Published in final edited form as:

Electrophoresis. 2011 August ; 32(16): 2075–2090.

Barcoding of *Giardia duodenalis* isolates and derived lines from an established cryobank by a mutation scanning-based approach

Matthew J. Nolan¹, Aaron R. Jex¹, Jacqui Upcroft², Peter Upcroft², and Robin B. Gasser¹

¹Department of Veterinary Science, The University of Melbourne, Parkville, Victoria 3052, Australia

²Queensland Institute of Medical Research, Brisbane, Queensland 4006, Australia

Abstract

We barcoded 25 *in vitro* isolates (representing 92 samples) of *Giardia duodenalis* from humans and other animals, which have been assembled by the Upcroft team at the Queensland Institute of Medical Research over a period of almost three decades. We used mutation scanning-coupled sequencing of loci in the triosephosphate isomerase (*tpi*), glutamate dehydrogenase (*gdh*) and beta-giardin (*bg*) genes, combined with phylogenetic analysis, to genetically characterise them. Specifically, the isolates (n = 14) of *G. duodenalis* from humans from Australia (AD113; BRIS/83/HEPU/106; BRIS/87/HEPU/713; BRIS/89/HEPU/1003; BRIS/91/HEPU/1279; BRIS/92/HEPU/1541; BRIS/92/HEPU/1590; BRIS/92/HEPU/2443; BRIS/93/HEPU/1706), Malaysia (KL/92/IMR/1106) and Afghanistan (WB), a cat from Australia (BAC2), a sheep from Canada (OAS1) and a sulphur-crested cockatoo from Australia (BRIS/95/HEPU/2041) represented assemblage A (sub-assemblage AI-1, AI-2 or AII-2); isolates (n = 10) from humans from Australia (BRIS/91/HEPU/1279; BRIS/92/HEPU/2342; BRIS/92/HEPU/2348; BRIS/93/HEPU/1638; BRIS/93/HEPU/1653; BRIS/93/HEPU/1705; BRIS/93/HEPU/1718; BRIS/93/HEPU/1727), Papua New Guinea (BRIS/92/HEPU/1487) and Canada (H7) represented assemblage B (sub-assemblage BIV); and an isolate from cattle from Australia (BRIS/92/HEPU/1709) had a match to assemblage E. Isolate BRIS/90/HEPU/1229 from a human from Australia was shown to represent a mixed population of assemblages A and B. These barcoded isolates (including stocks and derived lines) now allow direct comparisons of experimental data among laboratories and represent a massive resource for transcriptomic, proteomic, metabolic and functional genomic studies using advanced molecular technologies. (Nucleotide sequences reported in this paper are available in the GenBank database under accession nos. XXXXX-XXXXX).

Keywords

Giardia duodenalis; genetic characterization; barcoding; triose-phosphate isomerase; glutamate dehydrogenase gene and beta-giardin genes; mutation scanning; single-strand conformation polymorphism (SSCP); restriction endonuclease fingerprinting (REF)

1 Introduction

Giardia (Zoomastigophorea) is a group of anaerobic parasitic protists that infect the small intestines of terrestrial and marine vertebrates [1, 2]. Humans usually acquire infection following the ingestion of infective cysts (as few as 10 cysts³; ref. [3]), either directly by human-to-human transmission or indirectly from contaminated food or water [4-6]. Although giardial infection can be asymptomatic, the clinical manifestation of the disease (= giardiasis) in immuno-competent individuals includes watery diarrhoea, epigastric pain, nausea, headache and/or weight loss, with symptoms often appearing 6-15 days following the ingestion of cysts [7-9]. Clinical signs can vary among individuals and can be linked to the infective dose [10], genotype of *Giardia* [11-13] and/or host factors, such as age, developmental, nutritional and/or immunological status [10, 14]. The infection is often acute and short-lived (days to weeks), and is usually eliminated following the development of an effective host immune response [8, 10, 15]. However, chronic disease can occur in young, elderly and immuno-compromised/-suppressed patients [10, 15, 16], causing malabsorption, severe weight loss and/or failure to thrive [8].

Within the genus *Giardia*, *G. duodenalis* (syn. *G. lamblia* and *G. intestinalis*) is the only species known to infect humans and ~ 40 other species of mammalian host [17], is distributed globally and is the most common, non-bacterial and non-viral agent of diarrhoeal disease worldwide, causing an estimated 2.8×10^8 cases each year [14, 18]. Given the major medical importance of giardiasis in humans and other animals, the zoonotic nature and the apparently variable virulence/pathogenicity of some genotypes of *G. duodenalis* and the significant problems associated with drug resistance in *Giardia* populations [19-22], there has been a major effort by a number of laboratories around the world to establish and maintain *in vitro* cultures of *G. duodenalis* isolates and their derived stocks and lines, in order to explore their biological, biochemical and genetic features [6]. The availability of these banks of *G. duodenalis* provides a unique resource to researchers, but it is imperative that isolates, stocks and lines be defined genetically. Originally, *in vitro* cultured isolates were characterised extensively using conventional microscopic (*e.g.*, staining and immunological analyses) [23], biochemical (*e.g.*, multilocus enzyme electrophoresis) [24] and/or 'phenetic' electrophoretic tools [25-27]. Interestingly, in spite of the extensive experimental data available for many isolates or lines in these culture banks, few have been characterised systematically using common genetic markers (including the small subunit [SSU] of nuclear ribosomal RNA [rRNA], triose-phosphate isomerase [*tpi*], glutamate dehydrogenase [*gdh*], beta-giardin [*bg*] genes; refs. [28, 29]), now commonly utilised to assign assemblages (A-H) [6, 14, 30] by PCR-coupled methods. The characterization of individual isolates, stocks and derived lines employing these methods is important to allow direct genetic comparisons with those (cultured or derived from cysts from infected hosts) for which extensive sequence data are available in the published literature and in public gene databases (such as GenBank).

The Upcroft research team in the Queensland Institute of Medical Research (QIMR) assembled, over a period of almost three decades, a unique and large bank of cultured isolates, stocks and derived lines. Many of them have been very well characterised in terms of *in vivo* and *in vitro* growth rates, drug resistance [31], karyotype [32], metabolic activity [33], virulence and pathogenicity [11]. Here, we conducted a systematic genetic characterization (= barcoding) of DNAs from 25 *G. duodenalis* isolates (representing 92 samples) employing mutation scanning-coupled sequencing of three of the key loci (in *tpi*, *gdh* and *bg*), to allow future inter-laboratory comparisons of experimental results from, for example, transcriptomic, genomic, proteomic and metabolic studies.

2 Materials and methods

2.1 *Giardia duodenalis*

Here, we define an ‘isolate’ as a population of *Giardia* following one *in vitro* or *in vivo* passage; a ‘stock’ as an isolate following multiple, serial passages *in vitro* and *in vivo*, without an implication of homogeneity or characterisation (in this case, not characterised genetically employing the current approach); and, a ‘line’ as being derived from a stock following *in vitro* or *in vivo* maintenance under different physical conditions (*e.g.*, drug exposure) (*cf.*[34]).

A bank of 92 genomic DNA samples representing 25 *G. duodenalis* isolates, collected and routinely passaged in culture since 1983, was available (see Table 1). Most isolates were originally cultured from duodenal aspirates or excysted cysts derived from faecal samples from humans, but also from animals (including sheep, cattle, cat and sulphur-crested cockatoo) from Australia (Queensland, New South Wales, Victoria, South Australia and Western Australia), Papua New Guinea, Malaysia, Afghanistan and Canada. Some genomic DNAs were from isolate-derived lines of *G. duodenalis* induced to become drug resistant *in vitro* against albendazole [PubChem number CID 2082], metronidazole [CID 4173], furazolidone [CID 5323714] or quinacrine [CID 237] (PubChem: <http://pubchem.ncbi.nlm.nih.gov>) (Table 1).

2.2 Polymerase chain reaction (PCR) amplification of three independent genetic loci

Genomic DNA was subjected to PCR amplification of regions of the *tpi* (designated *ptpi*; ~530 bp), *gdh* (= *pgdh*; ~430 bp) and *bg* (= *pbg*; ~390 bp) genes. These three loci were employed because they have been reported to be well suited for the assigning of *G. duodenalis* isolates to assemblages and sub-assemblages [28, 35–37] and because substantial sequence data are publicly available for them in current databases for comparative analysis, in contrast to other genes, such as SSU and elongation-factor 1- α (*efl-1*). PCRs were carried out in 50 μ l volumes containing 10 mM Tris-HCl (pH 8.4), 50 mM KCl (Promega, Madison, USA), 3.0 mM MgCl₂, 200 μ M of each deoxynucleotide triphosphate (dNTP), 100 pmol of each primer and 1.25 U of GoTaq polymerase (Promega). The *ptpi* locus was amplified using primers AL3544 (forward; 5'-CCCTTCATCGGTGGTAACTT-3') and AL3545 (reverse; 5'-GTGGCCACCACTCCCGTGCC-3') [36], employing a cycling protocol of 94 °C/5 min (initial denaturation), followed by 35 cycles of 94 °C/20 s (denaturation), 55 °C/20 s (annealing), and 72 °C/20 s (extension) and a final extension of 72 °C for 10 min. The *pgdh* locus was amplified using primers GDHiF (forward; 5'-CAGTACAACCTCYGCTCTCGG-3') and GDHiR (reverse; 5'-GTTRTCCTTGACACATCTCC-3') [38]. This locus was amplified employing a cycling protocol of 94 °C/5 min (initial denaturation), followed by 35 cycles of 94 °C/20 s (denaturation), 62 °C/20 s (annealing), and 72 °C/20 s (extension) and a final extension of 72 °C for 10 min. The *pbg* locus was amplified using primers GBGF2 (forward; 5'-CTCGTTCGAGGACATCCG-3') and GBGR (reverse; 5'-CGAGCTTCGTGTTTGTGA-3'); primer GBGF2 is located between nucleotide positions 239 and 256, and GBGR is between positions 616 and 633 (linked to the reference sequence with GenBank accession number M36728; ref [39]). PCR of the latter locus employed a cycling protocol of an initial denaturation cycle of 94 °C/5 min, followed by 35 cycles of 94 °C/20 s (denaturation), 60 °C/20 s (annealing) and 72 °C/20 s (extension), and a final extension of 72 °C for 10 min.

