Multilocus genotyping of *Giardia duodenalis* in Malaysia

Choy Seow Huey\(^a\), Mohammed A.K. Mahdy\(^a,b,\ast\), Hesham M. Al-Mekhlafi \(^a,b\), Nabil A. Nasr\(^a\), Yvonne A.L. Lim\(^a\), Rohela Mahmund\(^a\), Johari Surin\(^a\)

\(^a\) Department of Parasitology, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia
\(^b\) Department of Parasitology, Faculty of Medicine, Sana’a University, Sana’a, Yemen

**A R T I C L E   I N F O**

Article history:
Received 18 December 2012
Received in revised form 6 April 2013
Accepted 9 April 2013
Available online 25 April 2013

Keywords:
Multilocus genotyping
*Giardia*
tpi
bg
gdh
Malaysia

**A B S T R A C T**

*Giardia duodenalis* is considered the most common intestinal parasite in humans worldwide. In Malaysia, many studies have been conducted on the epidemiology of giardiasis. However, there is a scarcity of information on the genetic diversity and the dynamics of transmission of *G. duodenalis*. The present study was conducted to identify *G. duodenalis* assemblages and sub-assemblages based on multilocus analysis of the glutamate dehydrogenase (*gdh*), beta-giardin (*bg*) and triose phosphate isomerase (*tpi*) genes. Faecal specimens were collected from 484 Orang Asli children with a mean age of 7 years and examined using light microscopy. Specimens positive for *Giardia* were subjected to PCR analysis of the three genes and subsequent sequencing in both directions. Sequences were edited and analysed by phylogenetic analysis. *G. duodenalis* was detected in 17% (84 of 484) of the examined specimens. Among them, 71 were successfully sequenced using at least one locus. Genotyping results showed that 30 (42%) of the isolates belonged to assemblage A, 32 (45%) belonged to assemblage B, while discordant genotype results were observed in 9 specimens. Mixed infections were detected in 43 specimens using a tpi-based assemblage specific protocol. At the sub-assemblies level, isolates belonged to assemblage A were all. High nucleotide variation found in isolates of assemblage B made subtyping difficult to achieve. The finding of assemblage B and the anthropogenic genotype All implicates human-to-human transmission as the most possible mode of transmission among Malaysian aborigines. The high polymorphism found in isolates of assemblage B warrants a more defining tool to discriminate assemblage B at the sub-assemblage level.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

*Giardia duodenalis* (syn. *G. intestinalis*; *Giardia lamblia*) is a flagellate enteric parasite that infects the intestinal tract of humans and a wide variety of other mammals through ingestion of infective cysts (van der Giessen et al., 2006). It is a major cause of non-viral/bacterial diarrhea affecting both developed and developing countries and more frequently encountered in areas with sub-standard environment, poor hygienic practices and inadequate water treatment system (Daly et al., 2010; Robertson et al., 2009). *G. duodenalis* infections are most often asymptomatic. For symptomatic cases, the degree of symptoms and severity varies among different individuals (Buret, 2008).

Knowing the genetic diversity of *Giardia* is an essential component in enhancing our understanding on the taxonomy, epidemiology and population genetics of this parasite. Therefore, molecular characterization and in the recent years, multilocus genotyping (MLG) have become the prevailing tool used for genotyping and subtyping of *G. duodenalis*. This has in turn facilitated in outbreak surveillance, contamination source-tracking, and unveiling zoo-notic potential, dynamic transmission, and relationship of genotypes and hosts (Feng and Xiao, 2011). Molecular analyses revealed that *G. duodenalis* isolated from humans and animals is a species complex with at least 8 distinct assemblages designated as A to H, which demonstrate similarity in morphologic characteristics but are phenotypically and genotypically heterogeneous (Plutzer et al., 2010). Of these, only assemblages A and B can cause infection in humans with the exception of a small fraction of cases where animal host specific assemblages C–F were reported in human (Sprong et al., 2009). Besides, both assemblages A and B are also capable of infecting animals. The subdivision into Al–All has provided greater insights on the dynamic of transmission of these assemblages. Sub-assemble All has been regarded as anthropogenic whereas Al and All are predominant in livestock and wildlife, respectively (Feng and Xiao, 2011; Nolan et al., 2010; Sprong et al., 2009; Thompson et al., 2000).

