Assemblage-Dependent Genetic Features of *Giardia intestinalis* in Humans

Amjad Hussein¹,², Kentaro Nakamoto³, Tomoko Arai³ & Masaharu Tokoro*³

¹Department of Medical Laboratory Sciences, Faculty of Medicine and Health Science, An-Najah National University, Nablus, Palestine. ²Analysis, Poison Control and Calibration Center, An-Najah National University, Nablus, Palestine. ³Department of Parasitology, Graduate School of Medical Science, Kanazawa University, 13-1 Takaramachi, Kanazawa 920-8640, Japan.

*Corresponding author: tokoro@med.kanazawa-u.ac.jp

Received: (28/9/2016), Accepted: (7/11/2016)

ABSTRACT

Within a significant genetic diversity consisting of 7 genotypes of *Giardia intestinalis*; assemblage A and B genotypes have been suggested to have different host specificities and variable pathogenicity in humans. Assemblages are highly variable with regard to the genetic features of each genotype. A total of 89 fecal samples were collected from patients who sought medical consultation for abdominal pain in the West Bank, Palestine. *Giardia* positive samples were assessed microscopically and where genotyped to assemblage level by direct sequencing and subcloning for both triosephosphate isomerase and β-giardin gene markers. *G. intestinalis* were detected in 12 samples; 9 samples were belonging to assemblage A and 3 samples were belonging to assemblage B. A single genotype was identified from all assemblage A samples, whereas 2 to 4 subgenotypes were confirmed in each of assemblage B samples. No mixed genotypes (inter-assemblages) were found in any sample. This study revealed the contrasting genetic diversity that supports the separate-species notion for assemblage A and B genotypes; it also highlights the need to investigate assemblage B genotype more thoroughly, as has been the case for assemblage A genotype.

Keywords: *Giardia Intestinalis*, Genotyping, Triosephosphate Isomerase, B-Giardin, Palestine.

INTRODUCTION

*Giardia intestinalis* (syn. *G. duodenalis* and *G. lamblia*) is the most common intestinal protozoan parasite in a wide range of vertebrates, including humans (1, 2). While *G. intestinalis* isolates from various hosts are morphologically indistinguishable, they show significant genetic diversity (2-7, 9), resulting in the naming of 7 (assemblages A–G) main genotypes (5, 10). Assemblages A and B, the major genotypes detected in humans, are capable of infecting a wide range of mammalian hosts, while the infectivity of other assemblages (C–G) appears to be restricted to particular hosts (5, 10, 11).

Recently, separate species names—“*G. intestinalis*” for assemblage A genotype and “*G. enterica*” for assemblage B genotype—have been proposed on the basis of the overall variation in *Giardia* (5, 10, 11). The considerable phenotypic differences between assemblages A and B, such as in vitro growth differ-
ferences or culture adaptations (12), metabolisms (13), and clinical manifestations (14, 15) are well known, and different levels of genotype diversity in assemblages A and B have also been reported from endemic areas (4, 16-18). However, their detection in various hosts organisms and their showing wide range of pathogenicity from asymptomatic carrier status to severe chronic diarrhea, in humans (5) has led to controversy, which may in turn have delayed the taxonomic reclassification of *G. intestinalis* genotypes. Moreover, there is a gap in the available data from the two genotypes. The genome project and some studies revealed the presence of sexual reproduction or an exchange of genetic material in this pathogen (19-21), but the extensive genetic conformation in assemblage B remains relatively unexplored.

Assemblage A and B genotypes exhibit variable genetic diversity depending on the studied gene marker, and each genotype contains subgenotypes with considerably lesser extent of polymorphism (1). The level of gene-marker resolution and discrepancy power is not completely consistent within the different genotypes (17), which implies that different gene markers can be selected and used with different genotypes. For example, the triosephosphate isomerase (TPI) gene marker achieves the highest resolution between assemblage A and B genotypes (1), but not between the other genotypes, such as the assemblage C and D genotypes (5).

This study aimed to investigate *G. intestinalis* genotypes in Palestine using TPI and β-giardin genes, correlate them to the reported genotypes all over the world, explore the genetic features of the genotypes from human sources and evaluate the notion of different species names for assemblages A and B.