2.3 Mutation scanning, sequencing and phylogenetic analyses

Single-strand conformation polymorphism (SSCP)-based methods were used for the analysis of amplicons, in order to achieve maximum mutation detection capacity [40]. For *pgdh* and

pbg amplicons (< 500 bp in size), direct SSCP analysis was used to scan for sequence variation within and among amplicons, according to an established method [40]. For *ptpi* amplicons (> 500 bp), restriction endonuclease fingerprinting (REF) [41, 42] was employed. For REF, 10 µl (100-150 ng) of each amplicon were digested in 20 µl with 10 U of the restriction endonuclease *HaeIII* (Promega) at 37 °C for 16 h. This endonuclease was selected based on its theoretical ability to cleave (at least once) *ptpi* sequences representing the *G. duodenalis* assemblages A-G. After digestion, the endonuclease was inactivated at 70 °C/5 min, and 4. µl of each digest run on a 1.5% agarose gel to demonstrate complete digestion. For each locus, 2 µl of each amplicon (undigested or digested) were mixed with 5 µl of DNA sequencing-stop solution (Promega), heat-denatured at 94 °C/30 min, snap-cooled on a freeze-block (-20 °C) and subjected to electrophoresis at 74 V for 6 h (*ptpi*) or 16 h (*pgdh* and *pbg*) at 7.4 °C in a GMA Wide Mini S-2x25 gel in a SEA 2000 rig (Elchrom Scientific AG) using TAE buffer (40 mM Tris base, 20 mM acetic acid, 1.0 mM EDTA, Bio-Rad, USA). A control sample (representing a known genotype) was included on each gel to ensure the reproducibility of profiles among gels.

Following SSCP-based analysis, selected amplicons (n = one to four) representing each distinct electrophoretic profile were column-purified (Wizard® PCR-Preps DNA Purification System, Promega), according to the manufacturer's instructions and then subjected to direct, automated sequencing (BigDye® Terminator v.3.1 chemistry, Applied Biosystems, USA), in both directions, using the same primers employed for PCR. The quality of each sequence was verified by comparison with its corresponding electropherogram using the program FinchTV v1.4.0 (<http://www.geospiza.com/finchtv.html>). Open reading frames (ORFs) were predicted using the program Transeq (<http://www.ebi.ac.uk/Tools/emboss/transeq/index.html>). International Union of Pure and Applied Chemistry (IUPAC) codes were used to indicate polymorphic positions in a sequence.

Prior to phylogenetic analysis, representative sequences determined herein (corresponding to GenBank accession nos. XXXXX-XXXXX [*ptpi*], XXXXX-XXXXX [*pgdh*], and XXXXX-XXXXX [*pbg*]) were aligned with 71 (*ptpi*), 74 (*pgdh*) and 60 (*pbg*) reference sequences representing *Giardia duodenalis* (assemblages A-H), *G. ardeae*, *G. microti* and *G. muris*, presently available in public databases (*i.e.* GenBank [<http://www.ncbi.nlm.nih.gov/>] and/or ZOOBank [43]; see Table 2 and Supplementary Tables 1-3). Alignments were made of sequences representing individual loci or concatenated sequence datasets ('*ptpi* + *pgdh* + *pbg*' and '*ptpi* + *pbg*'). For the phylogenetic investigation of *pgdh* sequence data, two analyses were conducted. The first analysis incorporated sequences from assemblages A-G (aligned over 395 positions), whereas the second incorporated a further nine published sequences representative of sequence diversity within assemblage H (aligned over 216 positions). Each alignment was constructed using the program Clustal X [44]. The resultant alignments were adjusted manually, and pairwise comparisons of nucleotide sequence differences were conducted using the program BioEdit [45]. Polymorphic positions were excluded from the calculation of differences, if a base was common to both sequences being compared.

Phylogenetic analysis of sequence data was conducted by Bayesian inference (BI) using Monte Carlo Markov Chain (MCMC) analysis in MrBayes 3.1.2 [46, 47]. The general time-reversible model of evolution was used for each analysis, allowing for a 'site-specific rates model', with the 'rates' varying according to a defined partition (*i.e.* codon position) (prset ratepr=variable). Posterior probabilities (pp) were calculated via 4,000,000 (*pgdh* and '*ptpi* + *pbg*'), 2,000,000 (*ptpi* and *pbg*) or 1,000,000 ('*ptpi* + *pgdh* + *pbg*') generations, utilising four simultaneous tree-building chains, with every 100th tree being saved. At this point, the standard deviation of split frequencies was < 0.01, and the potential scale reduction factor

(PSRF) approached one. A 50% majority rule consensus tree for each analysis was constructed based on the final 75% of trees generated by BI. The outgroups used in these analyses were *G. ardeae* and *G. muris* (separately or in combination as sequence data was available; see Table 2 and Supplementary Figs. 1-4).

3 Results

We conducted mutation scanning analysis of all amplicons ($n = 92 \times 3$) representing 25 *G. duodenalis* isolates (= 92 DNA samples from all stocks and derived lines; see Table 1) and then sequenced amplicons representative of all SSCP profiles displayed. Sequence analysis (for the three loci) showed that 24 isolates represented a single assemblage (A, B or E), and one isolate (BRIS/90/HEPU/1229) represented a mixed population of assemblages A and B, consistent with previous findings using multiple other techniques [26]. For the 24 mono-assemblage isolates (= 65 stocks or lines) the analyses defined seven distinct sequence types for *pgdh* and *pbg*, and six sequence types for *ptpi*. Polymorphic positions were detected in five of the 20 sequence types and included: $_{27}Y$, $_{41}R$, $_{102}Y$ and $_{210}R$ (XXXXX; isolate 41) and $_{12}R$, $_{18}Y$ and $_{237}R$ (XXXXX; isolate 83) (*ptpi*); $_{96}R$ (XXXXX; isolate 74) and $_{288}Y$, $_{324}R$ and $_{388}M$ (XXXXX; isolate 91/95) (*pgdh*); and $_{108}Y$, $_{177}Y$, and $_{255}Y$ (XXXXX; isolate 32) (*pbg*). For each sequence type, all combinations of individual nucleotides recorded at individual polymorphic positions were considered in conceptual translations. In no case did the nucleotide alterations cause a disruption in the predicted open reading frame (ORF), and amino acid alterations were inferred only at one position ($_{41}R$; N → S; accession no. isolate 41) in a *ptpi* sequence and one position ($_{388}M$; L → M; isolate 91/95) in a *pgdh* sequence.

Comparison of all sequence data determined herein with information available from public databases inferred *G. duodenalis* assemblage A for isolates from humans from Australia (AD113; BRIS/83/HEPU/106; BRIS/87/HEPU/713; BRIS/89/HEPU/1003; BRIS/90/HEPU/1229; BRIS/91/HEPU/1279; BRIS/92/HEPU/1541; BRIS/92/HEPU/1590; BRIS/92/HEPU/2443; BRIS/93/HEPU/1706), Malaysia (KL/92/IMR/1106) and Afghanistan (WB), a cat from Australia (BAC2), a sheep from Canada (OAS1) and a sulphur-crested cockatoo from Australia (BRIS/95/HEPU/2041); *G. duodenalis* assemblage B was inferred for isolates from humans from Australia (BRIS/90/HEPU/1229; BRIS/91/HEPU/1279; BRIS/92/HEPU/2342; BRIS/92/HEPU/2348; BRIS/93/HEPU/1638; BRIS/93/HEPU/1653; BRIS/93/HEPU/1705; BRIS/93/HEPU/1718; BRIS/93/HEPU/1727), Papua New Guinea (BRIS/92/HEPU/1487) and Canada (H7); and, *G. duodenalis* assemblage E was identified in a bovine isolate from Australia (BRIS/92/HEPU/1709) (see Tables 1 and 2).

To provide support for the initial genetic classification of genomic DNA samples, the sequences datasets were aligned to reference sequences and subjected to BI analysis. The resultant phylograms (see Supplementary Figs. 1-4) indicated that each assemblage (A-H) was essentially shown to represent a monophyletic group (with some exceptions; see Supplementary Figs. 2 and 3), and the sequence types defined here grouped with reference sequences from assemblages A, B or E, to the exclusion of all sequence types representing the other five recognized *G. duodenalis* assemblages and other species of *Giardia*. The two exceptions were represented in the analyses of *pgdh* sequence data. In the first analysis (see Supplementary Fig. 2), the seven *pgdh* sequence types defined were aligned to reference sequences representing *G. duodenalis* assemblages A-G, and *G. ardeae*; the BI analysis indicated that sequence types for assemblage A formed a paraphyletic group with sequences representative of assemblage F ($pp = 0.82$). A second analysis (see Supplementary Fig. 3), which incorporated a further nine reference sequences representative of sequence diversity within assemblage H, showed that two groups of assemblage A formed a polytomy with sequences representing assemblages E/F/G, with strong support ($pp = 1.00$).

