



Multi-locus Genotyping Reveals High Occurrence of Mixed Assemblages in *Giardia duodenalis* within a Limited Geographical Boundary

Avik Kumar Mukherjee¹, Sumallya Karmakar¹, Dibyendu Raj¹
and Sandipan Ganguly^{1*}

¹Department of Parasitology, National Institute of Cholera and Enteric Diseases, P-33, CIT Road, Scheme XM, Beliaghata, Kolkata 700010, India.

Authors' contributions

This work was carried out in collaboration between all authors. Author AKM has planned and executed the laboratory work, analyzed the data and prepared the draft for the manuscript. Authors SK and DR have helped in sample collection and DNA sequencing. Author SG has designed the work, corrected and approved the final version of the manuscript to publish. All authors read and approved the final submission.

Short Communication

Received 18th December 2012
Accepted 3rd March 2013
Published 22nd March 2013

ABSTRACT

Aim: To determine the common genotypes of *Giardia duodenalis* causing diarrhea in the study region and to assess the extent of genetic polymorphism among them.

Study Design: Stool samples were collected from the patients attending IDBG Hospital, Kolkata with diarrheal complaints through a systemic sampling technique and were screened for *Giardia duodenalis*. The *G. duodenalis* positive samples were subjected to molecular genotyping through 'PCR - Direct DNA sequencing' procedure. All the sequence data obtained were incorporated into MEGA 4 software for multiple alignment and validation followed by phylogenetic analysis. The genotyping data obtained are stored in Excel spreadsheets and incorporated into Epilnfo 3.1 for analyzing possible association of genotype outcome with common physical factors such as age, sex etc.

Place and Duration of Study: Department of parasitology, National Institute of Cholera and Enteric Diseases, Kolkata, India from July 2009 to November 2011.

Methodology: A total of 68 *Giardia duodenalis* positive stool samples were identified from

*Corresponding author: Email: sandipanganguly@gmail.com;

the diarrhea patients attending IDBG hospital in the city and were subjected to multi-locus genotyping. Fragments of β -giardin, *Glutamate-dehydrogenase* and *Triosephosphate-isomerase* genes of *Giardia* were amplified from those samples with specific primers and sequenced. All the sequences were analyzed using MEGA 4 software for obtaining the genotyping results.

Results: Multi-locus genotyping identified 13 isolates as assemblage A and 41 as assemblage B, whereas 14 of them could not be assigned in a particular group. Detailed phylogenetic analysis revealed that multiple genotypes were observed in those 14 isolates depending upon the marker loci.

Conclusion: The study could produce a preliminary idea about the *G. duodenalis* genotypes found in Kolkata city. High percentage of mixed assemblages in the study population also revealed the presence of genetic diversity among a small population of diarrheal patient within a limited geographical boundary. It has also hypothesized the possibility of inter-assemblage genetic exchange among *Giardia*.

Keywords: *Giardia*; genotyping; mixed assemblages; local isolates.

ABBREVIATION

IDBG: Infectious Diseases and Beliaghata General; NICED: National Institute of Cholera and Enteric Diseases; PCR: Polymerase Chain reaction; ELISA: Enzyme Linked Immuno-Sorbent Assay; MEGA 4: Molecular Evolutionary Genetics Analysis, Version 4; NJ: Neighbor Joining.

1. INTRODUCTION

Giardia is a genus of intestinal flagellates that infects a wide range of vertebrate hosts. The genus currently comprises six species namely *Giardia agilis*, *Giardia ardeae*, *Giardia psittaci*, *Giardia microti*, *Giardia muris* and *Giardia duodenalis* (syn. *intestinalis*) infecting amphibians, birds, rodents and mammals [1]. These species can be distinguished based on the morphology and ultrastructure of their trophozoites. Among them only *G. duodenalis* is responsible for giardiasis in human, although it can also infect other vertebrates [2]. According to the previous reports, the members of species *G. duodenalis* have negligible morphological variations but great genetic diversity. Therefore, they are regarded as a species complex [3] with eight distinct genetic groups or assemblages (A-H) [4]. Assemblage A and B can infect and multiply in humans and other mammals whereas the remaining assemblages show a much restricted host range. Likewise C and D mostly infect the canids, whereas E, F, G and H infect livestock, cats, rodents and marine vertebrates respectively [4,5].

