Infection, Genetics and Evolution 54 (2017) 128–137

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

New multilocus genotypes of *Giardia lamblia* human isolates

Clarissa Perez Faria [a,b,c], Graziela Maria Zanini [c], Gisele Silva Dias [c], Sidnei da Silva [c], Maria do Céu Sousa [a,b,*]

[a] CNC - Center for Neurosciences and Cell Biology, University of Coimbra, 3030-548 Coimbra, Portugal
[b] Faculty of Pharmacy, University of Coimbra, 3030-548 Coimbra, Portugal
[c] Laboratory of Parasitology, Evandro Chagas National Institute of Infectious Diseases, Oswaldo Cruz Foundation, Rio de Janeiro, 21040-900, Brazil

ARTICLE INFO

Article history:
Received 14 November 2016
Received in revised form 28 June 2017
Accepted 28 June 2017
Available online 29 June 2017

Keywords:
Giardia lamblia
Assemblage
Multilocus genotyping
DNA sequencing
β-giardin
Glutamate dehydrogenase
Triose phosphate isomerase

ABSTRACT

Giardia lamblia is considered a species complex, whose members show little differences in their morphology, but have remarkable genetic variability. The aim of this study was to identify inter- and intra-assemblage genetic variation in *G. lamblia* among patients in Rio de Janeiro. The parasitological study was performed on faeces, and DNA was extracted from the samples which tested positive for *G. lamblia*. The genetic assemblages and subtypes were determined *via* multilocus sequence typing (MLST) using β-giardin, triose phosphate isomerase and glutamate dehydrogenase gene loci. Fourteen assemblage A samples were successfully genotyped at the three MLST loci (*bg*/*tpi*/*gdh*). Two previously identified multilocus genotypes were found (AII-1 and AII-4), and two novel multilocus genotypes are proposed (AII-8, profile A2/A2/A4; AII-9, profile A3/A2/A2). Sequence analysis showed that assemblage B isolates have a higher nucleotide variation than those from assemblage A. Novel assemblage B sequences are described and most (66.7%) have heterogeneous nucleotides, which prevent the definition of multilocus genotypes. This is the first time that MLST has been used to characterise *G. lamblia* isolates in human clinical samples from Rio de Janeiro. In addition, MLST has enabled the detection of novel subtypes in both assemblages and the description of two novel multilocus genotypes in assemblage A. This study provides new insights into the genetic diversity of assemblage A and shows that MLST should be used to characterise *G. lamblia* both in Brazil and globally.

1. Introduction

Giardia lamblia (syn. *Giardia intestinalis*, *Giardia duodenalis*) is an intestinal protozoa found in a wide range of hosts including humans, and domestic and wild animals. This parasite is considered a species complex with at least eight distinct assemblages (labeled A to H), but only assemblages A and B have been detected in both humans and a wide range of other mammalian hosts. The other assemblages are likely to be host-specific: C and D are mainly found in dogs and other canids, E in hoofed livestock, F in cats, G in rodents and H in marine mammals (Monis et al., 1999, 2003; Thompson, 2004; Lasek-Nesselquist et al., 2010). Molecular genetic analyses have shown that *G. lamblia* assemblage A is divided into three sub-assemblages: AI, mainly zoonotic; AII, commonly anthroponotic, although it has been reported in a few studies on animals, and AIII, found only in animals. Assemblage B, which is divided into two sub-assemblages, BIII and BIV, is predominantly found in humans and much less commonly in animals (Sprong et al., 2009; Ryan and Cacciò, 2013).

PCR-based molecular analysis techniques have been used to investigate the genetic diversity of *G. lamblia*. Initially, most of the molecular epidemiological studies were based on the analysis of a single locus, but parasite assemblages are currently distinguished by multilocus genotyping (MLG) tools. Small subunit ribosomal RNA (*ssu rRNA*) (Hopkins et al., 1997; Appelbee et al., 2003), β-giardin (*bg*) (Cacciò et al., 2002; Lalle et al., 2005), glutamate dehydrogenase (*gdh*) (Read et al., 2004; Cacciò et al., 2008) and triose phosphate isomerase (*tpi*) (Sulaiman et al., 2003; Bertrand et al., 2005) genes are the four genetic markers commonly used in genotyping and subtyping *G. lamblia* in many host species and geographical locations. As the amount of sequence information increased, incongruent genotyping results (lack of concordance in the assignment of isolates to a specific assemblage) began to be reported. Thus, the analysis of a single locus was not considered sufficient to reliably determine the distribution of assemblages across the population (Cacciò et al., 2008; Sprong et al., 2009; Almeida et al., 2010; Huey et al., 2013). Therefore, the use of MLG tools became imperative.

Based on MLG analysis of the *bg*, *gdh*, and *tpi* genes, Cacciò et al. (2008) proposed a subtype nomenclature system with 10 different MLGs for assemblage A isolates (AI-1, AI-2, AII-1, AII-2, AII-3, AII-4, AII-5, AII-6, AII-7, AIII-1). The same was not possible for assemblage B because of the greater diversity encountered between and within

* Corresponding author at: CNC - Center for Neurosciences and Cell Biology, University of Coimbra, 3030-548 Coimbra, Portugal.
E-mail address: mcsousa@ci.uc.pt (M.C. Sousa).

http://dx.doi.org/10.1016/j.meegid.2017.06.028
1567-1348/

these isolates. A precise nomenclature for the MLGs of assemblage B isolates was impossible to be proposed due the presence of heterogeneous templates (overlapping nucleotide peaks at specific positions in the sequence).

The distribution of assemblages A and B varies greatly in Brazil, largely because of the enormous expanse of territory involved. While earlier reports demonstrated the prevalence of assemblage A (Souza et al., 2007; Volotão et al., 2007), recent works have found a prevalence of assemblage B infection (Kohli et al., 2008; Santos et al., 2012; Colli et al., 2015; Oliveira-Arbex et al., 2015), and others have found no significant difference between the distribution of the assemblages (Durigan et al., 2014; David et al., 2015; Nunes et al., 2016; Faria et al., 2016).

The aim of this study was to identify inter- and intra-assemblage genetic variation in *G. lamblia* based on multilocus sequence typing (MLST) of the *bg*, *tpi* and *gdh* genes and to define MLGs for assemblages A and B in patients with giardiasis attending a referral hospital in Rio de Janeiro, Brazil.

2. Materials and methods

2.1. Study population, collection of faecal samples and laboratory methods

A total of 65 faecal samples positive for *G. lamblia* were collected between January 2011 and February 2015 from patients attended in Evandro Chagas National Institute of Infectious Diseases (INI/FIOCRUZ), a referral hospital in infectious diseases in Rio de Janeiro, Brazil.

Stool samples were collected by the patient in plastic disposable flasks without preservatives and maintained at 4 °C until laboratory analysis on the same day. The flasks were labeled with the name, collection date and the hospital number. The parasitological tests were conducted at the Parasitology Laboratory of INI, a laboratory certified by the College of American Pathologists, by experienced technologists. For diagnosis of *G. lamblia*, the fresh specimens were analysed by means of centrifugation sedimentation (de Carli, 2001) and centrifugal flotation in zinc sulphate solution (Faust et al., 1938). The slides were then observed under the microscope (Nikon Eclipse E200, magnification of 10 and 40×).

In addition, clinical and epidemiological data (age and gender) were collected from the hospital's database. The patients were classified as symptomatic if presented abdominal pain, asthenia, cough, diarrhea, fever, headache, myalgia, rapid weight loss, vertigo and/or vomiting. According to World Health Organization (WHO) criteria diarrhea is defined as the passage of three or more unformed stools per day or more frequent passage than is normal for the individual (WHO, 2013).