**MATERIALS AND METHODS**

**Fecal samples and microscopic examination**

A total of 89 fecal samples were collected from patients who sought medical consultation for abdominal pain in the West Bank, Palestine, in 2006. Of these, 69 samples had been used in a previous study (16). All samples were preserved in 2.5% (w/v) potassium dichromate at 4°C and subsequently processed for cyst purification using the sucrose centrifugal flotation method, as previously described (16). Purified cysts were stored at –20°C in phosphate buffer saline (PBS; pH 7.2) until further analysis. *Giardia*-positive samples were assessed microscopically by using a drop of the sucrose surface layer. Each clinical sample was collected after obtaining informed consent from the patients, and each sample was on-site labeled with a unique identification number to protect the patients’ personal information.

**DNA preparation and PCR amplification**

DNA was extracted mainly from the cysts in PBS solution using the QIAamp DNA Mini Kit (Qiagen, Tokyo, Japan) as described previously (16). The original fecal samples and the QuickGene DNA tissue kit S (Fujifilm Corporation, Tokyo, Japan) were also used, whenever needed, according to the manufacturer’s instructions, after performing three cycles of freezing at –80°C and thawing at 95°C. The extracted genomic DNA was concentrated by ethanol precipitation and preserved at –20°C until use as a DNA template aqueous solution. PCR amplifications were carried out in a 20-µl reaction mixture containing a 1–5 μl DNA template solution, 0.1 mM each of deoxynucleoside triphosphate, 1x PrimeSTAR buffer, 0.25 μM of primers and 0.25 U of
PrimeSTAR HS DNA Polymerase (Takara Bio Inc., Shiga, Japan). PCR was performed using the previously described primers (Table 1) of TPI (2, 5) and β-giardin (3) genes. New primer sets (Table 1) were designed according to the preliminary obtained sequence data and the published conserved regions of TPI (specific to assemblage A and B genotypes, mainly assemblage B genotypes) and β-giardin nucleotide sequences; this was done to cover a wider range of newly reported genotypes and increase the likelihood of amplifying the expected genotypes in each sample. Whenever needed, different primer combinations (Table 1) were used in nested or seminested PCR to amplify TPI and β-giardin genes. The following basic cycling parameters were performed: initial denaturation at 98°C for 1 min, followed by 30–45 cycles of 98°C for 30 s, 55–65°C for 5 s, 72°C for 1 min and a final extension at 72°C for 5 min. The reaction products were subjected to electrophoresis on a 2.0% agarose gel, stained with ethidium bromide and visualized under UV light. PCR fragments were then purified from the agarose gel by using Quantum Prep Freeze N Squeeze DNA Gel Extraction Spin Columns (BioRad Laboratories, CA, USA), according to the manufacturer’s instructions.

**Table 1:** TPI and β-giardin oligonucleotides used in this study.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5′ to 3′)</th>
<th>Target (bp)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPI16F</td>
<td>CCCTTCATCGGGYGGTAAC CACTGGCCAAGTTYTTCRCA</td>
<td>668</td>
<td>5</td>
</tr>
<tr>
<td>TPI16R</td>
<td>AAATYATGCCTGCTCGTGC CAAACCTTITCCGCAAACC</td>
<td>610</td>
<td>2</td>
</tr>
<tr>
<td>AL3543</td>
<td>TGCTCGTCGYCCCTTTCATCG TCTCGCAGTTRCTYCCATTGCC</td>
<td>665</td>
<td>This study*</td>
</tr>
<tr>
<td>AL3546</td>
<td>GAGGCCGCCCTGGATCTTCGAGACGAC</td>
<td>753</td>
<td>3</td>
</tr>
<tr>
<td>BGPalF1</td>
<td>CCGCAGTGCGACGYGAGAC TGGATCTTCGAGACGTCC</td>
<td>726</td>
<td>This study*</td>
</tr>
<tr>
<td>BGPalR1</td>
<td>GAGCCCGACGACCTCACCCCGCAGTGC GAGGCCGCCCTGGATCTTCGAGACGAC</td>
<td>753</td>
<td>3</td>
</tr>
</tbody>
</table>

*For data analysis, 623-bp TPI and 687-bp β-giardin gene sequences were used.

**DNA sequencing and subcloning**

All purified PCR products were directly sequenced with appropriate primers in both directions using the BigDye terminator cycle sequencing ready reaction kit (Applied Biosystems, CA, USA) on an ABI PRISM 310 Genetic Analyzer as described previously (16). The purified PCR fragments of three samples—PalH4, PalH6, and PalH9—were cloned into the EcoRV-digested pBluescript II SK(+) plasmid vector (Stratagene, CA, USA), using the blunting kination ligation kit (Takara Bio Inc, Shiga, Japan). The recombinant plasmids were transformed into Escherichia coli DH5α (Stratagene) and screened on Luria Broth (LB) agar plates supplemented with 100 mg/L of ampicillin. The clones were picked up as E. coli DH5α colonies on the plate and cultured overnight in the LB supplemented with 100 mg/L of ampicillin; they were then subjected to plasmid purification using the Qiagen Plasmid
Mini Kit. The full-length sequences were determined using T3, T7, and appropriate internal sequencing primers. The clones were confirmed by at least two independent PCR products amplified using the same or different primer sets. Therefore, unpaired sequences with some single nucleotide polymorphisms (SNPs) from a single PCR were excluded, since they could include PCR-generated artifacts.