A phylogenetic analysis of concatenated sequence data for *ptpi* and *pbg* allowed all 92 DNA samples to be linked to known sub-assemblages of *G. duodenalis* (AI-1, AI-2, AII-2, AIII, BIII and BIV) [48] by comparison to 37 genotypes, which also included *G. duodenalis* assemblages C-G, and *G. muris* as the outgroup. This analysis (Fig. 1) linked isolates from humans (AD113; BRIS/83/HEPU/106; BRIS/87/HEPU/713; BRIS/90/HEPU/1229; BRIS/91/HEPU/1279; BRIS/92/HEPU/1590; KL/92/IMR/1106; WB), sheep (OAS1) and a sulphur-crested cockatoo (BRIS/95/HEPU/2041) to sub-assemblage AI-2 (0.2%; 1 nt), reported previously from cattle [48]. The isolate from a cat (BAC2) grouped with sub-assemblage AI-1 (0.4%; 5 nt), whereas other isolates from humans (BRIS/89/HEPU/1003; BRIS/92/HEPU/2443; BRIS/92/HEPU/1541; BRIS/93/HEPU/1706) matched sub-assemblage AII-2 (0.0-0.4%; 0-3 nt over 858 alignment positions). The isolates from humans (BRIS/90/HEPU/1229; BRIS/93/HEPU/1705; BRIS/93/HEPU/1718; BRIS/93/HEPU/1727; BRIS/92/HEPU/1487; BRIS/92/HEPU/2342; BRIS/92/HEPU/2348; H7; BRIS/91/HEPU/1279; BRIS/93/HEPU/1638; BRIS/93/HEPU/1653) grouped closely genetically with sub-assemblage BIV and genotypes from various non-human primates (0.2-3.7%; 1-31 nt). One isolate from cattle (BRIS/92/HEPU/1709) formed a monophyletic group (pp = 1.00) with assemblage E (0.6-1.1%; 5-9 nt), also recorded previously in cattle or sheep [49].

4 Discussion

The present study genetically reclassified *G. duodenalis* isolates from a cryobank, originally established and curated by the Upcroft team over a period of nearly 30 years. This bank contains 25 isolates, and a range of derived stocks and lines, which had been well characterised in terms of clinical history, growth characteristics, virulence/pathogenicity, biochemistry, drug susceptibility/resistance and/or karyotype [11, 31-33]. Our systematic PCR-based barcoding categorized all stocks and derived lines by comparison with selected reference data available in current gene databases. We defined 20 genetic variants of *G. duodenalis*, which have been assigned unequivocally to assemblages A, B or E, based on phylogenetic analysis. Twelve sequence types (XXXX-XXXX) represented *G. duodenalis* assemblages A and B, and were the same as those reported previously from humans, livestock and companion animals, wildlife, and environmental and wastewater samples from around the world (see Table 3). The other eight types (XXXX-XXXX) matched assemblages A, B or E, and represent new sequence types.

Polymorphic nucleotide positions were detected in *ptpi*, *pgdh* and *pbg* sequence types for a subset of five samples, and were more common in sequences representing assemblage B than those with matches to assemblage A or E. To assess whether sequence polymorphism was due to the presence of multiple assemblages in a stock/line or a sequencing error, the IUPAC code at each site was substituted with each of the bases (separately), and all possible combinations of nucleotides at each position in each sequence were inferred [50]. None of the combinations of nucleotides (for all polymorphic positions) caused a disruption in the predicted open reading frame (ORF) of the sequence or a change in assemblage-assignment, suggesting that the mutations represented true sequence variation. These findings are consistent with other studies [48-52], indicating that variation exists within or among *Giardia* individuals, as a consequence of genetic or allelic divergence within an isolate or line. This conclusion is supported by evidence demonstrating that allelic sequence variability (0.5%) in the genome sequence of isolate GS representing assemblage B [53] is greater than for WB (< 0.01%) of assemblage A [54]. Evidence from other studies also supports the proposal that more allelic variability occurs in individual isolates or stocks representing assemblage B [48, 52, 55] than those of assemblage A [43]. Future studies should attempt to establish whether this variability is present within or among individual bi-

and quadrinucleate *G. duodenalis* trophozoites and cysts, respectively, using mutation scanning-based sequencing.

Despite polymorphism in some sequence types, and in contrast to previous “multilocus genotyping” (MLG) studies [38, 48, 56-58], there was no discrepancy in assemblage assignment among the three loci for the 24 mono-assemblage isolates examined in the present study. Phylogenetic analysis of ‘*ptpi* + *pbg*’ indicated that the 11 genotypes (see Table 2 and Fig. 1) were assigned unequivocally to assemblages A, B or E. Current samples characterized and assigned here to assemblage A grouped with *G. duodenalis* genotypes reported previously from humans as well as those from cats, livestock (cattle and sheep) and wild (e.g., *Capreolus capreolus*) and zoo-kept (e.g., *Nasua nasua*, *Oryx dammah*, *Pecari tajacu*) animals (Fig. 1). Also samples from humans assigned here to assemblage B grouped with *G. duodenalis* genotypes reported previously from humans, various non-human primates, rodents and carnivores (e.g., *Ursus malayanus*, *Docilchotis patagonum*, *Pan troglodytes*, *Colobus guereza* and *Hylobates lar*) (see Table 2). Also, genotypes linked to *in vitro* cultured lines of *G. duodenalis* grown in albendazole, metronidazole, furazolidone and/or quinacrine (genotype 1 and 7) could neither be differentiated genetically from the associated stocks (i.e. BRIS/83/HEPU/106, BRIS/87/HEPU/713, WB, OAS1 and BAC2) nor did they form specific phylogenetic relationships with each other or other reference genotypes [48]. This may suggest that although MLG has been shown to be a highly discriminatory approach for resolving *G. duodenalis* assemblages [49, 59], the loci employed herein (i.e. *tpi*, *gdh* and *bg*) may not reflect important phenotypic traits, such as drug resistance.

SSCP-coupled analyses of multiple genetic loci will be useful for assessing the population genetic richness and diversity within *Giardia* and the epidemiology of giardiasis and, particularly, zoonotic transmission. However, a potential limitation relates to the current paucity of published sequence data for other suitably ‘variable’ loci, including genes encoding -giardin and *gdh*. In addition, many previous studies have used phenetic methods, such as PCR-based restriction fragment length polymorphism (PCR-RFLP) or specific PCR detection, rather than sequencing, which are relatively rapid to perform, but do not yield sequence data and do not allow direct comparisons among studies. Unlike phenetic approaches, SSCP-coupled sequencing yields genetic data and has been shown to detect point mutations for amplicons of appropriate sizes (up to 500 bp for direct SSCP and up to 1.9 kb for REF) [40-42, 60]. Therefore, mutation scanning provides a sensitive tool for the direct analysis of subtle genetic variation within and among populations of *G. duodenalis* isolated from animals. Moreover, this tool can be employed routinely to ensure the purity of (mono-assemblage) stocks and lines maintained in the laboratory and too continually monitor (during *in vitro* passage) for any potential cross-over contamination during experimentation.

In conclusion, the genetic barcoding of samples from the present *G. duodenalis* cryobank, the ability to demonstrate that an isolate, stock or line represents a single assemblage following *in vitro* passage and to rapidly classify any new *G. duodenalis* populations from animals sets the scene for a wide range of future genomic, proteomic, metabolic and glycomic studies of *G. duodenalis* using advanced technologies. For instance, direct comparisons of genome sequences from current stocks/lineages could be conducted using next-generation sequencing technologies (such as Illumina) to establish whether there is concordance with the assemblage assignment achieved here using three genetic loci. Comprehensive comparisons among genome sequences representing the assemblages available here (i.e. A, B and E) will also provide insights into levels of synteny and the extent of genomic rearrangements, and have the potential to discover new genetic features or structures for isolates and derived lines with unique biological, biochemical and/or

karyotypic characteristics or traits (such as the presence or absence of chromosomal bands 3 and 4 for Bris/83/HEPU/106 and Bris/91/HEPU/1279; refs. [27, 61]). Importantly, genomic datasets will also provide a solid foundation for functional genomic, transcriptomic and proteomic studies. For example, temporal changes in transcription profiles (transcriptome) and gene expression (proteome) could be explored comparatively for different isolates/stocks/lines with differing phenotypic traits (*e.g.*, rapid *versus* slow growth in culture; drug resistant *versus* drug susceptible; virulent *versus* avirulent; pathogenic *versus* apathogenic) under well-controlled culture conditions, in order to provide hints as to the genetic basis of the phenotypic differences observed *in vitro* and/or *in vivo*. Isolate Bris/95/HEPU/2041, for instance, would be an ideal candidate to explore at genomic and transcriptomic levels. Originally isolated from a sulphur-crested cockatoo [62], this isolate has been shown experimentally to: (1) produce chronic infections in outbred Quackenbush Swiss mouse pups and inbred C3H/HeJ mice, by comparison with Bris/83/HEPU/106 and Bris/91/HEPU/1279 from humans, (2) result in reduced weight gain in mice (mimicking a failure-to-thrive syndrome in children), (3) have an intensity of infection in mice which is more than three to four times that observed for some other *G. duodenalis* isolates (*i.e.* Bris/83/HEPU/106 and Bris/91/HEPU/1279), and (4) contain duplications of chromosomes 3 (or 4) and 6, based on clamped homogeneous electrical field (CHEF)-based southern blotting [62, 63]. In addition, isolate Bris/93/HEPU/1706 would be another ideal candidate to explore genetically. This *G. duodenalis* isolate, from a human patient admitted to hospital with haematemesis, is known to contain a plasmid which is present in high copy number [17]. This plasmid has two (302 bp) repeat elements with a remarkable similarity to some prokaryotic transcriptional regulatory sequences, but does not encode identifiable virulence or drug resistance genes [17]. Interestingly, the copy number of this plasmid decreased following serial passage *in vitro*, and it is unstable in *Escherichia coli* but integrates transposable and other sequence elements from the genome of this bacterium [17]. Understanding the origin, structure and role of this enigmatic plasmid would be very interesting from a fundamental perspective and might reveal an entirely new aspect of the genetics (*e.g.*, recombination and exchange) of *G. duodenalis*. Clearly, comparative genetic studies of stocks and derived lines with differing phenotypic traits linked to the parasite-host interplay, virulence and pathogenicity, together with host genetics, will be critical to understanding giardiasis.

