Flecha, M.J. *et al.* (2015) 'Detection and molecular characterisation of Giardia duodenalis, Cryptosporidium spp. and Entamoeba spp. among patients with gastrointestinal symptoms in Gambo Hospital, Oromia Region, southern Ethiopia', *Tropical medicine & international health*, 20(9), pp. 1213–1222.

doi:10.1111/tmi.12535

(Article begins on next page)

Detection and molecular characterisation of *Giardia duodenalis*, *Cryptosporidium* spp. and *Entamoeba* spp. among patients with gastrointestinal symptoms in Gambo Hospital, Oromia Region, southern Ethiopia

María J. Flecha[1], Cynthia M. Benavides[1], Gabriel Tissiano[2], Abraham Tesfamariam[2], Juan Cuadros[1], Aida de Lucio[3], Begoña Bailo[3], Lourdes Cano[3], Isabel Fuentes[3] and David Carmena[3]

1 *Hospital Príncipe de Asturias, Alcalá de Henares, Madrid, Spain*
2 *Gambo Rural General Hospital, Shasehmane, Ethiopia*
3 *Centro Nacional de Microbiología, Majadahonda, Madrid, Spain*

Summary

OBJECTIVES To assess the prevalence and genetic diversity of the enteric protozoa species *G. duodenalis*, *Cryptosporidium* spp. and *Entamoeba histolytica* in individuals with gastrointestinal symptoms compatible with infections by these pathogens seeking medical attention in a rural area in southern Ethiopia.

METHODS A total of 92 stool samples were initially screened by direct microscopy and immunochromatography and further confirmed by molecular methods. *G. duodenalis*-positive samples were molecularly characterised by multilocus genotyping of the glutamate dehydrogenase and β-giardin genes of the parasite. PCR and DNA sequence analysis of the gene encoding the 60-kDa glycoprotein was used for the subtyping of *Cryptosporidium* isolates. Detection and differential diagnosis of *E. histolytica*/*dispar* were conducted by real-time PCR.

RESULTS PCR-based prevalences were 10.9% for *G. duodenalis*, 1.1% for *Cryptosporidium* spp. and 3.3% for *Entamoeba* spp. Seven (four novel and three known) subtypes of *G. duodenalis* assemblage B were identified at the *GDH* locus and 5 (one novel and four known) at the *BG* locus. A novel variant of *C. hominis* subtype IbA9G3 was also identified. Two *Entamoeba* isolates were assigned to *E. dispar* and an additional one to *E. histolytica*.

CONCLUSION Although preliminary, our results strongly suggest that giardiasis, cryptosporidiosis and amoebiasis represent a significant burden in Ethiopian rural population.

keywords intestinal protozoan, *Giardia*, *Cryptosporidium*, *Entamoeba*, gastrointestinal symptoms, neglected disease initiative, Ethiopia

Introduction

Giardia duodenalis, *Cryptosporidium* spp. and *Entamoeba histolytica* are considered the most common and important causes of protozoan-diarrhoeal disease in humans, being responsible for a significant amount of morbidity and mortality worldwide. In developing countries, about 200 million people have symptomatic giardiasis, with some 500 000 new cases reported each year [1]. *Cryptosporidium* accounts for up to 20% of all cases of childhood diarrhoea [2], whereas invasive amoebic infection by *E. histolytica* affects 50 million people worldwide each year, resulting in 40 000–100 000 deaths annually [3]. In addition, giardiasis, cryptosporidiosis and amoebiasis may result on severe diarrhoea, impaired intestinal absorptive function and malnutrition, conditions often associated with poor cognitive function and failure to thrive in early childhood [4, 5]. Due to their associated disease burden and elevated socio-economic consequences, both *Giardia* and *Cryptosporidium* were included in WHO's Neglected Disease Initiative in 2004 [6].

Giardia duodenalis, *Cryptosporidium* spp. and *E. histolytica*/*dispar* are endemic in Ethiopia. Infections by these enteric pathogens have been recurrently documented in epidemiological studies targeting individuals with diarrhoea seeking medical care [7–9] and community [10], paediatric [11–14], adult HIV patient [15, 16] and prison inmate [17] populations. Reported infection rates for giardiasis, cryptosporidiosis and amoebiasis in the country were in the range of 2% to 35%, depending

on the setting and the population under evaluation. However, it must be noticed that these figures are very likely underestimations of the actual disease burdens because (i) most surveys were based on the examination of a single stool sample, so infected subjects with intermittent (oo)cyst shedding may be missed, and (ii) diagnosis was primarily by conventional microscopy, a technique compromised by relatively low sensitivities even if concentration procedures are also carried out.

Molecular tools have greatly contributed to improve our understanding of the taxonomy, population genetics and epidemiology of *G. duodenalis*, *Cryptosporidium* spp. and *Entamoeba* spp. in human and animal populations. Commonly used markers for species/genotypes differentiation include the small-subunit rRNA and the 60-kDa glycoprotein genes (among others) in *Cryptosporidium* spp. and *Entamoeba* spp. and the genes coding for glutamate dehydrogenase, beta-giardin and triosephosphate isomerase in *G. duodenalis* [18, 19]. Compared with prevalence data, molecular information on the genotypic diversity of *G. duodenalis*, *Cryptosporidium* spp. and *Entamoeba* spp. in Ethiopian human populations is far scarcer. Assemblages A, B and F (alone or in combination) are the only *G. duodenalis* genotypes causing human infections reported to date [20]. Within the *Cryptosporidium* genus, 97% of the isolates characterised to date have been assigned to *C. parvum*, with *C. hominis* being responsible for the remaining infections [7]. Differentiation between *E. histolytica* and *E. dispar* has not been attempted at the molecular level hitherto.

Given this scenario, we present here a preliminary assessment of the current prevalence and genetic diversity of *G. duodenalis*, *Cryptosporidium* spp. and *E. histolytica* in individuals with symptoms compatible with gastrointestinal infection by these pathogens attending a rural hospital in southern Ethiopia.

