoea in human patients in Belgium2009. 1161-8 p.
75. Cooper MA, Adam RD, Worobey M, Sterling CR. Population genetics provides evidence for recombination in Giardia. Curr Biol. 2007;17(22):1984-8.
76. Ankarklev J, Lebbad M, Einarsson E, Franzen O, Ahola H, Troell K, et al. A novel high-resolution multilocus sequence typing of Giardia intestinalis Assemblage A isolates reveals zoonotic transmission, clonal outbreaks and recombination. Infect Genet Evol. 2018;60:7-16.

77. Durigan M, Ciampi-Guillardi M, Rodrigues RCA, Greinert-Goulart JA, Siqueira-Castro ICV, Leal DAG, et al. Population genetic analysis of Giardia duodenalis: genetic diversity and haplotype sharing between clinical and environmental sources. MicrobiologyOpen. 2017;6(2):e00424.
78. Choy SH, Mahdy MAK, Al-Mekhlafi HM, Low VL, Surin J. Population expansion and gene flow in Giardia duodenalis as revealed by triosephosphate isomerase gene. Parasites & Vectors. 2015;8:454.
79. Andersson JO. Double peaks reveal rare diplomonad sex. Trends Parasitol. 2012;28(2):46-52.
80. Lasek-Nesselquist E, Welch DM, Thompson RC, Steuart RF, Sogin ML. Genetic exchange within and between assemblages of Giardia duodenalis. J Eukaryot Microbiol. 2009;56(6):504-18.
81. Durigan M, Abreu AG, Zucchi MI, Franco RM, de Souza AP. Genetic diversity of Giardia duodenalis: multilocus genotyping reveals zoonotic potential between clinical and environmental sources in a metropolitan region of Brazil. PLoS One. 2014;9(12):e115489.
82. Aurrecoechea C, Brestelli J, Brunk BP, Carlton JM, Dommer J, Fischer S, et al. GiardiaDB and TrichDB: integrated genomic resources for the eukaryotic protist pathogens Giardia lamblia and Trichomonas vaginalis. Nucleic Acids Res. 2009;37(Database issue):D526-30.
83. Adam RD, Dahlstrom EW, Martens CA, Bruno DP, Barbian KD, Ricklefs SM, et al. Genome sequencing of Giardia lamblia genotypes A2 and B isolates (DH and GS) and comparative analysis with the genomes of genotypes A1 and E (WB and Pig). Genome Biol Evol. 2013;5(12):2498-511.
84. Ankarklev J, Franzén O, Peirasmaki D, Jerlström-Hultqvist J, Lebbad M, Andersson J, et al. Comparative genomic analyses of freshly isolated Giardia intestinalis assemblage A isolates. BMC Genomics. 2015;16(1):697.
85. Franzen O, Jerlstrom-Hultqvist J, Castro E, Sherwood E, Ankarklev J, Reiner DS, et al. Draft genome sequencing of giardia intestinalis assemblage B isolate GS: is human giardiasis caused by two different species? PLoS Pathog. 2009;5(8):e1000560.
86. Morrison HG, McArthur AG, Gillin FD, Aley SB, Adam RD, Olsen GJ, et al. Genomic minimalism in the early diverging intestinal parasite Giardia lamblia. Science. 2007;317(5846):1921-6.
87. Wielinga C, Thompson RCA, Monis P, Ryan U. Identification of polymorphic genes for use in assemblage B genotyping assays through comparative genomics of multiple assemblage B Giardia duodenalis isolates. Mol Biochem Parasitol. 2015;201(1):1-4.
88. Jerlstrom-Hultqvist J, Franzen O, Ankarklev J, Xu F, Nohynkova E, Andersson JO, et al. Genome analysis and comparative genomics of a Giardia intestinalis assemblage E isolate. BMC Genomics. 2010;11:543.
89. Pollo SMJ, Reiling SJ, Wit J, Workentine ML, Guy RA, Batoff GW, et al. Benchmarking hybrid assemblies of Giardia and prediction of widespread intra-isolate structural variation. Parasit Vectors. 2020;13(1):108.