Importantly, controlled experiments could be undertaken to assess changes in transcription and/or expression in *G. duodenalis* upon exposure to common anti-giardial drugs (*e.g.*, albendazole, metronidazole, furazolidone and/or quinacrine) in resistant (*e.g.*, WB1B [induced]) compared with susceptible (*e.g.*, BRIS/83/HEPU/106) lines, in order to elucidate the genetic basis of resistance and, potentially, the mode of action of some compounds. Such experiments could be supported also by genome-wide single nucleotide polymorphism [SNP] and proteomic comparisons between resistant and susceptible lines to infer the relationship(s) between mutations and drug resistance in *G. duodenalis*. Another area of interest would be to explore essential genes and metabolic bottlenecks in the search for new drug target candidates to which novel chemical compounds could be designed. In this context, gene knock-down and knock-out approaches [64, 65] could be particularly useful tools to test essentiality and gene function in *G. duodenalis*, as already demonstrated for related flagellate protozoans, such as *Trypanosoma brucei* [66] and *T. congolense* [67]. Clearly, the amenability of some assemblages of *Giardia* (particularly A and B) to *in vitro* culture would support such research. Advances in genomic technologies and computer-assisted data analysis now allow also detailed explorations of the host-specificity and/or virulence/pathogenicity of *Giardia* species and assemblages as well as traits that relate to susceptibility or resistance to giardial infection in host animals. Importantly, with the availability of the human genome and the imminent genome sequences for a range of *G. duodenalis* samples, we will have access to a tremendously rich source of potential targets for novel chemotherapeutics, vaccines and diagnostic tools, based on a fundamental

understanding, at the molecular level, of the pathogens, their interplay with the host and disease. This context provides exciting prospects to conduct systems biological studies of key lines from the present *G. duodenalis* cryobank.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding from the Australian Research Council, Melbourne Water Corporation, The Australian Academy of Science and Fulbright Commission is gratefully acknowledged. Support from the Victorian Life Sciences Computation Initiative (VLSCI) and the IBM Collaboratory is gratefully acknowledged. ARJ holds a Career Development Award (CDA) from the National Health and Medical Research Council (NHMRC) of Australia. This study was also supported by a U01 Cooperative Research Agreement AI75527 from the National Institutes of Health, U.S.A, and by the National Health and Medical Research Council of Australia grant 496640. Prof. Andrew Thompson and other colleagues are gratefully acknowledged for originally supplying some *Giardia* isolates to the Upcroft research team.

Abbreviations

REF	restriction endonuclease fingerprinting
SSCP	single-strand confirmation polymorphism
<i>bg</i>	beta-giardin gene
<i>gdh</i>	glutamate dehydrogenase gene
<i>tpi</i>	triose-phosphate isomerase gene

References

1. Kulda, J.; Nohýnková, E. Parasitic Protozoa. Kreier, PP., editor. Academic Press; New York: 1978. p. 2-127.
2. Lasek-Nesselquist E, Welch DM, Sogin ML. Int J Parasitol. 2010; 40:1063–1074. [PubMed: 20361967]
3. Rendtorff RC. Am J Hyg. 1954; 59:209–220. [PubMed: 13138586]
4. Ortega YR, Adam RD. Clin Infect Dis. 1997; 25:545–549. [PubMed: 9314441]
5. Erlandsen, SL.; Weissner, S.; Ottenwaelter, C. Giardia: The cosmopolitan parasite. Olson, BE.; Olson, ME.; Wallis, PM., editors. CABI Publishing; Wallingford, UK: 2002. p. 3-14.
6. Thompson RC, Monis PT. Adv Parasitol. 2004; 58:69–137. [PubMed: 15603762]
7. Ish-Horowicz M, Korman SH, Shapiro M, Har-Even U, Tamir I, Strauss N, Deckelbaum RJ. Pediatr Infect Dis J. 1989; 8:773–779. [PubMed: 2512565]
8. Farthing, MJG. Giardia: From Molecules to Disease. Thompson, RCA.; Reynoldson, JA.; Lymbery, AJ., editors. CAB; Wallingford: 1994. p. 15-37.
9. Farthing MJ. J Pediatr Gastroenterol Nutr. 1997; 24:79–88. [PubMed: 9093992]
10. Faubert G. Clin Microbiol Rev. 2000; 13:35–54. [PubMed: 10627490]
11. Williamson AL, O'Donoghue PJ, Upcroft JA, Upcroft P. Int J Parasitol. 2000; 30:129–136. [PubMed: 10704595]
12. Homan WL, Mank TG. Int J Parasitol. 2001; 31:822–826. [PubMed: 11403774]
13. Read C, Walters J, Robertson ID, Thompson RC. Int J Parasitol. 2002; 32:229–231. [PubMed: 11812501]
14. Cacciò SM, Ryan U. Mol Biochem Parasitol. 2008; 160:75–80. [PubMed: 18501440]
15. Faubert GM. Parasitol Today. 1996; 12:140–145. [PubMed: 15275222]
16. Huang DB, White AC. Gastroenterol Clin North Am. 2006; 35:291–314. [PubMed: 16880067]

17. Upcroft J, Upcroft P. Bioessays. 1998; 20:256–263. [PubMed: 9631653]
18. Lane S, Lloyd D. Crit Rev Microbiol. 2002; 28:123–147. [PubMed: 12109771]
19. Upcroft JA, Upcroft P. Parasitol Today. 1993; 9:187–190. [PubMed: 15463750]
20. Harris JC, Plummer S, Lloyd D. Appl Microbiol Biotechnol. 2001; 57:614–619. [PubMed: 11778868]
21. Wright JM, Dunn LA, Upcroft P, Upcroft JA. Expert Opin Drug Saf. 2003; 2:529–541. [PubMed: 14585063]
22. Lalle M. Infect Disord Drug Targets. 2010; 10:283–294. [PubMed: 20429863]
23. Ey PL, Bruderer T, Wehrli C, Köhler P. Parasitol Res. 1996; 82:52–60. [PubMed: 8825446]
24. Mayrhofer G, Andrews RH, Ey PL, Chilton NB. Parasitology. 1995; 111:11–17. [PubMed: 7609985]
25. Ey PL, Khanna K, Andrews RH, Manning PA, Mayrhofer G. J Gen Microbiol. 1992; 138:2629–2637. [PubMed: 1362583]
26. Upcroft JA, Upcroft P. J Eukaryot Microbiol. 1994; 41:189–194. [PubMed: 7914131]
27. Upcroft JA, Boreham PF, Campbell RW, Shepherd RW, Upcroft P. Acta Trop. 1995; 60:35–46. [PubMed: 8546037]
28. Monis PT, Andrews RH, Mayrhofer G, Ey PL. Mol Biol Evol. 1999; 16:1135–1144. [PubMed: 10486969]
29. Cacciò SM, Beck R, Almeida A, Bajer A, Pozio E. Parasitology. 2010; 137:919–925. [PubMed: 20233487]
30. Cacciò SM, Thompson RC, McLauchlin J, Smith HV. Trends Parasitol. 2005; 21:430–437. [PubMed: 16046184]
31. Upcroft JA, Healey A, Murray DG, Boreham PF, Upcroft P. Parasitology. 1992; 104:397–405. [PubMed: 1641239]
32. Upcroft JA, Healey A, Upcroft P. Int J Parasitol. 1993; 23:609–616. [PubMed: 8225763]
33. Brown DM, Upcroft JA, Upcroft P. Mol Biochem Parasitol. 1995; 72:47–56. [PubMed: 8538699]
34. World Health Organization. Bull World Health Organ. 1978; 56:467–480. [PubMed: 308413]
35. Monis PT, Mayrhofer G, Andrews RH, Homan WL, Limper L, Ey PL. Parasitology. 1996; 112:1–12. [PubMed: 8587793]
36. Sulaiman IM, Fayer R, Bern C, Gilman RH, Trout JM, Schantz PM, Das P, Lal AA, Xiao L. Emerg Infect Dis. 2003; 9:1444–1452. [PubMed: 14718089]
37. Lalle M, Pozio E, Capelli G, Bruschi F, Crotti D, Cacciò SM. Int J Parasitol. 2005; 35:207–213. [PubMed: 15710441]
38. Read CM, Monis PT, Thompson RC. Infect Genet Evol. 2004; 4:125–130. [PubMed: 15157630]
39. Holberton D, Baker DA, Marshall J. J Mol Biol. 1988; 204:789–795. [PubMed: 3225852]
40. Gasser R, Hu M, Chilton NB, Campbell BE, Jex AR, Otranto D, Cafarchia C, Beveridge I, Zhu X. Nature Protocols. 2006; 1:3121–3128.
41. Liu Q, Sommer SS. Biotechniques. 1995; 18:470–477. [PubMed: 7779398]
42. Zhu XQ, Gasser RB. Electrophoresis. 1998; 19:1366–1373. [PubMed: 9694283]
43. Sprong H, Cacciò SM, van der Giessen JW. PLoS Negl Trop Dis. 2009; 3:e558. [PubMed: 19956662]
44. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. Nucleic Acids Res. 1997; 24:4876–4882. [PubMed: 9396791]
45. Hall TA. Nucleic Acids Symp Ser. 1999; 41:95–98.
46. Huelsenbeck JP, Ronquist F. Bioinformatics. 2001; 17:754–755. [PubMed: 11524383]
47. Ronquist F, Huelsenbeck JP. Bioinformatics. 2003; 19:1572–1574. [PubMed: 12912839]
48. Cacciò SM, Beck R, Lalle M, Marinculic A, Pozio E. Int J Parasitol. 2008; 38:1523–1531. [PubMed: 18571176]
49. Lebbad M, Mattsson JG, Christensson B, Ljungström B, Backhans A, Andersson JO, Svärd SG. Vet Parasitol. 2010; 168:231–239. [PubMed: 19969422]