Materials and methods

Ethics

The study protocol was reviewed and approved by the managing direction of the Gambo Rural General Hospital (GRGH). Written informed consent was not required because only fully anonymised clinical and epidemiological information on participating patients were used. However, verbal informed consent was obtained from all the participants for stool collection and examination.

Area of study

The survey was carried out in GRGH, located in the Oromia Region, 245 km from Addis Ababa, Ethiopia. The GRGH has 150 inpatient beds and receives approximately 37 000 outpatient visits annually. It serves a catchment population of approximately 94 600 people from the surrounding districts (*woredas*) of Kore and Arsi Negele, which comprise 13 wards (*kebeles*) or neighbourhood associations. The area is some 2200 m above sea level and has an annual temperature range from 13 to 30 °C. The main economic activity of the population is agriculture and animal husbandry.

Stool sample collection

A community-based cross-sectional study was conducted in the summer of 2014 (July–August) among people seeking medical attention at the GRGH with gastrointestinal symptoms compatible with illnesses caused by the enteric protozoan *G. duodenalis*, *Cryptosporidium* spp. and/or *E. histolytica*. Symptoms considered included acute (defined as the abrupt onset of at least 3 loose stools per day) or persistent (defined as an episode that lasts longer than 14 days) diarrhoea, abdominal pain, weight loss, cramps, fever, nausea and vomiting. A single fresh stool sample was collected per patient. Collected samples were labelled with anonymised study codes to prevent the patient identity privacy. Stool samples were kept at 4 °C and processed within 2 days of collection. No frozen or preserved samples were used in this study.

Direct microscopy

Following routine diagnostic procedures at GRGH, a portion of each stool sample was processed and direct saline (0.85% NaCl solution) wet mounts were microscopically examined at 100 magnification, switching to 400 magnification when structures morphologically compatible with trophozoites or cysts of *G. duodenalis* and *Entamoeba* spp. were suspected. Thin smears were also prepared and stained by Ziehl–Neelsen technique for the specific identification of *Cryptosporidium* oocysts. Examinations were conducted by experienced staff at GRGH.

Immunochromatographic test (ICT) for the detection of *Giardia*, *Cryptosporidium* and *Entamoeba*

A commercially available solid-phase qualitative ICT for the rapid simultaneous detection of *G. duodenalis*, *Cryptosporidium* spp. and *Entamoeba* spp. (Cer Test Biotec S.L., Zaragoza, Spain) was used according to the manufacturer's instructions. The assays were conducted at room temperature exclusively on fresh faecal samples. This technique was based on genus-specific monoclonal mouse antibodies directed against *Giardia*, *Cryptosporidium*

and *Entamoeba*. Claimed diagnostic sensitivities and specificities ranged from 97% to 99% for all three species.

DNA extraction and purification

Total DNA was extracted from all samples that tested positive by any of the screening tests (direct microscopy and/or ICT) used. A new, fresh aliquot (~220 mg) of each stool sample was homogenised in stool lysis buffer and incubated at 95 °C for 10 min. The DNA released from disrupted *Giardia* and *Entamoeba* cysts and *Cryptosporidium* oocysts was subsequently extracted using the QIAamp DNA Stool Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Purified DNA samples (200 µl) were stored at −20 °C and shipped to the National Centre for Microbiology, Spain, for further molecular analysis.

Molecular detection of *Giardia duodenalis*

A real-time PCR (RT-PCR) was initially used for the specific detection of *G. duodenalis* in faecal samples [21]. This assay targeted a 62-bp region of the small-subunit ribosomal RNA (*SSU* rRNA) gene of the parasite using the primer pair Gd-80F and Gd-127R and the probe described in Table 1. Amplification reactions were performed in a volume of 25 µl containing 3 µl of genomic DNA, 12.5 pmol of each primer, 10 pmol of probe and 1X TaqMan® Gene Expression Master Mix (Applied Biosystems, California, USA). We adopted the amplification protocol for TaqMan® recommended by the manufacturer, consisting on an initial hold step of 2 min at 55 °C and 15 min at 95 °C followed by 45 cycles of 15 s at 95 °C and 1 min at 60 °C. Appropriate positive, negative and inhibition controls were routinely included in each round of RT-PCR assays. Amplification and detection of parasitic DNA were performed on a Corbett Rotor-Gene 6000 RT-PCR cycler (Qiagen Corbett, Hilden, Germany). Rotor Gene 6000 Series software version 1.7 was used for data analysis. FAM fluorescence signal (510 nm) was measured at the end of the annealing step of each cycle. The ramping of the machine was 10 °C/s in every step.

Molecular characterisation of *Giardia duodenalis* isolates

Giardia duodenalis isolates confirmed by RT-PCR were subsequently analysed by multilocus genotyping using two gene loci: glutamate dehydrogenase (*GDH*) and ß-giardin (*BG*). The amplification of the *GDH* gene was performed by a semi-nested PCR with minor modifications [22]. The primer pair GDHeF and GDHiF (Table 1)

Table 1 List of oligonucleotides used for the molecular detection and characterisation of *Giardia duodenalis, Cryptosporidium* spp. and *Entamoeba histolytica/dispar* in this study