Kolkata (N 22°577242 E 88°398743) is a densely populated city of India, where the giardiasis in human caused by infection of *G. duodenalis* is a serious problem of public health. It is one of the major metro cities of eastern India and Southeast Asia with plenty of immigration and emigration of multi cultured people with varied socio economic condition. It has a favorable climatic condition for the growth of enteric pathogens. So, it is very much likely to find plenty of opportunistic enteric parasites in this city. Previously, the prevalence of *Giardia* in Kolkata has been studied by Mukherjee et al. [9] but no genotyping study was reported. Genotyping of local *Giardia* isolates could possibly reflect the scenario of an urbanized region of a developing country with similar demography. Assemblage type has been thought to be

associated with the risk factors of the disease [6], hence genotyping may help in better understanding of the parasite epidemiology in a particular setting or area [7]. Therefore, we decided to investigate the common genotypes of *Giardia duodenalis* circulating within the limited boundary of Kolkata.

Many genotyping studies based on single locus analysis of *G. duodenalis* have shown the presence of a particular assemblage in variety of hosts. Irrespective of the marker loci and the hosts the sequence obtained from *G. duodenalis* isolates appeared to be similar if not identical [8] and were regarded as possible transmission of the parasite among different animal host. Here we have used multi-locus genotyping for proper discrimination of assemblages. Detection of true mixed assemblages through combined analysis of three different loci is possible with the help of this process.

2. MATERIALS AND METHODS

The study has received an ethical clearance from the institutional (NICED) ethical committee. All the stool samples were collected as a part of an ongoing project between NICED and IDBG hospital through a “systemic sampling technique”. Single stool samples were taken from 700 subjects attending IDBG hospital who had previously tested for *Giardia duodenalis* via microscopy, PCR and ELISA following published protocols [9], and were kept at -80°C. Samples were selected from every fifth diarrhea patient in two randomly selected days of a week and enrolled in the study from July 2009 to November 2011; the subjects were identified depending on two criteria: a. the subject should be residing within the city boundary; b. at least 20 mL of stool sample are available at the time of enrolment i.e. before medication. The whole study remained unbiased of age and gender of the subjects. The DNA was extracted directly from the positive stools using StoolDNAMiniKit (QIAGEN, USA) according to the manufacturer’s protocol. A portion of β -giardin (βg) [511 bp] on 90 kb long contig ctg02-35, *Glutamate dehydrogenase* (*gdh*) [434 bp] on 231 kb long contig ctg02-15 and *Triose phosphate isomerase* (*tpi*) [530 bp] on 200 kb long contig ctg02-19 (www.Giardiadb.org), were individually amplified according to the previously described nested PCR protocols [10-12]. The *tpi* gene is mapped to chromosome no. 5 whereas *gdh* and βg are mapped to chromosome no. 4 [1]. The nested PCR products were separated in 1.5% (w/v) agarose gel and purified by gel cut purification process using High Pure PCR purification Kit (Roche, Germany) as per the manufacturer’s protocol. Bi-directional sequencing was performed with the respective purified products and nested PCR primers on an ABI 3100 automated sequencer by using the Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystem, USA). The sequences obtained from all three loci (βg , *gdh* and *tpi*) were validated using the database BLAST search (i.e. NCBI and *Giardiadb*) and were submitted to NCBI GenBank (ACC no. JF918436 – JF918523 & JN647526 – JN647641). The sequences from each locus were separately aligned by using ‘MEGA Version 4’ software [13,14] and were manually checked and edited. Previously reported sequences of the respective loci representing different *G. duodenalis* assemblages were included in the analysis to get a better resolution of the assemblage distribution. The extent of sequence diversity among the wild isolates based on the target loci was determined using the ‘Maximum Composite Likelihood’ method through the MEGA4 software. Three separate ‘Neighbor-Joining (NJ)’ phylogenetic trees [15] for each locus were generated using MEGA4 for the 14 unassigned isolates (Fig. 1). Previously reported sequences of the respective loci representing different *G. duodenalis* assemblages were also included along with closely related species such as *Giardia ardeae/muris* as an out-group for acquiring a better resolution and topology of the evolutionary tree. Reliability of the phylogenetic analysis was

assessed by using 'Maximum Composite Likelihood' model and bootstrap method with 1000 replicates [16,17].