90. Ortega-Pierres MG, Jex AR, Ansell BRE, Svard SG. Recent advances in the genomic and molecular biology of Giardia. Acta Trop. 2018;184:67-72.
91. Poxleitner MK, Carpenter ML, Mancuso JJ, Wang CJ, Dawson SC, Cande WZ. Evidence for karyogamy and exchange of genetic material in the binucleate intestinal parasite Giardia intestinalis. Science. 2008;319(5869):1530-3.
92. Xu F, Jerlstrom-Hultqvist J, Andersson JO. Genome-wide analyses of recombination suggest that Giardia intestinalis assemblages represent different species. Mol Biol Evol. 2012;29(10):2895-8.
93. Ramesh MA, Malik SB, Logsdon JM, Jr. A phylogenomic inventory of meiotic genes; evidence for sex in Giardia and an early eukaryotic origin of meiosis. Curr Biol. 2005;15(2):185-91.

94. Barash NR, Nosala C, Pham JK, McInally SG, Gourguechon S, McCarthy-Sinclair B, et al. Giardia Colonizes and Encysts in High-Density Foci in the Murine Small Intestine. mSphere. 2017;2(3):e00343-16.
95. Zierdt CH. Blastocystis hominis--past and future. Clinical Microbiology Reviews. 1991;4(1):61-79.
96. Zierdt CH, Rude WS, Bull BS. Protozoan characteristics of Blastocystis hominis. American Journal of Clinical Pathology. 1967;48(5):495-501.
97. Silberman JD, Sogin ML, Leipe DD, Clark CG. Human parasite finds taxonomic home. Nature. 1996;380(6573):398.
98. 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. Journal of Eukaryotic Microbiology. 2002;49(1):42-53.
99. Tan KSW. New Insights on Classification, Identification, and Clinical Relevance of Blastocystis spp. Clinical Microbiology Reviews. 2008;21(4):639-65.
100. Londono-Franco AL, Loaiza-Herrera J, Lora-Suarez FM, Gomez-Marin JE. [Blastocystis sp. frequency and sources among children from 0 to 5 years of age attending public day care centers in Calarca, Colombia]. Biomedica. 2014;34(2):218-27.
101. Cavalier-Smith T. A revised six-kingdom system of life. Biological Reviews. 1998;73(3):203-66.
102. ZIERDT CH. Cytochrome-Free Mitochondria of an Anaerobic Protozoan—Blastocystis hominis. Journal of Eukaryotic Microbiology. 1986;33(1):67-9.
103. Stechmann A, Hamblin K, Pérez-Brocal V, Gaston D, Richmond Gregory S, van der Giezen M, et al. Organelles in Blastocystis that Blur the Distinction between Mitochondria and Hydrogenosomes. Current Biology. 2008;18(8):580-5.
104. Tan K, Howe J, Yap EH, Singh M. Do Blastocystis hominis colony forms undergo programmed cell death?2001. 362-7 p.
105. Keenan TW, Zierdt CH. Lipid biosynthesis by axenic strains of Blastocystis hominis. Comparative Biochemistry and Physiology Part B: Comparative Biochemistry. 1994;107(4):525-31.
106. Zhang X, Zhang S, Qiao J, Wu X, Zhao L, Liu Y, et al. Ultrastructural insights into morphology and reproductive mode of Blastocystis hominis. Parasitology research. 2012;110(3):1165-72.
107. Tan T, Suresh K. Predominance of amoeboid forms of Blastocystis hominis in isolates from symptomatic patients. Parasitology research. 2006;98(3):189-93.
108. M Vassalos C, Spanakos G, Vassalou E, Papadopoulou C, Vakalis N. Differences in Clinical Significance and Morphologic Features of Blastocystis sp Subtype 32010. 251-8 p.
109. Kyaw-Thu M, Singh M, Howe J, C Ho L, Tan K, Q Chen X, et al. Development of Blastocystis hominis cysts into vacuolar forms in vitro1999. 103-8 p.