Target organism	Locus	Oligonucleotide	Sequence (5′–3′) and labels	Reference
Giardia duodenalis	*SSU* rRNA	Primer Gd-80F	GACGGCTCAGGACAACGGTT	[21]
Giardia duodenalis	*SSU* rRNA	Primer Gd-127R	TTGCCAGCGGTGTCCG	[21]
Giardia duodenalis	*SSU* rRNA	Probe	FAM–CCCGCGGCGGTCCCTGCTAG–BHQ1	[21]
Giardia duodenalis	*GDH*	Primer GDHeF	TCAACGTYAAYCGYGGYTTCCGT	[22]
Giardia duodenalis	*GDH*	Primer GDHiF	CAGTACACCTCYGCTCTCGG	[22]
Giardia duodenalis	*GDH*	Primer GDHiR	GTTRTCCTTGCACATCTCC	[22]
Giardia duodenalis	*BG*	Primer G7-F	AAGCCCGACGACCTCACCCGCAGTGC	[23]
Giardia duodenalis	*BG*	Primer G759-R	GAGGCCGCCCTGGATCTTCGAGACGAC	[23]
Giardia duodenalis	*BG*	Primer G99-F	GAACGAACGAGATCGAGGTCCG	[23]
Giardia duodenalis	*BG*	Primer G609-R	CTCGACGAGCTTCGTGTT	[23]
Cryptosporidium spp.	*GP60*	Primer AL-3531	ATAGTCTCCGCTGTATTC	[24]
Cryptosporidium spp.	*GP60*	Primer Al-3535	GGAAGGAACGATGTATCT	[24]
Cryptosporidium spp.	*GP60*	Primer AL-3532	TCCGCTGTATTCTCAGCC	[24]
Cryptosporidium spp.	*GP60*	Primer AL-3534	GCAGAGGAACCAGCATC	[24]
Entamoeba histolytica/dispar	*SSU* rRNA	Primer Ehd-239F	ATTGTCGTGGCATCCTAACTCA	[25]
Entamoeba histolytica/dispar	*SSU* rRNA	Primer Ehd-88R	GCGGACGGCTCATTATAACA	[25]
Entamoeba histolytica	*SSU* rRNA	Probe	FAM–TCATTGAATGAATTGGC CATTT–MGB	[26]
Entamoeba dispar	*SSU* rRNA	Probe	VIC–TTACTTACATAAATTGGC CACTTTG–MGB	[26]

BHQ, Black Hole Quencher; MGB, Minor Groove Binder. Fluorophores: FAM and VIC.

was used in the primary PCR with 5 μl of genomic DNA. Five microlitres of PCR product from the primary reaction was added to the secondary PCR containing the primer pair GDHiF and GDHiF (Table 1) to yield a 432-bp fragment. PCR mixtures consisted of 12.5 pmol of each primer, 200 μM of each dNTP, 1.5 mM $MgCl_2$, 2.5 units of Taq DNA polymerase and 1X Reaction Buffer in a final volume of 25 μl. The primary and secondary PCRs were carried out as follows: 1 cycle of 95 °C for 3 min, followed by 35 cycles of 95 °C for 30 s, 55 °C for 30 s and 72 °C for 1 min. A final extension of 72 °C for 7 min and a 4 °C hold was used.

The amplification of the *BG* gene was performed using a nested PCR with minor modifications [23]. The primer pair G7_F and G759_R (Table 1) was used in the primary PCR with 3 μl of genomic DNA. Three microlitres of PCR product from the primary reaction were added to the secondary PCR containing the primer pair G99_F and G609_R (Table 1) to yield a 511-bp fragment. PCR mixtures consisted of 10 pmol of each primer, 200 μM of each dNTP, 1.5 mM $MgCl_2$, 2.5 units of Taq DNA polymerase and 10X Reaction Buffer. The primary PCR was carried out with the following amplification condition: 1 cycle of 95 °C for 7 min, followed by 35 cycles of 95 °C for 30 s, 65 °C for 30 s and 72 °C for 1 min. A final extension of 72 °C for 7 min and a 4 °C hold was used. Cycling parameters for the secondary PCR were the same as above except that the annealing temperature was 55 °C.

Subgenotyping of *Cryptosporidium* isolates

Identification and subtyping of *Cryptosporidium* species was carried out by a nested PCR assay [24] to specifically amplify a fragment of the *GP60* gene. In the first round of PCR, 3 μl of DNA sample were amplified using the primer pair AL 3531 and AL 3535 (Table 1). Subsequently, 2 μl of the obtained PCR product was amplified in the secondary reaction using primers AL 3532 and AL 3534 (Table 1). Amplification reactions were performed in a volume of 50 μl with 1X Reaction buffer, 3 mM $MgCl_2$, 200 μM of each dNTP, 200 nM of each primer and 2.5 units of *Taq* DNA polymerase. Primary cycling conditions consisted of 5 min at 94 °C followed by 35 cycles of 45 s at 94 °C, 45 s at 59 °C and 1 min at 72 °C. A final extension of 72 °C for 10 min and a 4 °C hold was used. The secondary PCR was similar to that described for the primary PCR step with the exception that the annealing temperature was 50 °C.

All conventional *Giardia* and *Cryptosporidium* PCRs were carried out using BIOTAQ DNA polymerase (Bioline GmbH, Luckenwalde, Germany) on a 2720 thermal cycler (Applied Biosystems, California, USA). Appropriate positive and negative controls were routinely included in each round of PCR. PCR products were resolved on 2% D5 agarose gels (Conda, Madrid, Spain) stained with Pronasafe nucleic acid staining solution (Conda, Madrid, Spain). The amplicons of anticipated size (*Giardia GDH* and *BG* genes: ~432 bp and ~511 bp, respectively; *Cryptosporidium GP60* gene: ~890 bp) were sent for DNA sequencing in both directions using the same internal primer sets described in their respective PCR protocols.

Molecular detection of *Entamoeba histolytica* and *Entamoeba dispar*

Differential diagnosis between pathogenic *E. histolytica* and non-pathogenic *E. dispar* was achieved by RT-PCR using the primer pair and probe sequences described elsewhere [25] with some modifications [26]. Briefly, a RT-PCR was carried out to specifically amplify a 172-bp fragment of the *SSU* rRNA gene using the *E. histolytica*/*E. dispar*-specific primer pair and the TaqMan® probes described in Table 1. RT-PCR cycling conditions and settings were the same as for the detection of *G. duodenalis*. FAM and VIC fluorescence signals (510 nm for *E. histolytica* and 580 for *E. dispar*, respectively) were measured at the end of the annealing step of each cycle.