2.2. Ethics statement

The Research Ethics Committee Evandro Chagas National Institute of Infectious Diseases (INI/FIOCRUZ) approved the study (protocol number: 127.542). This project was in accordance with the Brazilian Ethical Resolutions, especially Resolution CNS 196/1996 and its complementary and the Code of Medical Ethics of 1988 (articles 122–1307). Written informed consent was obtained from all patients or legal guardians of patients younger than 18 years, prior to sample collection. The informed consent was provided after a detailed explanation of the objectives of the work. A term of privacy and confidentiality was signed by the researches for patients for whom it was not possible to obtain informed consent beforehand.

2.3. DNA extraction

The molecular analysis of *G. lamblia* was performed only on faecal samples without preservatives. Approximately 5 g of faecal sample was washed with distilled water, filtered through doubled gauze, and then centrifuged (100×*g* for 2 min). These steps were repeated two more times. The concentrated cysts were stored at −20 °C until DNA extraction. Samples collected in 2011 and 2012 were subjected to DNA extraction in 2013, whereas samples collected from 2013 were extracted regularly within one month of collection. DNA extraction was performed using the QIAamp DNA Stool mini Kit (Qiagen) according to the manufacturer's instructions. A new DNA extraction was performed in PCR-negative samples with minor modifications. In the first step, glass pearls and polyvinylpyrrolidone 10% solution was added and the incubation time was increased to one hour at 95 °C, and in the finals steps glycogen was added for DNA precipitation. Genomic DNA was preserved at −20 °C until analysis.

2.4. PCR analysis

Extracted DNA was subsequently analysed by PCR using three *G. lamblia* gene loci: β-giardin (*bg*), triose phosphate isomerase (*tpi*) and glutamate dehydrogenase (*gdh*). The amplification of the *bg* locus gene was performed using a semi-nested PCR (Cacciò et al., 2002) and a nested PCR was done for amplification of fragments of the *tpi* gene (Sulaiman et al., 2003). For the amplification of the *gdh* gene, a semi-nested PCR and a nested PCR were performed as described by Read et al. (2004) and Cacciò et al. (2008), respectively.

All reactions contained 12.5 μL of NZYTaq 2× Green Master Mix (Nzytech), 1 μL of each primer (10 pmol/μL), 1 μL of extracted DNA and 8.5 μL of sterile water, performing a final volume of 25 μL. PCR was carried out on the MJ Mini™ Thermal Cycler (BioRad). In all PCR reactions, *Giardia*-positive DNA sample (strain WB, clone 6 [ATCC 30957]) and nuclease-free distilled water were used as positive and negative controls, respectively. The PCR products were analysed on 1.5% agarose gels stained with ethidium bromide and visualized using a gel documentation system (Uvisave, Uvitec).

2.5. Multilocus sequencing and phylogenetic analysis

Positive secondary PCR products were purified using QIAquick PCR Purification kit (Qiagen), quantified (ND1000, NanoDrop) and sequenced with secondary PCR primers in both directions using the respective forward and reverse primers with an Applied Biosystems 3730 xL DNA Analyser (Applied Biosystems). Chromatograms and nucleotide sequences were analysed, edited using BioEdit Sequence Alignment Editor Programme (http://www.mbio.ncsu.edu/bioedit/bioedit.html), and were aligned with reference sequences retrieved from the GenBank using Clustal W. In order to determine the *Giardia* assemblage, sub-assemblage and subtype, BLAST software (http://www.ncbi.nlm.nih.gov/blast/) was used for comparison of the nucleotide sequences obtained in the present work with sequences available at GenBank. The phylogenetic analysis was performed in MEGA software v.6 (www.megasoftware.net) using neighbour-joining (NJ) and maximum likelihood (ML) algorithms with evolutionary distances calculated by Tamura-Nei method. Bootstrap analysis was applied to evaluate the reliability of clusters by using 1000 replicates. The sequences from the three loci (*bg*, *tpi* and *gdh*) were also concatenated. The accession numbers and the corresponding assemblages of the sequences obtained from the GenBank database are displayed in S1 Table. The novel sequences without ambiguous positions from this study were deposited in the GenBank under accession numbers KX085486-KX085495.

3. Results

Of the 65 samples positive for *G. lamblia*, 39 were successfully amplified and sequenced (Table 1). Sequence results showed that 18 samples belonged to assemblage A, 21 samples belonged to assemblage B and no mixed infections (A + B) were detected. Thirty-one (79.5%) samples were amplified at the *bg* locus, 32 (82.1%) at *gdh* locus and 37 (94.9%) at *tpi* locus. Twenty-nine samples were amplified in all loci, three were amplified in two loci and seven were only amplified at one locus (one was amplified at *bg* locus, one at *gdh*, and five at *tpi*). Among the

Table 1
Genotyping results of the 39 positive stool samples for *G. lamblia* based on three different loci using sequence analysis.

Assemblage	1 locus			2 loci		3 loci	Total
	bg	*tpi*	*gdh*	*bg* + *tpi*	*tpi* + *gdh*	*bg* + *tpi* + *gdh*	
A	1	3	–	–	–	14	18
B	–	2	1	1	2	15	21
Total	1	5	1	1	2	29	39

32 samples amplified in more than one locus, no incongruent assignment to assemblages ("assemblage swapping") was detected.

The majority of the patients were adults (69.2%) with an average of 32.54 ± 13.69 years (Mean ± SD; median = 32). The highest prevalence occurred between 30 and 39 years (33.3%) and there were more male than female patients (69.2% *versus* 30.8%). Of the 39 samples that were successfully amplified and sequenced, 14 (35.9%) showed clinical symptoms while 25 (64.1%) were asymptomatic (Table 2). When we studied the association of the symptoms with assemblages, no statistical association was observed.

3.1. Assemblage A sequences

All 18 samples classified as assemblage A corresponded to sub-assemblage AII. Among the 15 samples amplified at the *bg* locus, nine showed complete sequence identity with the previously described subtype A2 (KC8, accession number AY072723) and five were identical to subtype A3 (ISSGF7, accession number AY072724) (Table 3). One sample (INI 17) showed overlapping nucleotides (C - T/Y) at the positions 421 and 429, which could classify the sample as subtype A2 and/or A3 (Table 3).

Of the 17 samples successfully amplified at the *tpi* locus, all sequences showed 100% sequence identity with subtype A2 (JH, accession number U57897) except the sample INI 44 where a single nucleotide polymorphism (SNP) were observed (G - A at position 349) (Table 3). This sample showed complete identity with two previously described sequences, KF922901 and KF922892 from São Paulo/Brazil (Durigan et al., 2014) (Table 3) and is distinct from the six subtypes (A1-A6) described by Cacciò et al. (2008) and thus represents a new subtype. No overlapping nucleotides were observed at *tpi* sequences.

Fourteen samples were amplified at *gdh* locus, seven being identical to subtype A2 (Bris136, accession number AY178737) and the others were identical to subtype A4 (ECUST 2196, accession number JX994237) (Table 3). No sequences with novel substitutions or overlapping nucleotides, at any position, were observed at *gdh* sequences.

3.2. Assemblage B sequences

At the *bg* locus, 12 (75%) of 16 assemblage B sequenced samples exhibited no ambiguous nucleotides at any position in the chromatogram and six subtypes could be identified, one had already been described and five were novel (Table 4). Three samples (INI 54, 61 and 64) showed complete identity with a previously described sequence (KF736104, sub-assemblage BIII) (Table 4). The sequences of four samples (INI 27, 48, 50 and 67) were classified as new variants of BIII reference sequence (Ld18, AY072726) and the remaining five samples (INI 21, 49, 53, 57 and 68) were classified as new variants of BIV reference sequence (ISSGF4, AY072728). These nine samples showed novel substitutions at positions 393, 396, 456, 477, 525, 570, 606, 615, 714 and 717, along *bg* sequence and prevented the assigned to a specific subtype. All of them showed 99% identity with previous published sequences.