**Computer analysis of the sequenced data**

Homology search for the resultant sequences were carried out using the BLAST v2.2.18 program available at the DNA Data Bank of Japan (DDBJ) homepage (http://blast.ddbj.nig.ac.jp/top-e.html). Alignment analyses were performed using DNASIS-Mac v3.6 (Hitachi Software Engineering, Yokohama, Japan) and the ClustalW v1.83 program available at the DDBJ homepage (http://clustalw.ddbj.nig.ac.jp/top-e.html).

**Nucleotide sequence accession numbers**

The new nucleotide sequences have been assigned the following DDBJ/EMBL/GenBank accession numbers: AB480868, AB480869, AB480870, AB480871, AB480872, AB480873, and AB480874 for TPI gene and AB480875, AB480876, and AB480877 for β-giardin gene.

**RESULTS**

Out of the 89 samples, *G. intestinalis* were detected in 12 samples that were genotyped using the direct sequencing and the subcloning procedure of the TPI and β-giardin gene markers. Assemblage A were detected in 9 samples but the detected assemblage B were more complicated in their electropherograms analysis. Therefore, subcloning were used for the samples of the interesting mixed electropherograms profiles using the direct sequencing that were detected in 3 samples only, PalH4, PalH8, and PalH9. Interestingly, all of the mixed profiles were belonging to assemblage B genotype, but the remaining 9 samples belonged to the assemblage A-II genotype. None of the samples showed mixed electropherograms of A and B.

<table>
<thead>
<tr>
<th>Subgenotype</th>
<th>Clone frequency</th>
<th>TPI nucleotide position*</th>
</tr>
</thead>
<tbody>
<tr>
<td>L02116</td>
<td></td>
<td>581 633 704 707 839 846 926 980 1079 1106 1156 1172 1175</td>
</tr>
<tr>
<td>PalH4-clone 1</td>
<td>17/44</td>
<td>A T G T A A C T G A G C C</td>
</tr>
<tr>
<td>-clone 2</td>
<td>8/44</td>
<td>G T G T A A C T G G G C C</td>
</tr>
<tr>
<td>-clone 3</td>
<td>9/44</td>
<td>G C A C A C A A C C G G G C C</td>
</tr>
<tr>
<td>-clone 4</td>
<td>10/44</td>
<td>A C A C G A T T G A G C C</td>
</tr>
<tr>
<td>PalH8-clone 1</td>
<td>23/35</td>
<td>G C A C A A A C T G A A A C C</td>
</tr>
<tr>
<td>-clone 2</td>
<td>5/35</td>
<td>G C G C A A A C T A G G G C C</td>
</tr>
<tr>
<td>-clone 3</td>
<td>7/35</td>
<td>A T A C A C A A C T G G G C C</td>
</tr>
</tbody>
</table>

*Nucleotide position numbers are according to the reference (L02116). Underlined numbers indicate synonymously conserved positions.

Inter and intra isolates variability was noticed using both subcloned genes. Sample PalH4 showed 4 clones of both TPI (Table 2) and β-giardin (Table 3) genes. In sample PalH8, 3 and 2 clones were detected using TPI
(Table 2) and β-giardin (Table 3) genes, respectively. However, β-giardin gene only showed the variability in sample Pal H9 that revealed 3 clones but only one clone using TPI (Table 3). As seen above, the subgenotype numbers and frequencies in each sample, depending on gene markers (Table 2 and 3), was highly variable depending on the used primer sets. Therefore, we used multiple primers targeting different sequences (Table 1) on each locus to amplify the possible genotypes in each sample. Any combination of the primer sets could be a practical choice in overcoming the amplification problems often encountered in genotyping of G. intestinalis, even we cannot recommend special primers combination for the detection as sometimes the results are unpredictable, which is usually happen when running PCR. However, the maximum confirmed subgenotypes of both TPI and β-giardin genes were 4 subgenotypes in sample PalH4 (Tables 2 and 3).

