Available online at www.sciencedirect.com
ScienceDirect
International Journal for Parasitology 38 (2008) 1523–1531
www.elsevier.com/locate/ijpara

Multilocus genotyping of *Giardia duodenalis* reveals striking differences between assemblages A and B ☆,☆☆

S.M. Cacciò [a,*], R. Beck [a,b], M. Lalle [a], A. Marinculic [b], E. Pozio [a]

[a] *Department of Infectious, Parasitic and Immunomediated Diseases, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Rome, Italy*
[b] *Department of Parasitology and Parasitic Diseases, University of Zagreb, Heinzelova 55, 10000 Zagreb, Croatia*

Received 25 February 2008; received in revised form 14 April 2008; accepted 17 April 2008

Abstract

Giardia duodenalis is a widespread parasite of mammalian species, including humans. Due to its invariant morphology, investigations of aspects such as host specificity and transmission patterns require the direct genetic characterisation of parasites from faecal samples. We performed a sequence analysis of four genes (ssrRNA, β-giardin, glutamate dehydrogenase and triose phosphate isomerase) of 61 human isolates and 29 animal isolates. The results showed that multilocus genotypes (MLGs) can be readily defined for *G. duodenalis* isolates of assemblage A but not for assemblage B. Indeed, for assemblage A isolates, there was no evidence of intra-isolate sequence heterogeneity, and congruent genotyping results were obtained at the four genetic loci investigated. Sequence comparison and phylogenetic analysis showed that human-derived and animal-derived MLGs are different, and further indicated the presence of a new sub-assemblage (referred to as "AIII"), which was found exclusively in wild hoofed animals. On the other hand, there were variable levels of intra-isolate sequence heterogeneity (i.e., the presence of two overlapping nucleotide peaks at specific positions in the chromatograms, or "heterogeneous templates") in assemblage B isolates from humans and animals, and this prevented the unambiguous identification of MLGs. Furthermore, in five human isolates and one non-human primate isolate, the assignment to assemblage B was problematic, given that one of the four markers supported an assignment to assemblage A. These findings raise concerns about the interpretation of genotyping data based on single markers, and indicate the need to understand the mechanisms that are responsible for the differences between *G. duodenalis* assemblages A and B.

Keywords: *Giardia duodenalis*; Multilocus genotyping; Assemblage A; Assemblage B; Sequence heterogeneity; Zoonotic transmission

1. Introduction

To accurately discriminate among *Giardia duodenalis* isolates, genetic characterisation is necessary, given that the parasite's morphology does not vary. In the past 10 years, and especially since the introduction of DNA amplification techniques, genetic characterisation has been extensively used to evaluate the genetic variability within *G. duodenalis*, to assess the role of animals in the epidemiology of human infection, and to develop methods for tracing sources of infection (Cacciò et al., 2005; Smith et al., 2006). However, most of these studies have been based on the analysis of single markers, and the genetic variability and usefulness of the different genetic loci in identifying isolates have not been systematically evaluated. Indeed, the "sub-assemblages" within assemblages A and B (referred to as "AI", "AII", "BIII" and "BIV") have mainly been identified based on isoenzyme analysis, as has the extent of genetic variability among isolates (Monis et al., 2003). However, isoenzyme analysis cannot be applied routinely to clinical or environmental samples because it requires the in vitro or in vivo propagation of *Giardia* spp.

☆ *Note.* Supplementary data associated with this article.
☆☆ *Note.* Nucleotide sequence data reported in this paper are available in the GenBank, EMBL and DDBJ databases under the accession numbers EU637578 to EU637593.
* Corresponding author. Tel.: +39 06 4990 2484; fax: +39 06 4990 3561.
E-mail address: simone.caccio@iss.it (S.M. Cacciò).

0020-7519/$34.00 © 2008 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijpara.2008.04.008

In the present study, we performed an extensive characterisation of human and animal isolates of *G. duodenalis* by PCR and direct sequencing of portions of the ssrRNA, β-giardin (*bg*), triose phosphate isomerase (*tpi*) and glutamate dehydrogenase (*gdh*) genes, which are the four genetic markers commonly used for genotyping studies. The objectives of the study were to determine the intra-assemblage level of genetic variation at the different loci and to define multilocus genotypes (MLGs) for assemblages A and B of human and animal origins.

2. Materials and methods

2.1. Source of isolates

The isolates were collected over a 3-year period in Italy, Africa and Croatia. The human isolates were from Italy ($n = 28$) and Africa ($n = 33$). The animal isolates were from: domestic animals [four calves, one domestic pig and two water buffalos (*Bubalus bubalis*), all from Italy, and three domestic cats from Italy and Croatia]; wildlife [eight fallow deer (*Dama dama*) from Italy and one wild boar (*Sus scrofa*) from Croatia]; and captive non-human primates [seven Barbary macaques (*Macaca sylvanus*), one mandrill (*Mandrillus sphinx*) and two chimpanzees (*Pan troglodytes*), all from Italy)].

