Rapid identification of *Giardia duodenalis* assemblages in NSW using terminal-restriction fragment length polymorphism

A. J. ASHER*, L. S. WALDRON *and* M. L. POWER

Department of Biological Sciences, Macquarie University, Sydney, New South Wales 2109, Australia

(*Received 28 October 2011; revised 28 January 2012; accepted 9 February 2012; first published online 15 March 2012*)

SUMMARY

Humans are infected by 2 genetic assemblages (A and B) of *Giardia duodenalis*, a protozoan parasite that causes gastrointestinal disease. Sub-assemblages AI, AII, BIII and BIV are commonly identified in human cases. Detection requires amplification of *G. duodenalis* loci. Subsequent DNA sequencing or restriction fragment length polymorphism (RFLP) identifies sub-assemblages but is expensive (DNA sequencing) or insensitive (RFLP). This study investigated a fluorescence-based detection method, using terminal-restriction fragment length polymorphism (T-RFLP) of the glutamate dehydrogenase gene to characterize human infections. Clinical samples ($n=73$), positive for *Giardia* were collected in New South Wales, Australia, and were used to evaluate T-RFLP detection. The accuracy and sensitivity of T-RFLP detection was established by comparison to DNA sequencing and RFLP. Sub-assemblage assignment by T-RFLP identified BIV as the common subtype in N.S.W cases, whilst AI, AII and BIII were also detected. When compared to DNA sequencing and RFLP, analysis by T-RFLP was a reliable and reproducible method. Automated fluorescent detection enabled accurate sizing of restriction fragments and provided a sensitive alternative to RFLP. Discrimination of sub-assemblages by T-RFLP was comparable to DNA sequencing, but was efficient and inexpensive. The protocol described here provides a rapid and sensitive diagnostic tool for routine sample screenings in epidemiological research.

Key words: *Giardia duodenalis*, T-RFLP, *gdh*, rapid identification, clinical samples.

INTRODUCTION

Giardia duodenalis is a protozoan parasite of mammals that causes giardiasis, a disease characterized by diarrhoea, nausea, fatigue, vomiting and weight loss (Roxstrom-Lindquist *et al*. 2006; Caccio and Ryan, 2008). Previously, human infection has been estimated at 280 million cases per year, worldwide (Lane and Lloyd, 2002; Caccio and Ryan, 2008). Parasite transmission via the fecal-oral route contributes to a high prevalence in areas of poor sanitary and hygiene conditions (Savioli *et al*. 2006; Caccio and Ryan, 2008). Mammals are infected by 8 morphologically indistinguishable genetic assemblages of *G. duodenalis* (syn. *G. intestinalis, G. lamblia*) (Thompson *et al*. 2000; Read *et al*. 2004; Nesselquist *et al*. 2010). These are referred to as assemblages A through to H and are only identifiable through molecular analyses (Read *et al*. 2004; Caccio and Ryan, 2008; Nesselquist *et al*. 2010). Assemblages A and B can be further differentiated into several discrete sub-assemblages (Read *et al*. 2004). Humans are infected by assemblages A and B with sub-assemblages AI, AII, BIII and BIV most commonly identified (Read *et al*. 2004; Monis *et al*. 1999).

* Corresponding author: Tel: +612 9850 9259. Fax: +612 9850 8245. E-mail: amy.asher@mq.edu.au

The epidemiology of human giardiasis is poorly understood due to complexity in species identification (Smith *et al*. 2006). Molecular data are required to document the occurrence of human infective sub-assemblages, evaluate transmission cycles and infection sources, and understand disease aetiology (Monis and Thompson, 2003; Caccio and Ryan, 2008). Identification of assemblages and sub-assemblages is typically achieved by DNA sequencing and PCR-restriction fragment length polymorphism (RFLP) (Read *et al*. 2004; Lebbad *et al*. 2008). These methods are commonly used in *Giardia* epidemiological research (Read *et al*. 2004; Traub *et al*. 2005; Yang *et al*. 2010). DNA sequencing is a robust, highly sensitive and reliable method that can be applied to all diagnostic loci. However, the expense of sequence analysis often presents as a limitation to large-scale studies (Waldron *et al*. 2009). Rapid assemblage identification is achieved by RFLP of the diagnostic locus glutamate dehydrogenase (*gdh*) (Read *et al*. 2004). Variation within the gene differentiates *G. duodenalis* assemblages and human infective sub-assemblages AI, AII, BIII and BIV (Read *et al*. 2004). Although RFLP is a cost-effective alternative, the method is limited by gel electrophoresis. Low PCR yields as a result of low parasite numbers and low molecular weights of digestion products impair analysis (Waldron *et al*. 2009).

Parasitology (2012), **139**, 1005–1013. © Cambridge University Press 2012
doi:10.1017/S0031182012000388

https://doi.org/10.1017/S0031182012000388 Published online by Cambridge University Press

This study investigated fluorescence-based detection using terminal-restriction fragment length polymorphism (T-RFLP) as an alternative tool for *G. duodenalis* identification. Fluorescence-based detection offers several advantages over conventional methods. Sample analysis is automated by the use of capillary electrophoresis enabling high precision, high throughput analysis that is rapid and sensitive (Marsh, 1999; Waldron *et al*. 2009). Terminal-RFLP is similar to RFLP. DNA is digested with restriction enzymes to produce assemblage and sub-assemblage specific profiles from sequence polymorphisms (Marsh, 1999). Terminal-RFLP, however, incorporates a fluorescently labelled primer to amplify the target gene and genotyping is based on the size of only one restriction fragment, the terminal restriction fragment (Marsh, 1999; Waldron *et al*. 2009). Upon digestion with restriction endonucleases, only the terminal restriction fragment (T-RF), adjacent to the fluorescent primer maintains fluorescence and only the T-RF is detectable by capillary electrophoresis (Marsh, 1999). Initially developed to analyse complex genetic diversity within bacterial communities (Marsh, 1999), T-RFLP now has wide applications including the differentiation of *Cryptosporidium* species and subtype families (Waldron *et al*. 2009; Waldron and Power, 2011). The purpose of this study was to develop and evaluate T-RFLP for identification of *G. duodenalis* sub-assemblages that are of importance to human disease. In developing T-RFLP for *G. duodenalis*, our aim was to provide a diagnostic tool that is a rapid, cost-effective and a sensitive alternative to current molecular techniques.