50. Nolan MJ, Jex AR, Pangasa A, Young ND, Campbell AJ, Stevens M, Gasser RB. *Electrophoresis*. 2010; 31:287–298. [PubMed: 20084628]
51. Leonhard S, Pfister K, Beelitz P, Wielinga C, Thompson RC. *Vet Parasitol*. 2007; 150:33–38. [PubMed: 17913365]
52. Lalle M, Bruschi F, Castagna B, Campa M, Pozio E, Cacciò SM. *Trans R Soc Trop Med Hyg*. 2009; 103:834–838. [PubMed: 19477474]
53. Franzén O, Jerlström-Hultqvist J, Castro E, Sherwood E, Ankarklev J, Reiner DS, Palm D, Andersson JO, Andersson B, Svärd SG. *PLoS Pathog*. 2009; 5:e1000560. [PubMed: 19696920]
54. Morrison HG, McArthur AG, Gillin FD, Aley SB, Adam RD, Olsen GJ, Best AA, Cande WZ, Chen F, Cipriano MJ, Davids BJ, Dawson SC, Elmendorf HG, Hehl AB, Holder ME, Huse SM, Kim UU, Lasek-Nesselquist E, Manning G, Nigam A, Nixon JE, Palm D, Passamaneck NE, Prabhu A, Reich CI, Reiner DS, Samuelson J, Svärd SG, Sogin ML. *Science*. 2007; 317:1921–1926. [PubMed: 17901334]
55. Lebbad M, Ankarklev J, Tellez A, Leiva B, Andersson JO, Svärd S. *Acta Trop*. 2008; 106:44–53. [PubMed: 18325480]
56. Traub RJ, Monis PT, Robertson I, Irwin P, Mencke N, Thompson RC. *Parasitology*. 2004; 128:253–262. [PubMed: 15080083]
57. Langkjær RB, Vigre H, Enemark HL, Maddox-Hyttel C. *Parasitology*. 2007; 134:339–350. [PubMed: 17076923]
58. Levecke B, Geldhof P, Claerebout E, Dorny P, Vercammen F, Cacciò SM, Vercruysse J, Geurden T. *Int J Parasitol*. 2009; 39:1595–1601. [PubMed: 19523472]
59. Plutzer J, Ongerth J, Karanis P. *Int J Hyg Environ Health*. 2010; 213:321–333. [PubMed: 20619729]
60. Liu Q, Feng J, Sommer SS. *Biotechniques*. 1997; 23:836–839. [PubMed: 9383547]
61. Upcroft JA, Boreham PF, Upcroft P. *Int J Parasitol*. 1989; 19:519–527. [PubMed: 2777469]
62. Upcroft JA, McDonnell PA, Gallagher AN, Chen N, Upcroft P. *Parasitology*. 1997; 114:407–412. [PubMed: 9149411]
63. Upcroft JA, McDonnell PA, Upcroft P. *Parasitol Today*. 1998; 14:281–284. [PubMed: 17040784]
64. Gueiros-Filho FJ, Beverley SM. *Science*. 1997; 276:1716–1719. [PubMed: 9180085]
65. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. *Nature*. 1998; 391:806–811. [PubMed: 9486653]
66. Ngo H, Tschudi C, Gull K, Ullu E. *Proc Natl Acad Sci U S A*. 1998; 95:14687–14692. [PubMed: 9843950]
67. Inoue N, Otsu K, Ferraro DM, Donelson JE. *Mol Biochem Parasitol*. 2002; 120:309–313. [PubMed: 11897138]
68. Capon AG, Upcroft JA, Boreham PF, Cottis LE, Bundesen PG. *Int J Parasitol*. 1989; 19:91–98. [PubMed: 2651342]
69. Townson SM, Laqua H, Upcroft P, Boreham PF, Upcroft JA. *Trans R Soc Trop Med Hyg*. 1992; 86:521–522. [PubMed: 1475822]
70. Upcroft JA, Mitchell RW, Upcroft P. *Int J Parasitol*. 1993; 23:785–792. [PubMed: 8300289]
71. Upcroft JA, Healey A, Upcroft P. *J Eukaryot Microbiol*. 1994; 41:639–642. [PubMed: 7866388]
72. Mayrhofer G, Andrews RH, Ey PL, Albert MJ, Grimmond TR, Merry DJ. *Parasitology*. 1992; 105:255–263. [PubMed: 1454424]
73. Nash TE, McCutchan T, Keister D, Dame JB, Conrad JD, Gillin FD. *J Infect Dis*. 1985; 152:64–73. [PubMed: 2409186]
74. Upcroft JA, Campbell RW, Upcroft P. *Parasitology*. 1996; 112:309–313. [PubMed: 8728994]
75. Uji, A.; Wallis, PM.; Wenman, WM. *Advances in Giardia research*. Wallis, PM.; Hammond, BR., editors. University of Calgary Press; Calgary: 1988.
76. Beck R, Sprong H, Bata I, Lucinger S, Pozio E, Cacciò SM. *Vet Parasitol*. 2011; 175:40–46. [PubMed: 20970259]
77. Cacciò SM, De Giacomo M, Pozio E. *Int J Parasitol*. 2002; 32:1023–1030. [PubMed: 12076631]
78. Guy RA, Xiao C, Horgen PA. *J Clin Microbiol*. 2004; 42:3317–3320. [PubMed: 15243104]

79. Mowatt MR, Weinbach EC, Howard TC, Nash TE. *Exp Parasitol.* 1994; 78:85–92. [PubMed: 8299763]
80. Lasek-Nesselquist E, Welch DM, Thompson RC, Steuart RF, Sogin ML. *J Eukaryot Microbiol.* 2009; 56:504–518. [PubMed: 19883439]
81. Wielinga C, Ryan U, Andrew Thompson RC, Monis P. *Int J Parasitol.* 2011; 41:495–503. [PubMed: 21176781]
82. Trout JM, Santín M, Greiner E, Fayer R. *Vet Parasitol.* 2004; 124:179–186. [PubMed: 15381298]
83. Feng Y, Ortega Y, Cama V, Terrel J, Xiao L. *Parasitol Res.* 2008; 103:87–92. [PubMed: 18338181]
84. Suzuki J, Murata R, Kobayashi S, Sadamasu K, Kai A, Takeuchi T. *Parasitology.* 2011; 138:493–500. [PubMed: 21040620]
85. Teodorovic S, Braverman JM, Elmendorf HG. *Eukaryot Cell.* 2007; 6:1421–1430. [PubMed: 17557879]
86. Abe N, Tanoue T, Noguchi E, Ohta G, Sakai H. *Parasitol Res.* 2010; 106:733–736. [PubMed: 20054561]
87. Lasek-Nesselquist E, Bogomolni AL, Gast RJ, Welch DM, Ellis JC, Sogin ML, Moore MJ. *Dis Aquat Organ.* 2008; 81:39–51. [PubMed: 18828561]
88. Trout JM, Santín M, Fayer R. *J Parasitol.* 2003; 89:1254–1255. [PubMed: 14740923]
89. Sulaiman IM, Jiang J, Singh A, Xiao L. *Appl Environ Microbiol.* 2004; 70:3776–3780. [PubMed: 15184191]
90. Baruch AC, Isaac-Renton J, Adam RD. *J Infect Dis.* 1996; 174:233–236. [PubMed: 8656003]
91. Yee J, Dennis PP. *J Biol Chem.* 1992; 267:7539–7544. [PubMed: 1559991]
92. Souza SL, Gennari SM, Richtzenhain LJ, Pena HF, Funada MR, Cortez A, Gregori F, Soares RM. *Vet Parasitol.* 2007; 149:258–264. [PubMed: 17900812]
93. van der Giessen JW, de Vries A, Roos M, Wielinga P, Kortbeek LM, Mank TG. *Int J Parasitol.* 2006; 36:849–858. [PubMed: 16701663]
94. Pelayo L, Nunez FA, Rojas L, Furuseth Hansen E, Gjerde B, Wilke H, Mulder B, Robertson L. *Ann Trop Med Parasitol.* 2008; 102:585–595. [PubMed: 18817599]
95. Abe N, Kimata I, Tokoro M. *Jpn J Infect Dis.* 2005; 58:57–58. [PubMed: 15728998]
96. Robertson LJ, Forberg T, Hermansen L, Gjerde BK, Langeland N. *J Infect.* 2007; 55:79–88. [PubMed: 17368548]
97. Johnston AR, Gillespie TR, Rwego IB, McLachlan TL, Kent AD, Goldberg TL. *PLoS Negl Trop Dis.* 2010; 4:e683. [PubMed: 20485494]
98. Guy RA, Payment P, Krull UJ, Horgen PA. *Appl Environ Microbiol.* 2003; 69:5178–5185. [PubMed: 12957899]
99. Winkworth CL, Learmonth JJ, Matthaei CD, Townsend CR. *Appl Environ Microbiol.* 2008; 74:5100–5105. [PubMed: 18567681]
100. Khan SM, Debnath C, Pramanik AK, Xiao L, Nozaki T, Ganguly S. *Vet Parasitol.* 2010; 171:41–47. [PubMed: 20356678]
101. Geurden T, Thomas P, Casaert S, Vercruysse J, Claerebout E. *Vet Parasitol.* 2008; 155:142–145. [PubMed: 18565678]
102. Robertson LJ, Forberg T, Hermansen L, Hamnes IS, Gjerde B. *J Wildl Dis.* 2007; 43:576–585. [PubMed: 17984252]
103. Thompson RC, Smith A, Lymbery AJ, Averis S, Morris KD, Wayne AF. *Vet Parasitol.* 2010; 170:207–211. [PubMed: 20211528]
104. Geurden T, Levecke B, Cacciò SM, Visser A, De Groote G, Casaert S, Vercruysse J, Claerebout E. *Parasitology.* 2009; 136:1161–1168. [PubMed: 19631012]
105. Matsubayashi M, Kimata I, Abe N. *J Vet Med Sci.* 2005; 67:337–340. [PubMed: 15805742]
106. Di Giovanni GD, Betancourt WQ, Hernandez J, Assadian NW, Flores Margez JP, Lopez EJ. *Int J Environ Health Res.* 2006; 16:405–418. [PubMed: 17164167]
107. Robertson LJ, Gjerde BK, Furuseth Hansen E. *Vet Parasitol.* 2010; 171:140–145. [PubMed: 20381251]

