# Whole genome sequencing of clinical isolates of *Giardia lamblia*

**K. Hanevik[1,2], R. Bakken[1], H. R. Brattbakk[1,3], C. S. Saghaug[1] and N. Langeland[1,4]**

1) *Department of Clinical Science, University of Bergen,* 2) *National Center for Tropical Infectious Diseases, Department of Medicine,* 3) *Center for Medical Genetics and Molecular Medicine, and* 4) *Department of Medicine, Haukeland University Hospital, Bergen, Norway*

## Abstract

Clinical isolates from protozoan parasites such as *Giardia lamblia* are at present practically impossible to culture. By using simple cyst purification methods, we show that *Giardia* whole genome sequencing of clinical stool samples is possible. Immunomagnetic separation after sucrose gradient flotation gave superior results compared to sucrose gradient flotation alone. The method enables detailed analysis of a wide range of genes of interest for genotyping, virulence and drug resistance.



**Keywords:** Clinical isolates, *Giardia*, immunomagnetic separation, sucrose gradient flotation, whole genome sequencing
**Original Submission:** 1 June 2014; **Revised Submission:** 13 August 2014; **Accepted:** 15 August 2014
Editor: G. Greub
**Article published online:** 29 October 2014

**Corresponding author**: K. Hanevik, University of Bergen, Department of Clinical Science, 8th floor, Lab-building, N-5021 Bergen, Norway
**E-mail:** kurt.hanevik@med.uib.no

Resistance in the intestinal protozoan parasite *Giardia lamblia* (syn. *G. intestinalis, G. duodenalis*) against metronidazole is increasingly seen in clinical practice, but clinically relevant mechanisms of resistance are largely unknown [1]. Many issues of *Giardia* virulence and pathogenicity are also unresolved [2]. Further progress in this field is hampered by the difficulty of culturing clinical samples [3]. Genotyping of *Giardia* for studies of epidemiology, symptomatology, virulence and drug resistance has therefore been limited to polymerase chain reaction (PCR) amplification and partial sequencing of one, or a few, specific genes [4,5].

Recently, genomes of three culturable isolates, representative of *G. lamblia* assemblages infective to humans (A1, A2 and B), have been sequenced [6–8]. The aim of the present study was to evaluate whether it is possible to obtain a satisfactory coverage and depth with whole genome sequencing of cysts purified directly from patients' stool samples.

*Giardia* cyst purification was performed in three clinical samples, one assemblage A2 (sample 1) and two assemblage B (sample 2 and 3), using 4 g of faeces from each. Samples were washed by repeated centrifugation at 500 × *g* with $dH_2O$ until the supernatant was clear. Then sucrose gradient flotation (SF) was done by placing the remaining pellet material on top of a 1 M sucrose solution before centrifuging at 600 × *g* for 10 minutes. The supernatant was collected and went through one repeated SF cycle before the remaining material above the sucrose gradient was washed again, and half of it became the SF fraction from which DNA was extracted (Fig. 1a). The other half was further purified by immunomagnetic separation (IMS) with magnetic beads coupled to antibodies against *Giardia* cyst wall protein-1 (Waterborne Inc., New Orleans, LA) (Fig. 1b). Cultured *Giardia* trophozoites (strain WB, ATCC 50803) were included for reference and internal control.

DNA was extracted using the Qiagen Stool Kit after five freeze/thaw cycles and was measured using Qubit 2.0 (Life Technologies, Grand Island, NY). *Giardia* genotype was determined by triosephosphate isomerase (tpi)-PCR and sequencing [9]. DNA quality was assessed by Agilent 2100 BioAnalyzer using the DNA 1000 and High Sensitivity DNA kit (Agilent Technologies, Santa Clara, CA) and was sequenced on a SOLiD 5500xl system after per-protocol fragmentation and random amplification. LifeScope Genomic Analysis Software v.2.5.1 for SOLiD Next-Generation Sequencing and the genomic.resequencing.frag workflow were used for mapping the resulting reads to the relevant *Giardia* reference genome: assemblage A1 (WB) [6], assemblage A2 (DH) [8] and B (GS) [7]. Mapped sequences were analysed using the software packages SAMtools (http://samtools.sourceforge.net/) and R (http://www.r-project.org/) to evaluate coverage and depth of the genome as a whole and in a selected set of three genes commonly used for genotyping (*tpi, glutamate dehydrogenase, β-giardin*) [5], four genes important in host–pathogen interactions (*arginine deiminase, ornithine carbamoyltransferase, fructose-bisphosphate aldolase, enolase*) [10] and seven metabolic genes involved in metronidazole function (*NADH oxidase, NADH ferredoxin oxidoreductase, pyruvate ferredoxin oxidoreductase, thioredoxin*



