Vol. 81: 39–51, 2008
doi: 10.3354/dao01931

DISEASES OF AQUATIC ORGANISMS
Dis Aquat Org

Published August 19

*Contribution to DAO Special 3: 'Marine vertebrate zoonoses'*

OPEN ACCESS

# Molecular characterization of *Giardia intestinalis* haplotypes in marine animals: variation and zoonotic potential

**Erica Lasek-Nesselquist[1, 2,*], Andrea L. Bogomolni[3], Rebecca J. Gast[3], David Mark Welch[2], Julie C. Ellis[4], Mitchell L. Sogin[2], Michael J. Moore[3]**

[1]Department of Ecology and Evolutionary Biology, Brown University, 80 Waterman Street, Providence, Rhode Island 02912, USA

[2]Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, 7 MBL Street, Woods Hole, Massachusetts 02543, USA

[3]Biology Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

[4]Department of Environmental and Population Health, Cummings School of Veterinary Medicine at Tufts University, North Grafton, Massachusetts 01536, USA

ABSTRACT: *Giardia intestinalis* is a microbial eukaryotic parasite that causes diarrheal disease in humans and other vertebrates worldwide. The negative effect on quality of life and economics caused by *G. intestinalis* may be increased by its potential status as a zoonosis, or a disease that can be transmitted from animals to humans. The zoonotic potential of *G. intestinalis* has been implied for over 2 decades, with human-infecting genotypes (belonging to the 2 major subgroups, Assemblages A and B) occurring in wildlife and domesticated animals. There are recent reports of *G. intestinalis* in shellfish, seals, sea lions and whales, suggesting that marine animals are also potential reservoirs of human disease. However, the prevalence, genetic diversity and effect of *G. intestinalis* in marine environments and the role that marine animals play in transmission of this parasite to humans are relatively unexplored. Here, we provide the first thorough molecular characterization of *G. intestinalis* in marine vertebrates. Using a multi-locus sequencing approach, we identify human-infecting *G. intestinalis* haplotypes of both Assemblages A and B in the fecal material of dolphins, porpoises, seals, herring gulls *Larus argentatus*, common eiders *Somateria mollissima* and a thresher shark *Alopias vulpinus*. Our results indicate that *G. intestinalis* is prevalent in marine ecosystems, and a wide range of marine hosts capable of harboring zoonotic forms of this parasite exist. The presence of *G. intestinalis* in marine ecosystems raises concerns about how this disease might be transmitted among different host species.

KEY WORDS: *Giardia intestinalis* · Zoonosis · Marine birds · Marine mammals · Thresher shark



## INTRODUCTION

*Giardia intestinalis* is one of the most ubiquitous enteric parasites in the world, capable of infecting virtually all mammals and even some other vertebrates, such as birds (Thompson 2004). Upon infection, humans as well as other animals (e.g. livestock and dogs) can display clinical symptoms such as diarrhea, malnutrition, nausea and failure to thrive (Thompson & Monis 2004). The life cycle of *G. intestinalis* is composed of 2 stages: (1) an inactive cyst that is ingested by the host through a fecal–oral route or through the consumption of contaminated water, and (2) an active trophozoite that attaches to the intestinal lumen of the host and divides by binary fission (Adam 2001).

*Giardia intestinalis* includes 7 genetically distinct subgroups or assemblages (A to G). Only Assemblages A and B have been documented in humans. These 2

*Email: erica_lasek-nesselquist@brown.edu

assemblages also infect other vertebrates while Assemblages C to G appear to be more host specific (Thompson et al. 2000). Because Assemblages A and B infect humans, wildlife such as beaver and deer, and domesticated animals such as dog, cat and livestock, *G. intestinalis* may represent a zoonosis (Thompson 2004). The presence of human-infecting forms in other animals does not prove that *G. intestinalis* is transmitted from animals to humans. However, a study of a tea-growing community in India, where dogs and humans live in close proximity to each other, revealed a significant correlation between the presence of *G. intestinalis* in dog and owner, which strongly suggests zoonosis (Traub et al. 2004).

The contamination of aquatic ecosystems with *Giardia intestinalis* from untreated or improperly treated wastewater and agricultural runoff is an increasing concern. Cysts have been detected in the water columns of both fresh water and marine environments, in shellfish, which serve as bioindicators of water contamination, and even in feces of marine mammals such as seals, sea lions and whales (Fayer et al. 2004, Hughes-Hanks et al. 2005, Robertson 2007). The effect of this pathogen on these aquatic environments is largely undetermined and although the potential for zoonoses exists, the transmission of this parasite between marine and human hosts has not been explored.

To gain insight into the diversity and zoonotic potential of *Giardia intestinalis* in marine environments, we used a multi-locus sequencing approach on fecal samples from 9 marine animal species. To our knowledge, we provide the first molecular data of *G. intestinalis* in marine vertebrates. Through the use of replicate PCR amplifications on multiple loci we show that this parasite has a seemingly unrestricted presence in marine environments, with several animals containing multiple haplotypes of *G. intestinalis* in their fecal material.

## MATERIALS AND METHODS

**Fecal collection.** Stranded marine mammals and birds and those from fisheries bycatch were collected with the assistance of staff from the New England Aquarium, University of New England Marine Animal Rehabilitation Center, the National Oceanographic Atmospheric Administration (NOAA) Northeast Fisheries Science Center (NEFSC) Observer Program, the Seabird Ecological Assessment Network (SEANET, www.tufts.edu/vet/seanet/), the Massachusetts Audubon Society and the authors. Large cetaceans were necropsied at the site of stranding (usually on a beach), and a subset of birds were frozen before sampling. Other animals were necropsied at Woods Hole Oceanographic Institution (WHOI) within 4 to 48 h post mortem (stored at 4°C overnight). Gut contents and fecal material were sampled at the time of necropsy using sterile materials. Dolphins and porpoises were obtained in Massachusetts. Beached common dolphins *Delphinus delphis* were from Skaket Beach in Orleans (sample no. 10; Table 1), Lieutenant Island in Wellfleet (samples no. 11 and 15; Table 1), or bycaught near Fall River (sample no. 14; Table 1). Beached white-sided dolphins *Lagenorhynchus acutus* were obtained from Beach Point (samples no. 16 and 17; Table 1) and Chipman's Cove in Wellfleet (sample no. 18; Table 1). A Risso's dolphin *Grampus griseus* was obtained from Thumpertown Beach in Eastham (sample no. 19; Table 1). Harbor porpoises *Phocoena phocoena* were collected from the coast guard station of Sandwich, directly South of Cape Cod and directly North of Cape Cod (samples no. 20 through 22; Table 1).