In Malaysia, giardiasis is an endemic disease and is associated with malnutrition among children in the rural areas resulting in stunting, wasting and vitamin A deficiency (Al-Mekhlafi et al.,...
The prevalence of human *Giardia* infection varies between 0.2% and 29.2% (Lim et al., 2008; Noor Azian et al., 2007). Most of the epidemiological studies detected *Giardia* on the basis of microscopic examination without employing molecular approach. Data on genotypes of *G. duodenalis* up to the assemblage level remains scarce. In a previous genotyping study using SSU rRNA locus, one specimen was identified as assemblage A in 42 specimens and the rest were assemblage B (Madyh et al., 2009). In a study on immunocompromised patients, assemblage A was identified in four of the microcopy-positive *Giardia* specimens using tpi gene (Lim et al., 2011). Assemblage A was also isolated from environmental samples including recreational lake water and water bodies in a zoo (Lim et al., 2009a,b). In addition, genotyping study was conducted on animals and assemblages A and E were detected among goats (Lim et al., 2013). However, subtyping of assemblages A and B in the previous studies was not conducted which limit our understanding on the transmission dynamics and the source of infection of giardiasis in this country. Thus, the present study was aimed to identify *G. duodenalis* assemblages and sub-assemblages based on multiloci genes which included gdh, bg and tpi genes to attain better understanding of the genetic diversity and transmission of giardiasis in Malaysia. The study also aimed at determining the occurrence of mixed infections using primers targeting tpi gene specific for assemblages A and B.

### 2. Materials and methods

#### 2.1. Source of samples

A total of 484 children (51% males, 49% females) with the mean age of 7 years were involved in the present study. The sample collection was conducted from April to September 2011 in 13 Orang Asli villages located in Lipis district, Pahang state, Peninsular Malaysia. The study protocol was approved by The Medical Ethics Committee of University of Malaya Medical Center, Kuala Lumpur, Malaysia under the MEC Ref. No.: 788.74. Informed consents were obtained from the participants’ parents or guardians prior to the collection. Single sample was collected from each participant. Upon receipt, the samples were transferred to the Department of Parasitology, University of Malaya, preserved in 2.5% potassium dichromate and stored in cold room (4°C) until further analysis.

#### 2.2. Microscopy

The faecal samples were concentrated based on formal-ether technique. Briefly, a small amount of stool samples (pea size) was emulsified with 7 ml of 10% formalin and sieved through a double-layered gauze and collected in a beaker. The suspension was transferred into 15 ml conical centrifuge tube and topped up with three ml of diethyl ether. The centrifuge tube with the suspension was capped and mixed by shaking before centrifuging for 5 min at 2500 rpm. The top three layers (ether, debris and froth) were removed and the sediment was examined by light microscope under x100 and x400 magnification for the presence of *Giardia* cysts and other intestinal parasites after staining with iodine.

#### 2.3. Molecular analysis

DNA was extracted directly from samples positive for *Giardia* cyst using PowerSoil™ DNA Isolation Kit (MoBio Laboratories Inc., Carlsbad, California) following the manufacturer’s instructions. Elution step was accomplished by adding reduced volume of solution C6 (10 mM Tris) to obtain a final volume of 50 µl and the DNA was stored in freezer under –20°C.

A partial sequence of gdh gene (≈432-bp) was amplified using semi-nested PCR as described by Read et al. (2004). For primary PCR, the forward primer GDHeF (5’-TCA ACG TTA AYC GYG GYT TCC GT-3’) and the reverse primer GDHir (5’-GTT RTC CTT GCA CAT CTC C-3’) were used. In the secondary PCR, the forward primer GDHIF (5’-CAG TAC AAT TCC CTT GCC GG-3’) and the reverse primer GDHir were used. Primary and secondary PCR reactions were performed in a 50 µl PCR master mix comprising 0.5 µM of each primer (Bioneer Q-Oligos, Korea), 2.5 U of HotStarTaq™ Plus DNA Polymerase (Qiagen, Hilden, Germany), 1× PCR buffer (Qiagen, Hilden, Germany), 200 µM of dNTP (Fermentas, Ontario, Canada), 1.5 mM MgCl2 (Qiagen, Hilden, Germany), 5% dimethyl sulfoxide (Sigma–Aldrich, USA) and 0.4 mg/ml BSA (New England Biolabs, Ipswich, USA). 2 µl of DNA template were used in both amplifications that were run in the MyCycler thermal cycler (Bio-Rad, Hercules, USA) under the following conditions: Initial activation at 95°C for 5 min, 1 cycle at 94°C for 2 min, 56°C for 1 min and 72°C for 2 min, followed by 55 amplification cycles at 94°C for 30 s, 56°C for 20 s, 72°C for 45 s, and a final extension at 72°C for 7 min. For secondary PCR, the number of cycles was reduced to 33.

A partial sequence of bg (≈530-bp) was amplified using nested-PCR protocol according to Sulaiman et al. (2003). Primary PCR was run using forward primer A3543 (5’-AAA TIA TGC CTG CTC TGC G-3’) and reverse primer A3546 (5’-CAA ACC TTI TCC GCA AAC C-3’). For secondary PCR, forward primer A3544 (5’-CCC TTC TTC GGT AAC TTC GGT AAC TT-3’) and reverse primer A3545 (5’-GTG GAC ACC ACI CCC GTG CC-3’) were used. Primary and secondary PCRs were performed in a 50 µl PCR mix comprising 0.2 µM of each primer (Bioneer Q-Oligos, Korea), 1 U of HotStarTaq™ Plus DNA Polymerase (Qiagen, Hilden, Germany), 1× PCR buffer (Qiagen, Hilden, Germany), 200 µM of dNTP (Fermentas, Ontario, Canada), 1.5 mM MgCl2 (Qiagen, Hilden, Germany), and 0.2 mg/ml BSA (New England Biolabs, Ipswich, USA). 2 µl of DNA template were used and the prepared master mix was incubated in the MyCycler thermal cycler (Bio-Rad, Hercules, USA) under the following conditions: Initial hot start at 95°C for 5 min, 35 amplification cycles at 94°C for 45 s, 50°C for 45 s (58°C for secondary PCR), 72°C for 60 s and a final extension at 72°C for 10 min.