At the *tpi* locus, 12 (60%) of 20 samples displayed unambiguous nucleotides (Table 4). Six different subtypes were identified: nine samples corresponded to three previously reported subtypes and three samples were novel subtypes. The sequence of five samples (INI 33, 43, 49, 53 and 56) showed 100% identity with the sub-assemblage BIII reference sequence (Bah12, AF069561) (Table 4). One sample (INI 50) showed complete identity with the published sequence AY368165 (BIII/isolate 2434) and three samples (INI 48, 65 and 69) had sequences identical to BIV isolate Ad-19 (AF069560). The sequences of INI 21 and INI 55 showed 99% identity with the BIV isolate GS/M (L02116) and Ad19, respectively. The sequence of sample INI 64 corresponded to a mixture of sub-assemblages BIII and BIV, and seven SNPs were observed along the sequence.

Among the 18 sequences obtained at the *gdh* locus, only five (27.8%) exhibited no double peaks and four different subtypes were identified (Table 4). One sample (INI 26) was identical to BIII isolate Bah 12c14 (EF685684) and other sample (INI 33) showed complete sequence identity with BIII isolate FCQ 21 (AY178756) (Table 4). INI 48 sequence was identical to sub-assemblage BIV reference sequence (Ad28, AY178738) or to JX972186 (which was found in a horse in Colombia). Two samples (INI 27 and 65) showed 99% identity with the BIV reference sequence (Ad28, AY178738), with only one SNP (G - A at position 354). Moreover, the two latter isolates showed 100% identity with two previously described sequences (KT334248 and KT334251), however with just 52% of coverage.

3.3. Multilocus analysis

A total of 14 assemblage A samples were successfully genotyped at the three MLST loci (*bg*/*tpi*/*gdh*) (Table 5). Two previously identified MLGs were found: AII-1, profile A2/A2/A2, in four samples (INI 47, 51,

Table 2
Association between the *G. lamblia* assemblages and symptomatology.

Assemblage	Symptoms	
	Yes	No
A	5	13
B	9	12
Total	14	25

Table 3
Polymorphisms at the β-giardin, triose phosphate isomerase and glutamate dehydrogenase loci of assemblage A samples compared to reference sequences obtained from GenBank.

Sub-assemblage	Subtype	Isolate (ID)	Nucleotide position			
β-giardin			421	429	567	690
AI	A1	**WB**	C	T	C	A
AII	A2	**KC8**, INI 19, 44, 47, 51, 52, 58, 59, 60, 66 55,858,59,60,66			T	G
AII	A3	**ISSGF7**, INI 11, 23, 39, 41, 45	T	C	T	G
AII	A2 and/or A3	INI 17	Y	Y	T	G
Triose phosphate isomerase			129	349	399	
AI	A1	**WB**	T	G	C	
AII	A2	**JH**, INI 7, 11, 17, 19, 23, 25, 38, 41, 45, 47, 51, 52, 58, 59, 60, 66	C		T	
AII	Novel A2	INI 44[a]	C	A	T	
Glutamate dehydrogenase			699	753	807	831
AI	A1	**Portland1**	T	C	C	C
AII	A2	**Bris136**, INI 11, 41, 45, 47, 51, 58, 60	C	T	T	T
AII	A4	**ECUST 2196**, INI 19, 17, 23, 44, 52, 59, 66		T	T	T

Nucleotide substitutions are numbered from the ATG codon of each gene, empty spaces indicate identity to the A1 reference sequence (GenBank accession numbers: AY258617 for *bg*, L02120 for *tpi* and M84604 for *gdh*). Isolates in bold are reference sequences from the GenBank.

[a] Sequence previously observed by Durigan et al., 2014 (GenBank acc. No. KF922901 and KF922892).

Table 4
Polymorphisms at the β-giardin, triose phosphate isomerase and glutamate dehydrogenase loci of assemblage B samples compared to reference sequences obtained from GenBank.

Isolate (ID)	*bg*																
	393	396		456		477	525		570		606	615		714		717	
BIII (Ld18)	A	C		C		T	T		C		C	C		C		C	
BIV (ISSGF4)							C		T		T	T		–		–	
BIII (KF736104), INI 54, 61, 64	G	T															
INI 21, 68	G	T					C		T		T	T		T		T	
INI 27, 48						C											
INI 50	G	T												T		T	
INI 49, 53, 57	G	T					C		T		T			T		T	
INI 67	G	T		T													
	tpi																
	39	45	91	165	168	210	217	271	280	297	304	393	402	429	438	483	504
BIII (Bah12), INI 33, 43, 49, 53, 56	–	T	C	C	C	G	G	C	A	A	A	C	A	G	T	A	C
BIV (Ad19), INI 48, 65, 69	A		T	T	T	A								A			
BIV (GS/M)	A		T	T	T	A											
INI 21	A		T	T	T												T
BIII (isolate 2434), INI 50	G										G	T					
INI 55	A		T	T	T	A	A							A			
INI 64	G	C	T	T	T			T	G	G			G		C	G	
	gdh																
	309	354	429	447	519		540	561	612	699	723		807	825	867	921	
BIII (Bah12c14), INI 26	–	–	T	T	C		C	C	G	T	C		G	A	C	G	
BIII (FCQ21), INI 33	Y	G			Y					C					Y		
BIV (Ad28), INI 48	T	G	C	C	C		T	T	A	C	T		T	G		A	
INI 27, 65	T	A	C	C	C		T	T	A	C	T		T	G		A	

Nucleotide substitutions are numbered from the ATG codon of each gene, empty spaces indicate identity with the BIII reference sequence (GenBank accession numbers: AY072726 [Ld18] for *bg*, AF069561 [Bah12] for *tpi* and EF685684 [Bah12c14] for *gdh*). Isolates in bold are reference sequences from the GenBank. The dash (−) indicates deletion.

58, 60); AII-4, profile A3/A2/A4, in one sample (INI 23). Two novel MLGs were identified: AII-8 (proposed), profile A2/A2/A4, in four samples (INI 19, 52, 59, 66); AII-9 (proposed), profile A3/A2/A2, in three samples (INI 11, 41, 45). One sample (INI 17), with overlapping nucleotides, could be identified as AII-4 and/or on the new proposed MLG, AII-8. One novel assemblage A MLG was observed (INI 44), and could not be classified according to Cacciò et al. (2008) nomenclature proposal.

A total of 15 assemblage B samples were sequenced at all three genes. Due the high number of heterogeneous nucleotides at *gdh* gene (13 isolates), only one isolate (INI 48) had no double peak at any position and were classified as a novel assemblage B MLG (Table 5).

The classification of the 21 assemblage B isolates into sub-assemblages is shown in Table 6. Mixed infection within the sub-assemblages BIII + BIV was identified in one sample (INI 64) at *tpi* locus. Incongruent genotyping results were observed in an intra-assemblage level, five isolates (INI 21, 27, 48, 64 and 69) were classified BIII at *bg* locus and as sub-assemblage BIV at the others loci.

Table 5
Multilocus genotyping results of assemblages A and B from 15 samples based on sequencing data from the β-giardin, triose phosphate isomerase and glutamate dehydrogenase loci.

Assemblage	Isolate (ID)	Subtype			MLG
		bg	*tpi*	*gdh*	
A	INI 47, 51, 58, 60	A2	A2	A2	AII-1
	INI 23	A3	A2	A4	AII-4
	INI 19, 52, 59, 66	A2	A2	A4	AII-8 (proposed)
	INI 17	A2 or A3	A2	A4	AII-4 or AII-8
	INI 11, 41, 45	A3	A2	A2	AII-9 (proposed)
	INI 44	A2	Novel A2	A4	Novel AII
B	INI 48	Novel BIII	BIV	BIV	Novel B

3.4. Heterogeneous sequencing profile in assemblage B

Regarding the assemblage B isolates, the number of sequences showing heterogeneous nucleotides was much higher than among assemblage A samples. Additionally, the percentage of samples with double peaks varied among the three different genes, being higher in *gdh* locus (13 of 18; 72.2%), followed by *tpi* (8 of 20; 40%) and *bg* (4 of 16; 25%) loci (Table 7).

At the *bg* gene, double peaks occurred randomly at eight positions (C or T/Y). Two samples (INI 10 and INI 69) were classified as new variants

Table 6
Multilocus sequencing typing of *G. lamblia* assemblage B isolates.