MATERIALS AND METHODS

Clinical samples

In total, 73 clinical samples positive for *Giardia* were sourced from pathology laboratories in Sydney, Australia. Sample screenings at pathology companies included a Remel ProSpecT *Giardia*/*Cryptosporidium* microplate immunoassay (Thermo Fisher Scientific© 2010) and microscopy to visualize *Giardia* cysts in positive fecal smears. Upon collection, samples were irradiated at Macquarie University using a Colbalt^{60} source (1500 Ci, 30 min) to allow for processing external to PC2 laboratory. All samples were assigned a Macquarie University *Giardia* (MQG) number and stored at 4 °C for DNA extraction.

DNA extraction and PCR amplification

DNA was extracted from 100 mg of clinical sample using an ISOLATE Fecal DNA Kit (Bioline, Sydney, Australia), following the manufacturer's instructions. Samples were confirmed as *gdh* competent by PCR amplification following the semi-nested protocol as outlined by Read *et al*. (2004). Primary and secondary reactions (25 μl volumes) were prepared using the GC-RICH PCR System, dNTPack (Roche Diagnostics, Indianapolis, IN, USA) using the reaction conditions previously described (Sulaiman *et al*. 2003; Asher *et al*. 2012). Two positive controls containing DNA extracted from laboratory cultivated *G. duodenalis* trophozoites; QIMR stocks BRIS/91/HEPU/1279 and BRIS/87/HEPU/713 (Upcroft *et al*. 1995) were included. PCR reactions were performed in an Eppendorf Mastercycler (Eppendorf, North Ryde, Australia). *Gdh* amplicons were resolved by agarose gel electrophoresis (2% w/v, 110 V for 30 min) in TBE (Tris, EDTA, boric acid, pH 8) with SYBR safe (Invitrogen, Mulgrave, Australia) using a HyperLadder™ II DNA marker (Bioline, Sydney, Australia) to estimate amplicon size (432 bp fragment).

DNA sequencing, analyses and phylogenetic inference

To establish a set of reference samples (representatives of AI, AII, BIII and BIV) for T-RFLP development and evaluation, a subset of clinical samples ($n=40$) were genotyped by DNA sequencing. Two trophozoite DNA controls were also included. Samples were amplified as described above and secondary *gdh* amplicons were purified using a QIAquick PCR Purification Kit (Qiagen, Melbourne, Australia), following the manufacturer's instructions. Products were sequenced in the forward direction on a 3130×1 genetic analyser (Applied Biosystems, Foster City, CA, USA) using the standard run protocol for a 50 cm, 16 capillary array using a BigDye™ terminator kit (Applied Biosystems). *Giardia duodenalis* subtypes were identified by aligning nucleotide sequences to previously described *gdh* DNA sequences retrieved from the NCBI GenBank database (http://www.ncbi.nlm.gov/genbank/index.html) for sub-assemblages AI, AII, BIII and BIV (Accession numbers; L40509, L40510, AF069059, L40508) (Monis *et al*. 1996, 1999); and Assemblages C, D, E, F and G (Accession numbers; U60982, U60986, AY178740, AF069057, AF069058) (Ey *et al*. 2002 *unpublished observations* Monis *et al*. 1998, 1999). For assemblage H only partial sequences of the 432 bp amplicon were available in GenBank, and assemblage H could not be included in sequence analyses. DNA sequences were trimmed in GeneiousPRO version 5.0.3 (Biomatters Ltd, Auckland, New Zealand). Trimmed sequences were aligned using Clustal W (Thompson *et al*. 1994) in *MEGA* version 5 (Tamura *et al*. 2011). For phylogenetic analyses, nucleotide substitution models were tested for maximum likelihood in *MEGA*5. Akaike Information Criterion corrected (AICc) values were used to determine the optimal parameters. A phylogenetic tree was

https://doi.org/10.1017/S0031182012000388 Published online by Cambridge University Press

constructed using maximum likelihood (Tamura 3-parameter distance model with the gamma distribution parameter) and bootstrap analysis (1000 replicates). A *gdh* DNA sequence for *Giardia ardae* (Accession no. AF069060) (Monis *et al*. 1999) was used as outgroup.

Design and development of T-RFLP protocol

T-RFLP design was based on the previously described *gdh* RFLP protocol (Read *et al*. 2004). Nucleotide sequences (*gdh*) for sub-assemblages AI, AII, BIII and BIV were used to identify T-RFs and calculate expected sizes. *Nla IV* and *Rsa I* restriction sites were mapped within the secondary primer region, and terminal restriction profiles were determined for each sub-assemblage.

The reference group of samples (see above) comprised 4 clinical samples and 2 trophozoite DNA controls, identified as sub-assemblages AI, AII, BIII and BIV from DNA sequencing. The reference group was cloned to overcome potential for mutliple T-RFs from mixed infections in clinical samples. Cloning was performed using the pGEM plasmid vector (Promega, Madison, USA) system and α-Select Gold Efficiency *E. coli* cells (Bioline, Sydney, Australia). Recombinant plasmids were purified using a QIAprep Miniprep Kit (Qiagen, Australia), sequenced and sub-assemblages confirmed as previously described.