In addition, the first PCR product of the reaction described by Sulaiman et al. (2003) underwent further amplification using a set of separate A (Geurden et al., 2007) and B (Geurden et al., 2009) assemblage-specific primers. Presence of mixed infection was detected by visualizing the occurrence of bands in the agarose gel at 332 bp for assemblage A amplified using primers AssA5 (5’-CGT ACA CTT TTC GTC-3’) and AssAR (5’-AGC AAC ATG CAC CTT CTG-3’) and at 400 bp for assemblage B amplified using primers AssB5 (5’-GTT GTG TTC CTT CCC TTT-3’) and AssBR (5’-CTG GCC CAT AGG CAA TTA CA-3’). The PCR reaction mix consisted of 0.2 µM (0.4 µM for assemblage B) of each primer (Bioneer Q-Oligos, Korea), 1.25 U of HotStarTaq™ Plus DNA Polymerase (Qiagen, Hilden, Germany), 1× PCR buffer (Qiagen, Hilden, Germany), 200 µM of dNTP (Fermentas, Ontario, Canada), 1.5 mM MgCl2 (Qiagen, Hilden, Germany) and 0.1 mg/ml BSA (New England Biolabs, Ipswich, USA) to a final volume of 25 µl. 1 µl of DNA template was added for assemblage A and 2 µl was added for assemblage B for the PCR amplifications following the cycle parameter: Initial hot start at 95°C for 5 min, initial denaturation at 94°C for 10 min, and 35 amplification cycles at 94°C for 45 s, 50 s for 45 s (58°C for secondary PCR), 72°C for 60 s and a final extension at 72°C for 10 min.

A partial sequence of tpi gene (≈511-bp) was amplified using PCR protocols described by Caccio et al. (2002) and Lalle et al. (2005). The primers for primary PCR were G7 (5’-AAG GCC GAC ACC CTC ACC CGC AGT GC-3’) and G7S5 (5’-GAG GCC GCC CTG GAT GTT CGG GAC GAC-3’). For secondary PCR, BGS11F (5’-GAA CGA AGG AGA TCT CAG TGG TCC G-3’) and BGS11R (5’-CTC GAC GAG CT TCG TGT T-3’) were used. The PCR master mix consisted of...
0.4 μM of each primer (Bioneer Q-Oligos, Korea), 2.5 U HotStarTaq® Plus DNA Polymerase (Qiagen, Hilden, Germany), 1× PCR buffer (Qiagen, Hilden, Germany) and 200 μM dNTP (Fermentas, Ontario, Canada) in a total volume of 50 μl. 2 μl of DNA template were added and the mixture was run in the MyCycler thermal cycler (Bio-Rad, Hercules, USA) under the following conditions: Initial hot start at 95 °C for 5 min, initial denaturation at 95 °C for 15 min, followed by 35 amplification cycles at 95 °C for 30 s, 65 °C for 30 s (55 °C for secondary PCR), 72 °C for 60 s, and a final extension at 72 °C for 7 min.

In all the PCR reactions, a Giardia-positive DNA specimen and distilled water were used as positive and negative control. The PCR products were analyzed using 2% agarose gel (Vivantis) electrophoresis stained with SYBR® Safe DNA (Invitrogen, Auckland, New Zealand).

### Table 1
The distribution of assemblages A and B based on the different loci and mixed infection.

<table>
<thead>
<tr>
<th>Genes</th>
<th>A</th>
<th>B</th>
<th>A + B</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>tpi</td>
<td>34 (58)</td>
<td>25 (42)</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>gdh</td>
<td>7 (18)</td>
<td>31 (82)</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>b-gardin</td>
<td>18 (64)</td>
<td>10 (36)</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>MLGb</td>
<td>3 (27)</td>
<td>8 (73)</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>tpi-Mixed</td>
<td>1 (2)</td>
<td>23 (34)</td>
<td>43 (64)</td>
<td>67</td>
</tr>
</tbody>
</table>