Isolate (ID)	Sequencing		
	bg	*tpi*	*gdh*
INI 10	BIII	N/A	N/A
INI 21	BIII	BIV	
INI 26		BIII	
INI 27	BIII	N/A	BIV
INI 28	N/A	N/A	N/A
INI 33		BIII	BIII
INI 43	N/A	BIII	N/A
INI 48	BIII	BIV	BIV
INI 49, 53	BIII	BIII	N/A
INI 50	BIII	BIII	N/A
INI 54, 61, 67	BIII	N/A	N/A
INI 55		BIV	
INI 56			BIII
INI 57	BIII	BIII*	BIII*
INI 64	BIII	BIII + BIV	N/A
INI 65		BIV	BIV
INI 68	BIII	N/A	N/A
INI 69	BIII*	BIV	N/A

Empty spaces indicate isolates that were not sequenced; N/A indicate isolates that could not be assigned to a specific sub-assemblage; and asterisks (*) indicate isolates containing heterogeneous nucleotides.

of BIII. Nonetheless two isolates (INI 28 and 43) could not be classified in sub-assemblage due to the great number of heterogeneous nucleotides (Table 7).

Overlapping nucleotides were observed at 25 positions in the *tpi* gene and the majority occurred at positions 39, 91, 165, 168, 210 and 429 (positions used to differentiate BIII and BIV). Forty-five positions presented double peaks at *gdh* gene, most of them occurred at positions 309, 357, 360, 429, 447, 519, 540, 546, 612, 699, 723, 807, 825, 834, 867 and 933 (nine of these positions have been proposed for sub-assemblage differentiation). The majority of samples could not be assigned to a specific sub-assemblage at the *tpi* (INI 10, 27, 28, 54, 61, 67 and 68) and *gdh* (INI 10, 28, 43, 49, 50, 53, 54, 61, 64, 67, 68 and 69) loci due the presence of ambiguous nucleotides in specific positions. Only one sequence (INI 57) can be classified as BIII sub-assemblage in both loci (Table 7).

3.5. Phylogenetic analysis

The tree obtained from the concatenated *bg*, *tpi* and *gdh* sequences confirmed the monophyletic status of assemblages A and B (Fig. 1), and presented similar results in comparison with the phylogenetic analysis of individual genes (Figs. 1S, 2S and 3S). The sequences distinctly grouped together into the expected assemblages A and B. Dendrograms were able to differentiate *G. lamblia* assemblages, sub-assemblages and subtypes, with high bootstrap support.

The maximum likelihood and neighbour joining trees placed all sequences classified as assemblage A in one cluster with sub-assemblage AII sequences references, with high bootstrap values. Additionally, the analyses showed that the MLGs identified in our samples were identical or very closely related to previously described MLGs (AII-1 and AII-4) in human isolates from Brazil and Europe. The proposed MLGs (AII-8 and AII-9), as well as the new MLG (INI 44, novel AII), clustered with MLGs previously described (Cacciò et al., 2008). The only assemblage B isolate (INI 48) in which the MLG can be performed has been grouped with isolates from assemblage BIV and was classified as a new MLG.

4. Discussion

This is the first study to use multilocus sequence typing (MLST) of the *bg*, *tpi* and *gdh* genes to characterise *G. lamblia* in Rio de Janeiro/Brazil. We have looked at the genetic variability of *G. lamblia* isolates, identified the inter- and intra-assemblage level of genetic variation and described two novel MLGs in assemblage A.

Sequence analysis showed that assemblage B isolates have a higher genetic polymorphism than those from assemblage A, with heterogeneous nucleotides found in fourteen of twenty-one isolates. The presence of double peaks and SNPs hindered the classification of sub-assemblages and subtypes, and, in most cases, prevented the unambiguous identification of MLGs. These results show that the sequenced isolates displayed mainly novel polymorphisms, and that the assemblage B isolates exhibited many more subtypes than those from assemblage A, which is in agreement with results from various other research groups (Cacciò et al., 2008; Lebbad et al., 2008; Minetti et al., 2015a). At the *bg* locus, we identified five new subtypes and one subtype previously described in Poland from a roe deer and a red deer (Stojecki et al., 2015). With the *tpi* gene, three novel subtypes were detected and three subtypes that have been previous reported (Monis et al., 1999; Sulaiman et al., 2004).Four different subtypes have been identified at the *gdh* locus, one new one and three that have previously been described, one of which was detected in a horse from Colombia (Santín et al., 2013).

Assemblage A sequences across the three studied loci showed 100% homology with previous published sequences. Ambiguous nucleotides in the chromatogram and sequences containing SNPs were observed in two isolates, and were identical with previously described sequences (Lebbad et al., 2011; Durigan et al., 2014;). The sequence of isolate INI 44 was classified as a novel variant of subtype A2 since a single point mutation suffices for the description of a new subtype (Cacciò and Ryan, 2008; Monis et al., 2009; Sprong et al., 2009). Furthermore, an SNP at position 699 (C - T) of the *gdh* gene was detected in two samples (BfR156 and BfR166) and was enough to classify the sequences as new (novel AII) (Broglia et al., 2013).

In 2008, Cacciò et al. identified 10 different MLGs for assemblage A on the basis of the combination of three genetic markers (*bg*, *tpi* and *gdh*) and proposed the standardization of the nomenclature. The sequences/subtypes found in our work have already been reported, but the combinations of them at the three MLST loci are different from those previously described. We detected four assemblage A MLGs: two previously described ones (AII-1 and AII-4) and two novel MLGs. Our group therefore, suggests following the numbering scheme of MLGs proposed by Cacciò et al. (2008) by naming these novel combinations as AII-8 and AII-9. Various other authors have also observed new MLGs with different combinations of *bg*, *tpi* and *gdh* (Geurden et al., 2009; Lebbad et al., 2010, 2011; Broglia et al., 2013), and two previous surveys (Minetti et al., 2015a; Soba et al., 2015) report the same combinations of our proposed MLGs (AII-8 and AII-9).

Samples classified as assemblage A were all typed as AII, consistent with previous studies conducted in Brazil (Souza et al., 2007; Santos et al., 2012; Durigan et al., 2014; David et al., 2015; Colli et al., 2015; Oliveira-Arbex et al., 2015; Faria et al., 2016). However, our results contrast with data obtained in Rio de Janeiro (Volotão et al., 2007), where 97% of the samples were identified as sub-assemblage AI and assemblage B was not detected. These findings should be analysed carefully since genotyping data was based on a single marker (*bg*). Some studies suggest that assemblage A sequences might be preferentially amplified by the *bg* gene, while sequences classified as assemblage B are not amplified or are less amplified in this locus (Broglia et al., 2013; Durigan et al., 2014). Our group recently detected the presence of assemblage B in Rio de Janeiro city, and observed a change in the genetic profile over the years, with assemblage A being prevalent until 2012 whereas an increased number of cases of assemblage B was observed from 2013 to 2015 (Faria et al., 2016). In addition to changes over time in the frequency of infections with different assemblages, we also noted a switchover in the sub-assemblages, where sub-assemblage AI was not detected and sub-assemblage AII became more frequent. These findings could explain the discrepancy between our results and those reported by Volotão et al. (2007) on Rio de Janeiro The use of several molecular markers (MLGs) certainly allowed us to considerably increase the characterization of *G. lamblia* isolates and to reliably determine the distribution of assemblages in the population of Rio de Janeiro. Additionally, our results suggest that zoonotic transmission is low or absent in Rio de Janeiro, at least in relation to assemblage A, where sub-assemblage AI was not detected.