Fecal samples were collected from live-caught herring gulls *Larus argentatus* at Kent Island, New Brunswick, Canada. Herring gulls were captured during egg incubation using either a walk-in nest trap (a chicken wire cage with an opening on the bottom and an entrance on one side) or a drop-down trap (chicken wire cage propped up on one side by a wooden peg attached to a line) (sample no. 1 through 6; Table 1). Fresh fecal samples were collected by placing each

Table 1. Marine animals sampled for *Giardia intestinalis* Assemblages A and B and their assemblage profiles. Samples 1 through 11 were examined at 4 loci (*gdh*, *mlh*, *tpi5'* and *tpi3'*); samples 12 through 23 were examined only at *gdh* due to limited DNA

| Marine animal sampled | Sample name | Assemblage(s) |
|---|---|---|
| 1 *Larus argentatus* | Gull 1 | A & B |
| 2 *Larus argentatus* | Gull 3 | A & B |
| 3 *Larus argentatus* | Gull 11 | A & B |
| 4 *Larus argentatus* | Gull 12 | A & B |
| 5 *Larus argentatus* | Gull 13 | A & B |
| 6 *Larus argentatus* | Gull 16 | A & B |
| 7 *Somateria mollissima* | Eider NE2 | A & B |
| 8 *Somateria mollissima* | Eider NE3 | A & B |
| 9 *Phoca groenlandica* | Seal 10A | A & B |
| 10 *Delphinus delphis* | Dolphin 13 | A & B |
| 11 *Delphinus delphis* | Dolphin 316 | A & B |
| 12 *Somateria mollissima* | Eider NE4 | A |
| 13 *Phoca vitulina* | Seal 24A | B |
| 14 *Delphinus delphis* | Dolphin 17B | B |
| 15 *Delphinus delphis* | Dolphin 23A | B |
| 16 *Lagenorhynchus acutus* | Dolphin 20A | B |
| 17 *Lagenorhynchus acutus* | Dolphin 30A | B |
| 18 *Lagenorhynchus acutus* | Dolphin 33 | B |
| 19 *Grampus griseus* | Dolphin 14A | B |
| 20 *Phocoena phocoena* | Porpoise 26A | B |
| 21 *Phocoena phocoena* | Porpoise 37A | A & B |
| 22 *Phocoena phocoena* | Porpoise 38B | B |
| 23 *Alopias vulpinus* | Shark 68R | B |

bird in a plastic box just prior to release; most birds responded to box placement by voiding their cloacas almost immediately. Fecal samples were transferred to sterile cryovials using plastic sterile Pasteur pipettes or syringes. To avoid contamination, the liner at the bottom of the box was replaced between each bird.

Fecal samples from live seals and birds were also collected from beaches in Massachusetts on Nantucket Island (samples no.7, 8 and 12; Table 1), Ram Island at Marian (sample no. 9; Table 1), and directly east of Cape Cod (sample no. 13). Visual identifications and photographs of the species present at each beach were made before approaching the animals and collecting feces. Animals were identified as harbor seal *Phoca vitulina*, harp seal *P. groenlandica* or common eider *Somateria mollissima*. Feces were collected from the sand surface using sterile spoons and 50 ml centrifuge tubes and transported back to the laboratory on ice. All fecal samples were stored frozen at −80°C until nucleic acid extraction. Fecal and gut contents from a thresher shark *Alopias vulpinus* (sample no. 23; Table 1) were obtained from an animal caught off of Martha's Vineyard.

**Sample screening by beta-giardin PCR amplification.** Nucleic acids were extracted using either the Qiagen stool kit (Valencia) with maximum incubation times for cyst disruption and proteinase K digestion or the MoBio PowerSoil DNA extraction kit (Carlsbad). All samples were initially screened for the presence of *Giardia* using the beta-giardin primers GGL639-658/GGR789-809 (Mahbubani et al. 1992). These primers were selected because they should amplify all members of the genus, and the small size of the amplicon should amplify efficiently. We applied the primers in a 2-round amplification protocol using the same primers in both reactions. All reactions were set up in 25 µl volumes containing 10 mM Tris-HCl, 40 mM KCl, 2.5 mM MgCl2, 0.2 mM dNTP, 0.25 M primer, 0.4 M BSA, 0.25 µl (1.25 units) Go-Taq DNA polymerase (Promega), and 1 µl of sample DNA. One microliter of the first reaction was used as template for the second. Amplification parameters were 94°C for 2 min, followed by 94°C for 30 s, 56°C for 30 s, 72°C for 1 min and a hold at 4°C. The first amplification was carried out for 25 cycles, and the second amplification was 40 cycles. When tested on a dilution series of *Giardia lamblia* genomic DNA, the lowest concentration detected was $100 \times 10^{-15}$ g, the equivalent of approximately 10 cells per reaction. We recognize that our fecal samples may be less sensitive than this due to a combination of inhibitory substances that copurify with DNA and the lack of concentration of cysts from the fecal material.

**PCR, cloning and sequencing of 4 loci.** DNA from fecal samples that tested positive for beta-giardin was amplified with primers designed to glutamate dehydrogenase (*gdh*), *mlh1* (a gene involved in DNA repair after recombination in other eukaryotes; Ramesh et al. 2005), a 5′ region of triose-phosphate isomerase (*tpi5′*), and a 3′ region of triose-phosphate isomerase (*tpi3′*). Although the 2 regions of *tpi* overlap, the presence of multiple different *Giardia interstinalis* sequences for both *tpi3′* and *tpi5′* within a single fecal sample prevents assembly of the 2 regions into single contigs and, therefore, the 2 regions are treated as separate loci. Each gene (*gdh*, *mlh1* and *tpi*) is present in a single copy in the haploid genome of *G. lamblia* WB (Morrison et al. 2007). Both *gdh* and *tpi* have been used in previous genotyping efforts (Wielinga & Thompson 2007). We used a nested PCR approach to amplify each locus; primer sequences are specified in Table 2. For *tpi5′* we used the degenerate primers designed by Sulaiman et al. (2003); for *tpi3′* we used primers designed to be specific to Assemblage A in one set of reactions and primers designed to be specific to Assemblage B in a second set of reactions. Similarly, we amplified *gdh* using a set of Assemblage A specific primers and, in a separate set of reactions, a set of primers designed to Assemblage B supplemented with a pair of more general primers. For *mlh*, the initial amplification was performed with either a pair of Assemblage B specific primers or a combination of Assemblage B specific forward and *Giardia* specific reverse primers followed by amplification with an internal pair of Assemblage B specific primers. Assemblage A specific primers were designed using the *Giardia lamblia* genome database (www.Giardiadb.org); Assemblage B specific primers were created from alignments of previously sequenced Assemblage B axenic human isolates (E. Lasek-Nesselquist unpubl.). The more general *gdh* primers were designed to conserve regions of a *Giardia intestinalis*–*Spironucleus vortens gdh* sequence alignment (*G. intestinalis* and *S. vortens* belong to the monophyletic group Diplomonadida). Final product sizes ranged from 284 to 846 bp in length (Table 2). Although our primers may not detect other *G. intestinalis* assemblages that might be present in these marine animals, we are interested in zoonoses for this particular study, and therefore focus on the A and B assemblages.

Reactions using external primers were set up in 25 µl volumes containing 10 mM Tris-HCl, 50 mM KCl, 2.5 mM $MgCl_2$, 0.08 mM dNTP, 0.25 µM primer, 0.11 µl (0.55 units) *Taq* DNA polymerase (Promega) and 1 to 5 µl of sample DNA. PCR conditions were an initial melting at 95°C for 2 min followed by 30 cycles of 95°C for 30 s, 50 to 55°C for 30 s and 72°C for 60 to 90 s, proceeded by a final extension at 72°C for 7 min. Products were separated on agarose gels and amplicons of the expected size were gel-extracted using the Qiagen gel extraction kit. Two microliters of extracted amplifica-