Data analysis

Raw sequencing data in both forward and reverse directions were viewed using the Chromas Lite version 2.1 sequence analysis program (http://chromas-lite.software.informer.com/2.1/). The presence of double peaks was also verified. The BLAST tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to compare nucleotide sequences with sequences deposited in the NCBI, GiardiaDB (http://giardiadb.org/giardiadb/) and CryptoDB (http://cryptodb.org/cryptodb/) databases. The resulting DNA consensus sequences were aligned to reference sequences using MEGA version 6.0 to determine *G. duodenalis* assemblages and subassemblages, and *Cryptosporidium* species and subtypes. Phylogenetic analyses, based on the neighbour-joining method, were performed using the same software [27].

Results

Infection rates of *G. duodenalis*, *Cryptosporidium* spp. and *Entamoeba* spp.

A total of 92 subjects attending the GRGH with gastrointestinal symptoms were initially screened for the presence

of *G. duodenalis*, *Cryptosporidium* spp. and *Entamoeba* spp. by direct microscopy and ICT and further confirmed by molecular methods. The male:female ratio was 1.09. The age range was from 0.5 to 80 years (mean, 15.7; SD, 16.8). Among the 92 stool samples examined, 10 (10.9%) tested positive for *G. duodenalis*, one (1.1%) for *Cryptosporidium* species and three (3.3%) for *Entamoeba* species. Overall, 14 (15.2%) of the samples analysed were positive for any of the three enteric protozoan species studied. A mixed *Giardia–Entamoeba* infection was recorded in a single stool sample. Five (50%) of the 10 patients diagnosed with giardiasis presented persistent diarrhoea.

Molecular characterisation of *Giardia duodenalis* isolates

Among 10 RT-PCR products that tested positive for *G. duodenalis*, seven were successfully amplified at the *GDH* locus and five at the *BG* locus, respectively. DNA sequencing of the *GDH* PCR products and subsequent sequence analysis revealed the presence of subassemblages BIII and BIV in three (42.9%) and four (57.1%) of the obtained isolates, respectively. No obvious mixed infections of *G. duodenalis* BIII + BIV could be demonstrated, although this possibility could not be completely ruled out since chromatograms displayed a number of double peaks. All five isolates at the *BG* locus were unmistakably characterised as assemblage B. The assemblages and subtypes of the *G. duodenalis* isolates identified in the present study according to gender and age of the infected patients are shown in Table 2.

Multiple sequence alignment analysis of the three isolates (C21, A27 and A38) assigned to *G. duodenalis* subassemblage BIII with the corresponding reference sequence (GenBank accession number AF069059) for the partial *GDH* gene allowed the identification of a 403-bp fragment, equivalent to positions 44–446 of AF069059, which was cleanly read from the raw sequence data for each sequence. Alignment analysis revealed that our isolates differ by three to five single nucleotide polymorphisms (SNPs) from reference sequence AF069059, some of them exhibiting double peaks at the chromatogram profile (Table 3). In addition, nucleotide sequences obtained differ also by at least three SNPs with previously published sequences. Nucleotide sequences of the partial *GDH* gene of new variants of *G. duodenalis* subassemblage BIII were submitted to GenBank under accession numbers KP026303 to KP026305.

Similarly, the four isolates (C1, C20, A13 and A66) identified as *G. duodenalis* subassemblage BIV were also

Table 2 Subassemblages of *Giardia duodenalis*, species/subtypes of *Cryptosporidium* spp. and species of *Entamoeba* spp. isolates characterised by PCR and sequencing in patients attending the Gambo Rural General Hospital, Ethiopia

Isolate	Gender	Age	*G. duodenalis*	*Cryptosporidium* spp.	*Entamoeba* spp.
C1	Female	4	Subassemblage BIV	–	–
C7	Male	53	–	–	*E. dispar*
C20	Female	2	Subassemblage BIV	–	–
C21	Female	2	Subassemblage BIII	–	–
A2	Male	<1	–	–	*E. dispar*
A13	Female	3	Subassemblage BIV	–	–
A16	Male	1	–	*C. hominis* IbA9G3	–
A27	Male	14	Subassemblage BIII	–	–
A38	Male	20	Subassemblage BIII	–	–
A66	Female	4	Subassemblage BIV	–	–
A67	Female	<1	–	–	*E. histolytica*

Table 3 Single nucleotide polymorphisms detected in *Giardia duodenalis* subassemblage BIII isolates at the *GDH* locus in this study

	Nucleotide at position of reference sequence AF069059										
Isolate/Sequence	60	87	99	147	204	219	237	309	372	375	405
AF069059	T	C	C	T	C	T	T	C	G	C	A
Isolate C21 (KP026303)	.	.	.	.	.	.	Y	T	A	Y	R
Isolate A27 (KP026304)	.	.	Y	C	.	C	.	.	.	.	.
Isolate A38 (KP026305)	Y	T	.	.	T	.	.	.	.	.	.

R: A/G; Y: C/T.

aligned with reference sequence L40508 for the partial *GDH* gene. A 405-bp fragment, corresponding to positions 78–482 of L40508, was clearly identified from all raw sequences. Alignment analysis showed that nucleotide sequences differ by one to six SNPs compared to reference sequence L40508 (Table 4). A single SNP in isolate A66 was identified as a double peak during chromatogram reading (Table 4). Furthermore, nucleotide sequences of isolates C1 and A13 were identical to those published in GenBank under accession numbers EF507682.1 and HQ616623.1 (among others), respectively. In contrast, nucleotide sequences of isolates C20 and A66 were distinct from any other published isolate at the *GDH* locus and thus represent new subtypes of *G. duodenalis* subassemblage BIV. Nucleotide sequences of the partial *GDH* gene of the two known and the two novel variants of *G. duodenalis* subassemblage BIV described in this study were deposited in the GenBank under accession numbers KP026306 to KP026309.

Figure 1 shows the phylogenetic tree obtained with the neighbour-joining analysis of *GDH* sequences from all the subtypes of assemblage B obtained in this study and the representative reference sequences taken from the NCBI database. Both BIII and BIV isolates formed closely related but independent clusters that were sister groups to assemblages C and D.