MLG approaches improve our understanding of transmission dynamics of human giardiasis, determine possible associations between symptoms and assemblages and help the study of potential zoonotic transmission (Huey et al., 2013; Ryan and Cacciò, 2013; Minetti et al., 2015a; Thompson and Ash, 2016). However, most of the studies in Brazil were based on the analysis of a single marker, and it is only in the last few years that the MLST protocol has been applied (Durigan et al., 2014; Colli et al., 2015; David et al., 2015; Oliveira-Arbex et al., 2015). For example, Colli et al. (2015) determined the prevalence of *G. lamblia* assemblages in Paraná (Southern Brazil) and reported the connection among different links in the epidemiology chain using the MLG tools.

Comparing our results with other studies on Brazil, we see that reported data on samples from Paraná (Colli et al., 2015) and from São Paulo (David et al., 2015; Oliveira-Arbex et al., 2015) showed low levels of polymorphism with no detection of heterogeneous nucleotides, which was significantly different from the profile of our sequences. In contrast, another study conducted in São Paulo reported a high degree of genetic diversity, with the detection of mixed assemblages (A + B) in 25% of the samples, the observation of overlapping nucleotides and the identification of assemblage C in five human clinical samples

Table 7
List of heterogeneous positions in β-giardin, triose phosphate isomerase and glutamate dehydrogenase genes sequences of assemblage B.

Isolate (ID)	*bg*			
	393	396	477	525	570
BIII (Ld18)*	A	C	T	T	C
BIV (ISSGF4)*				C	T
INI 10	G	T		**Y**	
INI 28				C	**Y**
INI 43				C	T
INI 69	G	T	**Y**		

	tpi																					
					1		1		1	1		1	2		2	2		2				
	3	4		9	4		6		6	6		6	1		1	1		5				
	9	1		1	1		0		2	5		8	0		6	7		6				
BIII (Bah12)*	–	–		C	C		G		G	C		C	G		C	G		G				
BIV (Ad19)*	A	G		T						T		T	A									
INI 10	**R**	G		**Y**						**Y**		**Y**			**Y**							
INI 27	A	G		**Y**	**Y**				**R**	**Y**		**Y**	**R**									
INI 28	G	**R**		**Y**						**Y**		**Y**				**R**						
INI 54	**R**	G		**Y**			**R**			**Y**		**Y**	**R**									
INI 57	G	G																				
INI 61	**R**	G		**Y**						**Y**		**Y**						**R**				
INI 67	A	G		**Y**						**Y**		**Y**										
INI 68	**R**	G		**Y**						**Y**		**Y**	**R**									

	gdh																						
	297	309	353	354	357	360	370	390	402	429	447	486	492	498	519	540	546	561	570	576	582	597	612
BIII (Bah12c14)*	–	–	–	–	–	–	–	–	–	T	T	T	C	T	C	C	C	C	C	G	G	C	G
BIV (Ad28)*	C	T	C	G	T	G	G	C	C	C	C					T		T					A
INI 10	**Y**	C	C	G	**Y**	**R**	**R**	C	C						**Y**	**Y**	**Y**	**Y**					
INI 28	C	**Y**	C	G	**Y**	**R**	G	C	C						T		**Y**						
INI 43	C	**Y**	C	G	**Y**	**R**	G	C	C						T		**Y**						
INI 49	C	**Y**	C	G	**Y**	**R**	G	C	C	**Y**					T		**Y**						
INI 50	C	T	C	G	C	**R**	G	**Y**	C	**Y**	**Y**				**Y**	**Y**	**Y**						
INI 53	C	**Y**	**Y**	G	**Y**	**R**	G	C	**Y**	**Y**	**Y**	**Y**		**Y**	**Y**		**Y**	**Y**			**R**		**R**
INI 54	C	**Y**	C	G	**Y**	G	G	C	C	**Y**	**Y**				**Y**	**Y**			**Y**			**Y**	**R**
INI 57	C	C	C	G	T	G	G	C	C						T		**Y**						
INI 61	**Y**	**Y**	C	**R**	**Y**	**R**	G	C	C	**Y**	**Y**				**Y**	**Y**	**Y**	**Y**	**Y**			**Y**	**R**
INI 64	**Y**	**Y**	C	G	T	G	G	C	C	**Y**	**Y**		**Y**		**Y**	**Y**	**Y**						**R**
INI 67	C	C	C	G	T	G	G	C	C	**Y**	C				**Y**	**Y**				**R**			
INI 68	C	T	C	G	C	**R**	G	C	C	**Y**	**Y**				**Y**	**Y**	**Y**		**Y**			**Y**	**R**
INI 69	**Y**	**Y**	C	**R**	**Y**	**R**	G	C	C	**Y**	**Y**				**Y**	**Y**		**Y**					A

Nucleotide substitutions are numbered from the ATG codon of each gene, empty spaces indicate identity with the BIII reference sequence (GenBank accession numbers: AY072726 [Ld18] for *bg*, AF069561 [Bah12] for *tpi*, and EF685684 [Bah12c14] for *gdh*). Asterisks (*) indicates reference sequences obtained from GenBank. Heterogeneous positions are in bold and the dash (—) indicates deletion.

Table 7 (*continued*)

Isolate (ID)	*bg*																					
	570	606				615				636				714					717			
BIII (Ld18)*	C	C				C				C				C					C			
BIV (ISSGF4)*	T	T				T				C				–					–			
INI 10																						
INI 28	**Y**	**Y**				**Y**				**Y**				**Y**					**Y**			
INI 43	T	**Y**				**Y**								**Y**					**Y**			
INI 69																						
	tpi																					
	2	2		2	2		3	3	3		3	4		4	4		4	4		5	5	
	6	7		8	9		0	1	8		9	0		2	7		8	8		0	1	
	4	1		0	7		4	2	1		3	2		9	1		3	6		4	0	
BIII (Bah12)*	G	C		A	A		A	C	C		C	A		G	A		A	G		C	G	
BIV (Ad19)*														A								
INI 10															**M**							
INI 27														**R**								
INI 28	**R**														**M**			**R**				
INI 54														**R**						**Y**		
INI 57							**R**				**Y**											
INI 61								**Y**							**M**							
INI 67		**Y**		**R**	**R**				**Y**			**R**			C		G				A	
INI 68														**R**								
	gdh																					
	615	624	660	687	690	699	702	705	723	738	756	807	815	825	834	841	843	867	894	897	921	933
BIII (Bah12c14)*	A	A	C	C	G	T	C	T	C	G	T	G	A	A	C	C	C	C	C	C	G	T
BIV (Ad28)*						C			T			T		G							A	C
INI 10						**Y**			**Y**		**Y**	**K**		**R**								C
INI 28				**Y**		**Y**			**Y**	**R**		**K**			**Y**			**Y**				**Y**
INI 43						C						**K**	**R**		T			**Y**				C
INI 49						**Y**				**R**		**K**			**Y**			**Y**				C
INI 50						**Y**						T			T			T				C
INI 53						C			**Y**			**K**			**Y**			**Y**				**Y**
INI 54					**R**	C	**Y**	**Y**	T		C	**K**		**R**		**Y**						**Y**
INI 57						C						**K**	G		T			T				C
INI 61						C			**Y**		**Y**	**K**		**R**							**R**	**Y**
INI 64												T			T							C
INI 67	**R**	**R**	**Y**						**Y**			T		**R**	**Y**				**Y**	**Y**	**R**	**Y**
INI 68					**R**	**Y**		**Y**	**Y**	**R**	**Y**	T		**R**	**Y**			**Y**				C
INI 69												T		G			**Y**				**R**	**Y**