Table 2. Primers used for amplification of Assemblage A and B *Giardia intestinalis* sequences from marine animals. The open reading frame (ORF) identifications derive from the *Giardia lamblia* genome project (www.Giardiadb.org). Prefixes 'A' and 'B' denote Assemblage A and B specific primers. 'Diplo' and 'All' designate more general, non-assemblage specific primers used to increase the success of amplification for *gdh* and *mlh*, respectively. Assemblage B specifc *mlh* primers amplified both A and B assemblage sequences. Primers with an 'AL' prefix derive from Sulaiman et al. (2003) and amplified *tpi5'* locus for both assemblages. Product lengths are given in nucleotides (nt) and amino acid (aa) positions are based on complete coding sequences from *Giardia lamblia* genome, ATCC 50803

| Locus | Primers | Primer sequence | Product size (nt) | Position of aa |
|---|---|---|---|---|
| *gdh* (ORF 21942) | A_gdhF1 | 5′-GCCTGCCCAGACGATCGAGGAGC-3′ | 827 | 376–1200 |
| | A_gdhR1 | 5′-CGCAGCCCTGCTCGATCATA-3′ | | |
| | A_gdhF2 | 5′-CACGCTCCCGATGGGCGGCGGCAAG-3′ | | |
| | A_gdhR2 | 5′-CTCATGATGCCGCGGAGCTTC-3′ | | |
| | B_gdhF1 | 5′-CTTCGATCCTAAGGGCAAGTCGGAC-3′ | 704 | 412–1113 |
| | B_gdhR1 | 5′-CTCACGGAGTTCTGAGACATCTCGAGA-3′ | | |
| | B_gdhR2 | 5′-ACGGAAACACCACCGGCGTTGGCG-3′ | | |
| | Diplo_gdhF | 5′-CACGCTCCCGATGGGCGGCGGCAAG-3′ | | |
| | Diplo_gdhR | 5′-GARTTCTGGSACATCTCGAGGCCG-3′ | | |
| *mlh* (ORF 16149) | B_mlhF1 | 5′-CAGRGACCGTTTAACGCTGCCAAG-3′ | 846 | 127–969 |
| | B_mlhR1 | 5′-CACAGCCACCGGGTTCTGTGACAATC-3′ | | |
| | B_mlhF2 | 5′-CATTGAAAATGCRACTGACGCGGATG-3′ | | |
| | B_mlhR2 | 5′-CATCAATTACACTCTGAACAGCACTC-3′ | | |
| | All[e]_mlhF | 5′-CTGAYGCRGATGCWCCWRC-3′ | | |
| | All_mlhR | 5′-ATMCCTGCCTGAATGGTRCTA-3′ | | |
| *tpi5′* (ORF 93938) | AL_3543F | 5′-AAATIATGCCTGCTCGTCG-3′ | 490 | 37–525 |
| | AL_3546R | 5′-CAAACCTTITCCGCAAACC-3′ | | |
| | AL_3544F | 5′-CCCTTCATCGGIGGTAACTT-3′ | | |
| | AL_3545R | 5′-GTGGCCACCACICCCGTGCC-3′ | | |
| *tpi3′* (ORF 93938) | A_tpiF1 | 5′-CTATGTACGGGTCTTCGTAAG-3′ | 436 | 241–675 |
| | A_tpiR1 | 5′-CTGATTCCGTGGACGTCGTCAT-3′ | | |
| | A_tpiF2 | 5′-GAAGCCGTCAATATTCGGACAYTGICC-3′ | | |
| | A_tpiR2 | 5′-GCGTGGACTGGIGARACIAG-3′ | | |
| | B_tpiF1 | 5′-GAGCGTCGAGATGCTGCYGGACATG-3′ | 284 | 343–624 |
| | B_tpiR1 | 5′-GCAGTTACTCCCATTGGCAGAC-3′ | | |
| | B_tpiF2 | 5′-ATGAGCAGAGTGCTAAGAAGGCG-3′ | | |
| | B_tpiR2 | 5′-GCAGACCCTCCATAGATGATGCGG-3′ | | |

tion product served as DNA template for the nested PCR using internal primers with reactions conditions specified above. Gel-extracted products were cloned with the TOPO TA kit (Invitrogen) and plasmids were purified from positive clones using standard alkaline-lysis with a Biomek FX liquid handling robot (Beckman Coulter). Clones were sequenced in both directions using universal M13 primers, ABI BigDye 3.1 chemistry and a 3730xl 96 capillary array genetic analyzer (Applied Biosystems). The number of clones sequenced ranged from 24 to 96 per locus per sample.

**Sequence analysis.** A bioinformatics pipeline using the programs phred, cross_match, and phrap, translated chromatograms into base calls and associated quality scores, removed vector sequences and assembled forward and reverse reads into full length sequences for each of the cloned PCR amplicons (Ewing & Green 1998, Ewing et al. 1998). Sequences were aligned with ClustalW (Thompson et al. 1994, Maddison & Maddison 2001).

For each locus from each sample, sequence variants with high quality base calls (based on phred scores) were determined to be real (i.e. not due to PCR misincorporation error or PCR recombination) if one or more of the following criteria were met: (1) the variant was amplified multiple times from more than one PCR, (2) the variant was identical at all phylogenetically informative sites to an allele from a clonal line or a sequence variant that represented the only haplotype obtained from another sample, and (3) the variant constituted a majority of the sequence pool in a single PCR, where PCR recombinants could be easily identified due to low or nonexistent representation.

To confirm the accuracy of our sequence screening criteria, we performed additional PCR experiments on fecal samples using 2 techniques to minimize polymerase error and PCR recombination. The first technique involved reducing the number of amplification cycles to 30 for the initial reaction and 20 for the nested reaction. The second technique employed the use of

*PfuTurbo* DNA polymerase (Stratagene), which has a 3′ to 5′ proofreading activity, and 30 cycles for both initial and nested PCRs. Products from each reaction were cloned and 24 clones were sequenced in each direction. Sequences obtained with standard *Taq* polymerase amplification were no more variable than those obtained with either of these methods (data not shown).

Each confirmed sequence variant was considered to be a haplotype. While laboratory cultures or clonal lines can be assigned multi-locus haplotypes, it was not possible to accomplish this with our data because we amplified directly from fecal material without isolation of cysts. Therefore, *Giardia intestinalis* sequences might represent DNA derived from cysts of a single clone or from cysts of multiple genetically different individuals within a single host or sample.

Levels of intra- and inter-assemblage nucleotide variation and number of haplotypes were calculated using DnaSP v. 3.0 (Rozas & Rozas 1999). Phylogenies were reconstructed using parsimony searches in PAUP v. 4.0b10 (Swofford 2002), employing 10 random additions and TBR branch swapping, and assessed with 500 bootstrap replicates. For each locus, the bootstrap consensus tree represents the phylogeny. To assess the zoonotic potential of organisms present in marine animals we included in each analysis sequences from 2 Assemblage A and 3 Assemblage B human isolates that were grown clonally in axenic conditions. The human Assemblage A isolates represent 2 consistently defined subgroups: A1 and A2 (Wielinga & Thompson 2007). The Assemblage A1 isolate is referred to as Human A1 and is synonymous with the clonal isolate BAH2c2A (Meloni et al. 1995). The *gdh*–*mlh*–*tpi* haplotype for this isolate (accession nos. EF685690, EF687726 and EF688031) is found in 13 other human axenic A1 cultures, making it a good representative of this subgroup (note all accession numbers hereafter will be referred to in the order *gdh*–*mlh*–*tpi* and that for the human axenic isolates, *tpi* is comprised of both the *tpi5*′ and *tpi3*′ regions). The Assemblage A2 isolate, referred to here as Human A2, derives from the human clonal line BAH26c11A and has a haplotype (accession nos. EF685686, EF687722, and EF688048) representative of 3 other A2 human isolates. The 3 human Assemblage B isolates, here referred to as Human B1, B2 and B3, represent 3 different heterozygous clonal lines: (1) GS (ATCC no. 50583; accession nos. EF685679, EF687720, and EF688027 to EF688030); (2) BAH34c8B (accession nos. EF685680, EF687721 and EU662850, and EF688023); and (3) BAH12c14B (accession nos. EF685681 to EF685685, EF687716 to EF687719, and EF688024 to EF688026) (Meloni et al. 1995, Monis et al. 1999).