108. Fayer R, Santín M, Trout JM, DeStefano S, Koenen K, Kaur T. J Zoo Wildl Med. 2006; 37:492–497. [PubMed: 17315434]
109. Robertson LJ, Hermansen L, Gjerde BK, Strand E, Alvsvåg JO, Langeland N. Appl Environ Microbiol. 2006; 72:2212–2217. [PubMed: 16517674]
110. Cooper MA, Adam RD, Worobey M, Sterling CR. Curr Biol. 2007; 17:1984–1988. [PubMed: 17980591]
111. Kosuwin R, Putaporntip C, Pattanawong U, Jongwutiwes S. Gene. 2010; 449:1–8. [PubMed: 19796671]
112. Geurden T, Goossens E, Levecke B, Vercammen F, Vercruysse J, Claerebout E. J Zoo Wildl Med. 2009; 40:126–130. [PubMed: 19368251]
113. Alonso JL, Amoros I, Canigral I. Appl Microbiol Biotechnol. 2011; 89:1203–1211. [PubMed: 21061006]

Figure 1.

The genetic relationships of genotypes of *Giardia duodenalis* inferred from concatenated sequence data (Fig. 1A, 'ptpi + pgdh + pbg'; Fig. 1B, 'ptpi + pbg'), determined employing Bayesian inference (BI). Posterior probabilities are indicated at all major nodes. Fig. 1A is rooted at the midpoint. Genotype no. (*G. duodenalis* isolate): 1 (AD113; BRIS/83/HEPU/106; BRIS/87/HEPU/713; BRIS/90/HEPU/1229; BRIS/91/HEPU/1279; BRIS/92/HEPU/1590; KL/92/IMR/1106; W B; OAS1; BRIS/95/HEPU/2041), 2 (BRIS/90/HEPU/1229; BRIS/93/HEPU/1705; BRIS/93/HEPU/1718; BRIS/93/HEPU/1727; BRIS/92/HEPU/1487), 3 (BRIS/92/HEPU/1487), 4 (BRIS/92/HEPU/2342), 5 (BRIS/92/HEPU/2348), 6 (H7), 7 (BAC2), 8 (BRIS/91/HEPU/1279; BRIS/93/HEPU/1638; BRIS/93/HEPU/1653), 9 (BRIS/

89/HEPU/1003; BRIS/92/HEPU/2443), 10 (BRIS/92/HEPU/1541; BRIS/93/HEPU/1706), 11 from cattle (BRIS/92/HEPU/1709).

Table 1

List of 92 samples representing 25 *Giardia duodenalis* isolates (in grey) and associated stocks and derived lines (= 92 DNA samples) genetically characterised in the present study, together with phenotypic and epidemiological information, accession numbers for sequence types generated for each *tpi*, *gdh* and *bg* loci, and references ^a.

DNA sample no.	<i>Giardia duodenalis</i> isolate	Host	Locality <i>g</i>	Phenotype	Accession nos. for loci			References
					<i>tpi</i>	<i>pgdh</i>	<i>pbg</i>	
Stock / Line								
BRIS/83/HEPU/106								
1	25/08/94 - 160 ng/ul	Human	Australia (QLD)		(tpi_3) A	(gdh_3) A	(bg_3) A	[68]
2	10/10/94				(tpi_3) A	(gdh_3) A	(bg_3) A	[68]
3	26/01/00			Grown in 915 μM Mz ^b	(tpi_3) A	(gdh_3) A	(bg_3) A	[68]
BRIS/87/HEPU/713								
4	MORGAN - 11/03/92 - 5 μM	Human	Australia (QLD)	Grown in 5 μM Mz after selection in 100 μM Mz	(tpi_3) A	(gdh_3) A	(bg_3) A	[69]
5	MORGAN - 19/03/92 - 50 μM			Grown in 50 μM Mz after selection in 100 μM Mz	(tpi_3) A	(gdh_3) A	(bg_3) A	[69]
6	MORGAN - 01/04/92				(tpi_3) A	(gdh_3) A	(bg_3) A	[69]
BRIS/89/HEPU/1003								
7	STONE2 - 25/11/91	Human	Australia (QLD)	Contains plasmid DNA; from patient with bloody diarrhoea	(tpi_86) A	(gdh_86) A	(bg_3) A	
8	STONE1 - 13/12/91				(tpi_86) A	(gdh_86) A	(bg_3) A	
BRIS/90/HEPU/1229								
9	M3P2 ^c - 03/12/91	Human	Australia (QLD)		(tpi_11) B	(gdh_11) B	(bg_11) B	[26] cf. [70]
10	M3P4 - 04/02/92				(tpi_3) A	(gdh_3) A	(bg_3) A	[26] cf.[70]
11	M3P5 - 05/06/92				(tpi_11) B	(gdh_11) B	(bg_11) B	[26] cf.[70]
12	M4P2 - 17/09/91				(tpi_3) A	(gdh_3) A	(bg_3) A	[26] cf.[70]
13	M4P4 - 03/03/92				(tpi_3) A	(gdh_3) A	(bg_3) A	[26] cf.[70]
14	M4P5 - 27/05/92				(tpi_11) B	(gdh_11) B	(bg_11) B	[26] cf.[70]
15	M5P2 - 19/05/91				(tpi_11) B	(gdh_11) B	(bg_11) B	[26] cf.[70]
16	M5P2 - 20/05/92				(tpi_11) B	(gdh_11) B	(bg_11) B	[26] cf.[70]
17	M5P2 - 08/11/91				(tpi_11) B	(gdh_11) B	(bg_11) B	[26] cf.[70]
18	M5P4 - 04/02/92				(tpi_3) A	(gdh_3) A	(bg_3) A	[26] cf.[70]
19	M5P5 - 11/05/92				(tpi_13) A/B	(gdh_3) A	(bg_3) A	[26] cf.[70]
20	M6P5 - 15/06/92				(tpi_11) B	(gdh_11) B	(bg_11) B	[26] cf.[70]

DNA sample no.	<i>Giardia duodenalis</i> isolate	Host	Locality <i>g</i>	Phenotype	Accession nos. for loci			References
					<i>pipi</i>	<i>pgdh</i>	<i>pbg</i>	
21	M6P5 - 27/05/92				(tpi_11) B	(gdh_11) B	(bg_11) B	[26] <i>cf</i> [70]
22	M7P5 - 22/06/92				(tpi_11) B	(gdh_11) B	(bg_11) B	[26] <i>cf</i> [70]
23	M9P5 - 16/12/91				(tpi_11) B	(gdh_11) B	(bg_11) B	[26] <i>cf</i> [70]
24	M10P2 - 28/11/91				(tpi_11) B	(gdh_11) B	(bg_11) B	[26] <i>cf</i> [70]
25	M10P4 - 03/03/92				(tpi_3) A	(gdh_3) A	(bg_3) A	[26] <i>cf</i> [70]
26	M11P2 - 25/11/91				(tpi_11) B	(gdh_11) B	(bg_11) B	[26] <i>cf</i> [70]
27	M11P4 - 13/02/92				(tpi_3) A	(gdh_3) A	(bg_3) A	[26] <i>cf</i> [70]
28	M11P4 - 13/02/92				(tpi_3) A	(gdh_3) A	(bg_3) A	[26] <i>cf</i> [70]
29	M11P4 - 05/02/92				(tpi_3) A	(gdh_3) A	(bg_3) A	[26] <i>cf</i> [70]
30	M11P5 - 12/06/92				(tpi_13) A/B	(gdh_3) A	(bg_3) A	[26] <i>cf</i> [70]
31	M12P2 - 29/10/91				(tpi_3) A	(gdh_3) A	(bg_3) A	[26] <i>cf</i> [70]
32	M12P4 - 13/02/92				(tpi_3) A	(gdh_3) A	(bg_3) A	[26] <i>cf</i> [70]
33	M12P4 - 13/02/92				(tpi_3) A	(gdh_3) A	(bg_3) A	[26] <i>cf</i> [70]
34	M12P4 - 05/02/92				(tpi_3) A	(gdh_3) A	(bg_3) A	[26] <i>cf</i> [70]
35	M12P5 - 16/06/92				(tpi_11) B	(gdh_11) B	(bg_11) B	[26] <i>cf</i> [70]
36	BRIS/91/HEPU/1279	Human	Australia (QLD)		(tpi_83) B	(gdh_11) B	(bg_11) B	[27] <i>cf</i> [71]
37	M6P1 - 14/04/92				(tpi_83) B	(gdh_11) B	(bg_11) B	[27] <i>cf</i> [71]
38	M6P1 - 08/08/01				(tpi_83) B	(gdh_11) B	(bg_11) B	[27] <i>cf</i> [71]
39	M10P1 - 15/11/91				(tpi_83) B	(gdh_11) B	(bg_11) B	[27] <i>cf</i> [71]
40	BRIS/92/HEPU/1541	Human	Australia (QLD)	Contains plasmid DNA and Rep-like sequence	(tpi_86) A	(gdh_86) A	(bg_86) A	[27] <i>cf</i> [70]
41	BRIS/92/HEPU/1590	Human	Australia (QLD)		(tpi_3) A	(gdh_3) A	(bg_3) A	[27]
42	BRIS/92/HEPU/2348	Human	Australia (QLD)		(tpi_11) B	(gdh_75) B	(bg_11) B	Unpublished
43	BRIS/92/HEPU/2443	Human	Australia (QLD)		(tpi_86) A	(gdh_86) A	(bg_3) A	Unpublished
44	BRIS/93/HEPU/1638	Human	Australia (QLD)		(tpi_83) B	(gdh_11) B	(bg_11) B	[27]