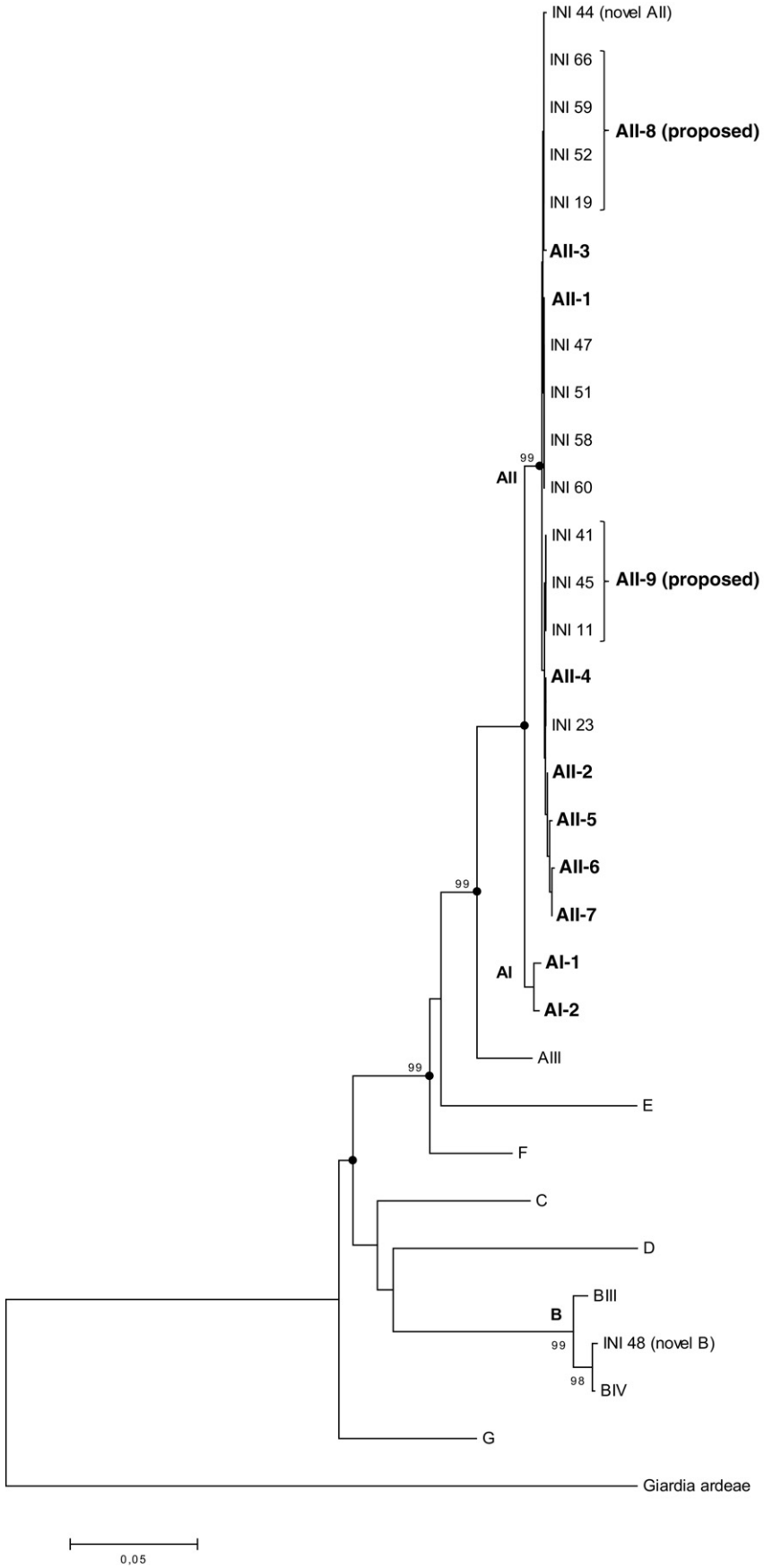

Fig. 1. Phylogenetic relationship of *G. lamblia* isolates constructed by neighbour-joining analysis of concatenated *bg*, *tpi* and *gdh* nucleotide sequences. Multilocus genotypes from a previous study (Cacciò et al., 2008) are indicated in bold. Symbol (●) indicates nodes branches conserved when the tree was reconstructed using maximum likelihood. Optimal nucleotide substitution model was Tamura-Nei, and only bootstraps values >80% are shown.

(Durigan et al., 2014). Discrepancies are likely to be found between the different studies accomplished in Brazil, mainly because the country is so vast and the cultural, behavioural and social diversity of the population so varied.

The successful amplifications of the three studied loci behaved differently. The *tpi* and *gdh* genes were amplified in 57% and 49% of samples, respectively, while the *bg* gene had the lowest rate of amplification (48%). In some studies, it has been observed that the *tpi* gene had the highest amplification success, followed by *gdh* and *bg* (Huey et al., 2013; Durigan et al., 2014; Oliveira-Arbex et al., 2015). Other groups have also reported differences in the amplification rate for different genetic loci (David et al., 2015; Minetti et al., 2015a). Several factors may influence the PCR efficiency, such as the quality and quantity of DNA templates and the presence of inhibitors (Wilke and Robertson, 2009). The stool samples were stored at −20 °C without preservatives before DNA extraction. Probably the freezing for several months (or years) may have led to degradation of the cysts/DNA of *G. lamblia*. Minetti et al. (2015b) had similar results, where the overall PCR amplification success of the samples varied accordingly to the year of collection and the time from collection to DNA extraction. When the DNA was extracted within one month of collection, the amplification success rate was higher compared with amplification of older specimens (Faria et al., 2016). Possibly the DNA region between the PCR primers and the genomic sequences were degraded, or there were nucleotide mismatches, which may cause a strong reduction or even a lack of amplification, especially considering that the three genes used were single copies (Broglia et al., 2013; Thompson and Ash, 2016). The presence of PCR inhibitors in stool samples, such as bilirubin, bile salts, hemoglobin, phenolic compounds, and complex polysaccharides, which are co-purified together during DNA extraction, also limited the success of amplification.

Congruent genotyping results were obtained at the three genetic loci investigated, which is in contrast with some previous reports (Cacciò et al., 2008; Sprong et al., 2009; Huey et al., 2013). However, discordant intra-assemblage assignment was observed in assemblage B isolates (INI 21, 27, 48, 64 and 69): the *bg* locus was not consistent with the other two loci (*tpi* and *gdh*). Inconsistent results were also observed at sub-assemblage level when these results were compared with data previously reported by our group (Faria et al., 2016) Six samples (INI 10, 50, 54, 61, 67 and 68) that were genotyped by PCR-RFLP as mixtures of sub-assemblages BIII and BIV were classified as sub-assemblage BIII by sequencing analysis in this study. Mixed patterns of alleles have been observed, which may be the result of recombination between different assemblage B subtypes (Lebbad et al., 2011).

Two principal mechanisms can explain both the lack of agreement in the assignment of assemblages in different loci and the occurrence of DNA sequences with overlapping nucleotides: allelic sequence heterozygosity (ASH) and mixed infections. *Giardia* has two diploid nuclei which may accumulate specific mutations independently, and this could generate ASH; in addition, mixed infections may lead to the presence of genetically different cysts in the same sample and can occur at the inter- and intra-assemblage levels. Although the occurrence of ASH between the two nuclei of a single cyst and also between different cysts has been demonstrated, the reasons for different ASH levels between assemblages A and B is not yet understood (Ankarklev et al., 2012). According to the literature, the sub-assemblage AI genome displays a very low level of ASH compared with assemblage B and sub-assemblage AII (Cacciò and Sprong, 2010; Ankarklev et al., 2012; Ankarklev et al., 2015). The occurrence of ASH complicates the assignment of specific subtypes, especially for assemblage B, and it is very difficult to distinguish between ASH and mixed infections. Another possible explanation for the presence of double peaks is the occurrence of meiotic recombination, which suggests the potential for sexual reproduction in *Giardia*, although the evidence of sex still elusive (Birky, 2010; Cacciò and Sprong, 2010).

It is important to emphasize that we did not find any mixed infections of assemblages A and B. Sometimes, the failure to detect mixed infections does not mean that they are not present, but that one assemblage or sub-assemblage has been preferentially amplified over another at one locus (Ryan and Cacciò, 2013). To prevent this, three different assemblage-specific primers (*tpi*, *gdh* and *orf*C4) were used (Faria et al., 2016), and it was confirmed that mixed infections were not present in the samples studied. It therefore seems that the occurrence of double peaks is not the result of mixed infections of assemblages A and B, but may be caused either by there being multiple subtypes in the same sample or the occurrence of ASH.