**Sequence data.** Sequences from this article have been deposited in the GenBank database under accession nos. EU362930 to EU362960 and EU362962 to EU362973 (*gdh* sequences), EU518499 to EU518525 (*mlh* sequences), EU518552 to EU518593 (*tpi5′* sequences) and EU518526 to EU518551 (*tpi3′* sequences).

## RESULTS

### Presence of Assemblage A and B sequence types in marine animals

Every sample that was positive for giardin contained sequences belonging to *Giardia intestinalis* Assemblages A and/or B (Table 1). All samples that were examined at all 4 loci contained sequences belonging to both assemblages (Table 1). Of the 12 samples only characterized at the *gdh* locus due to insufficient material, one contained sequences belonging to both assemblages, 10 contained one or more sequences belonging to Assemblage B and one contained a single sequence belonging to Assemblage A (Table 1). Though Assemblage B appears to have a higher prevalence in marine organisms than Assemblage A (22 versus 13 individuals), this may be due to incomplete screening or the presence of Assemblage A sequence variants that do not amplify with the primers we used.

### Genetic diversity of *Giardia intestinalis* from marine hosts

Our results indicate that within individual samples, there are often multiple haplotypes of each assemblage (A and B, Table 3). For samples subjected to the same depth of sequencing (in bold print in Table 3), herring gulls contained the greatest number of *Giardia intestinalis* haplotypes. In addition, many of the haplotypes found were identical to human isolates (Fig. 1).

Intra-assemblage variation for the entire dataset was also quite large (Table 4) and the number of haplotypes found for both assemblages surpasses the number previously recorded for datasets of equivalent or larger sample sizes (Wielinga & Thompson 2007). For the entire dataset, each locus identified an average of 21.25 haplotypes (sum of all haplotypes in Table 4/ number of loci [i.e. 85/4]). Even though the number of marine animals with Assemblage A haplotypes was lower than those with B, the total number of Assemblage A versus B haplotypes collected from our sampling efforts was roughly equivalent for each gene (Table 4). However, for each gene, the nucleotide diversity of Assemblage A was consistently lower than

Table 3. Number of Assemblage A and B haplotypes sequenced per locus for each sample. Samples in **bold** represent those probed using 4 locus PCR assay and correspond to samples 1 through 12 of Table 1. Samples not in bold are those only probed using *gdh* primers

| Sample | *gdh* A | *gdh* B | *mlh* A | *mlh* B | *tpi 5′* A | *tpi 5′* B | *tpi3′* A | *tpi3′* B |
|---|---|---|---|---|---|---|---|---|
| **Gull 1** | **1** | **0** | **0** | **3** | **0** | **1** | **1** | **0** |
| **Gull 3** | **1** | **0** | **2** | **0** | **3** | **3** | **3** | **1** |
| **Gull 11** | **1** | **0** | **1** | **1** | **4** | **1** | **2** | **0** |
| **Gull 12** | **2** | **2** | **4** | **4** | **2** | **2** | **2** | **3** |
| **Gull 13** | **2** | **1** | **2** | **1** | **1** | **2** | **2** | **2** |
| **Gull 16** | **2** | **2** | **2** | **2** | **3** | **3** | **2** | **0** |
| **Eider NE2** | **1** | **1** | **0** | **1** | **0** | **0** | **0** | **0** |
| **Eider NE3** | **0** | **1** | **0** | **0** | **0** | **4** | **3** | **0** |
| Eider NE4 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Seal 10A** | **3** | **3** | **0** | **0** | **0** | **3** | **0** | **0** |
| Seal 24A | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Dolphin 13** | **2** | **1** | **2** | **1** | **1** | **4** | **1** | **1** |
| Dolphin 14A | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Dolphin 17B | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Dolphin 23A | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Dolphin 20A | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Dolphin 30A | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Dolphin 33A | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Dolphin 316** | **0** | **0** | **0** | **1** | **0** | **3** | **1** | **1** |
| Porpoise 26A | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Porpoise 37A | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Porpoise 38B | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| Shark 68R | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |

that of Assemblage B, with estimates of nucleotide polymorphism ranging from 0.0037 to 0.0099 and 0.0051 to 0.0102 for pi ($\pi$) and theta ($\theta$), respectively, for Assemblage A and from 0.0052 to 0.0209 and 0.0072 to 0.0189 for $\pi$ and $\theta$ for Assemblage B (Table 4). All loci reflect large inter-assemblage genetic distances between Assemblage A and B, with values ranging from 12 to 14% (data not shown).

## Locus comparison

Based on samples where all 4 loci were sequenced (samples no. 1 through 11, Table 1 and samples in bold type, Table 3), the *mlh* locus detected the greatest number of haplotypes (16) while the 2 *tpi* loci detected the fewest number of haplotypes (13 and 18, respectively, for *tpi5′* and *tpi3′*). Although the differences in variation detected by *mlh* and *tpi* may simply reflect the greater length of the region examined for *mlh*, estimates of diversity per nucleotide were lower for *tpi* as well (Table 4).

Despite the somewhat different estimates of haplotype diversity each locus provides, each locus suggests that there is a considerable amount of variation within Assemblage A and B within marine hosts. For example, no sample contained a shared Assemblage B *mlh* haplotype and the majority of the haplotypes sequenced at each locus have never been sequenced before (Fig. 1).

## Phylogenies

We used *gdh*, *mlh*, *tpi5′*, and *tpi3′* sequences to reconstruct phylogenetic relationships among *Giardia intestinalis* haplotypes from marine animal fecal samples and haplotypes from 3 Assemblage B isolates and 2 Assemblage A isolates from humans (Fig. 1). At each locus, fecal samples displayed previously undescribed sequence variants (Fig. 1) and some clades were solely composed of sequences from marine hosts, such as the Assemblage B dolphin/porpoise/seal clade in the *gdh* tree (Fig. 1a) and the Assemblage B dolphin/gull/eider clade in the *tpi5′* tree (Fig. 1c). Sequences identical to haplotypes from human isolates were found in 18 of the 23 animals sampled (including at least one from each species), and several had more than one human-isolate haplotype (e.g. gull 3, Fig. 1). Not only did marine animals and humans have common *G. intestinalis* haplotypes, marine animals also shared haplotypes with each other, such as the *gdh* haplotype shared amongst porpoises 37A and 38B, dolphins 30A, 34A and 33A, and seal 10A (Fig. 1a). Both the number of haplotypes found and the number of shared haplotypes amongst marine hosts, regardless of species type, increased as a function of the depth of sequencing from each sample. For instance, gull and dolphin hosts contained the greatest number of *G. intestinalis* haplotypes in this dataset (Table 3) as well as the greatest number of shared haplotypes within and between