http://dx.doi.org/10.1016/j.cmi.2014.08.014

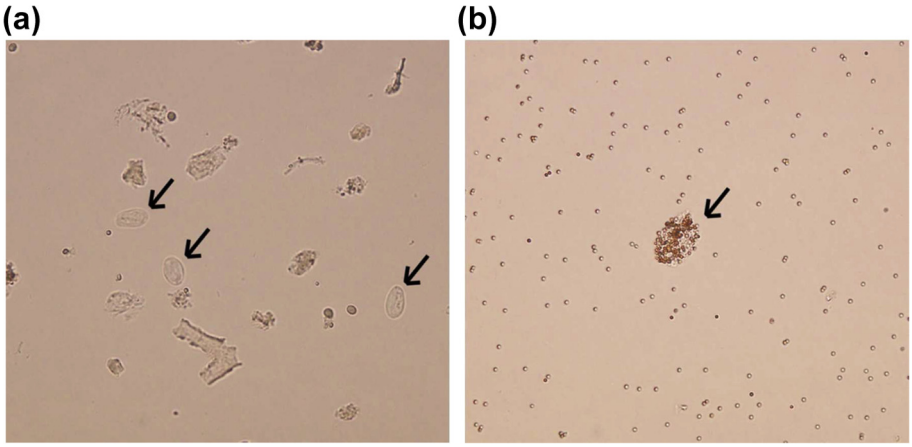


**FIG. 1.** (a) Three *Giardia* cysts (arrow) after sucrose gradient flotation. Some faecal debris and probably fungi are still present. (b) One *Giardia* cyst (arrow) after immunomagnetic separation. It is covered by magnetic beads coupled to cyst wall protein. Original magnification, 40×.

*reductase, nitroreductase family protein, malate dehydrogenase, nitroreductase Fd-NR2*) [11].

The DNA extraction yields were low for all three clinical samples. Levels of IMS purified samples were all below the Q-bit measurement range (<0.01 ng/μL) (Table 1). Between seven and ten amplification cycles were used before sequencing. Sequencing read lengths were between 25 and 75 bases.

Mapping data for the cultured control and purified clinical samples are listed in Table 1. When mapped against their matching reference genomes, reads from two out of three SF purified clinical samples revealed a 1× genome coverage of >87% and above 85% for all the clinical samples purified by IMS after SF. The IMS-purified samples had the highest depth, with 36.5%, 63.7% and 82.5% of the genome covered 10×.

Evaluations of single gene coverages are shown in Supplementary Tables S1 to S6. Within the IMS-purified samples, the 1× coverage of the complete length of the three genotyping genes was above 87.7% in sample 1, 100% in sample 2 and above 99.3% in sample 3. The corresponding figures for the SF-purified samples were between 11.5% and 100%. For the host–pathogen interaction genes, we also found the IMS-purified samples to cover over 99% of the complete gene sequences in samples 2 and 3, while sample 1 had a 1× coverage of 89.2 and above. Likewise, in the metabolism gene set, we found the complete genes to be covered 1× by 97.7% or more in all samples. The 5× coverage was above 50% for sample 1, above 98% for sample 2 and above 80% for sample 3.

By purifying *Giardia* cysts from clinical faecal samples, we demonstrated that it is possible to achieve good coverage of the *Giardia* genome as a whole and for genes of particular interest. We also show that cyst purification by sucrose flotation alone gives variable success, and that further purification with IMS gives better results, even with a very low amount of extracted DNA.

Coverage in sample 2 was good for both SF and IMS as a result of a high number of cysts in the faeces. Sample 3 was from a person with metronidazole refractory giardiasis and

**TABLE 1. Presequencing characteristics and sequence analysis results for three clinical *Giardia* isolates and one cultured reference isolate**

| | Cultured trophozoites (WB) | Patient sample 1 | | Patient sample 2 | | Patient sample 3 | |
|---|---|---|---|---|---|---|---|
| Assemblage | A1 | A2 | | B | | B | |
| Reference genome used, isolate (genome) | WB (AACB) | DH (AHGT) | | GS (ACGJ) | | GS (ACGJ) | |
| Purification method | Centrifugation only | SF | SF + IMS | SF | SF + IMS | SF | SF + IMS |
| Q-bit DNA quant (ng/μL) | 12.3 | < 0.01 | < 0.01 | 1.40 | < 0.01 | 1.41 | < 0.01 |
| DNA amplification cycles | 7 | 10 | 8 | 10 | 10 | 8 | 8 |
| BioAnalyzer DNA quality | OK | Not visible | Not visible | OK | OK | OK | OK |
| Mean fragment size (bp) | 281 | 267 | 262 | 280 | 269 | 288 | 292 |
| Q-bit DNA quant (ng/μL) | 0.026 | 0.154 | 0.141 | 0.059 | 0.259 | 0.145 | 0.211 |
| Total no. of reads (1000×) | 23 563 | 10 612 | 2245 | 21 362 | 16792 | 26 660 | 7009 |
| No. of reads mapped to reference (1000×) | 21 203 | 1170 | 1279 | 4365 | 8387 | 181 | 2147 |
| Reads mapped to reference, % | 90.0 | 11.0 | 57.0 | 20.4 | 49.9 | 0.7 | 30.6 |
| No. of genome bases covered 1× (1000×) | 10 988 | 9784 | 10 281 | 10 485 | 10 553 | 6155 | 10 313 |
| Genome covered to 1× depth (%) | 98.0 | 91.4 | 96.0 | 87.2 | 87.8 | 51.2 | 85.8 |
| Genome covered to 5× depth (%) | 97.3 | 51.5 | 79.1 | 82.7 | 85.0 | 1.9 | 79.3 |
| Genome covered to 10× depth (%) | 97.2 | 14.4 | 36.5 | 75.9 | 82.5 | 0.1 | 63.7 |
| Mean depth per base | 85.7 | 5.1 | 7.4 | 18.4 | 35.1 | 1.1 | 12.8 |
| Min–max reads per base | 0–1401 | 0–95 | 0–128 | 0–2101 | 0–4145 | 0–3539 | 0–1109 |