Recently, an increase in sequence information and the description of new sequences/subtypes have been reported (Ryan and Cacciò, 2013; Durigan et al., 2014; Minetti et al., 2015a; Nunes et al., 2016). Even isolates classified as assemblage A (for example, INI 17 and INI 44), that show less variability than B, could not be classified under the previous MLST nomenclature proposal (Cacciò et al., 2008). MLG is a useful subtyping tool for assemblage A but less valuable for assemblage B because of the high frequency of heterogeneous positions in the chromatogram. As suggested previously, a re-evaluation of the current molecular epidemiological methods and the development of new target regions of the genome with lower substitution rates may facilitate the study of assemblage B (Ankarklev et al., 2012; Wielinga et al., 2015).

A number of studies have been carried out with different research groups in various countries, reporting correlations between assemblages and symptoms (Pestechian et al., 2014; Puebla et al., 2014; Minetti et al., 2015b). However, in our study no significant association was found between the clinical manifestation and assemblages, probably because of the small number of genotyped samples.

This is the first report to use MLST to genetically characterise *G. lamblia* human isolates in Rio de Janeiro, Brazil. The identification of the novel subtypes in both assemblages and the description of novel assemblage A MLGs reflect the genetic distribution of the parasite. Our work provides a new insight into the genetic diversity of parasites and improves our understanding of the epidemiology of the disease by elucidating the dynamics of giardiasis in the population of Rio de Janeiro. It has shown that MLST should be applied in the characterization of *G. lamblia* in Brazil and elsewhere. Further studies on the molecular epidemiology are imperative, since the parasite was found to vary widely in Brazilian isolates and this knowledge is crucial for both veterinary and public health researchers.

Acknowledgments

The authors would like to thank Dr. André Almeida (Instituto Nacional de Saúde Dr. Ricardo Jorge, Porto, Portugal), Dr. Diogo Neves Proença (IMAR-CMA, University of Coimbra, Portugal), and Prof. Paula V Morais (IMAR-CMA, University of Coimbra, Portugal; Department of Life Sciences, FCTUC, University of Coimbra, Portugal) for the technical support. We are grateful to DNA sequencing platform of PDTIS/FIOCRUZ, Rio de Janeiro. This work was supported by FCT POCTI (FEDER) and COMPETE (PEst-C/SAU/LA000172013-2014); Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (245531/2012-2).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.meegid.2017.06.028.

References

Almeida, A., Pozio, E., Cacciò, C.M., 2010. Genotyping of *Giardia duodenalis* cysts by new real-time PCR assays for detection of mixed infections in human samples. Appl. Environ. Microbiol. 76 (6), 1895–1901.

Ankarklev, J., Svärd, S.G., Lebbad, M., 2012. Allelic sequence heterozygosity in single *Giardia* parasites. BMC Microbiol. 12, 65.

Ankarklev, J., Franzén, O., Peirasmaki, D., Jerlström-Hultqvist, J., Lebbad, M., Andersson, J., Andersson, B., Svärd, S.G., 2015. Comparative genomic analyses of freshly isolated *Giardia intestinalis* assemblage A isolates. BMC Genomics 16, 697.

Appelbee, A.J., Frederick, L.M., Heitman, T.L., Olson, M.E., 2003. Prevalence and genotyping of *Giardia duodenalis* from beef calves in Alberta, Canada. Vet. Parasitol. 112, 289–294.

Bertrand, I., Albertini, L., Schwartzbrod, J., 2005. Comparison of two target genes for detection and genotyping of *Giardia lamblia* in human feces by PCR and PCR-restriction fragment length polymorphism. J. Clin. Microbiol. 43, 5940–5944.

Birky, C.W., 2010. *Giardia* sex? Yes, but how and how much? Trends Parasitol. 26, 70–74.

Broglia, A., Weitzel, T., Harms, G., Cacció, S.M., Nöckler, K., 2013. Molecular typing of *Giardia duodenalis* isolates from German travellers. Parasitol. Res. 112, 3449–3456.

Cacciò, S.M., Ryan, U., 2008. Molecular epidemiology of giardiasis. Mol. Biochem. Parasitol. 160, 75–80.

Cacciò, S.M., Sprong, H., 2010. *Giardia duodenalis*: genetic recombination and its implications for taxonomy and molecular epidemiology. Exp. Parasitol. 124, 107–112.

Cacciò, S.M., De Giacomo, M., Pozio, E., 2002. Sequence analysis of the β-giardin gene and development of a polymerase chain reaction-restriction fragment length polymorphism assay to genotype *Giardia duodenalis* cysts from human faecal samples. Int. J. Parasitol. 32, 1023–1030.

Cacciò, S.M., Beck, R., Lalle, M., Marinculic, A., Pozio, E., 2008. Multilocus genotyping of *Giardia duodenalis* reveals striking differences between assemblages A and B. Int. J. Parasitol. 38, 1523–1531.

de Carli, G., 2001. Parasitologia clínica. Seleção de métodos e técnicas de laboratório para o diagnóstico de parasitoses humanas. 1st ed. Atheneu, São Paulo.

Colli, C.M., Bezagio, R.C., Nishi, L., Bignotto, T.S., Ferreira, É.C., Falavigna-Guilherme, A.L., Gomes, M.L., 2015. Identical assemblage of *Giardia duodenalis* in humans, animals and vegetables in an urban area in southern Brazil indicates a relationship among them. PLoS One 10, e0118065.

David, É.B., Guimarães, S., Oliveira, A.P., Oliveira-Sequeira, T.C.G., Nogueira Bittencourt, G., Moraes Nardi, A.R., Martins Ribolla, P.E., Bueno Franco, R.M., Branco, N., Tosini, F., Bella, A., Pozio, E., Cacciò, S.M., 2015. Molecular characterization of intestinal protozoa in two poor communities in the state of São Paulo, *Giardia* sex. 8 pp. 1–12.

Durigan, M., Abreu, A.G., Zucchi, M.I., Franco, R.M.B., de Souza, A.P., 2014. Genetic diversity of *Giardia duodenalis*: multilocus genotyping reveals zoonotic potential between clinical and environmental sources in a metropolitan region of Brazil. PLoS One 9, e115489.

Faria, C.P., Zanini, G.M., Dias, G.S., da Silva, S., Sousa, M.C., 2016. Molecular characterization of *Giardia lamblia*: first report of assemblage B in human isolates from Rio de Janeiro (Brazil). PLoS One 11 (8), e0160762.

Faust, E.C., D'Antoni, J.S., Odom, V., Miller, M.J., Peres, C., Sawitz, W., Thomen, L.F., Tobie, J., Walker, J.H., 1938. Critical study of clinical laboratory technics for the diagnosis of protozoan cysts and helminth eggs in feces. Am. J. Trop. Med. Hyg. s1–18, 169–183.

Geurden, T., Goossens, E., Levecke, B., Vercammen, F., Vercruysse, J., Claerebout, E., 2009. Occurrence and molecular characterization of *Cryptosporidium* and *Giardia* in captive wild ruminants in Belgium. J. Zoo Wildl. Med. 40, 126–130.

Hopkins, R.M., Meloni, B.P., Groth, D.M., Wetherall, J.D., Reynoldson, J.A., Thompson, R.C., 1997. Ribosomal RNA sequencing reveals differences between the genotypes of *Giardia* isolates recovered from humans and dogs living in the same locality. J. Parasitol. 83, 44–51.

Huey, C.S., Mahdy, M.A.K., Al-Mekhlafi, H.M., Nasr, N.A., Lim, Y.A.L., Mahmud, R., Surin, J., 2013. Multilocus genotyping of *Giardia duodenalis* in Malaysia. Infect. Genet. Evol. 17, 269–276.

Kohli, A., Bushen, O.Y., Pinkerton, R.C., Houpt, E., Newman, R.D., Sears, C.L., Lima, A.A.M., Guerrant, R.L., 2008. *Giardia duodenalis* assemblage, clinical presentation and markers of intestinal inflammation in Brazilian children. Trans. R. Soc. Trop. Med. Hyg. 102, 718–725.