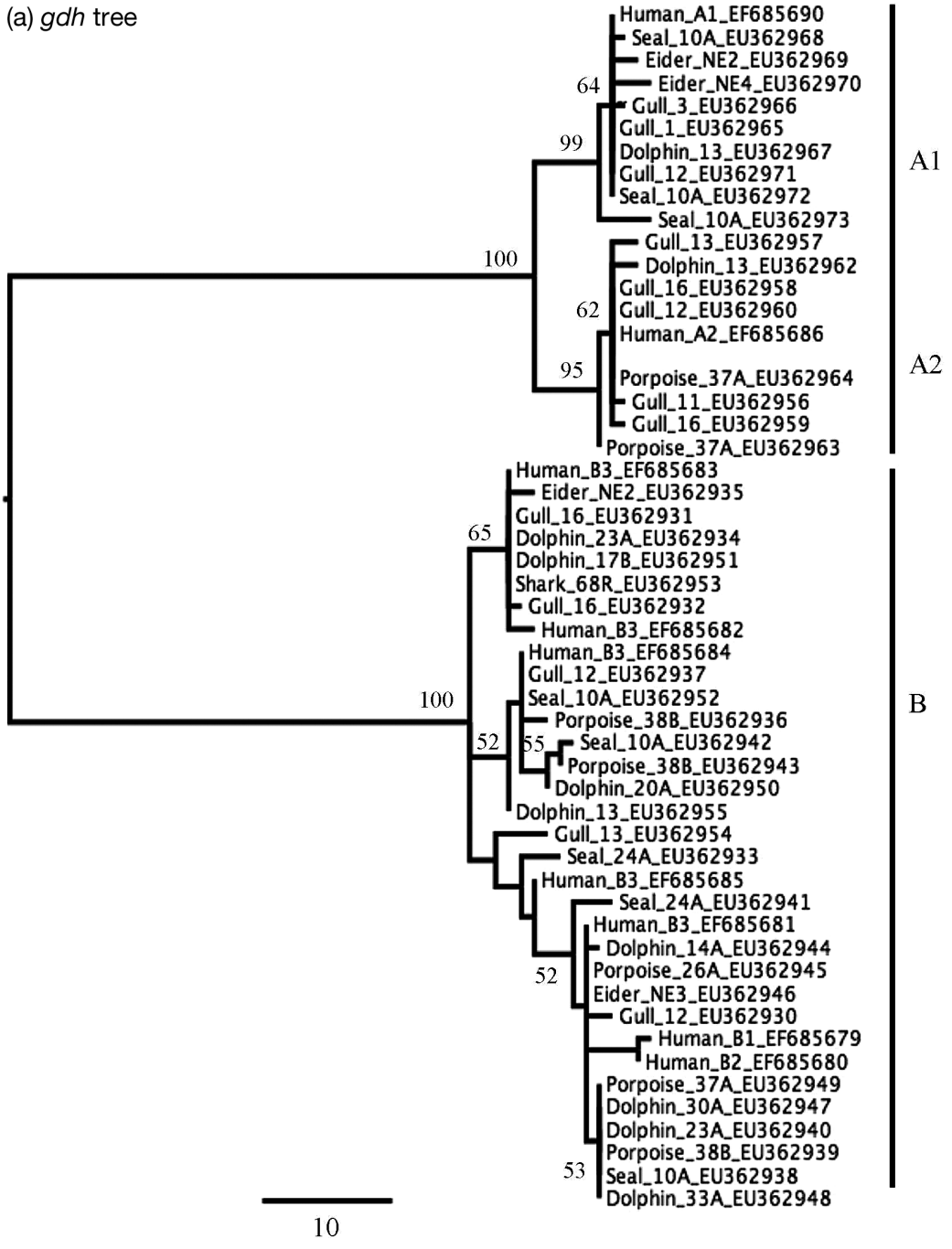


Fig. 1. Phylogenetic trees representing relationships among *Giardia intestinalis* from marine animals and humans based on (a) *gdh*, (b) *mlh*, (c) *tpi5'* and (d) *tpi3'* sequences. Each tree describes 2 well-resolved assemblages, A and B, with Assemblage A being further divided into A1 and A2 clades. Consensus trees were built from parsimony searches of 500 bootstrap replicates. Branch lengths correspond to the number of steps. Bootstrap values less than or equal to 50 are not shown

their respective populations (Fig. 1). Gulls and dolphins were also the animals most extensively examined for *G. intestinalis*.

Phylogenies revealed that Assemblage A is divided into 2 clades/subgroups separated by small genetic distances (Fig. 1). These 2 clades represent previously documented subgroups of Assemblage A (Wielinga & Thompson 2007), termed cluster A1 and A2, which were consistently resolved for all loci, albeit with low support for all genes except *gdh* (Fig. 1a). Only at the *tpi5'* locus did there appear to be an intermediate form of the A1 and A2 clades (Fig. 1c).

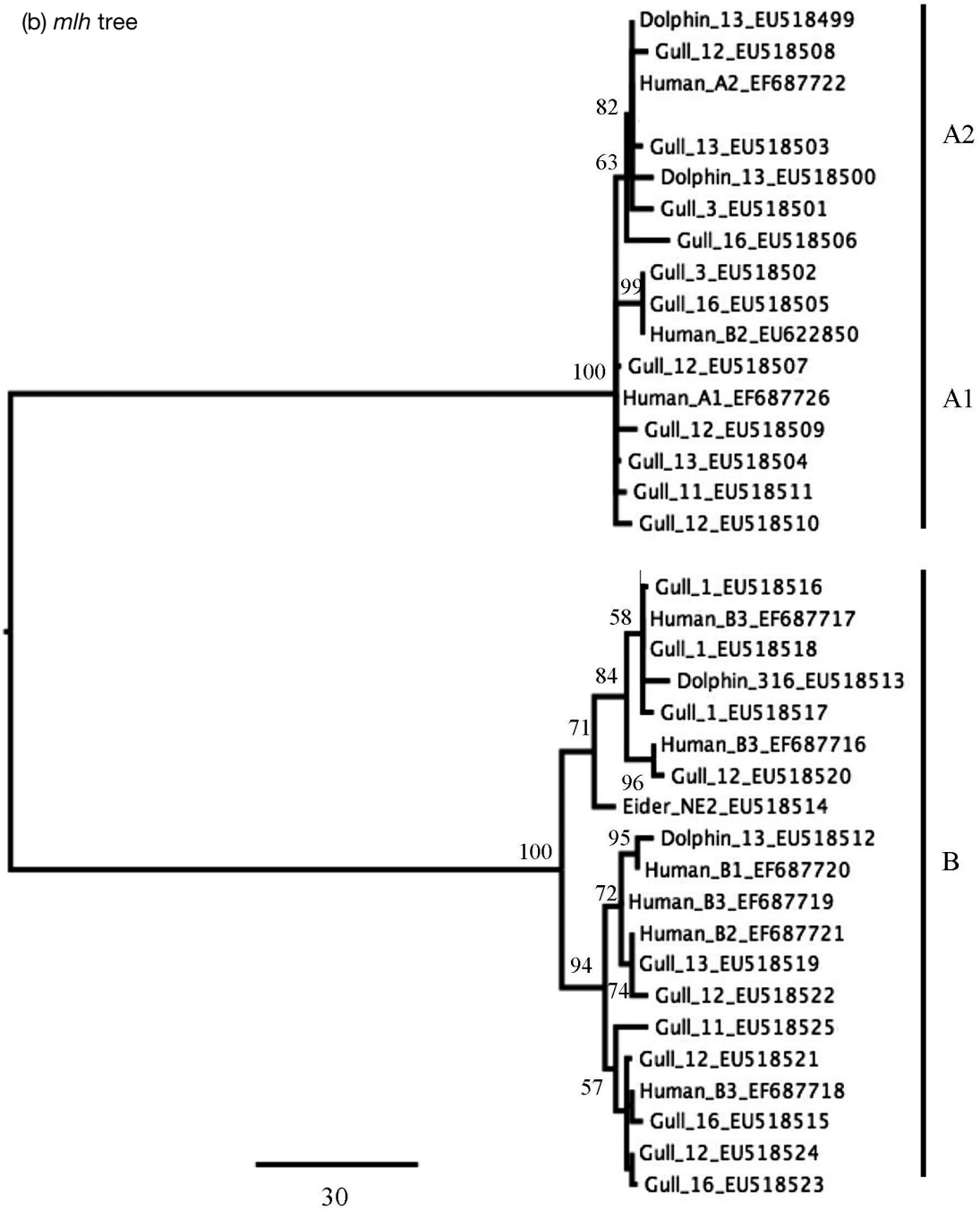


Fig. 1 (continued)