Gdh PCR was optimized to incorporate a fluorescent primer and generate a fluorescent PCR product. The reverse primer GDHiR was labelled with a 5′6-carboxyfluorescein (6-FAM) (Sigma-Aldrich, Sydney, Australia) and used in the secondary reaction. A gradient PCR (temperature range 52–60 °C, 2 °C increments) was performed using positive DNA controls to determine the optimal annealing temperature for the fluorescent primer.

Purified fluorescent PCR products were digested separately using 2 restriction enzymes *Nla IV* (New England Biolabs) and *Rsa I* (Roche Diagnostics, Indianapolis, IN, USA) according to the protocol developed by Read *et al*. (2004).

Capillary electrophoresis for detection of fluorescent terminal restriction fragments

Digested fluorescent *gdh* amplicons were analysed by capillary electrophoresis at the Macquarie University Sequencing Facility, as previously described (Waldron *et al*. 2009) using the DNA size standard LIZ 500 (Applied Biosystems). Outputs were analysed using GeneScan software, version 4.0 (Applied Biosystems). A complete *gdh* PCR sequence (432 bp) was included in each run to control for peak shifts that may arise from run to run variability (Waldron *et al*. 2009).

Comparison of electrophoretic detection methods for identification of restriction fragments

Detection of diagnostic restriction fragments for identification of *G. duodenalis* assemblages is achieved by T-RFLP and RFLP analyses. These methods differ by the electrophoretic technique used to detect base pair lengths of restriction fragments. To examine the sensitivity of T-RFLP for *Giardia* identification and to compare T-RFLP to RFLP, fluorescent PCR amplification and enzymatic digestion was conducted on an additional 33 clinical samples. Digested products were separated by agarose gel electrophoresis (3·5% w/v, 100 V for 50 min) using 10 μl of product. These samples were also analysed by capillary electrophoresis (T-RFLP) and results were compared.

RESULTS

Identification of G. duodenalis *assemblages by DNA sequencing and phylogenetic analysis*

Prior to evaluating T-RFLP a subset of samples with known *gdh* sequences was needed to develop the method. DNA sequencing was conducted on 40 clinical samples and phylogenetic analysis was performed to confirm genotyping. Sequencing was successful for the 40 clinical samples selected. For 36 of these, *gdh* nucleotide sequences were 100% homologous to GenBank references for sub-assemblages AII and BIV. Trophozoite DNA controls were identified as AI. Four clinical samples (MQG95, MQG103, MQG127, MQG136) contained several single nucleotide polymorphisms (Table 1). Nucleotide sequence searches in the GenBank database identified these samples as sub-assemblages BIII or BIV (Table 1). MQG136 could not be resolved below the assemblage level due to sequence diversity.

Phylogenetic analysis grouped clinical samples with GenBank references for sub-assemblages AII, BIII and BIV (Fig. 1). The phylogenetic tree consisted of 2 major clusters and 7 distinct branches representing assemblages A to G. Human infective assemblages A and B formed 2 highly divergent groups, as depicted by their phylogenetic distance in the tree. Samples containing higher nucleotide sequence diversity, MQG103 and MQG127 were grouped together in a clade that was sister to sub-assemblage BIV. Sample MQG136 was positioned internally to sub-assemblages BIII and BIV. Bootstrap analyses (1000 replicates) showed strong support for clustering of clinical samples within Assemblages A and B. All *gdh* sequences generated in this study were submitted to NCBI GenBank database (http://www.ncbi.nlm.gov/genbank/index.html) under Accession numbers JQ700387–JQ700436.

https://doi.org/10.1017/S0031182012000388 Published online by Cambridge University Press

Table 1. Nucleotide sequence diversity observed in DNA sequences for 4 clinical samples

	GenBank sequences		Clinical samples			
Position (bp)[a]	AF069059 (BIII)	L40508 (BIV)	MQG95	MQG103	MQG127	MQG136
183	T	T	T	T	T	C
345	C	C	T	C	C	T
366	C	T	C	T	T	C
387	C	T	C	C	C	C
396	C	C	C	T	T	C
423	C	C	C	T	C	C
438	G	A	G	A	A	G
Similarity (%) to GenBank sequences			99·7 (BIII)	99·2 (BIV)	99·4 (BIV)	98·6 (BIII) 98·6 (BIV)

[a] Position (bp) refers to the published sequence for L40508.

Development and evaluation of T-RFLP protocol

T-RF lengths (bp) were determined using *gdh* nucleotide sequences that were retrieved from GenBank. Both restriction endonucleases were required to differentiate T-RFs at the assemblage and sub-assemblage level. At the assemblage level the expected T-RFs were 131 bp (assemblage A) and 288 bp (assemblage B), whilst the predicted T-RFs for sub-assemblages were; 146 (AI), 69 (AII), 131 (BIII), and 428 (BIV).

Eight clones containing the recombinant plasmid for sub-assemblages AI, AII, BIII and BIV were generated for T-RFLP development, and were used as representative controls of known genotypes. T-RFLP on cloned sequences showed that T-RFs were consistent with those expected for assemblages A and B, and sub-assemblages AI, AII, BIII and BIV. T-RFs were within 3–4 base pairs of the predicted size (Fig. 2). This result was in agreement with variation observed for the internal control (undigested *gdh* PCR amplicon) detected at 428 bp.