All the subgenotypes from the cloned TPI gene were unique, and no overlapping of the same subgenotype was found among the subgenotypes (Table 2). However, an overlap between identical subgenotypes was noticed in the β-giardin gene: PalH4-clone 2 was identical to PalH8-clone 1; PalH4-clone 3, PalH8-clone 2, and PalH9-clone 2 were identical clones; and PalH4-clone 4 was identical to PalH9-clone 3 (Table 3). To the best of our knowledge this is the first report to notice the intra-assemblage (subgenotypes) level overlapping as it is noticed on β-giardin gene marker level but not TPI gene within the same sample subgenotypes (Table 3).

Detailed analyses of the SNPs in these clones provided greater insight into the genetic diversity as a function of genotype and subgenotype features. Most of the TPI and β-giardin subgenotype SNPs were synonymous: 10 of the 13 SNPs in TPI gene (Table 2) and 6 of the 7 SNPs in β-giardin gene (Table 3) were synonymous. The resultant 7 TPI and 9 β-giardin subgenotypes had 13 SNPs (2.1%) of the analyzed 623-bp (Table 2) sequence and 7 SNPs (1.0%) of the analyzed 687-bp (Table 3) sequence, respectively. This result indicates that the TPI gene (2.1%) distance is being twice that of the β-giardin (1.0%) distance within assemblage B subgenotypes. While the situation is reversed within assemblage A subgenotypes, β-giardin showed the maximum discriminative power (data not shown). The present study result demonstrated that it is impossible to prove that two isolates are identical or genetically different, as determined with the identical β-giardin gene clones and unique TPI gene clones from assemblage B genotype samples. However, this gene-marker confirmation requires further investigation, as it can be observed only in the closely related subgenotypes at the intra-assemblage level, and probably only within assemblage B genotype.

**DISCUSSION**

Developing countries are not only more vulnerable to G. intestinalis infection, but they could also provide greater insight and information on G. intestinalis genotypes (6, 16-19, 22-26). Previously, we studied G. intestinalis mixed-subgenotype samples using a part of the isolates analyzed in this study and focused on the repeated and cumulative infections of G. intestinalis under the endemic conditions found in developing countries such as Palestine (16). In this study, we further analyzed the samples using the subcloning procedure that targets multiple gene loci and evaluated the assemblage-dependent characteristics of gene
diversity.

In consistence with the previous studies (3, 4, 8, 17, 22, 23, 25-28) mixed-genotype electropherogram profiles were noticed in assemblage B genotypes. The detection of assemblage A genotypes with the higher prevalence without mixed-genotype is also reported (29, 30). Higher assemblage B prevalence with the mixed profiles were documented (22, 23, 25-27) and also assemblage A and B mixed profiles were also reported (18, 25, 31-33). It seems assemblage B is more prevalent in low living standards with the symptomatic cases and the less symptomatic cases in assemblage A. In this study, we used the subcloning approach to reveal the genotypes in the mixed electropherogram profile samples. Interestingly, both TPI and β-giardin genes revealed 4 subgenotypes in sample PalH4 (Tables 2 and 3) even TPI gene possesses a higher degree of sequence polymorphism than β-giardin gene. Although mixed infections cannot be excluded; such result may correlate with the allelic heterogeneity of the target gene marker based on the genome tetraploid (1, 34) or multi-copy genes (9, 35) in the basic G. intestinalis trophozoite stage. Indeed, real-time PCR quantification indicated 4 copies of β-giardin gene in the basic G. intestinalis trophozoite stage (36).

Table (3): Variation in β-giardin nucleotide sequences.

<table>
<thead>
<tr>
<th>Subgenotype†</th>
<th>Clone frequency</th>
<th>β-giardin nucleotide position*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>218</td>
</tr>
<tr>
<td>M36728</td>
<td></td>
<td>G</td>
</tr>
<tr>
<td>PalH4-clone 1</td>
<td>(4/28)</td>
<td>G</td>
</tr>
<tr>
<td>-clone 2</td>
<td>(3/28)</td>
<td>G</td>
</tr>
<tr>
<td>-clone 3</td>
<td>(8/28)</td>
<td>A</td>
</tr>
<tr>
<td>-clone 4</td>
<td>(13/28)</td>
<td>A</td>
</tr>
<tr>
<td>PalH8-clone 1</td>
<td>(18/21)</td>
<td>G</td>
</tr>
<tr>
<td>-clone 2</td>
<td>(3/21)</td>
<td>A</td>
</tr>
<tr>
<td>PalH9-clone 1</td>
<td>(14/23)</td>
<td>A</td>
</tr>
<tr>
<td>-clone 2</td>
<td>(5/23)</td>
<td>A</td>
</tr>
<tr>
<td>-clone 3</td>
<td>(4/23)</td>
<td>A</td>
</tr>
</tbody>
</table>