## DISCUSSION

### Zoonoses in marine animals

All *Giardia intestinalis* sequences recovered from the 23 marine hosts examined, including birds, mammals and fish, belong to Assemblages A or B, which are considered zoonotic (Table 1). Marine animals not only harbor potentially zoonotic forms of *G. intestinalis*, but also often show the presence of multiple genetically distinct zoonotic sequence types at all 4 loci (Table 3, Fig. 1). More importantly, all 4 loci yield *G. intestinalis* haplotypes from marine hosts that are genetically identical to human isolates, strengthening the argument for zoonoses in marine animals (Fig. 1). Although the sample size is too low to be conclusive, gulls appear to share the greatest number of *G. intestinalis* haplotypes in common with humans, which implies that the risk of

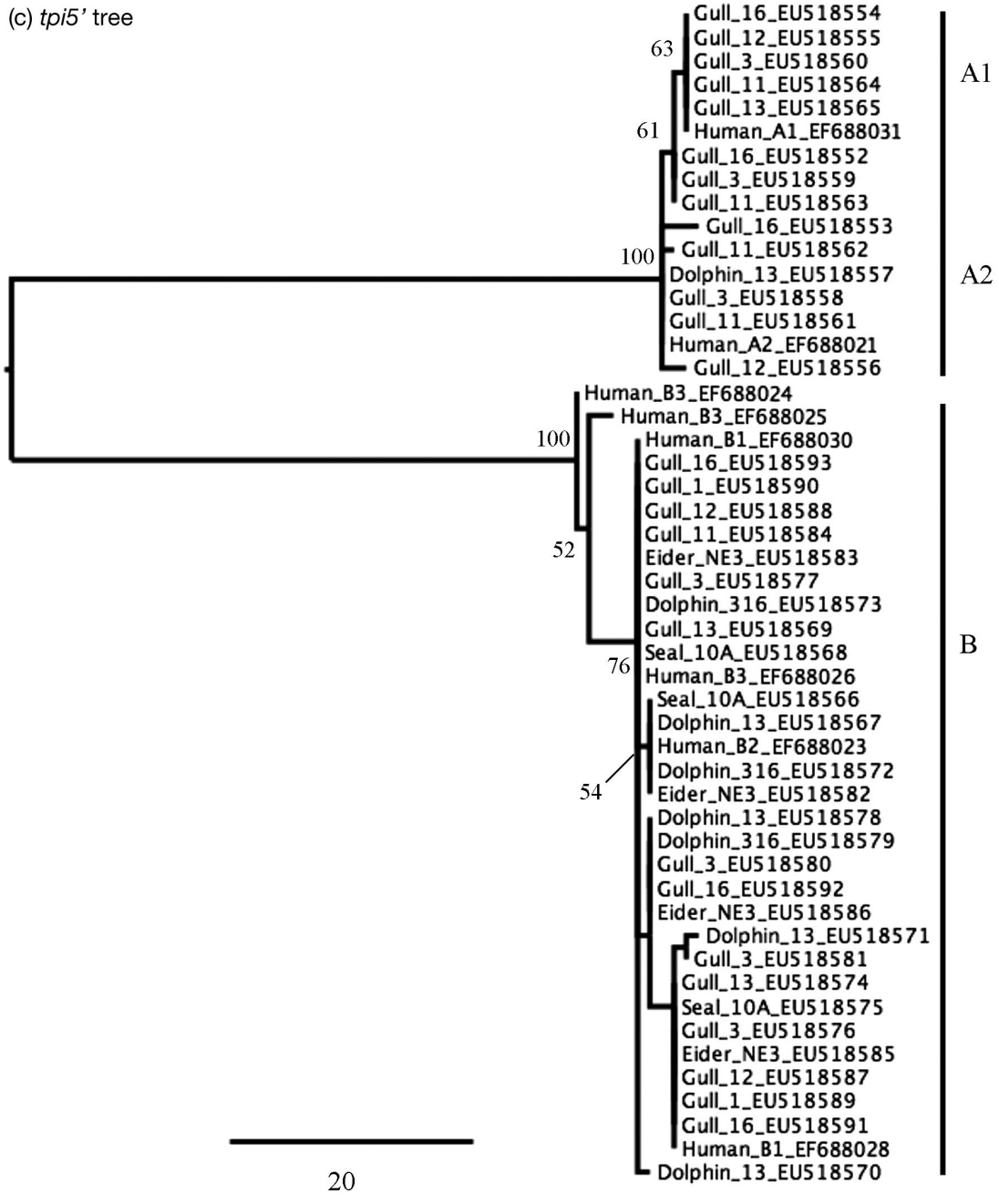


Fig. 1 (continued)

zoonosis is reflected in the degree to which humans associate with other animals, where humans more closely associate with gulls than any of the other animals considered in this study. In other words, the degree to which wildlife interacts with humans and encounters human waste (such as gulls feeding at dump sites or sewage treatment plants) may increase the chances of serving as a reservoir for human disease.

Although suggestive of infection, presence of Assemblage A and B sequences in the fecal material of marine animals does not conclusively demonstrate that they serve as hosts for zoonoses. For example, seabirds can serve as mechanical vectors, transferring cysts between terrestrial and marine environments (Graczyk et al. 2008), or *Giardia intestinalis* cysts might passively transfer through the digestive tract of marine