### Table 2
The distribution of assemblages A and B based on the different loci and mixed infection.

<table>
<thead>
<tr>
<th>Isolate</th>
<th>tpi</th>
<th>gdh</th>
<th>bg</th>
<th>tpi-Mixed</th>
<th>Isolate</th>
<th>tpi</th>
<th>gdh</th>
<th>bg</th>
<th>tpi-Mixed</th>
</tr>
</thead>
<tbody>
<tr>
<td>TG2.5</td>
<td>A (A2)</td>
<td>A (A2)</td>
<td>A (A3)</td>
<td>B</td>
<td>KM13.2</td>
<td>B</td>
<td>B</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>TB13.3</td>
<td>A (A2)</td>
<td>A (A2)</td>
<td>–</td>
<td>A + B</td>
<td>CK7.1</td>
<td>B</td>
<td>B</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>SI14.1</td>
<td>A (A2)</td>
<td>A (A2)</td>
<td>–</td>
<td>–</td>
<td>SE13.1</td>
<td>B</td>
<td>B</td>
<td>–</td>
<td>B</td>
</tr>
<tr>
<td>ST15.1</td>
<td>A (A2)</td>
<td>A (A2)</td>
<td>–</td>
<td>A + B</td>
<td>TB15.1</td>
<td>B</td>
<td>B</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>KK16.4</td>
<td>A (A2)</td>
<td>–</td>
<td>A (A2)</td>
<td>–</td>
<td>SE9.3</td>
<td>B</td>
<td>B</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>KK16.5</td>
<td>A (A2)</td>
<td>–</td>
<td>A (A2)</td>
<td>A + B</td>
<td>KN6.4</td>
<td>B</td>
<td>–</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>TB7.3</td>
<td>A (A2)</td>
<td>A (A2)</td>
<td>A (A2)</td>
<td>A + B</td>
<td>TB9.1</td>
<td>B</td>
<td>–</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>SE10.1</td>
<td>A (A2)</td>
<td>–</td>
<td>A (A3)</td>
<td>B</td>
<td>UM3.1</td>
<td>B</td>
<td>–</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>SE17.1</td>
<td>A (A2)</td>
<td>–</td>
<td>A (A2)</td>
<td>A</td>
<td>SE4.8</td>
<td>B</td>
<td>–</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>TB15.3</td>
<td>A (A2)</td>
<td>–</td>
<td>A (A2)</td>
<td>A + B</td>
<td>SE27.6</td>
<td>B</td>
<td>–</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>TC5.2</td>
<td>A (A2)</td>
<td>–</td>
<td>A (A2)</td>
<td>A + B</td>
<td>SE27.7</td>
<td>B</td>
<td>–</td>
<td>–</td>
<td>B</td>
</tr>
<tr>
<td>PD1.1</td>
<td>A (A2)</td>
<td>–</td>
<td>A (A3)</td>
<td>A + B</td>
<td>KK2.3</td>
<td>B</td>
<td>–</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>PD11.2</td>
<td>A (A2)</td>
<td>–</td>
<td>A (A2)</td>
<td>A + B</td>
<td>KK2.5</td>
<td>B</td>
<td>–</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>SE29.3</td>
<td>A (A2)</td>
<td>–</td>
<td>A (A2)</td>
<td>–</td>
<td>KM7.3</td>
<td>–</td>
<td>B</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>KK11.3</td>
<td>A (A2)</td>
<td>–</td>
<td>A (A2)</td>
<td>A + B</td>
<td>KM6.1</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>B</td>
</tr>
<tr>
<td>KM4.1</td>
<td>A (A2)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>KM15.2</td>
<td>–</td>
<td>B</td>
<td>–</td>
<td>B</td>
</tr>
<tr>
<td>TB3.3</td>
<td>A (A2)</td>
<td>–</td>
<td>–</td>
<td>A + B</td>
<td>ST13.2</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>B</td>
</tr>
<tr>
<td>TB13.2</td>
<td>A (A2)</td>
<td>–</td>
<td>–</td>
<td>A + B</td>
<td>ST20.3</td>
<td>–</td>
<td>–</td>
<td>B</td>
<td>–</td>
</tr>
<tr>
<td>SE23.2</td>
<td>A (A2)</td>
<td>–</td>
<td>–</td>
<td>A + B</td>
<td>KM6.4</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>B</td>
</tr>
<tr>
<td>KK16.5</td>
<td>A (A2)</td>
<td>–</td>
<td>–</td>
<td>A + B</td>
<td>CK1.4</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>CK1.2</td>
<td>A (A2)</td>
<td>–</td>
<td>–</td>
<td>B</td>
<td>SB4.1</td>
<td>–</td>
<td>B</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>CK10.2</td>
<td>A (A2)</td>
<td>–</td>
<td>–</td>
<td>A + B</td>
<td>ST8.1</td>
<td>A (A2)</td>
<td>B</td>
<td>A (A2)</td>
<td>B</td>
</tr>
<tr>
<td>SB1.1</td>
<td>A (A2)</td>
<td>–</td>
<td>–</td>
<td>A + B</td>
<td>KN7.2</td>
<td>A (A2)</td>
<td>B</td>
<td>–</td>
<td>B</td>
</tr>
<tr>
<td>SE25.4</td>
<td>A</td>
<td>–</td>
<td>–</td>
<td>A + B</td>
<td>SM6.2</td>
<td>A (A2)</td>
<td>B</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>ST10.3</td>
<td>–</td>
<td>–</td>
<td>A (A2)</td>
<td>–</td>
<td>KM3.2</td>
<td>A (A2)</td>
<td>B</td>
<td>–</td>
<td>B</td>
</tr>
<tr>
<td>TB12.3</td>
<td>–</td>
<td>–</td>
<td>A (A2)</td>
<td>–</td>
<td>KM12.3</td>
<td>A (A2)</td>
<td>B</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>KK7.5</td>
<td>–</td>
<td>–</td>
<td>A (A3)</td>
<td>–</td>
<td>TB11.2</td>
<td>A (A2)</td>
<td>B</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>SE29.4</td>
<td>–</td>
<td>–</td>
<td>A (A2)</td>
<td>–</td>
<td>SB15.1</td>
<td>A (A2)</td>
<td>B</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>KM15.1</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>PD20.1</td>
<td>B</td>
<td>B</td>
<td>A (A3)</td>
<td>A + B</td>
</tr>
<tr>
<td>TB2.3</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>SE18.4</td>
<td>A</td>
<td>B</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>TB4.1</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>A + B</td>
<td>KK6.1</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>TB5.1</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>A + B</td>
<td>CK20.1</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>A + B</td>
</tr>
<tr>
<td>CK20.2</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>A + B</td>
<td>ST1.3</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>B</td>
</tr>
<tr>
<td>SE7.1</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>A + B</td>
<td>ST3.2</td>
<td>–</td>
<td>–</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>SE12.1</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>ST21.1</td>
<td>–</td>
<td>–</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>SE24.2</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>TB2.5</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>B</td>
</tr>
<tr>
<td>KN7.4</td>
<td>B</td>
<td>B</td>
<td>–</td>
<td>A + B</td>
<td>SE13.2</td>
<td>–</td>
<td>–</td>
<td>B</td>
<td>B</td>
</tr>
</tbody>
</table>