Evalutation of T-RFLP: comparison to DNA sequencing and RFLP

Fourty clinical samples were initially genotyped by DNA sequencing, and of these 39 were screened by T-RFLP. One sample could not be re-amplified by fluorescent PCR and did not undergo T-RFLP analysis. Genotyping by T-RFLP matched DNA sequencing data for 38 (97%) of the 39 samples screened by both methods (Table 2). T-RFLP detected a mixed template in one clinical sample which was not detected by DNA sequencing.

An additional 33 clinical samples were genotyped using T-RFLP and RFLP. For RFLP, many restriction fragments could not be resolved by eye following gel electrophoresis. High PCR yields did not produce clear restriction profiles. The sizes of restriction fragments below 100 bp could not be determined and fragments that were of similar sizes could not be distinguished. We were able to genotype 28 (85%) of the 33 clinical samples by RFLP. Following digestion with *Nla IV* and gel-electrophoresis restriction profiles for 5 samples contained a large fragment at approximately 120 bp. Diagnostic fragments below 100 bp could not be differentiated and RFLP profiles could not be accurately determined. For these samples capillary electrophoresis (T-RFLP) detected T-RFs at 65 bp and was consistent with sub-assemblage AII. For sub-assemblage BIV, a 4 bp difference between the full *gdh* nucleotide sequence (432 bp) and a large restriction fragment for BIV (428 bp) could not be differentiated by gel electrophoresis. A novel restriction profile was generated for one clinical sample. For this sample T-RFLP and RFLP data were consistent. Profiles did not correspond to any *G. duodenalis* assemblage or sub-assemblage and the sample was not identified.

Diversity of G. duodenalis *subtypes in human samples from NSW*

Seventy-two clinical samples were successfully genotyped (Table 2). Assemblage B was detected in 61 samples (85%) and of these 54 samples were sub-assemblage BIV, 3 were sub-assemblage BIII and 4 clinical samples generated multiple T-RFs consistent with mixed BIII/BIV templates (Table 2). Assemblage A was detected in 11 clinical samples (15%) with sub-assemblages AI ($n=1$) and AII ($n=10$).

Unique T-RFLP profiles were generated for 17 assemblage B clinical samples. Profiles contained additional T-RFs that were observed as low intensity secondary fluorescent peaks in the electropherograms (fluorescence intensity approximately $\leqslant 20\%$ of primary genotyping peak). These minor peaks did not correspond to any other *G. duodenalis* assemblage. Secondary peaks that were consistent with a mixed

https://doi.org/10.1017/S0031182012000388 Published online by Cambridge University Press

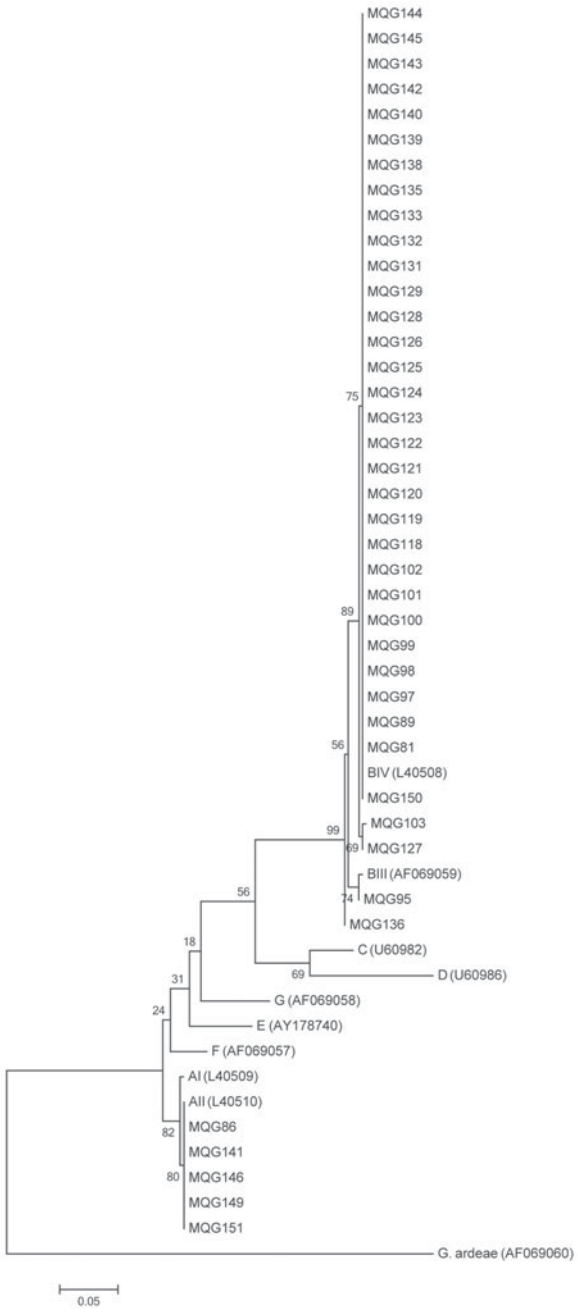


Fig. 1. Phylogenetic tree of *Giardia duodenalis gdh* sequences inferred by maximum likelihood. The phylogenetic tree shows *gdh* sequences (359 bp) for 40 clinical samples and 10 reference sequences obtained from GenBank. Clinical samples are divided into 2 main clusters of assemblages A and B. Bootstrap values (%) supporting the topology of the tree are placed to the left of each node.

template or that were of high fluorescent intensity were investigated further by comparing T-RFLP results to sequencing data.