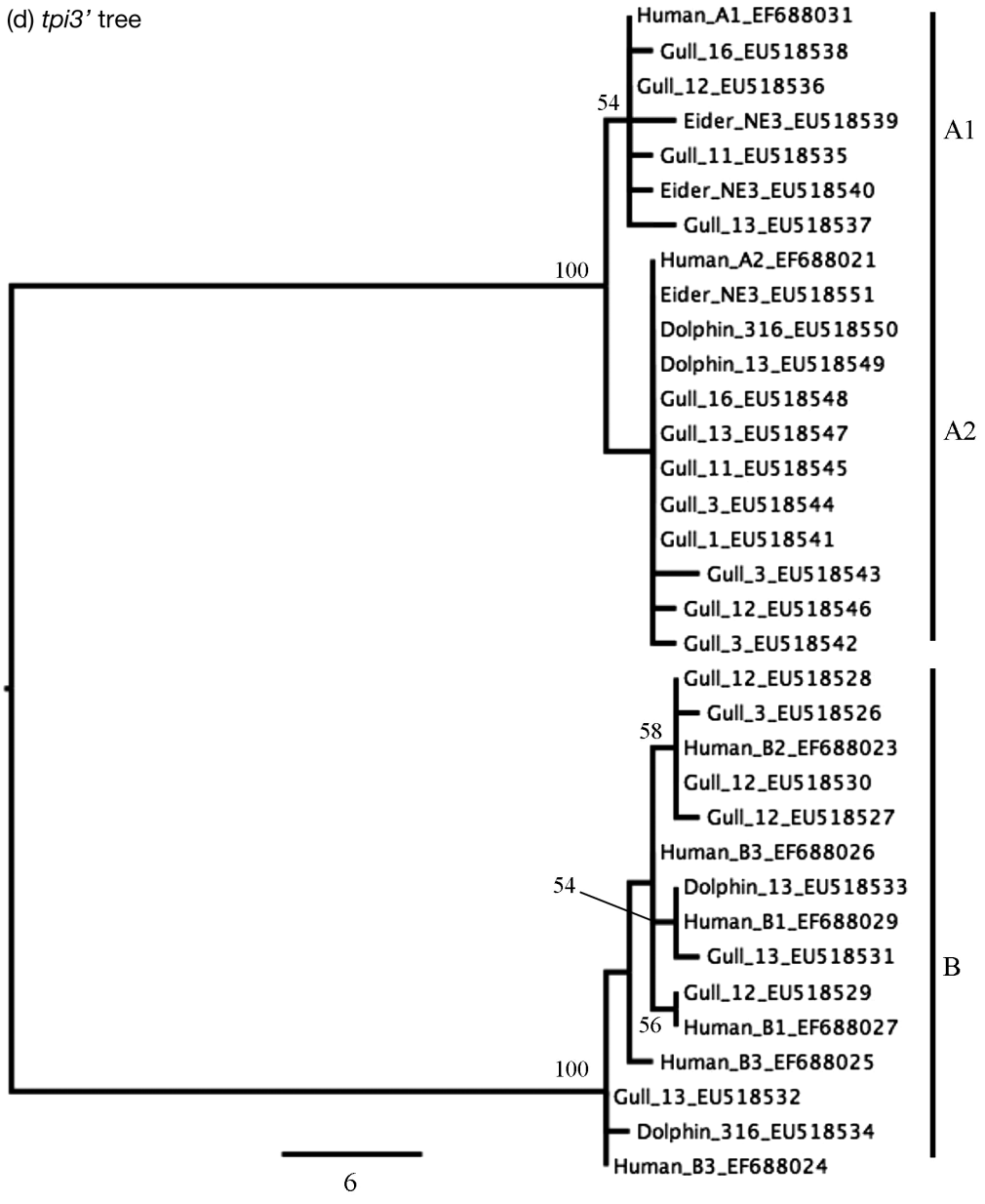


Fig. 1 (continued)

animals without infecting the animal itself. Future studies that determine the viability and concentration of cysts within marine animal feces will provide clues as to whether marine animals serve as mechanical vectors or hosts to zoonoses. For example, low cyst concentrations within feces might be associated with either passive transfer or infection with marine animals, while high cyst concentrations suggest infection and growth of *G. intestinalis* with these organisms.

Regardless of whether marine animals passively transfer *Giardia intestinalis* or actually harbor infection, the fact that marine animals are not restricted from acquiring any single zoonotic variant has many health implications for humans, including a greater risk of acquiring this parasite from marine hosts. The high incidence of shared *G. intestinalis* haplotypes among different marine species suggests that this parasite is easily transmitted between different marine communities and may contribute to a rapid spread of this parasite within and between marine and terrestrial environments. If the different haplotypes recovered from a single marine animal represent genetically distinct isolates (and if *G. intestinalis* is completing its life cycle within marine animals) there is also the opportunity for genetic exchange, which increases virulence in some parasites such as *Toxoplasma* (Grigg & Suzuki

Table 4. Total number of Assemblage A and B haplotypes found per locus and 2 estimates of nucleotide polymorphism, pi (π) and theta (θ). Assemblage B displays a greater amount of genetic variation than Assemblage A regardless of the locus examined or the method used to calculate variation (π vs. θ)

| Locus | Assemblage A | | | Assemblage B | | |
|---|---|---|---|---|---|---|
| | No. | π | θ | No. | π | θ |
| *gdh* | 12 | 0.0099 | 0.0102 | 16 | 0.0125 | 0.0115 |
| *mlh* | 12 | 0.0094 | 0.0152 | 14 | 0.0209 | 0.0189 |
| *tpi5′* | 6 | 0.0037 | 0.0051 | 7 | 0.0052 | 0.0072 |
| *tpi3′* | 10 | 0.0056 | 0.0088 | 8 | 0.0084 | 0.0104 |

2003). Both the health risk that marine animals infected with *G. intestinalis* pose to humans and the effects that this parasite has on marine species require further evaluation.

Our data suggest that anthropogenic activity could increase the role of marine animals as reservoirs of disease in humans. In the case of *Giardia intestinalis*, marine animals might acquire this parasite from contaminated human waste or other biological vectors (e.g. marine birds). The disease then passes back to humans through exposure to contaminated fecal material from marine animals or from interaction with seabirds transmitting the parasite from a marine to a terrestrial environment. Although the concentration of *G. intestinalis* cysts in the water required to pose a serious threat to humans needs to be assessed, the contamination of marine environments with human diseases should be viewed with concern considering that a flock of seabirds can deposit $1.0 \times 10^7$ cysts into the water during an average visitation and ingesting as few as 10 cysts can lead to infection in humans (Graczyk et al. 2008).

Agricultural run-off provides the opportunity for marine animals to acquire other *Giardia intestinalis* assemblages, such as those that infect livestock (Assemblage E) or dogs (Assemblages C and D). Because our research focused on zoonoses, we did not evaluate the status of Assemblages C through G (which are considered non-zoonotic) in marine animals. However, to gain a more complete picture of the potential host range and distribution of these assemblages, future investigations should employ Assemblage C through G specific primers to determine the presence of these groups within marine animals.

## Genetic variation

There are 2 levels of variation that require explanation in this study: (1) haplotype variation within and between samples and (2) intra- and inter-assemblage variation taken from the entire dataset. The locus analyzed influences the amount of variation observed at both levels. The differences in variation that each locus estimates are in turn influenced by the degree to which each gene is evolutionarily conserved, the depth of sequencing and the length of the gene fragment examined.

Using 4 loci to analyze fecal samples revealed a previously unrecorded genetic complexity to *Giardia intestinalis* populations within and between wildlife species. The prevalence of multiple *G. intestinalis* haplotypes identified within a sample regardless of the species indicates that this might be a common characteristic for wildlife harboring this parasite and marine animals are particularly susceptible to acquiring *G. intestinalis*.

The discovery of multiple alleles within clonal human lines revealed that some *Giardia intestinalis* isolates do not contain a completely homogenous genome (Teodorovic et al. 2007). This supports the hypothesis that while many of the different haplotypes we found within a single sample most probably originate from a population of genetically distinct homozygous individuals, some are probably alleles from heterozygous individuals. Although we cannot determine whether the haplotypic variation within a sample represents different clonal lines of *G. intestinalis* or a single highly heterozygous individual due to the nature of this study, if each sample contained a single isolate, then rampant genetic exchange between individuals would be required to account for the distribution of haplotypes of different loci (Fig. 1). Thus, at least some of the diversity observed is probably due to the presence of multiple isolates within each sample. However, the identification of Assemblage A alleles within Assemblage B human (Teodorovic et al. 2007) isolates suggests that some of the co-occurrences of A and B haplotypes within our marine samples originate from assemblage hybrids, potentially produced by genetic exchange. This would alter our current perception of whether genetic exchange occurs in *G. intestinalis* and how frequently it takes place, but alternative methods, such as *in situ* hybridizations of isolated cysts, are needed to test this hypothesis.