110. 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;94(6):391-6.
111. Lee LI, Chye TT, Karmacharya BM, Govind SK. Blastocystis sp.: waterborne zoonotic organism, a possibility? Parasites & Vectors. 2012;5:130-.
112. Li L-H, Zhou X-N, Du Z-W, Wang X-Z, Wang L-B, Jiang J-Y, et al. Molecular epidemiology of human Blastocystis in a village in Yunnan province, China. Parasitology International. 2007;56(4):281-6.
113. Stensvold CR, Suresh GK, Tan KS, Thompson RC, Traub RJ, Viscogliosi E, et al. Terminology for Blastocystis subtypes--a consensus. Trends Parasitol. 2007;23(3):93-6.
114. Ramírez JD, Sánchez A, Hernández C, Flórez C, Bernal MC, Giraldo JC, et al. Geographic distribution of human Blastocystis subtypes in South America. Infection, Genetics and Evolution. 2016;41:32-5.

115. 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;14:164.
116. Stenzel DJ, Boreham PF. Blastocystis hominis revisited. Clinical Microbiology Reviews. 1996;9(4):563-84.
117. Cirioni O, Giacometti A, Drenaggi D, Ancarani F, Scalise G. Prevalence and clinical relevance of Blastocystis hominis in diverse patient cohorts. European journal of epidemiology. 1999;15(4):387-91.
118. Abe N, Nagoshi M, Takami K, Sawano Y, Yoshikawa H. A survey of Blastocystis sp. in livestock, pets, and zoo animals in Japan. Veterinary Parasitology. 2002;106(3):203-12.
119. Boreham PF, Stenzel DJ. Blastocystis in humans and animals: morphology, biology, and epizootiology. Advances in parasitology. 1993;32:1-70.
120. Scanlan PD. Blastocystis: past pitfalls and future perspectives. Trends Parasitol. 2012;28(8):327-34.
121. Stensvold CR, Clark CG. Current status of Blastocystis: A personal view. Parasitol Int. 2016;65(6 Pt B):763-71.
122. 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(3):R29.
123. Poirier P, Wawrzyniak I, Vivarès CP, Delbac F, El Alaoui H. New Insights into Blastocystis spp.: A Potential Link with Irritable Bowel Syndrome. PLoS Pathogens. 2012;8(3):e1002545.
124. Anderson TJ. The dangers of using single locus markers in parasite epidemiology: Ascaris as a case study. Trends Parasitol. 2001;17(4):183-8.
125. Abe N, Wu Z, Yoshikawa H. Molecular characterization of Blastocystis isolates from primates. Vet Parasitol. 2003;113(3-4):321-5.
126. Clark CG. Extensive genetic diversity in Blastocystis hominis. Mol Biochem Parasitol. 1997;87(1):79-83.
127. Yoshikawa H, Nagono I, Yap EH, Singh M, Takahashi Y. DNA polymorphism revealed by arbitrary primers polymerase chain reaction among Blastocystis strains isolated from humans, a chicken, and a reptile. J Eukaryot Microbiol. 1996;43(2):127-30.
128. Arisue N, Hashimoto T, Yoshikawa H. Sequence heterogeneity of the small subunit ribosomal RNA genes among blastocystis isolates. Parasitology. 2003;126(Pt 1):1-9.
129. Ho LC, Armiugam A, Jeyaseelan K, Yap EH, Singh M. Blastocystis elongation factor-1alpha: genomic organization, taxonomy and phylogenetic relationships. Parasitology. 2000;121 (Pt 2):135-44.
130. Scicluna SM, Tawari B, Clark CG. DNA barcoding of blastocystis. Protist. 2006;157(1):77-85.
131. Ramirez JD, Sanchez LV, Bautista DC, Corredor AF, Florez AC, Stensvold CR. Blastocystis subtypes detected in humans and animals from Colombia. Infect Genet Evol. 2014;22:223-8.