DNA sample no.	<i>Giardia duodenalis</i> isolate		Host	Locality <i>g</i>	Phenotype	Accession nos. for loci			References
	Stock / Line					<i>pipi</i>	<i>pgdh</i>	<i>pbg</i>	
45	BRIS/93/HEPU/1653	Human	Australia (QLD)			(tpi_83) B	(gdh_11) B	(bg_11) B	[27]
46	BRIS/93/HEPU/1705	Human	Australia (QLD)			(tpi_11) B	(gdh_11) B	(bg_11) B	[27]
47	BRIS/93/HEPU/1718	Human	Australia (QLD)			(tpi_11) B	(gdh_11) B	(bg_11) B	[27]
48	BRIS/93/HEPU/1727	Human	Australia (QLD)			(tpi_11) B	(gdh_11) B	(bg_11) B	[27]
49	BRIS/93/HEPU/1706	Human	Australia (NSW)	Contains plasmid DNA		(tpi_86) A	(gdh_86) A	(bg_86) A	[27]
50	BRIS/92/HEPU/2342	Human	Australia (NSW)			(tpi_11) B	(gdh_74) B	(bg_11) B	Unpublished
51	AD113	Human	Australia (SA)	Fast growing		(tpi_3) A	(gdh_3) A	(bg_3) A	[72]
52	BRIS/92/HEPU/1487	Human	Papua New Guinea			(tpi_11) B	(gdh_11) B	(bg_11) B	Unpublished
53	M1P1 - 01/07/92					(tpi_11) B	(gdh_11) B	(bg_11) B	Unpublished
54	M2P1 - 21/07/92					(tpi_11) B	(gdh_11) B	(bg_11) B	Unpublished
55	M3P1 - 10/07/92					(tpi_11) B	(gdh_11) B	(bg_11) B	Unpublished
56	M3P1 - 17/11/92					(tpi_11) B	(gdh_11) B	(bg_11) B	Unpublished
57	M4P1 - 16/07/92					(tpi_11) B	(gdh_11) B	(bg_11) B	Unpublished
58	M5P1 - 01/07/92					(tpi_11) B	(gdh_11) B	(bg_11) B	Unpublished
59	M6P1 - 10/07/92					(tpi_11) B	(gdh_11) B	(bg_11) B	Unpublished
60	M7P1 - 21/07/92					(tpi_11) B	(gdh_11) B	(bg_11) B	Unpublished
61	M8P1 - 23/11/91					(tpi_11) B	(gdh_11) B	(bg_11) B	Unpublished
61	M8P1 - 26/06/92					(tpi_11) B	(gdh_11) B	(bg_32) B	Unpublished
62	KL/92/IMR/1106	Human	Malaysia	Asymptomatic human patient		(tpi_3) A	(gdh_3) A	(bg_3) A	
63	05/06/95 - Asymp - 150 ng/ μ L	Human	Afghanistan			(tpi_3) A	(gdh_3) A	(bg_3) A	
64	WB	Human	Afghanistan			(tpi_3) A	(gdh_3) A	(bg_3) A	
65	27/07/99					(tpi_3) A	(gdh_3) A	(bg_3) A	[68]cf [73]
66	1B - 19/09/97					(tpi_3) A	(gdh_3) A	(bg_3) A	[68]cf [73]

DNA sample no.	<i>Giardia duodenalis</i> isolate	Host	Locality <i>g</i>	Phenotype	Accession nos. for loci			References
					<i>pipi</i>	<i>pgdh</i>	<i>pbg</i>	
65	1B - 19/09/97				(tpi_3) A	(gdh_3) A	(bg_3) A	[68] <i>cf</i> [73]
66	1B - 22/03/99				(tpi_3) A	(gdh_3) A	(bg_3) A	[68] <i>cf</i> [73]
67	1B - 22/03/99				(tpi_3) A	(gdh_3) A	(bg_3) A	[68] <i>cf</i> [73]
68	1B - 14/06/01				(tpi_3) A	(gdh_3) A	(bg_3) A	[68] <i>cf</i> [73]
69	1B - 24/03/92 - 50 μ M			Grown in 50 μ M Mz after selection in 100 μ M Mz	(tpi_3) A	(gdh_3) A	(bg_3) A	[69] <i>cf</i> [73]
70	CLONE 16 - 03/11/97				(tpi_3) A	(gdh_3) A	(bg_3) A	<i>cf</i> [73]
71	CLONE 18 - 03/11/97				(tpi_3) A	(gdh_3) A	(bg_3) A	<i>cf</i> [73]
72	CLONE 19 - 06/11/97				(tpi_3) A	(gdh_3) A	(bg_3) A	<i>cf</i> [73]
73	CLONE 28 - 06/11/97				(tpi_3) A	(gdh_3) A	(bg_3) A	<i>cf</i> [73]
74	12/08/94 - 5M				(tpi_3) A	(gdh_3) A	(bg_3) A	[69] <i>cf</i> [73]
75	28/07/95 - 200 ng/ μ L				(tpi_3) A	(gdh_3) A	(bg_3) A	[68] <i>cf</i> [73]
76	30/08/94 - QUIN - 50 ng/ μ L				(tpi_3) A	(gdh_3) A	(bg_3) A	[74] <i>cf</i> [73]
77	16/03/95 - 5M ALB - 150 ng/ μ L			Grown in 5 μ M Qn ^d Grown in 5 μ M Mz after high dose selection + 1 μ M Az ^e ; Mz/Az resistant	(tpi_3) A	(gdh_3) A	(bg_3) A	[69] <i>cf</i> [73]
H7 Human Canada								
78	07/08/01 - 10 ng/ μ L				(tpi_11) B	(gdh_82) B	(bg_82) B	[68] <i>cf</i> [75]
BRIS/92/HEPU/1709 Cattle (<i>Bos taurus</i>) Australia (QLD)								
79	8/06/93				(tpi_91) E	(gdh_91) E	(bg_91) E	[27]
80	8/06/93				(tpi_91) E	(gdh_91) E	(bg_91) E	[27]
81	29/05/95 - 100 ng/ μ L				(tpi_91) E	(gdh_91) E	(bg_91) E	[27]
82	10/08/01				(tpi_91) E	(gdh_91) E	(bg_91) E	[27]
OAS1 Sheep (<i>Ovis aries</i>) Canada								
83	13/08/91 - 10 μ M F			Grown in 10 μ M Fz ^f	(tpi_3) A	(gdh_3) A	(bg_3) A	[68] <i>cf</i> [75]
84	29/08/91 - 10 μ M F			Grown in 10 μ M Fz	(tpi_3) A	(gdh_3) A	(bg_3) A	[68] <i>cf</i> [75]
85	2/09/91				(tpi_3) A	(gdh_3) A	(bg_3) A	[68] <i>cf</i> [75]
86	24/03/92				(tpi_3) A	(gdh_3) A	(bg_3) A	[68] <i>cf</i> [75]
BAC2 Cat (<i>Felis catus</i>) Australia (WA)								
87	13/08/91 - 25F			Grown in 25 μ M Fz	(tpi_41) A	(gdh_3) A	(bg_41) A	[68]
88	29/08/1991				(tpi_41) A	(gdh_3) A	(bg_41) A	[68]

DNA sample no.	<i>Giardia duodenalis</i> isolate	Host	Locality <i>g</i>	Phenotype	Accession nos. for loci			References
					<i>pipi</i>	<i>pgdh</i>	<i>pbg</i>	
89	01/04/92 - 25F	Sulphur-crested cockatoo (<i>Cacatua galerita</i>)	Australia (VIC)	Grown in 25 µM Fz; extra six months in high Fz concentrations	(tpi_41) A	(gdh_3) A	(bg_41) A	[68]
90	09/04/92 - 10F			Grown in 10 µM Fz	(tpi_41) A	(gdh_3) A	(bg_41) A	[68]
91	20/03/95				(tpi_3) A	(gdh_3) A	(bg_3) A	[62]
92	26/05/98				(tpi_3) A	(gdh_3) A	(bg_3) A	[62]

^aTen of 11 isolates (*i.e.* BRIS/83/HEPU/106 [DNA sample nos. 1-3]; BRIS/87/HEPU/713 [nos. 4-6]; BRIS/89/HEPU/1003 [nos. 7-8]; BRIS/91/HEPU/1279 [nos. 36-39]; BRIS/92/HEPU/1487 [nos. 52-61]; WB [nos. 63-77]; BRIS/92/HEPU/1709 [nos. 79-82]; OASI [nos. 83-86]; BAC2 [nos. 87-90]; and; BRIS/95/HEPU/2041 [nos. 91-92]), for which multiple stocks or lines (*i.e.* n = 2-15), including those grown in albendazole, furazolidone, metronidazole and quinacrine (represented by sample nos. 3-5, 69, 76, 77, 83, 84, 87, 89, 90) or reported to contain plasmid DNA (isolate BRIS/89/HEPU/1003), were available for barcoding. One isolate, BRIS/90/HEPU/1229 (nos. 9-35), representing lineages M3-M12, following multiple passages (P1-P5) represented a mixed population of *G. duodenalis* (assemblages A and B), consistent with previous findings (see [26]).

^bMz, metronidazole [PubChem number CID 4173].

^cRefers to the line and passage number; for example, BRIS/90/HEPU/1229 – M3P2 is the *in vitro* culture (line) derived from lineage 3 (M3), following two passages (P2) (see [26]).

^dQn, quinacrine [CID 237].

^eAz, albendazole [CID 2082].

^fFz, furazolidone [CID 5323714].