### 2.4. DNA sequencing and sequences analysis

Successfully amplified PCR products with the nested-PCR for each locus (except for tpi assemblage-specific PCR) were sequenced in both directions using Applied Biosystems 3730 x1 DNA Analyzer (Applied Biosystems, USA). The chromatograms and sequences generated from this study were viewed and assembled using BioEdit Sequence Alignment Editor Programme (http://www.mbio.ncsu.edu). Preliminary similarity comparison of the consensus sequence with the sequences in GenBank database was made using Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov). The isolate sequences were genotyped into assemblage and sub-assemblage using multiple alignments implemented by ClustalW (Thompson et al., 1994) with previously defined reference sequences retrieved from GenBank database. Phylogenetic analysis was performed in MEGA 5 (www.megasoftware.net) using neighbour-joining (NJ) algorithms with evolutionary distances calculated by Kimura-2 parameter method (Kimura, 1980) and 1000 bootstrap value. Sequences from this study were deposited in the GenBank under accession numbers KC313907-KC313948.

### 3. Results

#### 3.1. Identification of G. duodenalis assemblages and mixed assemblages

From a total of 484 participants, 84 (17.0%) were found infected with Giardia using microscopy. The 84 microscopy-pos-
tive specimens were analyzed at three loci (tpi, gdh, and bg) and a set of assemblage-specific PCRs was used to detect mixed infection. PCR amplicons from 78 specimens using at least one of the molecular markers were amplified and subsequently sequences were generated for 71 specimens. The alignment analysis of the sequences obtained from tpi, gdh, and bg genes with reference sequences from the GenBank database identified assemblages A and B in 30 (42%) and 32 (45%) specimens, respectively. Discordant genotype results were noted in 9 (13%) specimens. At tpi gene, PCR amplicons were produced from 70% of the examined specimens (59/84). Among them, assemblages A and B were found in 34 (58%) and 25 (42%) specimens, respectively. At gdh gene, 38 (45%) specimens were positive with assemblage B being more frequently observed compared to assemblage A (31 vs 7). PCR analysis of bg gene had the lowest detection rate [33% (28/84)]. The distribution of assemblage A and B for bg gene was 18 (64%) and 10 (36%), respectively. MLGs were successful in 11 specimens, which were identified as Assemblages A and B, in 3 and 8 specimens, respectively (Table 1). As mentioned above assemblage discordant results were observed in 9 specimens. One isolates (ST8.1) was identified as assemblage A at tpi and bg genes and as assemblage B at gdh gene. The specimen PD20.1 was identified as assemblage B at tpi and gdh genes and as assemblage A at bg gene. Seven specimens (KN7.2, SM6.2, KM3.2, KM12.3, TB11.2, SB15.1 and SE18.4) were found to harbour assemblage A at tpi gene and assemblage B at gdh gene. The tpi-mixed protocol showed that 6 of the 9 specimens harboured mixed assemblages A and B (Tables 2).