Raw sequence reads of the five isolates (C1, C20, A27, A38 and A66) obtained at the *BG* locus and characterised as *G. duodenalis* assemblage B were aligned with reference sequence AY072727.1. Clear reads were obtained for a 487-bp fragment equivalent to positions 103–589 of AY072727.1. Multiple sequence alignment analysis revealed that obtained nucleotide sequences differ only by one to two SNPs compared with reference sequence AY072727.1 (Table 5). A single SNP in isolate A27 was confirmed as a double peak at chromatogram examination (Table 5). Nucleotide sequences of isolates C20/A66, A27 and A38 exhibited 100% identity with those published in GenBank under accession numbers JX994238.1, JF918491.1 and JQ782392.1 (among others), respectively. Nucleotide sequences of the partial *BG* gene of the four known and single novel variants of *G. duodenalis* assemblage B reported here were deposited in the GenBank under accession numbers KP026310 to KP026314.

Figure 2 displays the phylogenetic tree constructed by the neighbour-joining method to compare the *BG* sequences characterised as assemblage B in this study with appropriate reference sequences obtained from the NCBI database. Our assemblage B isolates formed a well-supported cluster clearly differentiable from other *G. duodenalis* assemblages.

Table 4 Single nucleotide polymorphisms detected in *Giardia duodenalis* subassemblage BIV isolates at the *GDH* locus in this study

	Nucleotide at position of reference sequence L40508						
Isolate/Sequence	183	339	366	372	387	423	432
L40508	T	C	T	C	T	C	C
Isolate C1 (KP026306)	C	.	.	.	C	.	.
Isolate C20 (KP026307)	C	T	C	T	C	T	.
Isolate A13 (KP026308)	.	.	.	.	C	.	.
Isolate A66 (KP026309)	C	.	.	.	C	.	Y

Y: C/T.

Subgenotyping of *Cryptosporidium* isolates

Sequence analysis of the single *GP60* PCR product that tested positive for *Cryptosporidium* revealed the presence of *C. hominis* subtype IbA9G3. The gender and age of the patient infected with this *Cryptosporidium* isolate are shown in Table 2. Alignment analysis of the raw nucleotide sequence with reference sequence DQ665688 allowed the identification of a 825-bp fragment matching the full length of DQ665688. A new, unreported SNP (G to A) was observed at position 761. Nucleotide sequences of the partial *GP60* gene of this novel genotypic variant of *C. hominis* subtype IbA9G3 were submitted to GenBank under accession number KP026302.

Molecular differential diagnosis of *E. histolytica* and *E. dispar*

Among the three *Entamoeba* isolates found, two (66.7%) were assigned to *E. dispar* and the remaining to *E. histolytica* (33.3%). The *Entamoeba* species reported in this study according to gender and age of the infected patients are shown in Table 3.

Discussion

Besides bacterial and viral pathogens, the enteric protozoan *G. duodenalis*, *Cryptosporidium* spp. and *E. histolytica* remain a major cause of gastrointestinal illness in the developing world. Infections by *G. duodenalis* and *Cryptosporidium* spp. and to a lesser extent *E. histolytica*, have been significantly associated with persistent diarrhoea in children under 5 years of age

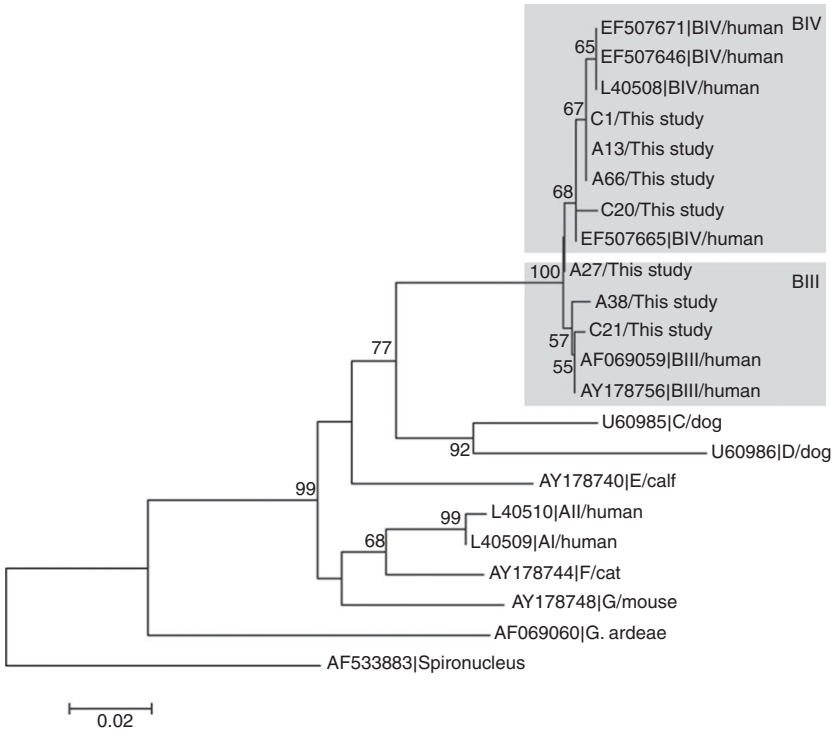

Figure 1 Evolutionary relationships among assemblages of *G. duodenalis* at the *GDH* locus inferred by a neighbour-joining analysis of the nucleotide sequence covering a 403-bp region (positions 80 to 482 of GenBank accession number L40510) of the gene. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 iterations) is indicated next to the branches. Bootstrap values lower than 50% were not displayed. The evolutionary distances were computed using the Kimura 2-parameter method. The rate variation among sites was modelled with a gamma distribution (shape parameter = 2). *Spironucleus vortens* was used as out-group taxa.

Table 5 Single nucleotide polymorphisms detected in *Giardia duodenalis* assemblage B isolates at the *BG* locus in this study

	Nucleotide at position of reference sequence AY072727.1		
Isolate/Sequence	183	309	573
AY072727.1	A	C	A
Isolate C1 (KP026310)	.	.	G
Isolate C20 (KP026311)	.	T	.
Isolate A27 (KP026312)	R	.	.
Isolate A38 (KP026313)	G	T	.
Isolate A66 (KP026314)	.	T	.