^gStates of Australia: QLD, Queensland; NSW, New South Wales; SA, South Australia; WA, Western Australia; VIC, Victoria.

Table 2

The eleven genotypes defined in the present study on the basis of concatenated *tpi*, *gdh* and *bg* sequence data, which represent all 25 *Giardia duodenalis* isolates and associated stocks and derived lines (= 92 DNA samples) genetically characterised in the present study, together with key reference sequences (see accession nos.) available from GenBank (<http://www.ncbi.nlm.nih.gov/>) and/or ZOOBank [43] for comparison, as well as epidemiological information and references.

Species	Inferred assemblage	Isolate/Genotype	Host origin	Locality	Accession numbers for loci			References
					(<i>tpi</i>)	(<i>pgdh</i>)	(<i>bg</i>)	
<i>Giardia duodenalis</i>								
A		Genotype 1	Human (<i>Homo sapiens</i>)	Afghanistan, Australia, Malaysia	(tpi_3) A	(gdh_3) A	(bg_3) A	Present study
		Genotype 9	Human	Australia	(tpi_86) A	(gdh_86) A	(bg_3) A	Present study
		Genotype 10	Human	Australia	(tpi_86) A	(gdh_86) A	(bg_86) A	Present study
		ISSGd141	Human	Africa	EU637593		EU637578	[48]
		ISSGd175	Human	Italy				[48]
		Genotype 7	Cat (<i>Felis catus</i>)	Australia	(tpi_41) A**	(gdh_3) A	(bg_41) A	Present study
		Swecat035	Cat	Sweden	EU781000		EU769204	[49]
		Swecat078	Cat	Sweden	EU781001		EU769205	[49]
		Swecat171	Cat	Sweden	EU781002		EU769206	[49]
		ISSCAT4	Cat	USA				[48]
		ISSCof1	Cattle (<i>Bos taurus</i>)	Italy				[48]
		NLR118	Roe Deer (<i>Capreolus capreolus</i>)	The Netherlands				
		ISSGdA334	Scimitar-horned oryx (<i>Oryx darroamah</i>)	Croatia				[76]
		Genotype 1	Sheep (<i>Ovis aries</i>)	Canada	(tpi_3) A	(gdh_3) A	(bg_3) A	Present study
		ISSGdA332	Collared peccary (<i>Pecari tajacu</i>)	Croatia				[76]
		ISSGdA696	Coatimundi (<i>Nasua nasua</i>)	Croatia				[76]
	B		Genotype 1	Sulphur-crested cockatoo (<i>Cacatua galerita</i>)	Australia	(tpi_3) A	(gdh_3) A	(bg_3) A
		Genotype 2	Human	Australia, Papua New Guinea	(tpi_11) B	(gdh_11) B	(bg_11) B	Present study
		Genotype 3	Human	Papua New Guinea	(tpi_11) B	(gdh_11) B	(bg_32) B**	Present study
		Genotype 4	Human	Australia	(tpi_11) B	(gdh_74) B**	(bg_11) B	Present study
		Genotype 5	Human	Australia	(tpi_11) B	(gdh_75) B	(bg_11) B	Present study
		Genotype 8	Human	Australia	(tpi_83) B**	(gdh_11) B	(bg_11) B	Present study
		Genotype 6	Human	Canada	(tpi_11) B	(gdh_82) B	(bg_82) B	Present study
	ISSGd177	Human	Italy				[48]	

	Species Inferred assemblage	Isolate/Genotype	Host origin	Locality	Accession numbers for loci			References
					(<i>pipi</i>)	(<i>pgdh</i>)	(<i>pbg</i>)	
C		Nij5	Human	The Netherlands			AY072725	[77]
		ISSGdA711	Barbary macaque (<i>Macaca sylvanus</i>)	Italy	EU637591		EU637580	[48]
		ISSGdA722	Barbary macaque	Italy	EU637589		EU637581	[48]
		ISSGdA297	Chimpanzee (<i>Pan troglodytes</i>)	Croatia				[76]
		Swegp138	Guinea pig (<i>Cavia porcellus</i>)	Sweden	EU781014		EU769209	[49]
		ISSGdA303	Malayan sun bear (<i>Ursus malayanus</i>)	Croatia				[76]
		ISSGdA809	Mantled Guereza (<i>Colobus guereza</i>)	Croatia				[76]
		ISSGdA748	Patagonian cavy (<i>Docilichotis patagonum</i>)	Croatia				[76]
		Swemon200	Vervet monkey (<i>Chlorocebus pygerythrus</i>)	Sweden	EU781015		EU769211	[49]
		ISSGdA810	White-handed gibbon (<i>Hyllobates lar</i>)	Croatia				[76]
D		ISSGdA777	Dog (<i>Canis familiaris</i>)	Croatia				[76]
		Sweddog074	Dog	Sweden	EU781006		EU768212	[49]
		Sweddog101	Dog	Sweden	EU781005		EU921646	[49]
		2643	Dog	USA	AY228641		AY545646	[36, 37]
		ISSGdA886	Dog	Croatia				[49]
E		Genotype 11	Cattle	Australia	(<i>tpi_91</i>) E	(<i>gdh_91</i>) E**	(<i>bg_91</i>) E	Present study
		ISSCof9	Cattle	Italy				
F		Swecattle541	Cattle	Sweden	EU781024		EU769217	[49]
		Swesheep026	Sheep	Sweden	EU781019		EU769215	[49]
		Swesheep034	Sheep	Sweden	EU781017		EU769216	[49]
		Swesheep036	Sheep	Sweden	EU781020		EU769222	[49]
		ISSGdA101	Cat	Italy				[48]
G		Swecat013	Cat	Sweden	EU781003		EU769219	[49]
		Swerat193	Rat (<i>Rattus norvegicus</i>)	Sweden				[76]
		Swerat195	Rat	Sweden	EU781013		EU769221	[49]
	<i>Giardia muris</i>	Ad-120	Mouse (<i>Mus musculus</i>)	Australia	AF069565		AY258618	[28, 78]

Table 3

The twelve sequence types characterised in the present study (with GenBank accession nos. XXXXXXXXXX), representing *G. duodenalis* assemblages A and B, which are identical to reference sequences reported previously from humans, livestock, companion and wildlife animals, and environmental and wastewater samples.

Accession nos.	Locus	Genotype	Host/source	Country (references)
Isolate 3	<i>ptpi</i>	1	Humans (<i>Homo sapiens</i>)	Afghanistan [79], Australia [80, 81], Egypt, Israel, Peru, Puerto Rico, USA [80]
			Cattle (<i>Bos taurus</i>)	USA [82, 83]
			Sheep (<i>Ovis aries</i>)	Australia [50]
			Cat (<i>Felis catus</i>)	Japan [84], USA [85]
			Dog (<i>Canis familiaris</i>)	Canada [80]
			Ferret (<i>Mustela furo</i>)	Japan [86]
			Guinea pig (<i>Cavia porcellus</i>)	USA [80]
			Beaver (<i>Castor canadensis</i>)	Canada [85]
			Herring gull (<i>Larus</i> sp.)	USA [87]
Isolate 11		2	White-tailed deer (<i>Odocoileus virginianus</i>)	USA [88]
			Humans	Australia [80, 81]
			Common dolphin (<i>Delphinus delphis</i>)	USA [2, 87]
			Common eider (<i>Somateria mollissima</i>)	USA [2, 87]
			Harp seal (<i>Pagophilus groenlandicus</i>)	USA [2, 87]
			Wastewater	USA [89]
Isolate 86		10	Herring gull	USA [2, 87]
			Humans	Australia [80], Japan [84], USA [85, 90]
			Cattle	USA [83]
			Common dolphin	USA [2, 87]
			Herring gull	USA [2, 87]
			Grey seal (<i>Halichoerus grypus</i>)	USA [2, 87]
Isolate 3	<i>pgdh</i>	1	Wastewater	USA [89]
			Humans	Australia [35, 80], Egypt, Israel, Peru, Puerto Rico [80], USA [91]
			Cat	Brazil [92], Japan [84]
			Dog	Canada [80]
Isolate 11		2	Beaver	Canada [80]
			Humans	Australia [35], Brazil [92], The Netherlands [93].
			Humans	Brazil [92]
Isolate 75		5	Humans	Brazil [92]
Isolate 82		6	Humans	Cuba [94 7056]
Isolate 86		10	Humans	Australia [35, 80], Brazil [92], Cuba [94], Japan [84, 95], Norway [96], The Netherlands [93], Uganda [97], USA [80]
Isolate 3	<i>pbg</i>	1	Human	Afghanistan [98], Australia [81], New Zealand [99], USA [39, 85]
			Cattle	India [100], New Zealand [99]
			Cat	Sweden [49] and USA [85]
			Beaver	Canada [85]

Accession nos.	Locus	Genotype	Host/source	Country (references)
Isolate 11	2		Goat (<i>Capra hircus</i>)	Belgium [101]
			Moose (<i>Alces alces</i>)	Norway [102]
			Common planigale (<i>Planigale maculata</i>)	Australia [103]
Isolate 82	6		Humans	Australia [81], Belgium [104], Japan [105], New Zealand [99]
			Cattle	New Zealand [99]
			Humans	USA USA [85, 106]
Isolate 86	10		Sheep	Norway [107]
			Beaver	Canada [85], USA [108]
			Barbary Macaque (<i>Macaca sylvanus</i>)	Italy [48]
			Cotton-top tamarin (<i>Saguinus oedipus</i>)	Sweden [49]
			Environmental samples	USA [85]
			Human	Italy [77], Norway [109], Peru [85, 110], Thailand [111]
			Cattle	Belgium [112]
			Arabian oryx (<i>Oryx leucoryx</i>)	Belgium [112]
			Bison (<i>Bison bison</i>)	Belgium [112]
			Wastewater	Mexico [106], Spain [113]