132. Malheiros AF, Stensvold CR, Clark CG, Braga GB, Shaw JJ. Short report: Molecular characterization of Blastocystis obtained from members of the indigenous Tapirape ethnic group from the Brazilian Amazon region, Brazil. Am J Trop Med Hyg. 2011;85(6):1050-3.
133. 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(4):497-509.
134. Ramirez JD, Florez C, Olivera M, Bernal MC, Giraldo JC. Blastocystis subtyping and its association with intestinal parasites in children from different geographical regions of Colombia. PLoS One. 2017;12(2):e0172586.
135. Stensvold CR. Blastocystis: Genetic diversity and molecular methods for diagnosis and epidemiology. Trop Parasitol. 2013;3(1):26-34.

136. Jiménez PA, Jaimes JE, Ramírez JD. A summary of Blastocystis subtypes in North and South America. Parasit Vectors. 2019;12(1):376.
137. Stensvold CR, Christiansen DB, Olsen KEP, Nielsen HV. Blastocystis sp. Subtype 4 is Common in Danish Blastocystis-Positive Patients Presenting with Acute Diarrhea. The American Journal of Tropical Medicine and Hygiene. 2011;84(6):883-5.
138. 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;1(5):167-78.
139. Santin M, Gomez-Munoz 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(1):205-12.
140. 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(1):85-93.
141. Stensvold CR, Nielsen HV, Molbak K, Smith HV. Pursuing the clinical significance of Blastocystis--diagnostic limitations. Trends Parasitol. 2009;25(1):23-9.
142. Villalobos G, Orozco-Mosqueda GE, Lopez-Perez M, Lopez-Escamilla E, Cordoba-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.
143. Villanueva-Garcia C, Gordillo-Chavez EJ, Lopez-Escamilla E, Rendon-Franco E, Munoz-Garcia CI, Gama L, et al. Clarifying the Cryptic Host Specificity of Blastocystis spp. Isolates from Alouatta palliata and A. pigra Howler Monkeys. PLoS One. 2017;12(1):e0169637.
144. 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(1):564.
145. Gentekaki E, Curtis BA, Stairs CW, Klimes V, Elias M, Salas-Leiva DE, et al. Extreme genome diversity in the hyper-prevalent parasitic eukaryote Blastocystis. PLoS Biol. 2017;15(9):e2003769.
146. Wawrzyniak I, Courtine D, Osman M, Hubans-Pierlot C, Cian A, Nourrisson C, et al. Draft genome sequence of the intestinal parasite Blastocystis subtype 4-isolate WR1. Genomics Data. 2015;4:22-3.
147. Andersen LO, Bonde I, Nielsen HB, Stensvold CR. A retrospective metagenomics approach to studying Blastocystis. FEMS Microbiol Ecol. 2015;91(7).
148. Perez-Brocal V, Clark CG. Analysis of two genomes from the mitochondrion-like organelle of the intestinal parasite Blastocystis: complete sequences, gene content, and genome organization. Mol Biol Evol. 2008;25(11):2475-82.
149. Wawrzyniak I, Roussel M, Diogon M, Couloux A, Texier C, Tan KS, et al. Complete circular DNA in the mitochondria-like organelles of Blastocystis hominis. Int J Parasitol. 2008;38(12):1377-82.
150. Morens DM, Folkers GK, Fauci AS. The challenge of emerging and re-emerging infectious diseases. Nature. 2004;430(6996):242-9.
151. Jacobsen KH, Ribeiro PS, Quist BK, Rydbeck BV. Prevalence of Intestinal Parasites in Young Quichua Children in the Highlands of Rural Ecuador. Journal of Health, Population, and Nutrition. 2007;25(4):399-405.
152. Xiao L. Molecular epidemiology of cryptosporidiosis: an update. Exp Parasitol. 2010;124(1):80-9.
153. Thompson RCA, Ash A. Molecular epidemiology of Giardia and Cryptosporidium infections. Infection, Genetics and Evolution. 2016;40:315-23.