Investigation of mixed T-RFs and sequence diversity

Sequencing data were available for 3 clinical samples (MQG123, MQG127, MQG136) and 1 cloned sequence (Clone 95-1) that contained multiple peaks in the T-RFLP profiles. Additional T-RFs for samples MQG123, MQG127 and MQG136 were consistent with mixed BIII/BIV templates. Investigation of DNA chromatograms revealed double peaks at position 366 of the published gene (L40508), which corresponded to an additional *Rsa I* restriction site, producing a T-RF for sub-assemblage BIII. The additional restriction site for sub-assemblage BIII could not be located in the chromatogram for sample MQG123 and DNA sequencing/T-RFLP data were not consistent for this sample.

Two reverse primer binding sites for GDHiR were identified in the DNA chromatogram for cloned sequence 95-1. These sites generated 2 secondary PCR products at 432 bp and 499 bp. Multiple T-RFs observed in the T-RFLP profile for clone 95-1 (*Nla IV* and *Rsa I*) were consistent with digestion of both PCR products and BIII restriction sites (Fig. 3).

DISCUSSION

This study investigated T-RFLP as a cost effective and alternative molecular diagnostic tool for *G. duodenalis* identification. *G. duodenalis* is globally distributed and poses a continuous threat to human populations (Caccio and Ryan, 2008). Understanding complex host-parasite relationships and disease transmission cycles requires identification of genetic subtypes and the development of reliable but inexpensive molecular tools for routine sample screening.

Analyses of 73 clinical samples showed that T-RFLP is a sensitive, robust and effective diagnostic tool for *G. duodenalis* identification. Fluorescent *gdh* amplification followed by enzyme digestion produced fluorescent T-RFs that were assemblage and sub-assemblage specific. Terminal-RFs were within 3-4 base pairs of the predicted size due to shifts in the detection system (Schutte *et al*. 2008; Waldron *et al*. 2009). Terminal-RFs are detected by capillary electrophoresis and base pair length is measured against a size standard (Marsh, 1999; Schutte *et al*. 2008). Similar to RFLP, factors such as temperature and the amount of DNA product can cause slight variations in the size that the T-RF is detected at. By including a *gdh* nucleotide sequence of known length to control for these shifts, the size of each T-RF could be accurately determined.

We evaluated the sensitivity and accuracy of T-RFLP by comparing the method with DNA sequencing and RFLP for the *gdh* gene. Consistent results between DNA sequencing and T-RFLP data demonstrated that fluorescent detection of the T-RF is a reliable and reproducible method. Of the 73 clinical samples used in this study, we were able to screen 72 by T-RFLP. Whilst DNA sequencing is still required to resolve ambiguous samples, we found that T-RFLP is a fast and effective genotyping tool to

https://doi.org/10.1017/S0031182012000388 Published online by Cambridge University Press