The tpi-based PCR protocol for mixed infection using assemblage-specific primers was conducted on the 84 Giardia-positive specimens. Of these, 67 specimens showed amplicons for G. duodenalis assemblages; 43 specimens had mixed infections with assemblages A and B, 23 specimens had single infection with assemblage B and one case was identified as assemblage A (Table 1 and 2). The specificity of this protocol was confirmed by sequencing selected cases representing assemblages A and B. When the chromatogram was scrutinized at the positions where nucleotide polymorphism could differentiate assemblage A and B, double peaks were observed in 14 of the mixed isolates.

3.2. Subtyping of G. duodenalis Assemblage A

Based on analysis targeting tpi gene, 34 of the assemblage A isolates (Table 2) were classified as subtype A2 based on the phylogenetic analysis (Fig. 1) and the substitution pattern (Table 3). The neighbor-Joining tree placed the four representative sequences (TB11.2, KK16.4, TB13.2 and SM6.2) in one cluster with All sequence references with high bootstrap support.

At gdh gene, all sequences of the assemblage A showed complete homologous to the reference sequence of subtype A2 (Accession No. L40510) except for isolate ST10.3 where substitution at two positions were observed (A–G at position 341 and T–G at position 485) (Table 3). Subtype A2 based on tpi or gdh belongs to sub-assemblage All (Caccio et al., 2008; Feng and Xiao, 2011).

At bg gene, 18 isolates were identified as assemblage A. Twelve were distinguished as A2 from A1 at position 606 (C–T) and from A3 at positions 460 (C–T) and 468 (T–C). Three were typed as A3 and the remaining 3 isolates have the subtype similar to A3 but different by one position at 460 (C–T). Phylogenetic analysis demonstrated that six isolates were typed as A3 (100% bootstrap value) while the remaining 12 isolates were grouped as A2 (Fig. 3). It should also be noted that subtypes A2 and A3 belong to sub-assemblage All (Caccio et al., 2008; Feng and Xiao, 2011).

3.3. Subtyping of G. duodenalis assemblage B

At tpi gene, 16 sequences representing 23 isolates which were identified as assemblage B, were multiple aligned with 5 references sequence representing BIV, BIV-like, BIII and BIII-like and analysed for BIV/BIII specific substitutions according to Wielinga and Thompson (2007). The 16 representative sequences had heterogeneous nucleotides at positions defining the sub-assemblies, making proposing specific sub-assemblies for assemblage B isolates not possible (Table 3). Phylogenetic analysis confirmed the mono-phyletic group of assemblage B (bootstrap = 99%). The Neighbor-Joining formed tree sub-clusters of assemblage B. However, sub-clustering was not supported by bootstrap values (Fig. 1).

At gdh gene, 15 sequences representing 30 isolates had high nucleotide variations, limiting the classification of assemblage B to sub-assemblies (Table 4). Similar to tpi gene, phylogenetic
4. Discussion

To the best of our knowledge, the present study was the first study of giardiasis using multi-locus genotyping (MLG) in Malaysia. In this study, the prevalence of giardiasis in the Orang Asli communities was 17.0%. This is comparable with a recent study conducted by Anuar et al. (2012) in Orang Asli communities in Selangor state, which showed an overall prevalence of giardiasis at 20.0% (n = 500).

Amplifications of the three PCR assays (tpi, gdh, and bg) were performed differently at different loci. Of the 84 Giardia microscopically-positive specimens, only 11 isolates were able to be amplified at three loci. The tpi locus achieved the highest percentage of amplicons produced (70%), followed by gdh (45%) and bg (33%). Similar occurrences were also reported in previous studies conducted by Anuar et al. (2012) in Orang Asli communities in Selangor state, which showed an overall prevalence of giardiasis at 20.0% (n = 500).

Table 3

<table>
<thead>
<tr>
<th>Isolates</th>
<th>GenBank accession no.</th>
<th>Nucleotide position from the start of the gene</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Assemblage A</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AI</td>
<td>L02120</td>
<td>T C</td>
</tr>
<tr>
<td>All (A2) JH</td>
<td>U57897</td>
<td>C T</td>
</tr>
<tr>
<td>All (A2) TB11.6</td>
<td>KC313923</td>
<td>C T</td>
</tr>
<tr>
<td><strong>Assemblage B</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIV Ad 19</td>
<td>AF069560</td>
<td>A T T C G G G T A C A A A G A A T A</td>
</tr>
<tr>
<td>BIV G5/M</td>
<td>L02116</td>
<td>... ... ... ... ... ... ... ... ... ... G</td>
</tr>
<tr>
<td>BIV-Like 7115</td>
<td>AY368167</td>
<td>... ... ... ... ... ... ... ... ... ... G</td>
</tr>
<tr>
<td>BIII 2434</td>
<td>AY368165</td>
<td>G C C C C G ... ... ... ... ... ... ... ... G A</td>
</tr>
<tr>
<td>BIII-Like 2436</td>
<td>AY368153</td>
<td>G C C T ... ... ... ... ... ... ... ... ... ... G</td>
</tr>
</tbody>
</table>

Numbers in bold represent nucleotide substitutions from the start of the gene, which differentiate between subtypes according to Weilinga and Thompson (2007). Dots (.) indicate nucleotides identity. Nucleotides in small letter indicate substitution not in the subtypes-defining positions.