Lalle, M., Pozio, E., Capelli, G., Bruschi, F., Crotti, D., Cacciò, S.M., 2005. Genetic heterogeneity at the β-giardin locus among human and animal isolates of *Giardia duodenalis* and identification of potentially zoonotic subgenotypes. Int. J. Parasitol. 35, 207–213.

Lasek-Nesselquist, E., Welch, D.M., Sogin, M.L., 2010. 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. 40, 1063–1074.

Lebbad, M., Ankarklev, J., Tellez, A., Leiva, B., Andersson, J.O., Svärd, S., 2008. Dominance of *Giardia* assemblage B in León, Nicaragua. Acta Trop. 106, 44–53.

Lebbad, M., Mattsson, J.G., Christensson, B., Ljungström, B., Backhans, A., Andersson, J.O., Svärd, S.G., 2010. From mouse to moose: Multilocus genotyping of *Giardia* isolates from various animal species. Vet. Parasitol. 168, 231–239.

Lebbad, M., Petersson, I., Karlsson, L., Botero-Kleiven, S., Andersson, J.O., Svenungsson, B., Svard, S.G., 2011. Multilocus genotyping of human giardia isolates suggests limited zoonotic transmission and association between assemblage B and flatulence in children. PLoS Negl. Trop. Dis. 5, 1–10.

Minetti, C., Lamden, K., Durband, C., Cheesbrough, J., Fox, A., Wastling, J.M., 2015a. Determination of *Giardia duodenalis* assemblages and multi-locus genotypes in patients with sporadic giardiasis from England. Parasit. Vectors 8, 444.

Minetti, C., Lamden, K., Durband, C., Cheesbrough, J., Platt, K., Charlett, A., O'Brien, S.J., Fox, A., Wastling, J.M., 2015b. Case-control study of risk factors for sporadic giardiasis and parasite assemblages in North West England. J. Clin. Microbiol. 53, 3133–3140.

Monis, P.T., Andrews, R.H., Mayrhofer, G., Ey, P.L., 1999. Molecular systematics of the parasitic protozoan *Giardia intestinalis*. Mol. Biol. Evol. 16, 1135–1144.

Monis, P.T., Andrews, R.H., Mayrhofer, G., Ey, P.L., 2003. Genetic diversity within the morphological species *Giardia intestinalis* and its relationship to host origin. Infect. Genet. Evol. 3, 29–38.

Monis, P.T., Caccio, S.M., Thompson, R.C.A., 2009. Variation in *Giardia*: towards a taxonomic revision of the genus. Trends Parasitol. 25, 93–100.

Nunes, B.C., Pavan, M.G., Jaeger, L.H., Monteiro, K.J.L., Xavier, S.C.C., Monteiro, F.A., Bóia, M.N., Carvalho-Costa, F.A., 2016. Spatial and molecular epidemiology of *Giardia intestinalis* deep in the Amazon, Brazil. PLoS One 8, e0158805.

Oliveira-Arbex, A.P., David, E.B., Oliveira-Sequeira, T.C.G., Bittencourt, G.N., Guimarães, S., 2015. Genotyping of *Giardia duodenalis* isolates in asymptomatic children attending daycare centre: evidence of high risk for anthroponotic transmission. Epidemiol. Infect. 1–11.

Pestechian, N., Rasekh, H., Rostami-nejad, M., Yousofi, H.A., 2014. Molecular identification of *Giardia lamblia*; is there any correlation between diarrhea and genotyping in Iranian population? Gastroenterol. Hepatol. 7, 168–172.

Puebla, L.J., Núñez, F.A., Fernández, Y.A., Fraga, J., Rivero, L.R., Millán, I.A., Valdés, L.A., Silva, I.M., 2014. Correlation of *Giardia duodenalis* assemblages with clinical and epidemiological data in Cuban children. Infect. Genet. Evol. 23, 7–12.

Read, C.M., Monis, P.T., Thompson, R.C.A., 2004. Discrimination of all genotypes of *Giardia duodenalis* at the glutamate dehydrogenase locus using PCR-RFLP. Infect. Genet. Evol. 4, 125–130.

Ryan, U., Cacciò, S.M., 2013. Zoonotic potential of *Giardia*. Int. J. Parasitol. 43, 943–956.

Santín, M., Cortés Vecino, J.A., Fayer, R., 2013. A large scale molecular study of *Giardia duodenalis* in horses from Colombia. Vet. Parasitol. 196, 31–36.

Santos, C.K.S., Grama, D.F., Limongi, J.E., Costa, F.C., Couto, T.R., Soares, R.M., Mundim, M.J.S., Cury, M.C., 2012. Epidemiological, parasitological and molecular aspects of *Giardia duodenalis* infection in children attending public daycare centers in southeastern Brazil. Trans. R. Soc. Trop. Med. Hyg. 106, 473–479.

Soba, B., Islamovic, S., Skvarc, M., Caccio, S.M., 2015. Multilocus genotyping of *Giardia duodenalis* (Lambl, 1859) from symptomatic human infections in Slovenia. Folia Parasitol. 6, 62.

Souza, S.L.P., Gennari, S.M., Richtzenhain, L.J., Pena, H.F.J., Funada, M.R., Cortez, A., Gregori, F., Soares, R.M., 2007. Molecular identification of *Giardia duodenalis* isolates from humans, dogs, cats and cattle from the state of São Paulo, Brazil, by sequence analysis of fragments of glutamate dehydrogenase (*gdh*) coding gene. Vet. Parasitol. 149, 258–264.

Sprong, H., Cacciò, S.M., Van Der Giessen, J.W.B., 2009. Identification of zoonotic genotypes of *Giardia duodenalis*. PLoS Negl. Trop. Dis. 3, 1–12.

Stojecki, K., Sroka, J., Cacciò, S.M., Cencek, T., Dutkiewicz, J., Kusyk, P., 2015. Prevalence and molecular typing of Giardia duodenalis in wildlife from eastern Poland. Folia Parasitol. 2015 (62), 42.

Sulaiman, I.M., Fayer, R., Bern, C., Gilman, R.H., Trout, J.M., Schantz, P.M., Das, P., Lal, A.A., Xiao, L., 2003. Triosephosphate isomerase gene characterization and potential zoonotic transmission of *Giardia duodenalis*. Emerg. Infect. Dis. 9, 1444–1452.

Sulaiman, I.M., Jiang, J., Singh, A., Xiao, L., 2004. Distribution of *Giardia duodenalis* genotypes and subgenotypes in raw urban wastewater in Milwaukee, Wisconsin. Appl. Environ. Microbiol. 70, 3776–3780.

Thompson, R.C.A., 2004. The zoonotic significance and molecular epidemiology of *Giardia* and giardiasis. Vet. Parasitol. 126, 15–35.

Thompson, R.C.A., Ash, A., 2016. Molecular epidemiology of *Giardia* and *Cryptosporidium* infections. Infect. Genet. Evol. 40, 315–323.

Volotão, A.C., Costa-Macedo, L.M., Haddad, F.S.M., Brandão, A., Peralta, J.M., Fernandes, O., 2007. Genotyping of *Giardia duodenalis* from human and animal samples from Brazil using β-giardin gene: a phylogenetic analysis. Acta Trop. 102, 10–19.

WHO, 2013. Diarrhoeal Disease. (Available: http://www.who.int/mediacentre/factsheets/fs330/en/. Accessed 22 July 2016).

Wielinga, C., Thompson, R.C.A., Monis, P., Ryan, U., 2015. Identification of polymorphic genes for use in assemblage B genotyping assays through comparative genomics of multiple assemblage B *Giardia duodenalis* isolates. Mol. Biochem. Parasitol. 201, 1–4.

Wilke, H., Robertson, L.J., 2009. Preservation of *Giardia* cysts in stool samples for subsequent PCR analysis. J. Microbiol. Methods 78, 292–296.