SF, sucrose gradient flotation; IMS, immunomagnetic separation.

had far fewer cysts. In SF-purified cysts, a variable amount of faecal debris, of similar specific weight to *Giardia* cysts, remains. This probably accounts for the large variability in the number of mapped reads and the resulting low coverage in some SF-purified samples. A sample with several *Giardia* genotypes would pose a problem for the choice of reference genome and the analysis, but it would also be a challenge for conventional PCR, or any sequencing method. The sequencing technique used here produces relatively short reads. Larger structural genome changes therefore cannot be discovered. The amplification steps in the procedure may also induce artefacts. Further optimization of cyst purification and newer sequencing methods with longer read lengths and no amplification steps are likely to improve these results considerably.

The successful genomewide sequencing directly from clinical isolates of *Giardia* opens possibilities for correlations between clinical disease, severity, susceptibility to antimicrobial drugs and genetic composition of the parasite. It is likely that other difficult-to-culture intestinal microorganisms of clinical importance, such as cryptosporidium, could be purified and sequenced in a similar fashion.

## Transparency declaration

This work was supported by the University of Bergen. All authors report no conflicts of interest relevant to this article.

## Acknowledgements

We acknowledge the technical support and service from the Genomics Core Facility at the Department of Clinical Science, University of Bergen. We are indebted to Christian Klotz and Toni Aebisher at the Robert Koch Institute, Berlin, for providing cultured WB trophozoites, and we thank Steinar Sørnes for important laboratory assistance.

## Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.cmi.2014.08.014.

## References

[1] Tian HF, Chen B, Wen JF. Giardiasis, drug resistance, and new target discovery. Infect Disord Drug Targets 2010;10:295–302.
[2] Cotton JA, Beatty JK, Buret AG. Host parasite interactions and pathophysiology in *Giardia* infections. Int J Parasitol 2011;41:925–33.
[3] Benere E, Geurden T, Robertson L, Van Assche T, Cos P, Maes L. Infectivity of *Giardia duodenalis* assemblages A and E for the gerbil and axenisation of duodenal trophozoites. Parasitol Int 2010;59:634–7.
[4] Ignatius R, Gahutu JB, Klotz C, Steininger C, Shyirambere C, Lyng M, et al. High prevalence of *Giardia duodenalis* Assemblage B infection and association with underweight in Rwandan children. PLoS Negl Trop Dis 2012;6(6):e1677.
[5] Lebbad M, Petersson I, Karlsson L, Botero-Kleiven S, Andersson JO, Svenungsson B, et al. Multilocus genotyping of human *Giardia* isolates suggests limited zoonotic transmission and association between assemblage B and flatulence in children. PLoS Negl Trop Dis 2011;5(8):e1262.
[6] Morrison HG, McArthur AG, Gillin FD, Aley SB, Adam RD, Olsen GJ, et al. Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia*. Science 2007;317(5846):1921–6.
[7] Franzen O, Jerlstrom-Hultqvist J, Castro E, Sherwood E, Ankarklev J, Reiner DS, et al. Draft genome sequencing of *Giardia intestinalis* assemblage B isolate GS: is human giardiasis caused by two different species? PLoS Pathog 2009;5(8):e1000560.
[8] Adam RD, Dahlstrom EW, Martens CA, Bruno DP, Barbian KD, Ricklefs SM, et al. Genome sequencing of *Giardia lamblia* genotypes A2 and B isolates (DH and GS) and comparative analysis with the genomes of genotypes A1 and E (WB and Pig). Genome Biol Evol 2013;5(12):2498–511.
[9] Sulaiman IM, Fayer R, Bern C, Gilman RH, Trout JM, Schantz PM, et al. Triosephosphate isomerase gene characterization and potential zoonotic transmission of *Giardia duodenalis*. Emerg Infect Dis 2003;9(11):1444–52.
[10] Palm JE, Weiland ME, Griffiths WJ, Ljungstrom I, Svard SG. Identification of immunoreactive proteins during acute human giardiasis. J Infect Dis 2003;187:1849–59.
[11] Raj D, Ghosh E, Mukherjee AK, Nozaki T, Ganguly S. Differential gene expression in *Giardia lamblia* under oxidative stress: significance in eukaryotic evolution. Gene 2014;535:131–9.