Table 4

<table>
<thead>
<tr>
<th>Isolates</th>
<th>GenBank accession no.</th>
<th>Nucleotide position from the start of the gene</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Assemblage A</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AI (Prtland1)</td>
<td>M84604</td>
<td>C C A A C T C</td>
</tr>
<tr>
<td>ST10.3</td>
<td>KC313925</td>
<td>? ? G T C T</td>
</tr>
<tr>
<td><strong>Assemblage B</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIV (Ad-7) (BAH-12)</td>
<td>AF069059</td>
<td>C T C T C C C G C G</td>
</tr>
<tr>
<td>SM1.2</td>
<td>KC313926</td>
<td>T ... ... C T ... ... ... ... ... ... ... ... ... A</td>
</tr>
</tbody>
</table>
| CK1.4 | KC313927 | T ... ... T ... ... ... ... ... ... ... ... ... ... ...
| KN7.3, TB11.2, KM8.1, SM6.2 | KC313928 | T ... ... C C T ... ... ... ... ... ... ... ... ... ... ...
| KM15.2, ST18.1, KM15.1, ST13.2, KM8.4, ST20.3, TB2.3 | KC313929 | T ... ... C C ... ... ... ... ... ... ... ... ...
| SE13.1 | KC313930 | T C ... ... T ... ... ... ... ... ... ... ... ...
| KM3.2 | KC313931 | T C ... ... C T ... ... ... ... ... ... ... ... ...
| TB15.1, SB15.1 | KC313932 | T C ... ... C C ... ... ... ... ... ... ... ...
| SE12.1, TB4.1 | KC313933 | T C ... ... C C T ... ... ... ... ... ... ...
| TB5.1, SE17.1 | KC313934 | T C ... ... C C T ... ... ... ... ... ... ...
| KM13.2, CK7.1, TB9.2, SE24.2, SE18.4 | KC313935 | T C ... ... C C T ... ... ... ... ... ... ...
| PD20.1, KN7.4 | KC313936 | ... ... C C T ... ... ... ... ... ... ...
| CK2.0 | KC313937 | ... ... C C ... ... ... ... ... ... ...
| KM12.3 | KC313938 | ... ... C C ... ... ... ... ... ... ...

Numbers in bold represent nucleotide substitutions from the start of the gene, which differentiate between subtypes according to Weilinga and Thompson (2007). Dots (.) indicate nucleotides identity. Question mark (?) indicates missing nucleotides. Accession numbers in bold are reference sequences from the GenBank.
Discordant genotyping results at the three loci (tpi, gdh, and bg) have been observed in 9 isolates (13%) in this study, which is in agreement with previous reports (Caccio et al., 2008; Sprong et al., 2009; Thompson et al., 2009; Yang et al., 2010a). An analysis of sequences from four genetic loci (tpi, gdh, bg and rDNA) has been conducted on 61 human isolates and 29 animal isolates of G. duodenalis (Caccio et al., 2008). In this study, incongruent assignment of five human isolates and one macaque isolates to G. duodenalis assemblages was reported (Caccio et al., 2008). Yang et al. (2010) genotyped 124 human isolates of G. duodenalis at gdh and rDNA genes, and reported incongruent genotyping results in five isolates, where three isolates were classified as assemblage A at rDNA gene and assemblage B at gdh gene, and two isolates were classified as assemblage B at rDNA gene and assemblage A at gdh gene. Inconsistency in genotyping results of G. duodenalis has also been reported among dogs and cats isolates, where 7 isolates (3 dogs and 4 cats) were classified as assemblage D at rDNA gene and assemblage B at gdh gene. One dog isolate was classified as assemblage C at rDNA and assemblage B at gdh gene (Read et al., 2004). Thompson et al. (2009) genotyped G. duodenalis isolated from 70 coyotes at rDNA and gdh, and reported four incongruent genotyping results.

The implication of this inconsistency has provoked questions about the strength and validity of genotyping relying on single locus and hence, the true picture of Giardia diversity has a high chance of being masked by the data generated from analysis using single marker (Feng and Xiao, 2011). Although the underlying mechanisms remain uncertain, Caccio and Ryan (2008) have reviewed a number of factors that could contribute to the incongruent assignment of assemblages by different makers. These factors include meiotic recombination, which suggests the potential of sexual reproduction in Giardia, introgression that involves backcrossing of hybrids with parental species and retention of ancestral polymorphism that could lead to presence of identical alleles in genetically distinct groups. Another contributing factor which would be discussed further in this context is the occurrence of mixed infections. With the use of the assemblage-specific assays, high prevalence (64%) of mixed infections was detected and further proven the biasness of using a standard PCR assay alone as the most plentiful assemblage would be amplified in preference (Wielinga and Thompson, 2007). Thus, the feature of mixed infections could be applied in the present study to explain the discordant genotype results.