3. RESULTS AND DISCUSSIONS

A total of 68 out of 700 *G. duodenalis* positive fecal samples were processed. Based on the cumulative sequence data of all three loci, 41 samples could be assigned as assemblage 'B' (60.2%) and 13 as assemblage 'A' (19.1%), while 14 (20.5%) isolates showed multiple assemblages depending on the marker loci (Supplementary Table 1).

Sequence analysis showed higher degree of diversity in the '*tpi*' loci in compare to other two (Table 1).

Table 1. Estimation of average sequence diversity

Target locus					
β giardin (<i>βg</i>)		Glutamate dehydrogenase (<i>gdh</i>)		Triose phosphate isomerase (<i>tpi</i>)	
<i>N</i>	<i>D</i> \pm <i>SE</i>	<i>N</i>	<i>D</i> \pm <i>SE</i>	<i>N</i>	<i>D</i> \pm <i>SE</i>
68	0.001 \pm 0.004	67	0.053 \pm 0.011	64	0.106 \pm 0.019

The number of base substitutions per site from averaging over all sequence pairs is shown. All results are based on the pairwise analysis of 'N' number of sequences. Standard error (SE) estimate(s) are shown besides the diversity (D) estimates and were obtained by a bootstrap procedure (1000 replicates). Analyses were conducted using the Maximum Composite Likelihood method in MEGA4. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). The tpi loci showed highest diversity in compare to the other two.

Detection of Assemblage 'B' and assemblage 'A' overlaps the results obtained from previous studies in southern India using single locus PCR-RFLP method [18]. However, non-specific sequence heterogeneity among same loci of identical assemblages led to the difficulty in assigning sub-assemblages [19]. In the previous reports from the same country, presence of mixed assemblages may be evident but, those were usually found through multiple assemblage specific bands in RFLP analysis of a particular locus [18]. This may occur due to presence of overlapping DNA sequence derived from two or more different isolates, although it was not proved. The unique finding of our study was, 14 (20.5%) isolates showed multiple assemblages depending on the marker loci. Clustal distribution with these 14 isolates in the NJ tree revealed a better picture where isolates were placed in distinct clusters of assemblage A and B (marked with red and blue bars respectively) supported by high bootstrap values (Fig. 1) but, the cluster selection (i.e. A or B) of the isolates were different for each loci. For example the isolate number GLI11 is positioned in assemblage A cluster in ' *β g*' and '*gdh*' tree (Fig. 1A & B) but falls in the other cluster in '*tpi*' tree (Fig. 1C). Again isolate GLI23 is positioned in assemblage A cluster in ' *β g*' and '*tpi*' tree (Fig. 1A & C) but is present in the assemblage B cluster in '*gdh*' tree (Fig. 1B).

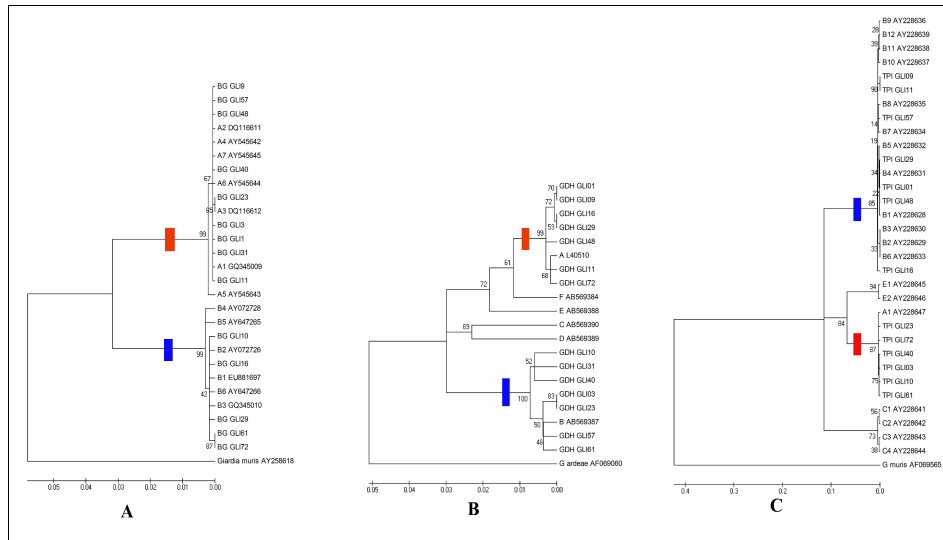


Fig. 1. Phylogenetic analysis of β -giardin, Glutamate dehydrogenase and Triose phosphate isomerase loci using the ‘MEGA version 4 program’