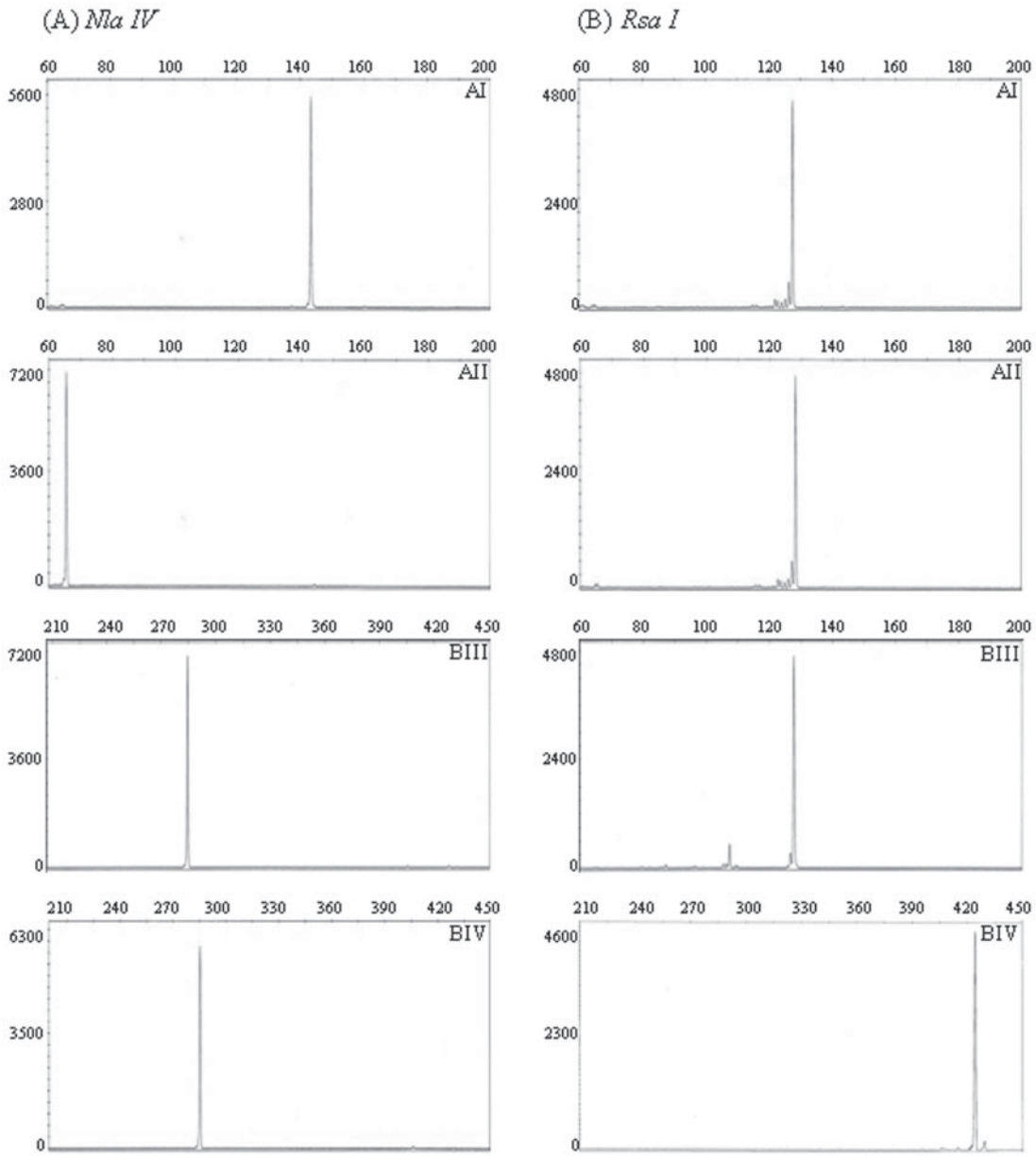


Fig. 2. Electropherograms showing assemblage and sub-assemblage specific T-RFs following digestion with *Nla IV* and *Rsa I*. Digestion with *Nla IV* generated T-RFs observed at 143 bp (AI), 65 bp (AII) and 285 bp (assemblage B). Sub-assemblages BIII and BIV are not differentiated by *Nla IV*. Digestion with *Rsa I* produced T-RFs observed at 128 bp (BIII), 424 bp (BIV) and 128 bp (assemblage A). A minor secondary peak at 108 bp is visible in the electropherogram for BIII (*Rsa I*) and is not typical of the BIII T-RFLP profile. All T-RFs are within 3 to 4 bp of the predicted sizes.

routinely screen samples, and a low cost alternative. Automated detection enabled high efficiency, high throughput processing (Marsh, 1999) and rapid identification of sub-assemblages. Although DNA sequencing is automated, additional thermal cycling and purification prior to capillary electrophoresis is required, and increases sample turn-around times by comparison. To evaluate the protocol, we chose to analyse *Nla IV* and *Rsa I* digestions for each sample separately. T-RFLP for both digestions per sample was performed at 33% of the DNA sequencing price but could be processed at approximately 16%, if digested products for each sample were combined for capillary electrophoresis.

In comparison with RFLP–gel electrophoresis, T-RFLP was a more sensitive and precise technique, requiring a lower concentration of template DNA and only 1 μl for analysis. Sub-assemblage assignment by RFLP was problematic when diagnostic restriction fragments were of similar base pair size, or when partial digestions produced ambiguous profiles. Low molecular weight restriction fragments were also not visible following RFLP and gel electrophoresis. In contrast these fragments were

https://doi.org/10.1017/S0031182012000388 Published online by Cambridge University Press

Table 2. Genotyping results for 73 giardiasis clinical samples

Genotyping method	Total number screened	Sub-assemblages detected					
		AI	AII	BIII	BIV	BIII/BIV	Unknown
DNA sequencing and T-RFLP[a]	40		5	1	31	3	
RFLP and T-RFLP	33	1	5	2	23	1	1
Total	73	1	10	3	54	4	1

[a] Of 40 clinical samples with DNA sequencing data, 39 were screened by T-RFLP.

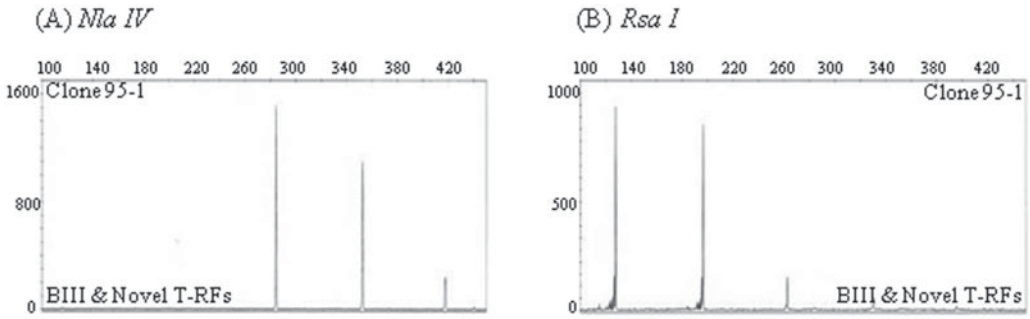


Fig. 3. Electropherograms showing multiple T-RFs consistent with nucleotide sequence diversity. Digestion of cloned sequence 95-1 with *Nla IV* (A) generated a primary T-RF observed at 285 bp and secondary T-RFs at 352 bp and 418 bp. Digestion with *Rsa I* (B) produced a primary T-RF observed at 128 bp and secondary T-RFs at 196 bp, 263 bp and 329 bp. Primary T-RFs are consistent with sub-assemblage BIII. Secondary T-RFs were generated by an additional reverse primer binding site in the nucleotide sequence.

clearly identified by fluorescent detection of the T-RF. Digital outputs of fluorescent peaks allowed accurate sizing, and T-RFs of similar base-pair length could be unambiguously differentiated.

For a number of samples in this study, evidence of mixed templates and/or sequence heterogeneity was detected by DNA sequencing and T-RFLP. Mixed templates have been previously identified in human cases; however, the mechanisms that generate heterogeneity are not well understood (Caccio and Ryan, 2008; Lebbad *et al*. 2008). *Giardia* trophozoites contain 2 nuclei, and it is unclear whether mixed templates and heterogeneity result from mixed infection or genetic exchange (Lebbad *et al*. 2008; Caccio and Sprong, 2010).