R: A/G.

[2, 28, 29] and, consequently, linked with impaired childhood development [4, 30]. Poverty, inadequate sanitation and hygiene practices, and restricted or unavailable access to adequate water supply and health facilities are the common denominator of these infections in endemic areas.

Compared to other sub-Saharan countries, Ethiopia has been recently demonstrated to bear a considerable burden of neglected tropical diseases, mainly attributable to leishmaniosis, trachoma, podoconiosis, ascariasis and leprosy [31]. Despite giardiasis, cryptosporidiosis and amoebiasis are also known to be widespread in the country [7, 8, 10–12, 16], including the Oromia Region [9, 14, 15], the public health and development impact of these enteric illnesses remains largely unknown. To gain insight and improve our understating on the epidemiological situation of *G. duodenalis*, *Cryptosporidium* spp. and *E. histolytica* infections in the later geographical area, in this study, we present preliminary prevalence and molecular data obtained from a relatively small population of individuals with gastrointestinal symptoms seeking medical attention at the GRGH. In a retrospective study targeting patients with diarrhoea examined at the same hospital setting, *G. duodenalis* (15.0%) and

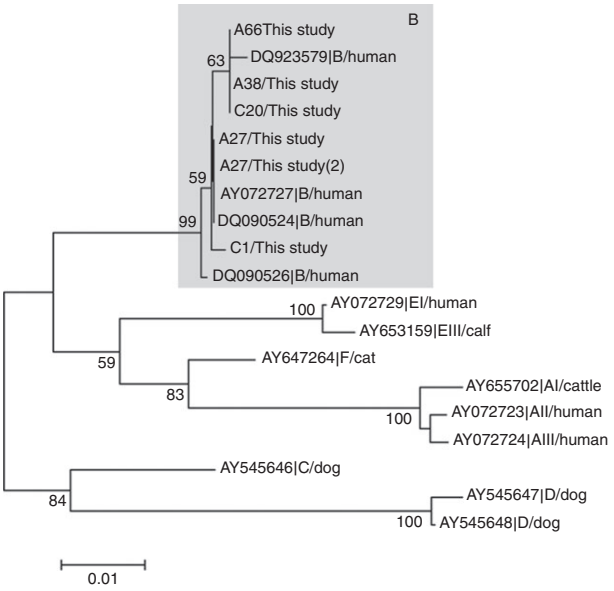

Figure 2 Evolutionary relationships among assemblages of *G. duodenalis* at the *BG* locus inferred by a neighbour-joining analysis of the nucleotide sequence covering a 487-bp region (positions 589 to 103 of GenBank accession number AY072727) of the gene. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 iterations) is indicated next to the branches. Bootstrap values lower than 50% were not displayed. The evolutionary distances were computed using the Kimura 2-parameter method. The rate variation among sites was modelled with a gamma distribution (shape parameter = 2). No out-group taxa were used as beta-giardin is a *Giardia*-specific cytoskeletal protein.

E. histolytica/dispar (5.4%) were the protozoan pathogens more frequently identified [9], although no genotyping analyses were conducted to assess the parasite species/subtypes circulating in the studied population. Our prevalence data are in agreement with those reported by Ramos *et al.* [9], with *G. duodenalis* and *E. histolytica/dispar* being found in 10.9% and 3.3% of the stool samples examined, respectively. *Cryptosporidium* spp. was also confirmed in an additional (1.1%) sample. Importantly, our prevalence figures may underestimate the true burden of these infections, as our diagnostic strategy was based on the combination of conventional microscopy and ICT as screening tests, whereas PCR-based methods were only used for confirmatory and typing purposes.

Molecular information regarding the diversity and frequency of *Giardia*, *Cryptosporidium* and *Entamoeba* species/genotypes currently circulating in Ethiopia is very limited. Among 59 human isolates of *G. duodenalis* obtained from both symptomatic and asymptomatic individuals from different regions of the country, 31 (53%) and 13 (22%) were characterised at the *BG* locus as assemblages A and B, respectively, whereas the remaining 15 (25%) isolates were typed as mixed infections A+F and A+B [20]. Interestingly, assemblage B was significantly more associated with symptomatic infection (nausea, abdominal pain and diarrhoea) than assemblage A. In our study, we use a multilocus genotyping approach based on the amplification of the *GDH* and *BG* loci to type *G. duodenalis* isolates both at the assemblage and at the subassemblage levels. Among the seven isolates that were successfully amplified at the *GDH* loci, multiple sequence analyses revealed three novel subtypes that were assigned to subassemblage BIII, and four (two novel and two known) subtypes characterised as subassemblage BIV. All five isolates amplified at the *BG* locus were allocated to assemblage B. To the best of our knowledge, these are the first human isolates of *G. duodenalis* typed at the subassemblage level in Ethiopia. These findings, together with the absence of assemblage A, are in agreement with those reported by Gelanew *et al.* [20], suggesting that assemblage B may be the predominant genotype in infected subjects with clinical manifestations in endemic areas.

Consistent with previous reports [32, 33], an elevated genetic polymorphism in assemblage B was observed at nucleotide level based on both *GDH* and *BG* genes. Meiotic recombination [34] or mixed subtype infections [35] have been proposed as potential mechanisms to explain this feature. In our study, most of the SNPs reported at both the *GDH* and the *BG* loci exhibited clear chromatogram readings in either forward and reverse directions, although a number of double peaks were also detected, particularly at the *GDH* locus. However, sequence alignment analyses do not seem to support the presence of mixed subtype infections. Regarding the two known subassemblages, BIV isolates based on the *GDH* gene found in our study, isolate C1 (GenBank accession number KP026306) showed 100% sequence identity to those reported in humans from Brazil (EF507682.1), Poland (DQ840541.1) and Sweden (HM136889.1). Isolate A13 (KP026308) was identical to those described in a number of primate species including brown howling monkeys (*Alouatta fusca*, e.g. HM134210.1) and saki monkeys (*Pithecia pithecia*, AB569387.1), Thomson's gazelles (KF443202.1) and humans from Japan (AB569386.1). Similarly, the four already known assemblage B isolates at the *BG* gene described in our survey (C20, A27, A38 and A66) have been previously reported in humans from Asian and European countries. Thus, isolates C20 and A66 (KP026311 and KP026314, respectively) exhibited 100% sequence identity with those reported in China