The isolates with assemblage A were all identified as A2, which belongs to sub-assemblage All. This is similar with other studies

**Fig. 2.** Phylogram of G. duodenalis genotypes constructed by neighbour-joining analysis, based on the nucleotide sequences of gdh retrieved from this study compared with reference sequences of known assemblages from Genbank. Bootstrap values obtained from 1000 replicates are indicated on branches in percentage.

**Fig. 3.** Phylogram of G. duodenalis genotypes constructed by neighbour-joining analysis, based on the nucleotide sequences of bg retrieved from this study compared with reference sequences of known assemblages from Genbank. Bootstrap values obtained from 1000 replicates are indicated on branches in percentage.
where all was the predominant sub-assemble in human (Bonhomme et al., 2011; Hussein et al., 2009; Lebbad et al., 2011). Sub-assemble A1 which was absent from this study was reported to have preference to cause infection in livestock or companion animals as opposed to all which was found mainly in human cases than animals (Sprong et al., 2009). The isolation of anthropogenic assemblages and sub-assemblages (B and AII) in this study implicates human as a potential source of the infection. In accordance with our findings, previous studies conducted in Orang Asli communities in Malaysia found that the presence of other family members infected with G. duodenalis was either the only or the main risk factor for giardiasis (Norhayati et al., 1998; Anuar et al., 2012). However, the postulation of human-to-human transmission is limited by the fact that animals were not included for analysis in this study.

On the other hand, sub-assemblages were not able to assign to isolates belonged to assemblage B due to high degree of nucleotides variation which was not uncommon among other molecular studies (Bonhomme et al., 2011; Caccio et al., 2008; Lalle et al., 2009; Lebbad et al., 2011; Levecke et al., 2009). This phenomenon led us back to the focus of mixed infection, which is known to take place at both the inter- and at the intra-assemble levels (Caccio and Ryan, 2008). Mixed infection can happen when a host ingests Giardia cysts of different genetic profiles or subsequent infection of an infected host by genetically different Giardia cysts. This is especially common in areas where giardiasis is endemic (Caccio and Ryan, 2008; Lebbad et al., 2011; Sprong et al., 2009). Therefore, mixed subtype infections (intra-assemble level) can stand as a factor that causes the high heterogeneity. Besides, high polymorphism in assemblage B can also be attributed to intra-assemble recombination and allelic sequence heterozygosity (ASH). In order to resolve the question whether the mixed genotype is due to mixed infections or ASH, Ankarklev et al. (2012) had conducted an investigation on single Giardia trophozoite and cyst. The result based on molecular analysis positively exhibited the occurrence of ASH at single cell level. However, high numbers of mixed sub-genotypes infections (demonstrated by variable sequence patterns) were also observed in different cysts isolated from the same sample indicating that mixed infection is also playing a role for the high heterogeneity.

In conclusion, MLG was conducted to analyze the genetic diversity of Giardia infections in Orang Asli communities in Malaysia. While assemblages A and B have almost equal frequency of infections, more than half of the isolates were mixed infections. The presence of assemblage B and sub-assemble AII in the samples suggest that the mode of transmission of giardiasis in Malaysia may be human-to-human. Nevertheless, the role of animals in the dynamic of transmission needs to be further investigated involving multilocus genotyping of parasites from animals and humans.

Acknowledgments

We are grateful to the Orang Asli children and their parents for their voluntary involvements and would like to acknowledge the financial support from UMRG grants RG302-11HTM and RG439/12HTM, and the student grant (PV063–2012A).

References


Table 5

<table>
<thead>
<tr>
<th>Assemblage A</th>
<th>Genbank accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>X85958</td>
</tr>
<tr>
<td>AII</td>
<td>AY072723</td>
</tr>
<tr>
<td>A3</td>
<td>KY75, TG2.5</td>
</tr>
<tr>
<td>A3</td>
<td>SE10.1</td>
</tr>
<tr>
<td>A3</td>
<td>PD20.1, PD1.1, TG13.2</td>
</tr>
<tr>
<td>B</td>
<td>AY072727</td>
</tr>
<tr>
<td>SB4.1</td>
<td>KC13939</td>
</tr>
<tr>
<td>TB2.1</td>
<td>KC13940</td>
</tr>
<tr>
<td>SE5.3</td>
<td>KC13941</td>
</tr>
<tr>
<td>TB4.1</td>
<td>KC13942</td>
</tr>
<tr>
<td>SE24.2</td>
<td>KC13943</td>
</tr>
<tr>
<td>TB5.1</td>
<td>KC13944</td>
</tr>
</tbody>
</table>

Table 5: Multiple alignments of bg sequences from this study with reference sequences obtained from GenBank, representing sub-assemblages of A and assemblages B.

Numbers in bold represent nucleotide substitutions from the start of the gene, which differentiate between subtypes according to Wielinga and Thompson (2007). Dots indicate nucleotides identity. Question mark (?) indicates missing nucleotides. The dash (–) indicates deletion.

A. Representative sequence of isolates identified as A2. Accession numbers in bold are reference sequences from the GenBank.