For this type of detection we found T-RFLP to be more informative than RFLP and DNA sequencing. Due to the low resolving power of RFLP and gel electrophoresis, we found that the higher sensitivity of capillary electrophoresis improved accuracy in detection and confidence in sub-assemblage assignment. Capillary electrophoresis has been previously applied to RFLP analyses (Nachamkin *et al*. 2001) but would produce complex electropherograms for *G. duodenalis*. Digestion of *gdh* sequences with *Nla IV* and *Rsa I* generates numerous restriction fragments, many of which are non-specific. Detection of these fragments by capillary electrophoresis would generate complex banding patterns and profiles containing mixed templates would be difficult to resolve. Although T-RFLP limits analysis of sequence diversity to the terminal restriction site, electropherograms are simple to interpret and provide a clear distinction between fluorescent peaks of different sub-assemblages.

Comparison of T-RFLP to DNA sequencing data identified double peaks in the DNA chromatograms of 2 clinical samples, which generated additional enzyme restriction sites and corresponded to more than one *G. duodenalis* sub-assemblage. Whilst double peaks in chromatograms can indicate mixed templates or heterogeneity, it can also be the product of PCR artifacts (Lebbad *et al*. 2008). Analysis by T-RFLP confirms that sequence diversity is present in the sample, through the production of terminal restriction fragments that would not otherwise appear.

Further modelling of the T-RFLP protocol has shown that whilst sub-assemblage AII can be differentiated from sub-assemblages AI, BIII and BIV, different primers and/or restriction enzymes are required to differentiate AII from other *G. duodenalis* assemblages. Whilst we were able to confirm the presence of sub-assemblage AII using additional molecular data, the protocol must be optimized to investigate zoonotic transmission.

Seventy-two clinical samples were genotyped in this research, and sub-assemblage BIV was most commonly detected (75% of samples), followed by AII (14% of samples). Our results correspond with

https://doi.org/10.1017/S0031182012000388 Published online by Cambridge University Press

previous research where prevalence of assemblage B has been documented in Western Australia during separate investigations of sporadic giardiasis (Hopkins *et al*. 1997; Read *et al*. 2002; Yang *et al*. 2010). Based on our analyses of *gdh* sequences, we have identified the presence of 4 *G. duodenalis* subassemblages causing human infection in NSW. Our results have provided preliminary data for NSW and future research is required to present a more accurate estimation of prevalence in this region.

In this study, *G. duodenalis* genotyping was conducted at the *gdh* locus. To confirm subassemblage assignment, analyses across several loci are usually required (Caccio and Ryan, 2008). The high cost of DNA sequencing limits the scope of many investigations (Waldron *et al*. 2009) and a range of diagnostic methods are often used to achieve multi-locus genotyping. We have described the first application of T-RFLP for the identification of *G. duodenalis* subtypes that are of importance to human disease. Due to the high diagnostic performance of the protocol, we expect that T-RFLP can be adapted to detect a wider range of assemblages and applied to other diagnostic loci. Detection of sequence diversity and the generation of unique sample profiles in this study have demonstrated potential application as a rapid and inexpensive source tracking tool in outbreak situations. The high precision and accuracy of this technique combined with the low cost and high throughput nature means that T-RFLP is an optimal tool for routine identification of *G. duodenalis*. Detection by T-RFLP should improve processing efficiency and provide a cost-effective means to screen larger sample sizes, generating the necessary molecular information required to understand the epidemiology of human disease.

ACKNOWLEDGMENTS

This research was conducted with financial assistance from Macquarie University. We would like to thank Dr Jacqui Uprcoft of the Queensland Institute of Medical Research for kindly providing trophozoites and also Douglass Hanly Moir Pathology, Symbion Health Laverty Pathology, Hunter Area Pathology Service, Westmead Hospital and John Hunter Hospital for their invaluable help in providing fecal samples.

FINANCIAL SUPPORT

This research received no specific funding but was supported by the Department of Biological Sciences, Macquarie University.

REFERENCES

Asher, A. J., Waldron, L. S. and Power, M. L. (2012). Evaluation of a PCR protocol for sensitive detection of *Giardia intestinalis* in human faeces. *Parasitology Research* **110**, 853–858. doi: 10.1007/s00436-011-2565-3.

Caccio, S. M. and Ryan, U. (2008). Molecular epidemiology of giardiasis. *Molecular and Biochemical Parasitology* **160**, 75–80. doi: 10.1016/j.molbiopara.2008.04.006.

Caccio, S. M. and Sprong, H. (2010). *Giardia duodenalis*: genetic recombination and its implications for taxonomy and molecular epidemiology. *Experimental Parasitology* **124**, 107–112. doi: 10.1016/j.exppara.2009.02.007.

Hopkins, R. M., Meloni, B. P., Groth, D. M., Wetherall, J. D., Reynoldson, J. A. and Thompson, R. C. A. (1997). Ribosomal RNA sequencing reveals differences between the genotypes of *Giardia* isolates recovered from humans and dogs living in the same locality. *Journal of Parasitology* **83**, 44–51.

Lane, S. and Lloyd, D. (2002). Current trends in research into the waterborne parasite *Giardia*. *Critical Reviews in Microbiology* **28**, 123–147.

Lebbad, M., Ankarklev, J., Tellez, A., Leiva, B., Andersson, J. O. and Svard, S. (2008). Dominance of *Giardia* assemblage B in Leon, Nicaragua. *Acta Tropica* **106**, 44–53. doi: 10.1016/j.actatropica.2008.01.004.

Marsh, T. (1999). Terminal restriction fragment length polymorphism (T-RFLP): an emerging method for characterising diversity among homologous populations of amplification products. *Current Opinion in Microbiology* **2**, 323–327.

Monis, P. T., Andrews, R. H., Mayrhofer, G. and Ey, P. L. (1999). Molecular systematics of the parasitic protozoan *Giardia intestinalis*. *Molecular Biology and Evolution* **16**, 1135–1144.