(JX994238.1), India (e.g. JF918489.1), Thailand (e.g. FJ971483.1) and Norway (DQ090523.1). In addition to these countries, isolate A27 (KP026312) has been also found infecting humans in New Zealand (EU274391.1). Of particular interest is isolate A38 (KP026313), whose nucleotide sequence contains two specific SNPs in position 183 (A to G) and position 309 (C to T) compared with reference sequence AY072727.1. This isolate was identical to that (JQ782392.1) identified in a Spanish traveller returning from a trip to India that was refractory to the prescribed treatment with tinidazole [36]. We speculate that the above-mentioned specific nucleotide substitutions in the *BG* sequence may be associated with refractory giardiasis and persistent diarrhoea. Therefore, it would be interesting to corroborate this hypothesis by genotyping *G. duodenalis* assemblage B isolates at the *BG* loci in a larger number of patients not responding to the conventional nitroimidazole therapy.

Typing of human isolates of *Cryptosporidium* spp. in Ethiopia has been conducted in a single molecular epidemiological study [7]. In this survey, zoonotic *C. parvum* was identified as the most prevalent (97%, 39/40) *Cryptosporidium* species circulating in the country, with anthroponotic *C. hominis* apparently playing a secondary role in the transmission of this enteric pathogen. In our study, we have confirmed a novel *C. hominis* IbA9G3 variant (KP026302) not previously documented.

The genus *Entamoeba* comprises six species, namely *E. histolytica*, *E. dispar*, *E. moshkovskii*, *E. coli*, *E. hartmanni* and *E. polecki*, of which only *E. histolytica* is pathogenic to humans. It is estimated that, depending on the geographical area, up to 90% of the amoebiasis cases documented worldwide are caused by *E. dispar*, with *E. histolytica* being responsible for a much smaller fraction of the infections [28, 37]. Therefore, obtaining accurate species prevalence data should be a priority, particularly in endemic regions. Our study is the first attempt to conduct differential molecular diagnosis of *E. histolytica* from morphologically identical but non-pathogenic *E. dispar* in Ethiopia. We have identified the presence of *E. histolytica* and *E. dispar* in one (33.3%) and two (66.7%) of the three cases of amoebiasis detected. The elevated prevalence of *E. histolytica* found may be explained by the fact that only patients with clinical manifestations were analysed.

In conclusion, we have shown in this preliminary molecular epidemiological survey that *G. duodenalis*, *C. hominis* and *E. histolytica/dispar* are frequently found in individuals with gastrointestinal symptoms from a rural endemic area in southern Ethiopia. Although our prevalence data must be interpreted with caution due to the limited number of samples tested, molecular and sequence analyses have allowed us to identify new *G. duodenalis* subtypes and *C. hominis* subgenotypes, confirming the elevated variability at the nucleotide level of these enteric pathogens. We believe these data will contribute to improve our current knowledge on the epidemiology, species diversity, genotype frequencies and transmission dynamics of *Giardia*, *Cryptosporidium* and *Entamoeba* in endemic areas of developing countries. More research should be conducted to confirm the epidemiological significance of the findings presented here.

Acknowledgements

We would like to thank all the staff at Gambo hospital for their collaboration, especially to Francisco Reyes. We also express our gratitude for his support to Miguel Górgolas, Director of the Master in Tropical Medicine and International Health of the UAM in Madrid. This project was partially funded by his Master and Research Projects.

References

1. WHO *The World Health Report – 1996, Fighting Disease Fostering Development*. World Health Organization: Geneva, Switzerland, 1996.
2. Mosier DA, Oberst RD. Cryptosporidiosis: a global challenge. *Ann N Y Acad Sci* 2000: **916**: 102–111.
3. Petri WA, Haque R, Lyerly D, Vines RR. Estimating the impact of amebiasis on health. *Parasitol Today* 2000: **16**: 320–321.
4. Berkman DS, Lescano AG, Gilman RH, Lopez SL, Black MM. Effects of stunting, diarrhoeal disease, and parasitic infection during infancy on cognition in late childhood: a follow-up study. *Lancet* 2002: **359**: 564–571.
5. Ijaz MK, Rubino JR. Impact of infectious diseases on cognitive development in childhood and beyond: potential mitigational role of hygiene. *Open Infect Dis J* 2012: **6**: 65–70.
6. Savioli L, Smith H, Thompson A. *Giardia* and *Cryptosporidium* join the 'Neglected Diseases Initiative'. *Trends Parasitol* 2006: **22**: 203–208.
7. Adamu H, Petros B, Hailu A, Petry F. Molecular characterization of *Cryptosporidium* isolates from humans in Ethiopia. *Acta Trop* 2010: **115**: 77–83.
8. Huruy K, Kassu A, Mulu A *et al.* Intestinal parasitosis and shigellosis among diarrheal patients in Gondar teaching hospital, northwest Ethiopia. *BMC Res Notes* 2011: **4**: 472.
9. Ramos JM, Rodríguez-Valero N, Tisiano G *et al.* Different profile of intestinal protozoa and helminthic infections among patients with diarrhoea according to age attending a rural hospital in southern Ethiopia. *Trop Biomed* 2014: **31**: 392–397.
10. Wegayehu T, Tsalla T, Seifu B, Teklu T. Prevalence of intestinal parasitic infections among highland and lowland