Although the number of Assemblage A haplotypes that each locus identified was roughly equivalent to the number of Assemblage B haplotypes, intra-assemblage variation was lower in Assemblage A and resolved into only 2 subgroups (A1 and A2). The fact that Assemblage B haplotypes did not consistently resolve into well-defined subgroups, with each locus presenting a different tree topology, suggests that (1) we have not fully characterized the genetic variation within this assemblage and (2) we cannot rely on a single gene to reconstruct accurate relationships

among *Giardia intestinalis* from different hosts. The large genetic distances that separate Assemblage A and B haplotypes reflect either an ancient divergence between these 2 groups or rapid evolution. Regardless, further research should be performed to understand how such genetically different individuals are capable of infecting virtually all mammals and other vertebrates while Assemblages C through G appear to be host specific.

Working with fecal and environmental samples presents many challenges. Failure to obtain a PCR product from a sample might reflect the absence of the parasite within the sample, inhibition of PCR by contaminants, or the use of inappropriate primers that do not detect the specific variant present within the sample. Using a multi-locus approach to examine fecal samples improved detection of *Giardia intestinalis* and allowed a more thorough assessment of genetic variation. For example, using only the *tpi5′* locus, one would conclude that the Assemblage B clade from marine animals is relatively undifferentiated. However, the *mlh* locus alters this conclusion, as each Assemblage B sequence isolated represents a different haplotype. Because the *mlh* locus showed the greatest amount of sequence divergence we encourage continuing its use as a more sensitive tool for genotyping than the loci currently used. Additionally, our study indicates that a robust and accurate assessment of genetic diversity within a sample requires cloning and sequencing of amplicons rather than direct sequencing of PCR product. While previous studies that relied on direct sequencing of PCR product only resolve one haplotype per sample (or a consensus of multiple sequences), we show that samples often harbor several haplotypes that can only be resolved through sequencing cloned amplicons. Also, replication of PCR amplifications is necessary to confirm the validity of each variant. Frequently, sequencing cloned amplification product from a replicate reaction yields new haplotypes, which emphasizes that fecal material is a complex, sometimes non-homogenous medium requiring thorough analysis. To begin to understand the population biology of *G. intestinalis*, we need to implement better approaches to detect diversity, such as sequencing clones, and better ways of resolving diversity, such as assigning multi-locus haplotypes to individual isolates. Assigning multi-locus haplotypes will require sequencing more loci in conjunction with other techniques, such as culturing and *in situ* hybridization of isolated cysts. Our results detail our attempts to address these issues and to call attention to *G. intestinalis* in previously unconsidered ecosystems, such as the oceans and in the animals residing in them.

*Acknowledgements*. We thank N. Wheelwright and R. Mauck at the Kent Island field station (Bowdoin College, Maine), K. Ampela for seal scat, G. Skomal for the thresher shark sample, the NOAA NEFSC Observer program, the Cape Cod Stranding Network, the Centers for Ocean and Human Health (COHH), and comments from 3 anonymous reviewers that improved the manuscript. This paper is a result of research funded under the following awards: NOAA Coastal Ocean Program award no. NA05NOS4781247, the NOAA Prescott Program award no. NA06NMF4390130, the COHH award no. NIEHS P50ES012742, and the National Science Foundation OCE award no. 0430724 given to Woods Hole Oceanographic Institution, Woods Hole, Massachusetts and the National Institutes of Health award no. AI0580C4 'Molecular Evolution of Eukaryotes,' given to the Marine Biological Laboratory, Woods Hole.

## LITERATURE CITED

Adam RD (2001) Biology of *Giardia lamblia*. Clin Microbiol Rev 14:447–475

Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res 8:186–194

Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 8:175–185

Fayer R, Dubey JP, Lindsay DS (2004) Zoonotic protozoa: from land to sea. Trends Parasitol 20:531–536

Graczyk TK, Majewska AC, Kellogg JC (2008) The role of birds in dissemination of human waterborne enteropathogens. Trends Parasitol 24:55–59

Grigg ME, Suzuki Y (2003) Sexual recombination and clonal evolution of virulence in *Toxoplasma*. Microbes Infect 5:685–690

Hughes-Hanks JM, Rickard LG, Panuska C, Saucier JR, O'Hara TM, Dehn L, Rolland RM (2005) Prevalence of *Cryptosporidium* spp. and *Giardia* spp. in five marine mammal species. J Parasitol 91:1225–1228

Maddison DR, Maddison WP (2001) MacClade 4: analysis of phylogeny and character evolution. Sinauer Associates, Sunderland, MA

Mahbubani MH, Bej AK, Perlin MH, Schaeffer FW 3rd, Jakubowski W, Atlas RM (1992) Differentiation of *Giardia duodenalis* from other *Giarda* spp. by using polymerase chain reaction and gene probes. J Clin Microbiol 30:74–78

Meloni BP, Lymbery AJ, Thompson RCA (1995) Genetic characterization of isolates of *Giardia duodenalis* by enzyme electrophoresis: implications for reproductive biology, population structure, taxonomy, and epidemiology. J Parasitol 81:368–383

Monis PT, Andrews RH, Mayrhofer G, Ey PL (1999) Molecular systematics of the parasitic protozoan *Giardia intestinalis*. Mol Biol Evol 16:1135–1144

Morrison HG, McArthur AG, Gillin FD, Aley SB and others (2007) Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia*. Science 28:1921–1926

Ramesh MA, Malik SB, Longsdon JMJ (2005) A phylogenomic inventory of meiotic genes; evidence for sex in Giardia and an early eukaryotic origin of meiosis. Curr Biol 15:185–191

Robertson LJ (2007) The potential for marine bivalve shellfish to act as transmission vehicles for outbreaks of protozoan infections in humans: a review. Int J Food Microbiol 120: 201–216

Rozas J, Rozas R (1999) DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15:174–175

Sulaiman IM, Fayer R, Caryn B, Gilman RH and others (2003) Triosephosphate isomerase gene characterization and potential zoonotic transmission of *Giardia duodenalis*. Emerg Infect Dis 9:1444–1452

Swofford DL (2002) PAUP*. Phylogenetic analysis using parsimony (*and other methods), v4. Sinauer Associates, Sunderland, MA

Teodorovic S, Baverman JM, Elmendorf HG (2007) Unusually low levels of genetic variation among *Giardia lamblia* isolates. Eukaryot Cell 6:1421–1430

Thompson RCA (2004) The zoonotic significance and molecular epidemiology of *Giardia* and giardiasis. Vet Parasitol 126:15–34

Thompson RCA, Monis PT (2004) Variation in *Giardia*: implications for taxonomy and epidemiology. Adv Parasitol 58:69–137

Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Thompson RCA, Hopkins RM, Homan WL (2000) Nomenclature and genetic groupings of *Giardia* infecting mammals. Parasitol Today 16:210–213

Traub RJ, Monis PT, Robertson I, Irwin P, Mencke N (2004) Epidemiological and molecular evidence supports the zoonotic transmission of *Giardia* among humans and dogs living in the same community. Parasitology 128: 253–262

Wielinga CM, Thompson RCA (2007) Comparative evaluation of *Giardia duodenalis* sequence data. Parasitology 134: 1–27

*Editorial responsibility: John Austin, Oldendorf/Luhe, Germany*

*Submitted: December 17, 2007; Accepted: April 30, 2008*
*Proofs received from author(s): July 21, 2008*