* Nucleotide position numbers are according to the reference (M36728). Underlined numbers indicate synonymously conserved positions.
† Subgenotypes identified in sample PalH8 were identical and overlapped with other sample subgenotypes: PalH8-clone 1 was identical to PalH4-clone 2, and PalH8-clone 2 was identical to PalH4-clone 3 and PalH9-clone 2. Two subgenotypes detected in sample PalH9 were identical and overlapped with other sample subgenotypes: PalH9-clone 2 was identical to PalH4-clone 3 and PalH8-clone 2, and PalH9-clone 3 was identical to PalH4-clone 4.

Considering the difficulties encountered in detecting multiple genotypes from a single sample, the differences in the subgenotype numbers and frequencies in each sample de-
pending on gene markers (Table 2 and 3) could be explained by the primer selection and/or the limitations of substitutions in target loci. In fact, it has been suggested that the exclusion of mixed genotypes in a single specimen cannot be confirmed without conducting repeated PCR analyses, with either a single or multiple primer sets (37); several reports have also indicated failure in getting amplicons for some genotypes when using a certain primer set but success when using a different one (8, 17).

The inter and intra isolates variation noticed on the different gene levels, even using multiple primers approach that were more apparent in the cloned TPI gene and the overlapping between the apparently identical subgenotypes in the β-giardin gene in the single samples is similar to the overlapping phenomena of the same genotypes from different samples reported in previous studies as a reflection of the discriminative level in the gene markers and/or mixed infections (4, 16, 17, 19, 22, 23, 25-27, 31, 38). However, the subgenotypes overlapping was noticed in the single sample isolates and were complicated, especially if used to confirm the subgenotypes using multiple gene markers (Table 3), by the detection of the multiple subgenotypes which may be related to assemblage-dependent features.

As noticed in the previous studies (2, 4, 7, 9, 16, 19, 22, 23, 27) there is prominent heterogeneity in assemblage B genotypes and most of the noticed SNPs were synonymous (Table 2 and 3). Previously, we reported 6 subgenotypes using glutamate dehydrogenase gene (GDH) that had 15 SNPs (3.8%) of the analyzed 393-bp sequence; nevertheless, there are a lower number of the detected subgenotypes using the GDH gene marker, the higher degree of polymorphism was noted among the GDH subgenotypes. Previously, Monis et al. (5) found that the inter-assemblage distance determined using the TPI gene is twice that of the GDH gene. The results of this study may suggest another specific intra-assemblage distances for assemblage A and B genotypes, while more data is required to make this assertion. It is clearly supported by the GDH gene distance (3.8%) being twice that of the TPI gene (2.1%) distance, which is in turn twice that of the β-giardin (1.0%) distance in assemblage B subgenotypes. The situation is almost reversed within assemblage A subgenotypes in consistence with the findings of a previous study (17).

The independent presence of assemblages A and B in patients, as mixed inter-assemblage profiles were never detected in any samples, and the existence of multiple genetically similar assemblage B subgenotypes in the single samples (Tables 2 and 3), in contrast to the single genotype in each of assemblage A-II samples, might indicate the special diversified characteristics of assemblage A and B genotypes. Nevertheless, a mixed inter-assemblage profile for assemblages A-I and B has also been reported (9); thus, we should pay special attention to such mixed inter-assemblage profiles while considering the notion of possible DNA recombination (19). Whether genetic mating could occur between assemblages A and B is an important question to consider when studying different species.

Although the exact separating mechanisms between assemblages A and B in endemic areas remains unknown, the contrasting features of genetic diversity and the independent presence of these genotypes in patients could support the notion that assemblages A and B be-
long to different species (6, 19).

CONCLUSIONS

Considering the comparatively well-documented status of assemblage A genotypes regarding extensive genetic conformations, the current study findings highlight the need to investigate assemblage B as thoroughly as has been the case of assemblage A.

ABBREVIATIONS

TPI, Triosephosphate Isomerase Gene; GDH, Glutamate Dehydrogenase Gene.

ACKNOWLEDGEMENTS

We would like to thank Moheb Abu-Zanat and Yousef Salam for their help in collecting the clinical samples. This work was supported by Grant 20406007 from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES


27) Flecha M, Benavides C, Tissiano G, Tesfamariam A, Cuadros J, de Lucio A et al. Detection and molecular characterisa-
Assemblage-Dependent Genetic Features of...