Faecal samples were concentrated by flotation on a sucrose gradient (specific gravity, 1.06), and *Giardia* spp. cysts were detected microscopically by immunofluorescence using a commercial kit (Merifluor, Meridian Bioscience, Cincinnati, Ohio, USA), following the manufacturer's instructions.

The following axenic strains of *G. duodenalis* of human origin were used as references: for assemblage A sub-assemblage AI, the strains WB and Ad-1 (Baruch et al., 1996; Monis et al., 1999); for assemblage A sub-assemblage AII, the strains Bris-162 and KC8 (Monis et al., 1996); for assemblage B sub-assemblage BIII, the strains BAH12 and Ld18 (Homan et al., 1998; Monis et al., 1999); and for assemblage B sub-assemblage BIV, the strains Ad28 and Nji5 (Homan et al., 1998; Monis et al., 1999).

2.2. Molecular analysis

DNA was extracted directly from faecal samples according to the procedure described by da Silva et al. (1999). In brief, an aliquot of each faecal sample was homogenised using the FP120 Fast Prep Cell disruptor (Savant, Thermo Electro Corporation, Woburn, Massachusetts, USA). The DNA, released after the lysis step, was purified using the Fast DNA extraction kit (Qbiogene, Illkirch Cedex, France).

Portions of the ssrRNA (292 bp), *tpi* (530 bp) and *bg* (511 bp) genes were individually amplified according to previously described protocols (Hopkins et al., 1997; Sulaiman et al., 2003; Lalle et al., 2005). The amplification of a fragment (530 bp) of the *gdh* gene was achieved using newly designed primers, Ghd1 (5′ TTCCGTRTYCAGTACAA CTC 3′) and Gdh2 (5′ ACCTCGTTCTGRGTGGCGCA 3′) for the primary amplification, and Gdh3 (5′ ATGACYG AGCTYCAGAGGCACGT 3′) and Gdh4 (5′ GTGGCGC ARGGCATGATGCA 3′) for the nested amplification. Identical conditions were used for the primary and secondary amplification of the *gdh* gene fragment: 35 cycles (94 °C for 30 s, 50 °C for 30 s and 72 °C for 1 min) in a T-personal thermocycler (Whatman-Biometra, Goettingen, Germany), with an initial hot start at 94 °C for 2 min and a final extension at 72 °C for 7 min.

The genes under investigation are unlinked in the *G. duodenalis* genome, at least in the assemblage A genome, which is desirable for genetic studies. The *tpi* gene is at position 95,921–96,694 on the 200-kb long contig ctg02-19; the *bg* gene is at position 55,484–56,302 on the 90-kb long contig ctg02_35; the *gdh* gene is at position 60,579–61,928 on the 231-kb long contig ctg02_15 (data taken from www.giardiadb.org). In previous studies, based on hybridization on chromosomes separated by pulsed-field gel electrophoresis, the *tpi* gene was mapped to chromosome 5; *gdh* and *bg* were mapped to chromosome 4; and the majority of ssrRNA gene copies were mapped to chromosome 1 (Adam, 2001).

PCR products were purified using spin columns (Qiagen, Milan, Italy) and sequenced from both strands. The sequences were edited using SeqMan 7.0 software (DNASTAR, Madison, Wisconsin) and aligned using Clustal X (Thompson et al., 1997). The phylogenetic analysis was performed using neighbour-joining (NJ) implemented in the MEGA program (Kumar et al., 2004, v. 3.1), and the "quartet puzzling method" implemented in the TREE-PUZZLE program (Strimmer and von Haeseler, 1996, v. 5.0). Bootstrap values were calculated by the analysis of 1000 replicates. Phylogenetic trees inferred from the analyses using these programs were drawn using the TreeView program (Page, 1996). The phylogenetic analysis also included sequences of the homologous genes from representative isolates of assemblage C (for *tpi*, Accession No. AY228641; for *bg*, AY545646; for *gdh*, U60982), E (for *tpi*, Accession No. AY228645; for *bg*, AY072729; for *gdh*, AY178741), and F (for *tpi*, Accession No. AF069558; for *bg*, AY647264; for *gdh*, AF069057). Also included were the sequences from *Giardia ardeae* (for *gdh*, Accession No. AF069060; for *tpi*, AF069564; no sequence was available for *bg*).

3. Results

3.1. Multilocus genotyping of isolates representing assemblage A

A total of 28 human isolates and 19 animal isolates (four from calves, one from a domestic pig, two from water buffalos, eight from fallow deer, one from a wild boar and three from domestic cats) were classified as assemblage A