Monis, P. T., Andrews, R. H., Mayrhofer, G., Mackrill, J., Kulda, J., Isaac-Renton, J. L. and Ey, P. L. (1998). Novel lineages of *Giardia intestinalis* identified by genetic analysis of organisms isolated from dogs in Australia. *Parasitology* **116**, 7–19.

Monis, P. T., Mayrhofer, G., Andrews, R. H., Homan, W. L., Limper, L. and Ey, P. L. (1996). Molecular genetic analysis of *Giardia intestinalis* isolates at the glutamate dehydrogenase locus. *Parasitology* **112**, 1–12.

Monis, P. T. and Thompson, R. C. A. (2003). *Cryptosporidium* and *Giardia*-zoonoses: fact or fiction? *Infection, Genetics and Evolution* **3**, 233–244. doi: 10.1016/j.meegid.2003.08.003.

Nachamkin, I., Panara, N. J., Li, M., Ung, H., Yuen, P. K., Kricka, L. J. and Wilding, P. (2001). Agilent 2100 bioanalyzer for restriction fragment length polymorphism analysis of the *Campylobacter jejuni* flagellin gene. *Journal of Clinical Microbiology* **39**, 754–757. doi: 10.1128/JCM.39.2.754-757.2001.

Nesselquist, E. L., Welch, D. M. and Sogin, M. L. (2010). The identification of a new *Giardia duodenalis* assemblage in marine vertebrates and a preliminary analysis of *G. duodenalis* population biology in marine systems. *International Journal for Parasitology* **40**, 1063–1074. doi: 10.1016/j.ijpara.2010.02.015.

Read, C. M., Monis, P. T. and Thompson, R. C. A. (2004). Discrimination of all genotypes of *Giardia duodenalis* at the glutamate dehydrogenase locus using PCR-RFLP. *Infection, Genetics and Evolution* **4**, 125–130. doi: 10.1016/j.meegid.2004.02.001.

Read, C., Walters, J., Robertson, I. D. and Thompson, R. C. (2002). Correlation between genotype of *Giardia duodenalis* and diarrhoea. *International Journal for Parasitology* **32**, 229–231.

Roxstrom-Lindquist, K., Palm, D., Reiner, D., Ringqvist, E. and Svard, S. G. (2006). *Giardia* immunity – an update. *Trends in Parasitology* **22**, 26–31. doi: 10.1016/j.pt.2005.11.005.

Savioli, L., Smith, H. and Thompson, A. (2006). *Giardia* and *Cryptosporidium* join the 'Neglected Diseases Initiative.' *Trends in Parasitology* **22**, 203–208. doi: 10.1016/j.pt.2006.02.015.

Schutte, U. M. E., Abdo, Z., Bent, S. J., Shyu, C., Williams, C. J., Pierson, J. D. and Forney, L. J. (2008). Advances in the use of terminal restriction fragment length polymorphism (T-RFLP) analysis of 16S rRNA genes to characterize microbial communities. *Applied Microbiology and Biotechnology* **80**, 365–380. doi:10.1007/s00253-008-1565-4.

Smith, H. V., Caccio, S. M., Tait, A., McLauchlin, J. and Thompson, R. C. (2006). Tools for investigating the environmental transmission of *Cryptosporidium* and *Giardia* infections in humans. *Trends in Parasitology* **22**, 160–167. doi: 10.1016/j.pt.2006.02.009.

Sulaiman, I. M., Fayer, R., Bern, C., Gilman, R. H., Trout, J. M., Schantz, P. M., Das, P., Lal, A. A. and Xiao, L. (2003). Triosephosphate isomerase gene characterization and potential zoonotic transmission of *Giardia duodenalis*. *Emerging Infectious Diseases* **9**, 1444–1452.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **10**, 2731–2739.

https://doi.org/10.1017/S0031182012000388 Published online by Cambridge University Press

Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, population-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673–4680. doi:10.1093/nar/22.22.4673.

Thompson, R. C. A., Hopkins, R. M. and Homan, W. L. (2000). Nomenclature and genetic groupings of *Giardia* infecting mammals. *Parasitology Today* **16**, 210–213.

Traub, R. J., Monis, P. T. and Robertson, I. D. (2005). Molecular epidemiology: a multidisciplinary approach to understanding parasitic zoonoses. *International Journal for Parasitology* **35**, 1295–1307. doi: 10.1016/j.ijpara.2005.06.008.

Upcroft, J. A., Boreham, P. F. L., Campbell, R. W., Shepherd, R. W. and Upcroft, P. (1995). Biological and genetic analysis of a longitudinal collection of *Giardia* samples derived from humans. *Acta Tropica* **60**, 35–46.

Waldron, L. S., Ferrari, B. C., Gillings, M. R., and Power, M. L. (2009). Terminal restriction fragment length polymorphism for identi fication of *Cryptosporidium* species in human feces. *Applied and Environmental Microbiology* **75**, 108–111. doi:10.1128/AEM.01341-08.

Waldron, L. S. and Power, M. L. (2011). Fluorescence analysis detects *gp60* subtype diversity in *Cryptosporidium* infections. *Infection, Genetics and Evolution* **11**, 1388–1395. doi:10.1016/j.meegid.2011.05.008.

Yang, R., Lee, J., Ng, J. and Ryan, U. (2010). High prevalence *Giardia duodenalis* assemblage B and potentially zoonotic subtypes in sporadic human cases in Western Australia. *International Journal for Parasitology* **40**, 293–297. doi: 10.1016/j.ijpara.2009.08.003.

https://doi.org/10.1017/S0031182012000388 Published online by Cambridge University Press