The evolutionary history was inferred using the Neighbour-Joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap tests (1000 replicates) are shown next to the branches. The evolutionary distances have been computed using the Maximum Composite Likelihood method and are shown in the units of the number of base substitutions per site. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). Assemblage A clusters are marked with RED bars and Assemblage B clusters are marked with BLUE bars. (1. A) Phylogenetic analysis of β -giardin locus. (1. B) Phylogenetic analysis of Glutamate dehydrogenase locus. (1. C) Phylogenetic analysis of Triose phosphate isomerase locus.

Homology analysis showed that the sequences within each assemblage A and B clusters for all the loci are mostly homologous (data not shown) taking all the isolates in account. Also in the combined phylogenetic analysis, these 14 isolates doesn't make any separate cluster and moreover they are evenly distributed in the respective cluster either assemblage A or B similar to Fig. 1. Hence, in this case the outcome of assemblage was solely depended on the marker loci (Supplementary Table 2). Association of genotype outcome with other physical factors such as Age, Sex and co-infection status was also checked but no particular association was observed (Supplementary File 1). From this finding it can be said that the differential taxon position of this 14 isolates could not be due to chance and it is a true example of mixed assemblage. Similar results are frequently reported in dogs where genotyping results changes with either host-adapted non-zoonotic assemblages C and D or as assemblage B with a zoonotic potential [17,20,21] depending upon the genetic markers used. This has a very high impact on the implication of the molecular epidemiological studies given that; different outcome is possible with usage of different markers, thus confirming the importance of multi locus genotyping in case of environmental samples. Two major reasons can be placed in favor of this outcome i.e. presence of mixed infection and occurrence of genetic recombination through sexual reproduction [22,23]. Previous reports of mixed assemblages relied on the presence of multiple peaks in a particular position in the chromatograms which lead to the ambiguous taxon positioning of the isolates in the phylogenetic tree which can be due to mixed infection but, in this case all the isolates with dual genotype are positioned perfectly in a particular cluster for a specific loci. Although, recent reports suggest towards considering this type of observations as mixed assemblage

infection, in spite of a thin probability of inter-assemblage recombination [24], still detailed molecular epidemiological study is required to find out the exact reason behind this unique finding. However, detection of high percentage of mixed genotype is evident, whether it is due to mixed infection or genetic recombination.

4. CONCLUSION

The study provided some useful information regarding the presence of common *Giardia duodenalis* genotypes in the city. High percentage of mixed assemblages in an endemic tropical city like Kolkata can explain the possibility of uncontrolled transmission of *Giardia*, leading to easy mixing of multiple genotypes. Other than the public health importance, abundance of mixed assemblages also enlightens the possibility of inter isolate genetic exchange among *Giardia* sp.

ACKNOWLEDGEMENTS

This study was jointly supported by a grant from the Okayama University Program of Founding Research Centre for Emerging and Re-emerging Infectious Disease, Ministry of Education, Culture, Sports, Science and Technology of Japan and Japan Health and Science Foundation, Govt. of Japan & Indian Council of Medical Research, Govt. of India. The authors acknowledge Dr. S. Chakraborty and Prof. S. Shinoda for their continuous constructive suggestions, support and critical review during this study. The authors also thank Council of Scientific and Industrial Research, India and University Grant Commission, India for providing fellowship to SK and DR during the study period. We also thank Dr. Mrinmoy Ghosh, Dr. Mihir K Bhattacharya for the arrangement of stool collection procedure, Ms. Punam Chowdhury for helping in the detection of *Giardia duodenalis* from stool sample and Mrs. Debarati Ganguly for her help in English language correction.

ETHICAL APPROVAL

The study was approved by National Institute of Cholera and Enteric Diseases ethical committee with clearance number A-1/2008-IEC dated 5th December 2008.

COMPETING INTEREST

Authors declare no competing interest for this work.