dwellers in Gamo area, South Ethiopia. *BMC Public Health* 2013b: **13**: 151.
11. Abossie A, Seid M. Assessment of the prevalence of intestinal parasitosis and associated risk factors among primary school children in Chencha town, Southern Ethiopia. *BMC Public Health* 2014: **14**: 166.
12. Ayalew D, Boelee E, Endeshaw T, Petros B. *Cryptosporidium* and *Giardia* infection and drinking water sources among children in Lege Dini, Ethiopia. *Trop Med Int Health* 2008: **13**: 472–475.
13. King JD, Endeshaw T, Escher E *et al.* Intestinal parasite prevalence in an area of Ethiopia after implementing the SAFE strategy, enhanced outreach services, and health extension program. *PLoS Negl Trop Dis* 2013: **7**: e2223.
14. Wegayehu T, Adamu H, Petros B. Prevalence of *Giardia duodenalis* and *Cryptosporidium* species infections among children and cattle in North Shewa Zone, Ethiopia. *BMC Infect Dis* 2013a: **13**: 419.
15. Adamu H, Wegayehu T, Petros B. High prevalence of diarrhoegenic intestinal parasite infections among non-ART HIV patients in Fitche Hospital, Ethiopia. *PLoS ONE* 2013: **8**: e72634.
16. Teklemariam Z, Abate D, Mitiku H, Dessie Y. Prevalence of intestinal parasitic infection among HIV positive persons who are naive and on antiretroviral treatment in Hiwot Fana Specialized University Hospital, Eastern Ethiopia. *ISRN AIDS* 2013: **2013**: 324329.
17. Mamo H. Intestinal parasitic infections among prison inmates and tobacco farm workers in Shewa Robit, north-central Ethiopia. *PLoS ONE* 2014: **9**: e99559.
18. Feng Y, Xiao L. Zoonotic potential and molecular epidemiology of *Giardia* species and giardiasis. *Clin Microbiol Rev* 2011: **24**: 110–140.
19. Ryan U, Fayer R, Xiao L. *Cryptosporidium* species in humans and animals: current understanding and research needs. *Parasitology* 2014: **141**: 1667–1685.
20. Gelanew T, Lalle M, Hailu A, Pozio E, Cacciò SM. Molecular characterization of human isolates of *Giardia duodenalis* from Ethiopia. *Acta Trop* 2007: **102**: 92–99.
21. Verweij JJ, Schinkel J, Laeijendecker D, van Rooyen MA, van Lieshout L, Polderman AM. Real-time PCR for the detection of *Giardia lamblia*. *Mol Cell Probes* 2003b: **17**: 223–225.
22. Read CM, Monis PT, Thompson RC. Discrimination of all genotypes of *Giardia duodenalis* at the glutamate dehydrogenase locus using PCR-RFLP. *Infect Genet Evol* 2004: **4**: 125–130.
23. Lalle M, Pozio E, Capelli G, Bruschi F, Crotti D, Cacciò SM. Genetic heterogeneity at the beta-giardin locus among human and animal isolates of *Giardia duodenalis* and identification of potentially zoonotic subgenotypes. *Int J Parasitol* 2005: **35**: 207–213.
24. Feltus DC, Giddings CW, Schneck BL, Monson T, Warshauer D, McEvoy JM. Evidence supporting zoonotic transmission of *Cryptosporidium* spp. in Wisconsin. *J Clin Microbiol* 2006: **44**: 4303–4308.
25. Verweij JJ, Oostvogel F, Brienen EA, Nang-Beifubah A, Ziem J, Polderman AM. Prevalence of *Entamoeba histolytica* and *Entamoeba dispar* in northern Ghana. *Trop Med Int Health* 2003a: **8**: 1153–1156.
26. Gutiérrez-Cisneros MJ, Cogollos R, López-Vélez R *et al.* Application of real-time PCR for the differentiation of *Entamoeba histolytica* and *E. dispar* in cyst-positive faecal samples from 130 immigrants living in Spain. *Ann Trop Med Parasitol* 2010: **104**: 145–149.
27. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 2013: **30**: 2725–2729.
28. Haque R, Huston CD, Hughes M, Houpt E, Petri WA. Amebiasis. *N Engl J Med* 2003: **348**: 1565–1573.
29. Muhsen K, Levine MM. A systematic review and meta-analysis of the association between *Giardia lamblia* and endemic pediatric diarrhea in developing countries. *Clin Infect Dis* 2012: **55**: S271–S293.
30. Tarleton JL, Haque R, Mondal D, Shu J, Farr BM, Petri WA. Cognitive effects of diarrhea, malnutrition, and *Entamoeba histolytica* infection on school age children in Dhaka, Bangladesh. *Am J Trop Med Hyg* 2006: **74**: 475–481.
31. Deribe K, Meribo K, Gebre T *et al.* The burden of neglected tropical diseases in Ethiopia, and opportunities for integrated control and elimination. *Parasit Vectors* 2012: **5**: 240.
32. 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**: 1523–1531.
33. Lebbad M, Petersson I, Karlsson L *et al.* Multilocus genotyping of human *Giardia* isolates suggests limited zoonotic transmission and association between assemblage B and flatulence in children. *PLoS Negl Trop Dis* 2011: **5**: e1262.
34. Cooper MA, Adam RD, Worobey M, Sterling CR. Population genetics provides evidence for recombination in *Giardia*. *Curr Biol* 2007: **17**: 1984–1988.
35. Hussein AI, Yamaguchi T, Nakamoto K, Iseki M, Tokoro M. Multiple-subgenotype infections of *Giardia intestinalis* detected in Palestinian clinical cases using a subcloning approach. *Parasitol Int* 2009: **58**: 258–262.
36. Requena-Méndez A, Goñi P, Lóbez S *et al.* A family cluster of giardiasis with variable treatment responses: refractory giardiasis in a family after a trip to India. *Clin Microbiol Infect* 2014: **20**: O135–O138.
37. Ali IK, Clark CG, Petri WA. Molecular epidemiology of amebiasis. *Infect Genet Evol* 2008: **8**: 698–707.

Corresponding Author David Carmena, Servicio de Parasitología, Centro Nacional de Microbiología, Instituto de Salud Carlos III, Ctra. Majadahonda-Pozuelo Km 2, 28220 Majadahonda, Madrid, Spain. Tel.: +34 918223775; Fax: +34 915097919; E-mail: dacarmena@isciii.es