REFERENCES

1. Adam RD. The *Giardia lamblia* genome. Int J Parasitol. 2000;30:475–84.
2. Thompson RC. The zoonotic significance and molecular epidemiology of *Giardia* and *Giardiasis*. Vet Parasitol. 2004;126(1-2):15–35.
3. Monis PT, Andrews RH, Mayrhofer G, Ey PL. Genetic diversity within the morphological species *Giardia intestinalis* and its relationship to host origin. Infect Genet Evol. 2003;3(1):29–38.
4. Lasek-Nesselquist E, Welch DM, Sogin ML. 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. 2010;1;40(9):1063-74.

5. Caccio SM, Ryan U. Molecular epidemiology of giardiasis. *Mol Biochem Parasitol.* 2008;160(2):75–80.
6. Laishram S, Kang G, Ajjampur SS. Giardiasis: a review on assemblage distribution and epidemiology in India. *Indian J Gastroenterol.* 2012;3-12
7. Thompson RC, Monis P, Giardia--from genome to proteome. *Adv Parasitol.* 2012;78:57-95.
8. Sprong H, Caccio SM, van der Giessen JW. Identification of zoonotic genotypes of *Giardia duodenalis*. *PLoS Neg Trop Dis.* 2009;3:e558.
9. Mukherjee AK, Chowdhury P, Bhattacharya MK, Ghosh M, Rajendran K Ganguly S. Hospital-based surveillance of enteric parasites in Kolkata. *BMC Research Notes.* 2009;2:110.
10. Lalle M, Pozio E, Capelli G, Bruschi F, Crotti D, Caccio SM. Genetic heterogeneity at the β -giardin locus among human and animal isolates of *Giardia duodenalis* and identification of potentially zoonotic sub-genotypes. *Int J Parasitol.* 2005;35(2):207-13.
11. Sulaiman IM, Fayer R, Bern C, Gilman RH, Trout JM, Schantz PM, Das P, Lal AA, Xiao L. Triosephosphate isomerase gene characterization and potential zoonotic transmission of *Giardia lamblia*. *Emerg Infect Dis.* 2003;9:1444-1452.
12. Caccio SM, Beck R, Lalle M, Marinculic A, Pozio E. Multilocus genotyping of *Giardia duodenalis* reveals striking differences between assemblage A and B. *Int J Parasitol.* 2008;38(13):1523-31.
13. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc Natl Acad Sci USA.* 2004;101(30):11030-11035.
14. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol.* 2007;24(8):1596-1599.
15. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol.* 1987;4(4):406–25.
16. Felsenstein J. Confidence limits on the phylogenies: an approach using bootstrap. *Evolution.* 1985;39:783–91.
17. Caccio SM, Sprong H. *Giardia duodenalis*: Genetic recombination and its implication for taxonomy and molecular epidemiology, *Exp Parasitol.* 2010;124(1):107-112.
18. Laishram S, Kannan A, Rajendran P, Kang G, Ajjampur SS. Mixed *Giardia duodenalis* assemblage infections in children and adults in South India. *Epidemiol Infect.* 2012;140(11):2023-7.
19. Ankarklev J, Svärd SG, Lebbad M. Allelic sequence heterozygosity in single *Giardia* parasites, *BMC Microbiology.* 2012;12:65
20. Read CM, Monis PT, Thompson RC. Discrimination of all genotypes of *Giardia duodenalis* at the glutamate dehydrogenase locus using PCR-RFLP. *Infect. Genet. Evol.* 2004;4(2):125–130.
21. Traub RJ, Monis PT, Robertson I, Irwin P, Mencke N, Thompson RC. Epidemiological and molecular evidence support the zoonotic transmission of *Giardia* among humans and dogs living in the same community. *Parasitology.* 2004;128(3):53–62.
22. Cooper MA, Adam RD, Worobey M, Sterling CR. Population genetics provides evidence for recombination in *Giardia*. *Curr. Biol.* 2007;17:1984–1988.
23. Teodorovic S, Braverman JM, Elmendorf HG. Unusually low levels of genetic variation among *Giardia lamblia* isolates. *Eukaryot. Cell.* 2007;6(8):1421–1430.

24. Takumi K, Swart A, Mank T, Lasek-Nesselquist E, Lebbad M, Caccio SM, Sprong H. Population-based analyses of *Giardia duodenalis* is consistent with the clonal assemblage structure. *Para Vect.* 2012;5:168.

© 2013 Mukherjee et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:

The peer review history for this paper can be accessed here:
<http://www.sciencedomain.org/review-history.php?iid=200&id=8&aid=1152>